



ISSN: 1646-9895

Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

D e z e m b r o 2 3 • D e c e m b e r 2 3



©RISTI 2023 <http://www.risti.xyz>

Nº 52

Edição / Edición

Nº 52, 12/2023

Tiragem / Tirage: 1000

Preço por número / Precio por número: 17,5€

Subscrição anual / Suscripción anual: 30€ (2 números)

ISSN: 1646-9895

Depósito legal:

Indexação / Indexación

Academic Journals Database, CiteFactor, Dialnet, DOAJ, DOI, EBSCO, GALE, IndexCopernicus, Index of Information Systems Journals, Web of Science, Latindex, ProQuest, QUALIS, SciELO, SCImago, Scopus, SIS, Ulrich's.

Publicação / Publicación

RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação

Pc. 9 de Abril, 26, 4200-422 Porto, Portugal

E-mail: aistic@gmail.com

Web: <http://www.risti.xyz>

Director

Álvaro Rocha, ISEG, Universidade de Lisboa, PT

Coordenadores da Edição / Coordinadores de la Edición

Vitor Santos, NOVA IMS, PT

Conselho Editorial / Consejo Editorial

Abel Méndez, Instituto Tecnológico de Costa Rica, CR

Alejandro Peña, Universidad EAFIT, CO

Alma Gómez-Rodríguez, Universidade de Vigo, ES

Ana Rita Calvão, ESTGA, Universidade de Aveiro, PT

Ania Cravero, Universidad de La Frontera, CL

António Abreu Silva, ISCAP, Instituto Politécnico do Porto, PT

Antonio Garcia, Universidad de Santiago de Compostela, ES

António Godinho, ISEC, Instituto Politécnico de Coimbra, PT

Antonio Jiménez-Martín, Universidad Politécnica de Madrid, ES

Arturo J. Méndez, Universidad de Vigo, ES

August Climent, Ramon LLull University, ES

Beatriz Rodríguez, Universidad de la Republica, UY

Borja Bordel, Universidad Politécnica de Madrid, ES

Brenda L. Flores-Rios, Universidad Autónoma de Baja California, MX

Bruno Miguel Ferreira Gonçalves, Instituto Politécnico de Bragança, PT

Carlos Alexandre Silva, Instituto Federal de Minas Gerais, BR

Carlos Carreto, Instituto Politécnico da Guarda, PT

Carlos Morais, Instituto Politécnico de Bragança, PT

Carlos Rompante Cunha, UNIAG & CeDRI & Instituto Politécnico Bragança, PT

Carlos Vaz de Carvalho, Instituto Politécnico do Porto, PT

Célio Marques, Instituto Politécnico de Tomar, PT

Ciro Martins, Universidade do Algarve, PT

Cristina Caridade, Instituto Politécnico de Coimbra, PT

Daniel Polónia, Universidade de Aveiro, PT

Dante Carrizo, Universidad de Atacama, CL

Edwin Cedeño Herrera, Universidad de Panamá, PA

Fábio Marques, ESTGA, Universidade de Aveiro, PT
Felipe Vasquez, Universidad de La Frontera, CL
Fernando Moreira, REMIT, IJP, Universidade Portucalense &
IEETA, Universidade de Aveiro, PT
Fernando Ribeiro, Instituto Politécnico de Castelo Branco, PT
Fernando Bandeira, PT
Filipe Cardoso, Instituto Politécnico de Viseu, PT
Flor Gomes de María Sánchez Aguirre, Universidad César Vallejo, PE
Francisco Javier Lena-Acebo, Universidad de Cantabria, ES
Gabriel Guerrero-Contreras, Universidad de Cádiz, ES
Gerardo González Filgueira, Universidade da Coruña, ES
Gloria Valencia, Universidad de las Fuerzas Armadas, EC
Guilllhermo Davila, Universidad de Lima, PE
Hélder Gomes, Universidade de Aveiro, PT
Helder Zagalo, Universidade de Aveiro, PT
Hélia Guerra, Universidade dos Açores, PT
Henrique Gil, Instituto Politécnico de Castelo Branco, PT
Henrique S. Mamede, Universidade Aberta, Portugal
Inês Domingues, ISEC, Instituto Politécnico de Coimbra, PT
Isaias Bianchi, Universidade Federal de Santa Catarina, BR
Isidro Navarro, Universidad Politécnica de Cataluña, ES
Ismael Etxeberria-Agiriano, Universidad del País Vasco, ES
Ivan Garcia, Universidad Tecnológica de la Mixteca, MX
Jeimy Cano, Universidad de los Andes, CO
João Emílio Almeida, ISTECS - Porto, PT
João Paulo Ferreira, Instituto Politécnico de Coimbra, PT
Joao Roberto de Toledo Quadros, CEFET/RJ, BR
Joao Tavares, Universidade do Porto, PT
Joaquim Reis, ISCTE - Instituto Universitário de Lisboa, PT
Jorge Eduardo Ibarra Esque, Universidad Autónoma de Baja California, MX
Jorge Hochstetter, Universidad de La Frontera, CL
José Álvarez-García, Universidad de Extremadura, ES
Jose Felipe Cocon Juarez, Universidad Autónoma del Carmen, MX

José Lousado, Instituto Politécnico de Viseu, PT
José Luis Pastrana Brincones, Universidad de Málaga, ES
Jose M. Molina, Universidad Carlos III de Madrid, ES
José Ribeiro, Politécnico de Leiria, PT
Jose Silvestre Silva, Academia Militar, PT
Josep Maria Marco-Simó, Universitat Oberta de Catalunya, ES
Juan Angel Contreras Vas, Universidad de Extremadura, ES
Juan M. Santos-Gago, Universidad de Vigo, ES
Juan Pablo DAmato, Universidad Nacional del Centro de la
Provincia de Buenos Aires, AR
Leila Weitzel, Universidade Federal Fluminense, BR
Leonardo Bermón Angarita, Universidad Nacional de Colombia, CO
Lilia Muñoz, Universidad Tecnológica de Panamá, PA
Lucila Romero, Universidad Nacional del Litoral, AR
Luis Álvarez Sabucedo, Universidad de Vigo, ES
Luis Enrique Sánchez Crespo, Universidad de Castilla-la Mancha, ES
Luisa María Romero-Moreno, Universidad Sevilla, ES
Luz María Hernández Cruz, Universidad Autónoma de Campeche, MX
Luz Sussy Bayona Oré, Universidad Nacional Mayor de San Carlos, PE
Marcelo Marciszack, Universidad Tecnológica Nacional, AR
Marcelo Zambrano Vizueté, Universidad Técnica Del Norte
Marco Painho, NOVA IMS, PT
Margarita Ramirez Ramirez, Universidad Autónoma de Baja California, MX
Maria Cristina Marcelino Bento, UNIFATEA, BR
María de la Cruz del Río-Rama, Universidad de Vigo, ES
Maria de los Milagros Gutierrez, Universidad Tecnológica Nacional, AR
Maria do Rosário Bernardo, Universidade Aberta, PT
Maria Hallo, Escola Politécnica Nacional, EC
Maria João Ferreira, Universidade Portucalense, PT
Maria João Gomes, Universidade do Minho, PT
Maria Sousa, ISCTE - Instituto Universitário de Lisboa, PT
Marisol B. Correia, ESGHT - Universidade do Algarve & CiTUR, PT
Maristela Holanda, Universidade de Brasília, BR

Martin Llamas Nistal, Universidad de Vigo, ES
Miguel Casquilho, Universidade de Lisboa, PT
Miguel Ramón González Castro, Centro Tecnológico Aimen, ES
Mirna Ariadna Muñoz Mata, CIMAT, MX
Nelson Rocha, Universidade de Aveiro, PT
Nuno Melão, Instituto Politécnico de Viseu, PT
Nuno Ribeiro, Universidade Fernando Pessoa, Portugal
Patricia Dias, Universidade do Estado de Minas Gerais, BR
Paula Prata, Universidade da Beira Interior, PT
Paulo Martins, Universidade de Trás-os-Montes e Alto Douro, PT
Paulo Pinto, Universidade Nova de Lisboa, PT
Paulo Rurato, Universidade Fernando Pessoa, Portugal
Pedro Araújo, Universidade da Beira Interior, PT
Pedro R. Palos- Sanchez, Universidad de Sevilla, ES
Pedro Sanz-Angulo, Universidad de Valladolid, ES
Pedro Sobral, Universidade Fernando Pessoa, Portugal
Pedro Sousa, Universidade do Minho, Portugal
René Faruk Garzozzi-Pincay, Universidad Estatal Península de Santa Elena, EC
Ruben Pereira, ISCTE – Instituto Universitário de Lisboa, PT
Rui Pedro Marques, Universidade de Aveiro, PT
Rui Silva Moreira, Universidade Fernando Pessoa, PT
Samuel Sepúlveda, Universidad de La Frontera, CL
Santiago Raul Gonzales Sanches, Universidad Cesar Vallejo, PE
Sara Balderas-Díaz, Universidad de Cádiz, ES
Sergio Araya Guzmán, Universidad del Bío-Bío, CL
Sergio F. Lopes, Universidade do Minho, PT
Sergio Gálvez Rojas, Universidad de Málaga, ES
Solange N Alves-Souza, Universidade de São Paulo, BR
Telmo Silva, DigiMedia, Universidade de Aveiro, PT
Teresa Guarda, Universidad Estatal Península de Santa Elena, EC
Thiago Dias, Centro Federal de Educação Tecnológica de Minas Gerais, BR
Vitor Carvalho, Instituto Politécnico do Cávado e do Ave, PT
Vitor Santos, Universidade Nova de Lisboa - NOVA IMS, PT

Wagner Tanaka Botelho, Universidade Federal do ABC, BR
Yamid Hernández Julio, Universidad del Sinú, CO

Índice / Index

EDITORIAL

Quinze anos de RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação: cinquenta números regulares e sessenta números especiais 1

Alvaro Rocha, Maria José Sousa

ARTIGOS / ARTICULOS / ARTICLES

Segmentación de tejido cerebral mediante redes neuronales convolucionales basadas en U-Net5

David Jiménez-Murillo, Stiven Alexander Henao-Garcés, Andrés Eduardo Castro-Ospina, Leonardo Duque-Muñoz, Daniel Nieto-Mora

Protótipo de Solução para Detetar e Sinalizar Defeitos em Pavimentos Rodoviários Baseado em Técnicas de Visão Computacional25

Miguel Gonçalves, Tomás Marques, Pedro D. Gaspar, Vasco N. G. J. Soares, João M. L. P. Caldeira

A relação entre a inovação tecnológica e o desempenho nos meios de hospedagem no contexto da pandemia da Covid-1945

Mariana Pires Lopes Vidal, Patrícia Oliveira de Freitas, Sandra Martins Lohn Vargas

Web scraping: Uso de plataformas de extracción de datos aplicadas a un sitio web sobre profesiones en México.....61

Felipe Cocón, Dámaris Pérez-Cruz, José Ángel Pérez-Rejón, Patricia Zavaleta-Carrillo, Ulises Barradas-Arenas, Rubí Gómez-Ramón, José Alonso Pérez Cruz

Perspectivas latinoamericanas del uso de las TIC en estudiantado universitario74

María Guadalupe Veytia Bucheli, Armando Sánchez Macías, Maritza Librada Cáceres Mesa, José Frank Vázquez Horta

Impacto de ChatGPT en la enseñanza: Un enfoque de aula invertida para fundamentos de programación97

Michael Apaza Calsin, Marco Aedo, Eveling Castro

As literacias mediática e digital e a participação cívica como tecnologias para dinamizar os territórios do interior de Portugal 113

Paulo Bruno Alves, Lídia Oliveira

Análisis bibliométrico de la educación virtual y rendimiento académico universitario periodo 1996-2021, utilizando datos Scopus.....126

Ignacio Romero Cruz

Impacto da Engenharia de Sistemas de Informação na Sociedade

Impact of Information Systems Engineering on Society

Vitor Santos ¹

vsantos@novaims.unl.pt

¹NOVA Information Management School (NOVA IMS), Campus de Campolide, 1070-312 Lisboa, Portugal.

DOI: 10.17013/risti.52.1-4

1. Introdução

A presente edição da RISTI é dedicada ao Impacto da Engenharia de Sistemas de Informação na Sociedade. É um tema abrangente e multidisciplinar, incluindo, entre outros aspetos, abordagens teóricas e metodológicas, sistemas de informação na saúde, na gestão das cidades e território e práticas de sistemas de informação específicas de educação. Procurou fazer-se um levantamento de ideias inovadoras, soluções desenvolvidas, trabalhos empíricos e estudos de caso neste domínio. Os oito artigos publicados no número 52 da RISTI enquadram-se neste pressuposto.

A seleção de artigos que se apresenta é o resultado de um exigente processo de avaliação das 91 propostas originalmente submetidas, provenientes de 14 países e de 3 continentes. Cada artigo foi avaliado por, pelo menos, três membros da Comissão Científica, resultando numa taxa de aceitação final de 8,8%. A qualidade evidenciada pelos oito artigos publicados é a face visível da exigência desse processo.

O enorme interesse que a comunidade académica demonstrou ter, desde o início, nesta publicação e a qualidade dos artigos científicos publicados constitui uma prova da sua relevância, um indicador de vitalidade da comunidade e uma indicação de que a RISTI continua a desempenhar um papel importante num dos maiores desafios que se coloca ao mundo: o desafio da inovação.

2. Estrutura

Este número da RISTI apresenta os seguintes artigos:

O primeiro artigo tem como objetivo analisar comparativamente o desempenho de três métodos de segmentação de tecidos em imagens produzidas por Ressonância Magnética (RM). Assim, é feita uma comparação entre uma rede neural convolucional baseada na arquitetura U-Net e dois outros métodos de segmentação de RM, Dipy e FSL. A comparação utiliza três conjuntos de dados publicamente disponíveis, MRBrainS13, MRBrainS18 e IBSR, que contêm volumes de RM com tecidos cerebrais segmentados manualmente. Os resultados demonstram que os métodos baseados em redes neurais convolucionais permitem uma melhor segmentação do tecido cerebral.

No segundo artigo os autores, apresentam um protótipo para avaliar e validar a utilização de técnicas de visão computacional na identificação de defeitos em passeios rodoviários no contexto de uma cidade inteligente. Inclui um estudo de avaliação de desempenho de três redes neurais convolucionais, YoloV4-Tiny, SSD MobileNet e RetinaNet, realizado neste contexto. Com base nos resultados do estudo, é descrita a arquitetura e o processo de implementação do protótipo. Salienta-se ainda que, comparativamente ao método atualmente utilizado pela Infraestruturas de Portugal para identificação de defeitos em calçada, esta abordagem promete ser mais ágil, eficaz e eficiente.

O terceiro artigo tem por objetivo discutir a relação entre inovação tecnológica e desempenho financeiro da hotelaria em regiões turísticas do Estado do Rio de Janeiro no contexto da pandemia da Covid-19. Descreve um interessante estudo quantitativo que revelou que os recursos tecnológicos podem influenciar positivamente o desempenho financeiro dos hotéis, fornecendo informações para que os gestores possam tomar melhores decisões.

O quarto artigo tem por foco a recolha de dados da Web. É apresentada uma análise exaustiva das principais ferramentas de recolha de dados da Web disponíveis no mercado e compara as suas características e funcionalidades. Adicionalmente, é selecionada uma ferramenta específica para demonstrar a sua utilização na obtenção de dados sobre as percentagens de licenciados em várias carreiras no México, bem como a distribuição relacionada com o sexo e os salários em vários estados do país.

O tema “Utilização das Tecnologias de Informação e Comunicação (TIC) na educação” é endereçado no quinto artigo. O artigo apresenta um estudo onde é descrito como os estudantes universitários de quatro países da América Latina (México, Guatemala, Equador e Bolívia) percebem a utilização das TIC na educação. Os resultados refletem diferenças estatisticamente significativas na utilização das TIC na educação, na utilização de programas e na literacia digital.

O sexto artigo aborda o Impacto do ChatGPT no ensino. Descreve uma experiência na adoção de uma abordagem de sala de aula invertida para o ensino de tópicos básicos de programação aplicada a estudantes universitários do primeiro ano, analisando as capacidades, potencial e limitações apresentadas por estas ferramentas de IA

As ações de promoção das literacias mediática e digital e da participação cívica dos habitantes nos territórios do interior de Portugal são o tema do sétimo artigo. Estes elementos dinamizadores do território são apresentados num quadro interpretativo dos territórios do interior tendo como cenário 21 Comunidades Intermunicipais, com muitos municípios e freguesias de baixa densidade.

Finalmente, no oitavo artigo apresenta-se uma análise descritiva da produção científica da educação virtual e do desempenho académico universitário utilizando uma metodologia bibliométrica. Os resultados indicam um crescimento exponencial de artigos, concluindo-se, ainda, que a educação virtual está associada aos avanços tecnológicos e que tende a suscitar interesse crescente por parte dos investigadores.

3. Agradecimentos

Gostaria de expressar o nosso agradecimento a todos os autores que submeteram o seu trabalho científico para ser avaliado para publicação nesta edição da RISTI; aos membros da Comissão Científica, pela criteriosa avaliação que fizeram dos artigos submetidos e às associações científicas AISTI e ITMA, promotoras da RISTI, assim como às Bases de Dados de Revistas Académicas como Academic Journals Database, Dialnet, DOAJ, DOI, EBSCO, GALE, Google Scholar, IndexCopernicus, Information Systems Journal, Latindex, ProQuest, QUALIS, SciELO, SCImago, Scopus, SIS, Ulrich's e Web of Science, entidades que contribuem para que a RISTI seja uma revista científica de referência internacional.

Segmentación de tejido cerebral mediante redes neuronales convolucionales basadas en U-Net

David Jiménez-Murillo¹, Stiven Alexander Henao-Garcés¹, Andrés Eduardo Castro-Ospina¹, Leonardo Duque-Muñoz¹, Daniel Nieto-Mora¹

davidjimenez215489@correo.itm.edu.co; stivenhenao236528@correo.itm.edu.co;
andrescastro@itm.edu.co; leonardoduque@itm.edu.co; danielnieto@itm.edu.co

¹ Grupo de investigación Máquinas Inteligentes y Reconocimiento de Patrones, Instituto Tecnológico Metropolitano, 050013, Medellín, Colombia.

DOI: 10.17013/risti.52.5-24

Resumen: La Resonancia Magnética (MRI, por sus siglas en inglés) es un método no invasivo que produce imágenes anatómicas volumétricas de alta resolución. Este método se utiliza comúnmente para diagnosticar diversas patologías cerebrales, como enfermedades desmielinizantes, demencia, enfermedades cerebrovasculares, epilepsia, entre otras. Sin embargo, la segmentación manual de tejidos en estas imágenes es una tarea compleja que requiere una cantidad significativa de tiempo y esfuerzo por parte de especialistas. Para abordar esto, se han propuesto diversas estrategias en la literatura que implican menos participación de expertos, tanto automáticas como semiautomáticas. En este estudio, se realiza una comparación entre una red neuronal convolucional basada en la arquitectura U-Net y otros dos métodos de segmentación de MRI, Dipy y FSL. La comparación utiliza tres conjuntos de datos disponibles públicamente, MRBrainS13, MRBrainS18 e IBSR, que contienen volúmenes de MRI con tejidos cerebrales segmentados manualmente. El rendimiento de cada método se evalúa usando el coeficiente de Dice, el índice de Jaccard, el área bajo la curva y la similitud estructural. Los resultados demuestran que los métodos basados en redes neuronales convolucionales logran mejorar la segmentación de tejido cerebral.

Palabras-clave: Imágenes de resonancia magnética, Redes neuronales convolucionales, segmentación de imágenes médicas, segmentación de tejido cerebral

Brain Tissue Segmentation Using U-Net-based Convolutional Neural Networks

Abstract: Magnetic Resonance Imaging (MRI) is a non-invasive method that produces volumetric, high-resolution anatomy images. This method is commonly used for diagnosing various brain pathologies, such as demyelinating diseases, dementia, cerebrovascular diseases, and epilepsy, among others. However, manually segmenting tissues in these images is a complex task that requires a significant amount of time and effort from specialists. To address this, various strategies have

been proposed in the literature that involve less expert involvement, both automatic and semi-automatic. In this study, a comparison is made between a convolutional neural network based on U-Net architecture and two other MRI segmentation methods, Dipy and FSL. The comparison uses three publicly available datasets, MRBrainS13, MRBrainS18, and IBSR, which contain MRI volumes with manually segmented brain tissues. The performance of each method is evaluated using four measures of similarity: The Dice coefficient, the Jaccard index, the area under the curve, and structural similarity. The results demonstrate that the methods based on convolutional neural networks achieve segmentations that are more similar to those performed manually by an expert compared to the other methods.

Keywords: Brain segmentation; convolutional neural network; deconvolution; medical image segmentation; magnetic resonance imaging.

1. Introducción

Las malformaciones en la estructura cerebral se refieren a alteraciones en la forma, tamaño o posición de la estructura cerebral, lo cual puede afectar la función neurológica de una persona y su salud en general (Aronica et al., 2012). Estas alteraciones pueden ser desarrolladas en una etapa embrionaria (malformaciones congénitas) o adquiridas a lo largo de la vida. Los cambios estructurales congénitos se deben usualmente a una disrupción en el desarrollo normal de la etapa embrionaria o fetal (Sarma & Pruthi, 2023), mientras que las malformaciones adquiridas corresponden usualmente a enfermedades, lesiones, circunstancias ambientales u otras causas externas (Peruzzo et al., 2016). En general, este tipo de condiciones conlleva a enfermedades como epilepsia refractaria causada por displasias corticales, consideradas como la causa principal de esta enfermedad (Sánchez Fernández et al., 2021), Alzheimer debido a contracciones en regiones del cerebro (Ito et al., 2019), esclerosis múltiple por pérdida de volumen cerebral (Clèrigues et al., 2023), microangiopatía causada por los espacios de Virchow-Robin (Rashid et al., 2022) y en el caso de malformaciones como las de Chianni, pueden aparecer múltiples síntomas como desequilibrio, tinnitus, dificultad para tragar, palpitaciones, apnea del sueño, debilidad muscular, fatiga crónica y tinciones dolorosas (Kumar et al., 2011).

El estudio de la estructura cerebral ha sido impulsado por el análisis de imágenes médicas. Específicamente, se han utilizados neuroimágenes para realizar estudios cuantitativos acerca de la estructura y funcionalidad del sistema nervioso central. Las imágenes por resonancia magnética (MRI), son uno de los tipos de neuroimágenes ampliamente utilizadas, estableciéndose como una herramienta valiosa para el diagnóstico de enfermedades neurológicas, y una gran ayuda al momento de evaluar la morfología cerebral (Castillo-Carranza et al., 2022; Hyman, 2003), permitiendo tanto la evaluación de la patología asociada a síntomas clínicos o cognitivos, y un diagnóstico diferencial (Zubrikhina et al., 2023). Mediante MRI es posible diferenciar estructuras del tejido cerebral como la materia blanca, materia gris y líquido fluido-espinal en tareas como clasificación, identificación y segmentación (Clèrigues et al., 2023; Jiang et al., 2023). La línea base para estas tareas se da de forma manual bajo el criterio y experticia de médicos especializados como radiólogos al ejecutar una inspección visual de las imágenes. Sin embargo, realizar este proceso de forma manual supone retos

en cuanto al manejo de grandes bases de datos, sesgos y equivocaciones asociadas a la diferencia de contrastes entre imágenes que suponen un reto incluso para los expertos dada la falta de criterios consistentes (Ito et al., 2019). En los últimos años, la segmentación de MRI cerebral ha sido conducido por técnicas de aprendizaje profundo. Este método consiste en la implementación de redes neuronales artificiales con gran volumen de capas y neuronas, las cuales logran extraer una jerarquía de características de imágenes de entrada sin formato. Algunos de los algoritmos de aprendizaje profundo conocidos son codificadores automáticos apilados, máquinas de Boltzmann profundas, redes neuronales profundas y redes neuronales convolucionales, siendo las últimas las más comúnmente aplicadas a la segmentación y clasificación de imágenes (Akkus et al., 2017; Toscano et al., 2023).

Se ha encontrado en el estado del arte un amplio rango de aplicaciones de algoritmos de aprendizaje profundo en la segmentación de estructuras en MRI. En (de Brebisson & Montana, 2015), propusieron una arquitectura de red neuronal profunda para segmentación automática, donde obtuvieron resultados promisorios de validación entrenando grandes redes, compuestas de decenas de millones de parámetros, con una cantidad relativamente pequeña de datos. En (Somasundaram & Kalaividya, 2016), emplearon un algoritmo de deconvolución Richardson-Lucy (RL) para mejorar la detección de límites y la calidad de imagen, tomando funciones de difusión de puntos tipo Gaussiano y extraer porciones del cerebro de imágenes de resonancia magnética coronal T1-W; también utilizaron los procesos de intensidad de umbralización, componente conectado y operaciones morfológicas para extracción cerebral. Este proceso ayudó a identificar el límite cerebral mejor que otros métodos. Para (Chang & Hsieh, 2017), surgía un reto más grande; separar el tejido cerebral de tejido no cerebral, ya que representaba un problema mayor al realizar la segmentación del cerebro, por ello, presentaron un algoritmo de extracción de cráneo basado en la combinación de análisis de características de textura de imagen y morfología matemática. Este algoritmo obtenía dos mapas de características de textura, donde uno correspondía a una máscara cerebrales y el otro a una máscara no cerebral. Posteriormente se realiza una serie de operaciones morfológicas para extraer el cerebro. En (Avants et al., 2011), presentaron Atropos, un algoritmo de segmentación de código abierto multivariado de n clases basado en Insight Toolkit (ITK). Para el desarrollo de este método, emplearon algoritmos de maximización de la esperanza (EM) con el modelado de las intensidades de clases, basadas en mezclas finitas paramétricas o no paramétricas. También, incorporaron mapas espaciales de probabilidad previa, mapas de etiquetas anteriores o modelado de campo aleatorio de Markov (MRF). En (Billot et al., 2020), presentaron un método de aprendizaje profundo denominado “SynthSeg”, donde aprovechaban un conjunto de segmentaciones de entrenamiento, para generar escaneos sintéticos de contrastes muy variables sobre la marcha durante el entrenamiento. Estos escaneos se producían utilizando el modelo generativo del marco de segmentación bayesiano clásico, con parámetros muestreados aleatoriamente para apariencia, deformación, ruido y campo de sesgo, y como resultado, obtuvieron una segmentación exitosa de cada contraste en los datos, con un rendimiento ligeramente mejor que la segmentación bayesiana clásica y tres órdenes de magnitud más rápido. Finalmente, en (Chen & Merhof, 2018), presentaron “MixNet”, una red neuronal convolucional profunda 2D semántica para segmentar la estructura del cerebro en imágenes de resonancia magnética multimodal,

reemplazando la capa convolucional tradicional por una capa convolucional dilatada, evitando el uso de capas agrupadas y capas deconvolucionales, reduciendo el número de parámetros de la red. Al final obtuvieron un coeficiente de Dice general de 84.7 % para materia gris, 87.3 % para sustancia blanca y 83.4 % para líquido cefalorraquídeo, con solo 4 sujetos como entrenamiento.

En el presente trabajo se presenta una red neuronal convolucional (CNN), basada en la arquitectura U-Net para la segmentación de imágenes y volúmenes de MRI de cerebro, obteniendo como resultados imágenes segmentadas con etiquetas correspondientes a: fondo, materia gris, materia blanca y fluido cerebroespinal. Posteriormente se comparan mediante cuatro métricas de desempeño los resultados obtenidos respecto a otros métodos de segmentación de MRI usados como Dipy y el método FAST de la librería FMRIB Software Library (FSL). Los resultados muestran que la segmentación de imágenes MRI empleando redes convolucionales, específicamente las redes U-Net 2D y U-Net 3D, son más efectivas que los demás métodos implementados, ofreciendo un mejor desempeño para segmentar las estructuras cerebrales sin necesidad de recurrir a módulos o software externos.

2. Metodología

La Figura 1 presenta el esquema metodológico para el análisis de MRI, el cual comprende los siguientes pasos. La realización de los estudios clínicos para la adquisición de las imágenes, la conformación de un banco de datos sobre condiciones específicas a tratar, es decir, las enfermedades o deformidades existentes en los pacientes, y los análisis realizados por el equipo médico experto que determina las condiciones clínicas que luego son utilizadas como etiquetas para entrenar los modelos. En este trabajo, se propone en la etapa del aprendizaje de máquina el uso de modelos de aprendizaje profundo, específicamente la red U-Net que se ha mencionado previamente. Con el entrenamiento de los modelos, luego se procede a validar con el conjunto de datos de prueba. El objetivo de este estudio es la segmentación precisa del área focal (materia gris y materia blanca), evitando resaltar otras zonas. Finalmente, en el esquema general se contempla la evaluación del modelo, en este caso se utilizaron métricas que determinan la similitud entre la segmentación del modelo y la segmentación dada por el experto.

2.1. Métodos de adquisición de imágenes médicas

Existen diversas tecnologías de adquisición de imagen médica, desde la obtención de imágenes empleando rayos X hasta llegar a los modernos dispositivos multimodales. Dentro de estas tecnologías se pueden encontrar seis de ellas como las más reconocidas a nivel general, empezando por las de ultrasonido, radiografías, tomografías, MRI, Gammacámara y finalmente la fluoroscopia.

La MRI se destaca al ser una herramienta muy potente para el diagnóstico anatómico debido a su alta resolución espacial y a que no es invasiva; lo cual la hace un examen por excelencia para el estudio de imágenes médicas. La MRI permite generar imágenes anatómicas tridimensionales en escala de grises, lo que es una gran ventaja a la hora de realizar procesamiento de estas (Hyman, 2003).

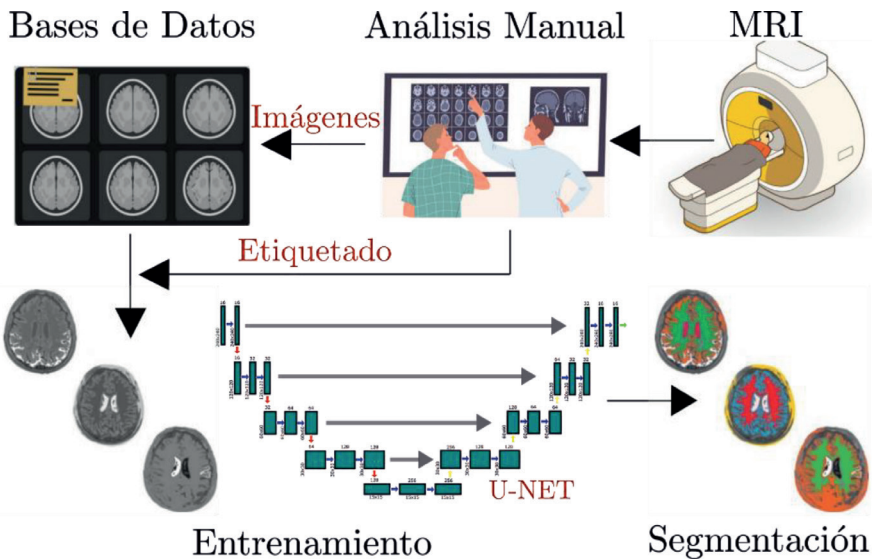


Figura 1 – Estructura metodológica propuesta para la segmentación de imágenes de resonancia magnética basado en aprendizaje profundo.

2.2. Red neuronal convolucional (CNN)

Las redes neuronales convolucionales (CNNs, por sus siglas en inglés) son un tipo de redes neuronales artificiales utilizadas frecuentemente para el análisis y reconocimiento de imágenes (Anwar et al., 2018; Ib et al., 2022; Liu Qing and Zhang, 2017). Estas redes son altamente efectivas en tareas de visión artificial, tales como clasificación y segmentación de imágenes, donde a diferencia de las redes neuronales convencionales que utilizan operaciones lineales entre las capas para obtener representaciones de las entradas, las CNN utilizan capas convolucionales, y a medida que la entrada pasa por cada una de las capas, estas van aplicando filtros a la imagen que logran abstraer información no lineal de los píxeles analizados, generando características discriminatorias.

Las redes neuronales convencionales son totalmente conectadas o fully connected (FC), es decir, todas las neuronas de una capa oculta están conectadas a todas las neuronas de las capas que le preceden y suceden (Alzubaidi et al., 2021). Este proceso genera una cantidad enorme de parámetros (también conocidos como los pesos de la red), debido a que cada capa que tenga la red va a tener conectada cada neurona entre sí, y a su vez, estas neuronas de cada capa se conectarán con cada píxel de la imagen, lo que supondrá un problema cuando se requiera entrenar una red con un volumen considerable de imágenes, ocasionando una inviabilidad por la enorme cantidad de pesos que se obtendrían.

Las redes neuronales convolucionales surgieron como una alternativa de solución a lo anterior. Estas redes cuentan con capas convolucionales y de agrupación que se alternan y finalizan con unas capas FC. Una capa convolucional consiste en filtros o kernels de tamaño $f \times f \times c$ cada uno, donde c hace referencia a la profundidad o el número de

canales (3 para imágenes a color, 1 para imágenes en grises). Adicionalmente se suele introducir una capa de agregación o pooling a la salida de una convolución que reduce el tamaño de las imágenes realizando la correspondiente operación de agrupación, comúnmente máximo o media (“MaxPooling” o “AvgPooling”) (Alzubaidi et al., 2021).

Dentro de las redes neuronales convolucionales, se han logrado desarrollar múltiples arquitecturas de redes CNN que han obtenido resultados destacados en diferentes áreas. Algunas de ellas son la LeNet (LeCun et al., 1998), AlexNet (Krizhevsky et al., 2012), VGGNet (Simonyan & Zisserman, 2014), U-Net (Ronneberger et al., 2015), GoogLeNet (Szegedy et al., 2015), entre otras. De estas, se empleará la arquitectura U-Net (Siddique et al., 2021) para desarrollar el algoritmo de segmentación de imágenes y volúmenes médicos.

2.3. Arquitectura U-Net

La red U-Net es una red convolucional para segmentación rápida y precisa de imágenes médicas. Existen dos arquitecturas U-Net; la primera arquitectura es la red U-Net 2D, compuesta de una ruta de contracción que sigue la arquitectura típica de una red convolucional; convoluciones 3×3 , unidades lineales rectificadas (ReLU), operaciones MaxPooling 2D y una ruta de expansión que consiste en un muestreo ascendente del mapa de características, seguido de convoluciones 2×2 (“convolución ascendente”) que reduce a la mitad el número de canales y convoluciones 3×3 , cada una seguida de una ReLU. En total, la red tiene 23 capas convolucionales (Ronneberger et al., 2015). La segunda arquitectura es la red U-Net 3D, en esta se usará el volumen completo (o subvolúmenes) y se usarán capas convolucionales, y operaciones de agrupación máxima tridimensionales (MaxPooling 3D), tanto para la ruta de contracción como para la ruta de expansión (Çiçek et al., 2016; Ronneberger et al., 2015).

2.4. Otros métodos de segmentación de MRI

Dipy: es una librería externa pero compatible con Python, por ende, es posible instalarla y utilizarla como una herramienta para el análisis de imágenes de resonancia magnética. Esta librería realiza la segmentación de 3 tejidos principalmente mediante la formulación bayesiana. El modelo de observación (término de probabilidad) se define como una distribución gaussiana y se usa un campo aleatorio de Markov (MRF) para modelar la probabilidad a priori de patrones dependientes del contexto de diferentes tipos de tejido del cerebro. La maximización de expectativas y los modos condicionales iterados se utilizan para encontrar la solución óptima (Zhang et al., 2001).

FSL: es una librería de herramientas para análisis de imágenes fMRI, MRI y DTI del cerebro (Jenkinson et al., 2012; Woolrich et al., 2009). Se puede acceder a sus funciones tanto por interfaz gráfica (GUI) como por línea de comandos. La librería contiene un algoritmo para extracción del cerebro a partir de imágenes de resonancia, denominado BET (Herramienta de Extracción del Cerebro, por sus siglas en inglés), el método “usa un modelo deformable que evoluciona para ajustarse a la superficie cerebral por medio de la aplicación de un conjunto de modelos locales adaptativos” (Smith, 2002). El algoritmo FAST (Herramienta de segmentación automatizada del grupo de análisis

FMRIB de la Universidad de Oxford, por sus siglas en inglés) de esta librería se utiliza para segmentar los diferentes tipos de tejido en imágenes 3D del cerebro utilizando un método basado en campos aleatorios de Markov y un algoritmo de maximización de esperanza asociado (Zhang et al., 2000).

2.5. Medidas de desempeño

1. Coeficiente de Dice: también conocido como índice de similitud de Sørensen se puede utilizar en el ámbito de la recuperación de información y en otras áreas. El coeficiente de dice mide la similitud entre conjuntos (Dice, 1945), de la forma indicada la ecuación (1).

$$DC = \frac{2(A \cap B)}{|A| + |B|} \quad (1)$$

donde A es un conjunto que representa la verdad fundamental y B representa la segmentación calculada. Ambas imágenes (conjuntos) son binario con valores 0 o 1 en cada uno de sus vóxeles o píxeles en el caso 2D (Shamir et al., 2019).

2. Índice de Jaccard: el índice de Jaccard mide la similitud en términos de la relación entre la intersección y la unión de dos conjuntos de muestras (Jaccard, 1912), de la forma indicada en la ecuación (2).

$$J = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

3. Area Bajo la Curva (AUC) de la curva ROC: La métrica AUC es una medida útil y ampliamente utilizada para evaluar el rendimiento de modelos de clasificación binaria y de clases múltiples. Sin embargo, no tiene en cuenta la salida numérica exacta de los modelos, sino que analiza cómo la salida clasifica los casos (Van Calster et al., 2008) permitiendo representar en un único valor el rendimiento del clasificador.

$$AUC = \frac{1 + TPR - FPR}{2} \quad (3)$$

donde True Positive Rate (TPR) representa las predicciones positivas correctas entre el número total de positivos y False Positive Rate (FPR) representa las predicciones positivas incorrectas entre el número total de negativos.

4. Similitud Estructural (SSIM): El SSIM es una medida de fidelidad de imagen que ha demostrado ser muy eficaz para medir la fidelidad de las señales. El enfoque SSIM fue motivado originalmente por la observación de que las imágenes naturales tienen señales altamente estructuradas con fuertes dependencias de vecindad. Estas dependencias llevan información útil sobre las estructuras de los objetos en la escena visual. El SSIM mide las distorsiones como una combinación de tres factores: pérdida de correlación, distorsión de la luminancia y distorsión del contraste (Ndajah et al., 2010).

3. Marco Experimental

3.1. Bases de datos

Se realizó una búsqueda de múltiples bases de datos de cerebro en diferentes repositorios, páginas web y desafíos relacionados con la segmentación de MRI. Principalmente, se examinaban bases de datos públicas que tuviesen gran cantidad de imágenes y heterogeneidad etaria, ya que esto garantizaría mejores resultados en el entrenamiento de la red. Finalmente, se escogieron tres bases de datos de segmentación de MRI, que se presentan a continuación:

1. La primera base de datos se extrajo del desafío (A. Mendrik, 2013; A. M. Mendrik et al., 2015), la cual contiene veinte escáneres cerebrales de resonancia magnética 3T de pacientes con edades mayores a los 50 años, con diabetes y controles pareados de diversos grados de atrofia y lesiones de la sustancia blanca. Esta base de datos se divide en cinco conjuntos de datos (escaneos MRI con segmentaciones manuales) para entrenamiento y quince conjuntos de datos (solo escaneos MRI) para pruebas. En nuestro caso son de interés solamente los cinco conjuntos de datos para entrenamiento, específicamente las imágenes ponderadas en T1 de cada conjunto de datos, ya que estos contienen su propia segmentación manual realizada por un experto. Las segmentaciones manuales están constituidas por ocho etiquetas desde la 1 hasta la 8, que corresponden a un tejido diferente del cerebro, pero solo fueron de interés las etiquetas 1, 3 y 5, correspondientes a la materia gris (GM), materia blanca (WM) y fluido cerebroespinal (CSF), respectivamente. Las etiquetas adicionales se tomaron como fondo en cada una de las pruebas.

Para cada paciente, se proporcionan las siguientes secuencias:

- T1_1mm: escaneo 3D ponderado en T1 (tamaño de vóxel: 1.0mm × 1.0mm × 1.0mm)
 - T1: escaneo 3D ponderado en T1 registrado en el T2 FLAIR (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
 - T1_IR: exploración de recuperación de inversión ponderada en T1 multicorte registrada en el T2 FLAIR (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
 - T2_FLAIR: escaneo FLAIR multicorte (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
2. MRBrainS18: La segunda base de datos se extrajo del desafío (Kuijf & Bennink, 2019), la cual contiene escáneres de 30 pacientes con edades mayores a los 50 años con diabetes, demencia, Alzheimer y controles pareados (con mayor riesgo cardiovascular) de diversos grados de atrofia y lesiones de la sustancia blanca. Estos datos consisten en 7 conjuntos de imágenes de MRI cerebral. Para nuestro estudio, son de interés las imágenes ponderadas en T1 de cada conjunto de datos, ya que estos contienen su propia segmentación manual realizada por un experto. Las segmentaciones manuales están constituidas por once etiquetas desde la 0 hasta la 10, que corresponden a un tejido diferente del cerebro, pero solo fueron de interés las etiquetas 1, 3 y 5, correspondientes a GM, WM y CSF,

respectivamente. Las etiquetas adicionales se tomaron como fondo en cada una de las pruebas.

Para cada paciente, se proporcionan las siguientes secuencias:

- T1: Secuencia 3D ponderada en T1 (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
 - T1-IR: Secuencia de recuperación de inversión ponderada en T1 de múltiples cortes (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
 - T2-FLAIR: Secuencia FLAIR T2 multicorte (tamaño de vóxel: 0.958 mm × 0.958 mm × 3.0 mm)
3. Repositorio de Segmentación Cerebral de Internet (IBSR): Esta es una base de datos de MRI del cerebro que incluye segmentaciones guiadas manualmente por un experto con el fin de desarrollar y evaluar métodos de segmentación. Los volúmenes fueron puestos en el dominio público por Rohlfing (Rohlfing, 2012) y pueden obtenerse en <http://www.nitrc.org/>.

La base de datos incluye estudios de 18 pacientes de los cuales 4 son mujeres y 14 hombres con edades que oscilan entre 7 y 71 años. Por cada paciente se encuentran MRI de la secuencia T1 con y sin cráneo, máscara del cerebro y segmentaciones de las estructuras cerebrales. Los volúmenes tienen una resolución de 256 × 256 × 128 px y un tamaño de vóxel de 0.938 × 0.938 × 1.5 mm.

En la Figura 2 se muestra un ejemplo de cada uno de los volúmenes T1 de las bases de datos seleccionadas, en la cual se muestran también las segmentaciones manuales de los tejidos cerebrales de nuestro interés: Materia Gris (GM) y Materia Blanca (WM).

Los tres conjuntos de datos fueron estandarizados a un espacio común de dimensión 240 × 240 × 48, buscando evitar el filtrado de información que afecte la validez de los resultados obtenidos por los métodos propuestos.

3.2. Métodos

Utilizando las bases de datos descritas anteriormente, se implementó el método de segmentación propuesto, para el cual se presentaron dos variantes, una basada en representaciones bidimensionales de las imágenes (U-Net 2D) y la otra que utiliza imágenes en 3D. Estos métodos fueron comparados con dos metodologías establecidas en el estado del arte con el fin de evidenciar el progreso en la tarea de segmentación.

1. Segmentación empleando redes convolucionales: El primer método consiste en la creación de una red convolucional basado en la arquitectura U-Net, para segmentar imágenes de resonancia magnética en 2 y 3 dimensiones; en ambos casos se usaron para entrenamiento cuatro de las cinco MRI con segmentaciones manuales que provee el conjunto de datos MRBrainS13, la MRI restante se usará para realizar evaluación de las redes entrenadas y las técnicas de comparación. Para la U-Net 2D, en la etapa de contracción se tienen como entrada, imágenes con dimensiones de 240 × 240 × 1, haciendo referencia a que son imágenes 2D en escala de grises, seguidas de dos convoluciones de 16 filtros cada una

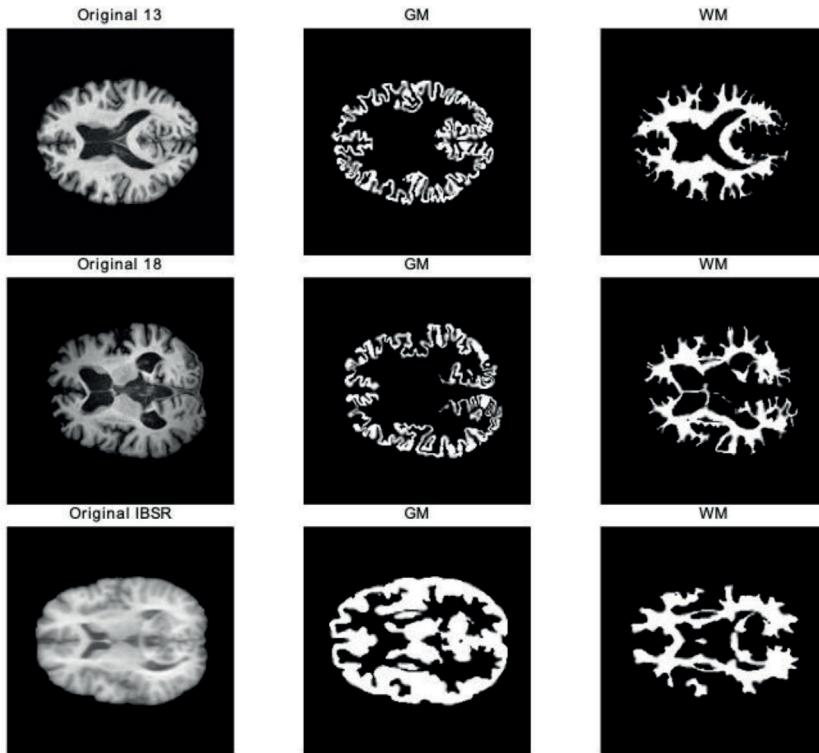


Figura 2 – Vista de ejemplo de un cerebro por cada base de datos y sus correspondientes segmentaciones manuales de materia gris y materia blanca

para empezar a extraer características de la imagen. Finalmente, se aplica un MaxPooling para reducir las características similares y a su vez, reducir a la mitad el tamaño de la imagen resultante.

Luego, se aplica este mismo procedimiento de dos convoluciones y un MaxPooling para tres niveles distintos de la red, pero la variación está en que, a cada nivel siguiente, se le aplica el doble de filtros del nivel anterior en cada convolución, es decir, en el primer nivel se aplicaron 16 filtros a la imagen, luego, en el siguiente nivel se aplicarán ya no 16 sino 32 filtros, y así sucesivamente para los demás niveles hasta llegar al fondo de la red, donde se realizan 2 convoluciones de 256 filtros sin el MaxPooling para extraer características más precisas.

Después de esta última etapa de contracción, se inicia la etapa de expansión, la cual tiene como objetivo regresar la imagen a su tamaño original. Acá se aplica el proceso inverso a la convolución, es decir la convolución transpuesta, que básicamente realiza un relleno en la imagen original seguido de unas capas convoluciones, aplicando los mismos filtros de la etapa de contracción, pero de forma descendente. A su vez, después de aplicar la convolución transpuesta, a la imagen también se le concatena con la imagen correspondiente de la ruta de contracción, con el objetivo de utilizar las características extraídas allí. Este

proceso mencionado anteriormente, se ejecuta para 3 niveles más de la etapa de expansión hasta el último nivel de la red, donde se realiza una última capa de convolución con un filtro de tamaño 1×1 para satisfacer los requisitos de predicción. El esquema completo de la red se puede apreciar en la Figura 3.

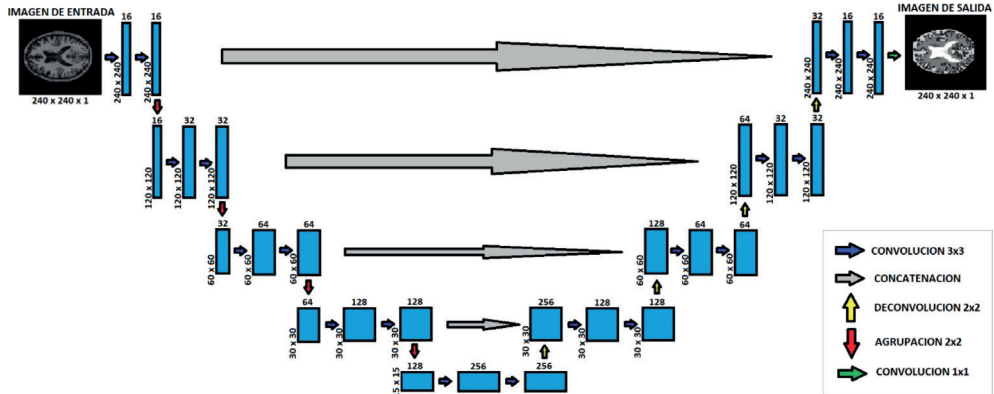


Figura 3 – Arquitectura de la red U-Net 2D

Luego de construir la red U-Net 2D, se realiza un pre-procesamiento de las imágenes de la base de datos, con la finalidad de tener imágenes estándar, con el tipo de dato correspondiente y las dimensiones correctas para proceder a realizar su entrenamiento. Para la red U-Net 3D se usaron subvolumenes de tamaño $80 \times 80 \times 16$, con el fin de aumentar sintéticamente el número total de datos de entrenamiento, reducir el costo computacional de la red y el número total de parámetros. Se usó la misma cantidad de filtros en cada capa que en la red U-Net 2D en su versión de tres dimensiones.

En el proceso de entrenamiento de las redes se utilizó el optimizador Adam, con tasa de aprendizaje establecida en 0.001, $\beta_1 = 0.9$ y $\beta_2 = 0.999$ (Kingma & Ba, 2014). La función de pérdida usada fue entropía cruzada binaria y se realizaron en total 100 épocas, almacenando la mejor época de entrenamiento basándose en el resultado de exactitud.

2. Segmentación empleando FSL: Este método se presenta como primera técnica de comparación. FSL es una librería para análisis estadístico de FMRI, MRI y DTI, en este caso, se utilizó el módulo para segmentación de tejido cerebral en materia gris (GM), materia blanca (WM) y CSF. Para la parametrización se utilizó salida por separado de las imágenes binarizadas para cada tipo de tejido, se realizó la estimación parcial del volumen, se seleccionaron 3 clases de salida correspondientes a los tejidos de interés y se utilizaron salidas tipo NIFTI. Adicionalmente, para el tipo de imagen de entrada se utilizó T1, lo que indica que se maximiza el contraste T1.
3. Segmentación empleando Dipy: Para este método se debe tener en cuenta que las imágenes de resonancia magnética deben estar sin el cráneo, de lo contrario, se deberá realizar un procesamiento adicional donde se remueva la parte no

cerebral de las imágenes, esto con la finalidad de obtener mejores resultados en la segmentación.

Luego de corroborar lo anterior, se ajustan algunos parámetros importantes como el número de clases, que para este caso son 3: GM, WM y CSF. También, el factor de suavizado de la imagen denominado beta, que ayuda a eliminar puntos de ruido de estas. Este parámetro oscila entre 0 y 0.5, que para nuestro caso se definió en 0.1, ya que es donde consigue un mejor rendimiento. Otro parámetro que también se puede redefinir es el número de iteraciones, ya que por defecto este viene en 100, sin embargo, es posible que se llegue a la convergencia en menos iteraciones dependiendo del modelo, por lo que no es necesario llegar o superar dicho número de iteraciones en algunos casos, así que se determinó ajustarlo para un número total de 20 iteraciones.

Después de asignar los parámetros definidos anteriormente, se procede a crear una instancia de la clase “TissueClassifierHMRF” con su respectivo método “classify” a las imágenes de resonancia magnética, donde al final, después del número de iteraciones asignadas, se obtendrá la imagen con las etiquetas esperadas de cada tejido del cerebro.

Buscando garantizar la reproducibilidad de este estudio, los códigos para el entrenamiento y uso de los métodos propuestos y métodos base se pueden encontrar en el siguiente repositorio: https://github.com/JovianPlanet/brain_segmentation.

4. Resultados

Luego del entrenamiento de las arquitecturas de U-Net presentadas, se procedió a la evaluación tanto de los modelos propuesto como los métodos comparativos. En esta etapa, se utilizaron métricas de evaluación basadas en teoría de conjuntos, el área bajo la curva (AUC) y la SSIM que se basa en la similitud estructural de imágenes. De este modo, se muestran a continuación los resultados estadísticos y ejemplos del desempeño de los modelos para realizar segmentación tanto de materia gris como de materia blanca.

Inicialmente, en la Figura 4 se presentan los resultados obtenidos para la segmentación de materia gris usando un diagrama de caja (boxplot). Para estos resultados se muestran los desempeños evaluados con las cuatro métricas implementadas. Los resultados muestran que, para cada método de evaluación, las arquitecturas propuestas (U-Net 2D y U-Net 3D) logran alcanzar puntajes por encima de 0.9 en la mayoría de las métricas, exceptuando el puntaje de Jaccard, el cual para todos los métodos analizados presenta una caída significativa. Sin embargo, los métodos basados en U-Net para el puntaje Jaccard mantienen una mediana cercana a 0.8, mientras que para los métodos de comparación se tiene mediana inferior a 0.6 en el caso de FSL e inferior a 0.4 para Dipy. En general se puede decir que, para la segmentación de materia gris, los métodos propuestos en este trabajo mantienen una varianza inferior a los métodos con los que se compara. Además, se tiene una mediana superior y menor variación de resultados entre las diferentes métricas.

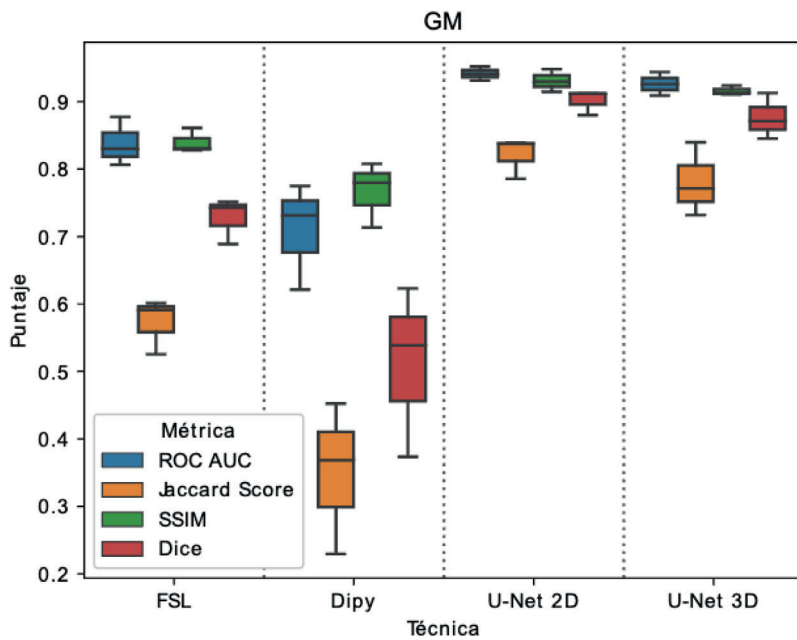


Figura 4 – Comparación de resultados para segmentación de materia gris

Adicionalmente, en la Figura 5 se muestran algunos resultados como ejemplo de la segmentación realizada por las técnicas propuestas y los métodos de comparación sobre un volumen del conjunto de datos MRBrainS18. En blanco se presenta el ground truth, y sobre él, utilizando transparencia, la delineación obtenida por cada método. La figura muestra que U-Net en ambas versiones consigue seguir los surcos de la estructura analizada, mientras que para las demás metodologías se marcan también ventrículos. Por otro lado, se percibe visualmente que las regiones de convergencia entre la etiqueta real y las salidas de los métodos usando U-Net son mayores que las obtenidas por FSL y Dipy. Esto se puede ver en los espacios blancos que quedan dentro de la estructura, siendo más notoria en Dipy que FSL.

En la Figura 6 se presentan los resultados del diagrama de caja obtenidos para la segmentación de materia blanca. Estos resultados muestran a diferencia de la materia gris, una menor variabilidad entre los diferentes métodos y métricas de evaluación, encontrando que la media de los puntajes se mantiene por encima de 0.6, mientras que para materia gris podían estar por debajo de 0.4, lo cual sugiere que la segmentación de materia blanca es una tarea menos exigente que la segmentación de materia gris. Al igual que en el experimento anterior, los métodos propuestos continúan teniendo una mediana alta en el puntaje para cada métrica. En el caso del puntaje Jaccard, al igual que con la materia gris, se nota una disminución significativa de la mediana y mayor

varianza en todos los métodos, aunque la caída sigue siendo menor para los métodos basados en U-Net.

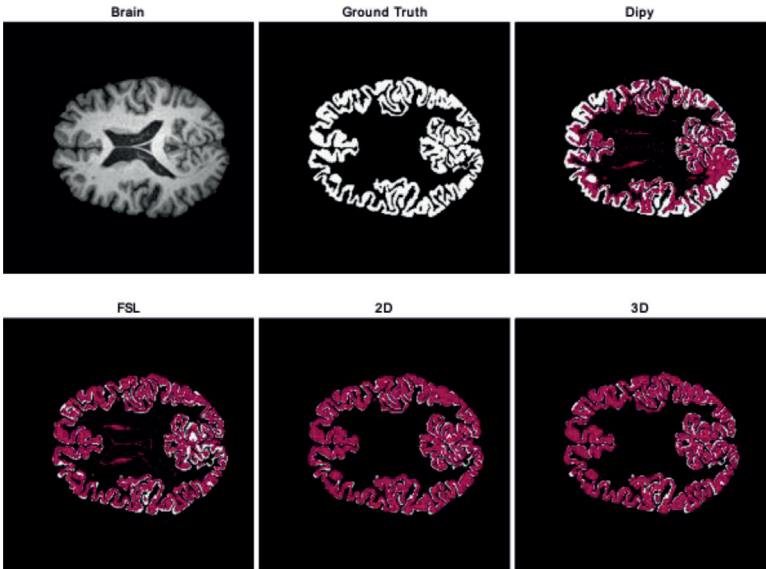


Figura 5 – Ejemplo de resultado para segmentación de materia gris sobre volumen de MRBrainS18.

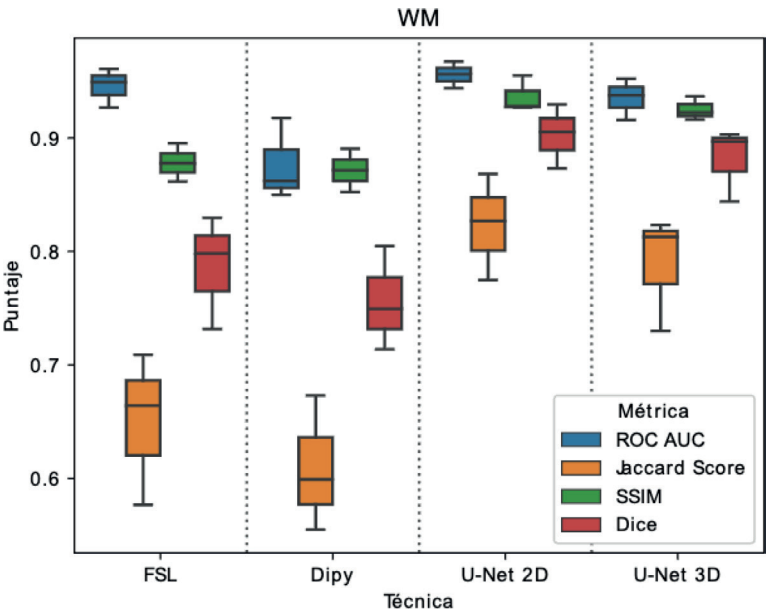


Figura 6 – Comparación de resultados para segmentación de materia blanca

Finalmente, un ejemplo de la segmentación de la materia blanca sobre un volumen del conjunto de datos IBSR, se muestra en la Figura 7. Se presentan los resultados para las técnicas contempladas, en blanco se presenta la etiqueta de comparación (Ground truth), y sobre esta, utilizando transparencia, la delineación obtenida por cada método. En este caso los resultados son consistentes con lo reportado en el experimento anterior, en donde Dipy presenta un gran número de espacios del tejido sin segmentar y FSL a pesar de que realiza una mejor segmentación, hace marcación errónea sobre componentes que están por fuera de la estructura de interés. Por otro lado, al analizar U-Net, se puede ver como los espacios no segmentados son reducidos (similares para ambas propuestas) y además se tiene una menor marcación de áreas fuera de la estructura, en donde se puede evidenciar un mejor rendimiento cuando es utilizado el modelo 2D.

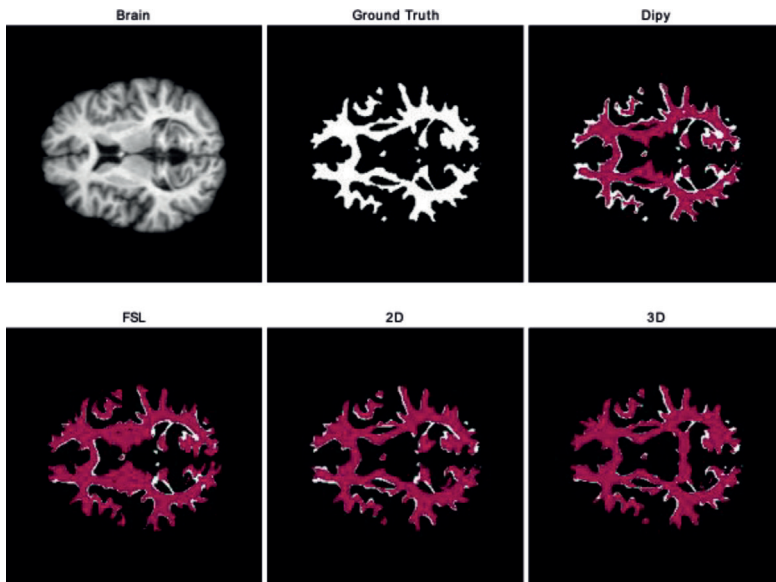


Figura 7 – Ejemplo de resultado para segmentación de materia blanca sobre volumen de IBSR

5. Conclusiones

En este trabajo se realizó la propuesta de una técnica de segmentación de tejido cerebral utilizando arquitecturas de redes neuronales, específicamente se trabajó con redes U-Net en dos versiones diferentes que recibían imágenes en 2D o volúmenes en 3D. Los resultados de esta investigación muestran como la técnica propuesta consigue mejorar el rendimiento de la tarea de segmentación de materia gris y blanca cuando se compara con metodologías recientes del estado del arte. De forma general se pudo evidenciar que las arquitecturas U-Net logran reducir significativamente la segmentación de componentes fuera de los tejidos de interés, es decir, hay una menor segmentación de los tejidos circundantes.

Al hacer la comparación entre el modelo 2D y 3D, se puede concluir que el modelo 2D es ligeramente superior al 3D cuando se observan los diagramas de cajas en todas las

medidas de evaluación. Adicionalmente, al ver las áreas de segmentación, se puede ver como el modelo 2D disminuye las áreas sin segmentar para materia gris principalmente, y para materia blanca se disminuyen las segmentaciones fuera del tejido. De forma general, el modelo 2D reduce tanto los falsos positivos como falsos negativos de la segmentación.

El principal reto para que la metodología propuesta pueda ser usada en línea se encuentra en la etapa de entrenamiento, debido a la sintonización iterativa de parámetros. Posteriormente en la etapa de evaluación, la velocidad de cómputo disminuye considerablemente para ser usada en tiempo real. Sin embargo, es importante considerar que los MRI a usarse se deben llevar al espacio de entrada de las redes, es decir, $240 \times 240 \times 48$, que permitan el uso individual en el eje axial para el modelo 2D y el uso de subvolumenes de dimensión $80 \times 80 \times 16$ en el modelo 3D.

Como trabajo futuro se encuentra validar el rendimiento de la arquitectura propuesta en otras estructuras cerebrales para evaluar la generalización del modelo. Por otro lado, aunque se tiene una mejora notoria en la tarea de segmentación, aún existen contornos que son difíciles de distinguir para el modelo, lo cual establece una ruta de trabajo en pro de implementar mejoras que permitan la segmentación de zonas al parecer son desafiantes para las distintas técnicas implementadas.

Agradecimientos

Este trabajo se enmarca en el proyecto de investigación P20214, financiado por el Instituto Tecnológico Metropolitano ITM de Medellín.

Referencias

- Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep learning for brain MRI segmentation: state of the art and future directions. *Journal of Digital Imaging*, 30(4), 449–459.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00444-8>
- Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, 42(11). <https://doi.org/10.1007/s10916-018-1088-1>
- Aronica, E., Becker, A. J., & Spreafico, R. (2012). Malformations of cortical development. *Brain Pathology*, 22(3), 380–401.
- Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., & Gee, J. C. (2011). An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics*, 9(4), 381–400.

- Billot, B., Greve, D., Van Leemput, K., Fischl, B., Iglesias, J. E., & Dalca, A. V. (2020). A learning strategy for contrast-agnostic MRI segmentation. *ArXiv Preprint ArXiv:2003.01995*.
- Castillo-Carranza, A. P., Bravo-Huivin, E. K., & Cieza-Mostacero, S. E. (2022). Revisión sistemática: La Aplicación del Deep Learning en el sector Salud entre los años 2018 a 2022 Systematic review: The Application of Deep Learning in the Healthcare.
- Chang, H.-H., & Hsieh, C.-C. (2017). Brain segmentation in MR images using a texture-based classifier associated with mathematical morphology. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 3421–3424.
- Chen, L., & Merhof, D. (2018). MixNet: Multi-modality Mix Network for Brain Segmentation. *International MICCAI Brainlesion Workshop*, 367–377.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 424–432.
- Clèrigues, A., Valverde, S., Salvi, J., Oliver, A., & Lladó, X. (2023). Minimizing the effect of white matter lesions on deep learning based tissue segmentation for brain volumetry. *Computerized Medical Imaging and Graphics*, 103(May 2022). <https://doi.org/10.1016/j.compmedimag.2022.102157>
- de Brebisson, A., & Montana, G. (2015). Deep neural networks for anatomical brain segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 20–28.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3), 297–302.
- Hyman, S. E. (2003). Diagnosing disorders. *Scientific American*, 289(3), 96–103.
- Ito, R., Nakae, K., Hata, J., Okano, H., & Ishii, S. (2019). Semi-supervised deep learning of brain tissue segmentation. *Neural Networks*, 116, 25–34. <https://doi.org/10.1016/j.neunet.2019.03.014>
- Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2), 37–50.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL (FMRIB Software Library). *Neuroimage*, 62(2), 782–790.
- Jiang, H., Diao, Z., Shi, T., Zhou, Y., Wang, F., Hu, W., Zhu, X., Luo, S., Tong, G., & Yao, Y. D. (2023). A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation. *Computers in Biology and Medicine*, 157(195), 106726. <https://doi.org/10.1016/j.compbiomed.2023.106726>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv Preprint ArXiv:1412.6980*.

- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- Kuijf, H. J., & Bennink, E. (2019). MRBrainS18 Grand challenge on MR brain segmentation at MICCAI 2018.
- Kumar, M., Rathore, R. K. S., Srivastava, A., Yadav, S. K., Behari, S., & Gupta, R. K. (2011). Correlation of diffusion tensor imaging metrics with neurocognitive function in Chiari i malformation. *World Neurosurgery*, 76(1–2), 189–194. <https://doi.org/10.1016/j.wneu.2011.02.022>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Liu, Q., Zhang, N. (2017). A Review of Image Recognition with Deep Convolutional Neural Network. In: *Intelligent Computing Theories and Application* (pp. 69–80). Springer International Publishing.
- Mendrik, A. (2013). Evaluation framework for MR brain image segmentation. MICCAI Grand Challenge.
- Mendrik, A. M., Vincken, K. L., Kuijf, H. J., Breeuwer, M., Bouvy, W. H., De Bresser, J., Alansary, A., De Bruijne, M., Carass, A., El-Baz, A. (2015). MRBrainS challenge: online evaluation framework for brain image segmentation in 3T MRI scans. *Computational Intelligence and Neuroscience*, 2015.
- Ndajah, P., Kikuchi, H., Yukawa, M., Watanabe, H., & Muramatsu, S. (2010). SSIM image quality metric for denoised images. In: *Proc. 3rd WSEAS Int. Conf. on Visualization, Imaging and Simulation*, 53–58.
- Peruzzo, D., Arrigoni, F., Triulzi, F., Righini, A., Parazzini, C., & Castellani, U. (2016). A framework for the automatic detection and characterization of brain malformations: Validation on the corpus callosum. *Medical Image Analysis*, 32, 233–242. <https://doi.org/10.1016/j.media.2016.05.001>
- Rashid, T., Liu, H., Ware, J. B., Li, K., Romero, J. R., Fadaee, E., Nasrallah, I. M., Hilal, S., Bryan, R. N., Hughes, T. M., Davatzikos, C., Launer, L., Seshadri, S., Heckbert, S. R., & Habes, M. (2022). Deep Learning Based Detection of Enlarged Perivascular Spaces on Brain MRI. *Neuroimage Reports*, 3(1), 100162. <https://doi.org/10.1016/j.ynirp.2023.100162>
- Rohlfing, T. (2012). Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy: Widely Used but Unreliable. *IEEE Transactions on Medical Imaging*, 31(2), 153–163. <https://doi.org/10.1109/tmi.2011.2163944>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241.

- Sánchez Fernández, I., Yang, E., Amengual-Gual, M., Barcia Aguilar, C., Calvachi Prieto, P., & Peters, J. M. (2021). Convolutional neural networks to identify malformations of cortical development: A feasibility study. *Seizure*, 91(May), 81–90. <https://doi.org/10.1016/j.seizure.2021.05.023>
- Sarma, A., & Pruthi, S. (2023). Congenital Brain Malformations- Update on Newer Classification and Genetic Basis. *Seminars in Roentgenology*, 6–27. <https://doi.org/10.1053/j.ro.2022.11.004>
- Shamir, R. R., Duchin, Y., Kim, J., Sapiro, G., & Harel, N. (2019). Continuous dice coefficient: a method for evaluating probabilistic segmentations. *ArXiv Preprint ArXiv:1906.11031*.
- Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2021.3086020>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155.
- Somasundaram, K., & Kalaividya, P. A. (2016). Brain portion segmentation from Magnetic Resonance Images (MRI) of human head scan using Richardson Lucy deconvolution and intensity thresholding. 2016 International Computer Science and Engineering Conference (ICSEC), 1–5.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.
- Toscano, I., Bravo, D. A. M., De-La-torre, M., Juárez, B. A., & Mireles, G. A. G. (2023). Automated analysis of SEM micrographs using deep learning. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, 2023(49), 100–114. <https://doi.org/10.17013/risti.49.100-114>
- Van Calster, B., Van Belle, V., Condous, G., Bourne, T., Timmerman, D., & Van Huffel, S. (2008). Multi-class AUC metrics and weighted alternatives. 2008 IEEE International Joint Conference on Neural Networks, 1390–1396.
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage*, 45(1), S173–S186.
- Zhang, Y., Brady, J. M., & Smith, S. (2000). Hidden Markov random field model for segmentation of brain MR image. *Medical Imaging 2000 Image Processing*, 3979, 1126–1137.

- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57.
- Zubrikhina, M. O., Abramova, O. V, Yarkin, V. E., Ushakov, V. L., Ochneva, A. G., Bernstein, A. V, Burnaev, E. V, Andreyuk, D. S., Savilov, V. B., Kurmishev, M. V, Syunyakov, T. C., Karpenko, O. A., Andryushchenko, A. V, Kostyuk, G. P., & Sharaev, M. G. (2023). Machine learning approaches to mild cognitive impairment detection based on structural MRI data and morphometric features. *Cognitive Systems Research*, 78(August 2022), 87–95. <https://doi.org/10.1016/j.cogsys.2022.12.005>

Protótipo de Solução para Detetar e Sinalizar Defeitos em Pavimentos Rodoviários Baseado em Técnicas de Visão Computacional

Miguel Gonçalves¹, Tomás Marques¹, Pedro D. Gaspar^{2,3},
Vasco N. G. J. Soares^{1,4}, João M. L. P. Caldeira^{1,4*}

g.miguel@ipcbcampus.pt; tomas.marques@ipcbcampus.pt; dinis@ubi.pt;
vasco.g.soares@ipcb.pt; jcaldeira@ipcb.pt

¹ Instituto Politécnico de Castelo Branco, Av. Pedro Álvares Cabral nº 12, 6000-084 Castelo Branco, Portugal

² C-MAST Center for Mechanical and Aerospace Science and Technologies, Universidade da Beira Interior, 6201-001 Covilhã, Portugal

³ Departamento de Engenharia Eletromecânica, Universidade da Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal

⁴ Instituto de Telecomunicações, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal

DOI: 10.17013/risti.52.25-44

Resumo: Este artigo apresenta um protótipo funcional para avaliar e validar a utilização de técnicas de visão computacional, na identificação de defeitos em pavimentos rodoviários, no contexto de uma cidade inteligente. É realizado um estudo de avaliação de desempenho de três redes neurais convolucionais, YoloV4-Tiny, SSD MobileNet e RetinaNet, aplicadas a este cenário. Partindo dos resultados observados, descreve-se a proposta e o processo de implementação do protótipo, que tem por base uma plataforma Raspberry Pi 4. O protótipo é sujeito a validação e testes funcionais. Comparativamente ao método atualmente utilizado pela Infraestruturas de Portugal, para a identificação de defeitos em pavimentos, esta abordagem é mais ágil, eficaz e eficiente, contribuindo para uma rápida deteção e notificação dos mesmos.

Palavras-chave: Cidades Inteligentes, Pavimentos Rodoviários, Deteção de Defeitos, Visão Computacional, Redes Neurais Convolucionais, Deteção de Objetos, Protótipo.

Prototype Solution for Detecting and Signaling Road Pavement Defects Based on Computer Vision Techniques

Abstract: This article presents a functional prototype to evaluate and validate the use of computer vision techniques to identify road pavement defects in the context of a smart city. A study is carried out to evaluate the performance of three convolutional neural networks, YoloV4-Tiny, SSD MobileNet and RetinaNet, applied to this scenario. Based on the results observed, the proposal and implementation process of the prototype is described, which is based on a Raspberry Pi 4 platform.

The prototype is subject to validation and functional tests. Compared to the method currently used by Infraestruturas de Portugal to identify pavement defects, this approach is more agile, effective, and efficient, contributing to their rapid detection and notification.

Keywords: Smart Cities, Road Pavement, Defect Detection, Computer Vision, Convolutional Neural Networks, Object Detection, Prototype.

1. Introdução

As entidades públicas têm vindo a procurar novas formas de tornar as cidades mais preparadas para o futuro, promovendo soluções inovadoras para resolver problemas sociais e melhorar a qualidade de vida dos cidadãos. Para a implementação destas soluções, são utilizadas tecnologias de informação e comunicação. Estas tecnologias permitem gerir os recursos e as infraestruturas das cidades. Grandes volumes de dados têm vindo a ser recolhidos através da crescente utilização de sensores, tecnologias de rede sem fios, dispositivos autónomos e aplicações móveis. Isto levou a que sistemas fossem amplamente instalados nas cidades, devido ao potencial de análise e extração de valor dos dados recolhidos por estas tecnologias, nascendo o conceito de “cidades inteligentes” (Yin et al., 2015).

Os pavimentos rodoviários são um dos elementos integrantes de qualquer cidade. Muitas vezes estes encontram-se deteriorados e necessitam de ser intervencionados. Tipicamente a manutenção e conservação dos pavimentos rodoviários está a cargo das entidades responsáveis, tais como a Infraestruturas Portugal (IP) (Infraestruturas de Portugal, 2023). No entanto, o processo de identificação e geolocalização dos pontos que necessitam de ser intervencionados é bastante moroso e difícil. Não obstante, é possível tirar partido dos veículos dos vários organismos de uma cidade, em deslocação diária entre diversos pontos. Estes podem facilitar a identificação das deficiências encontradas nos pavimentos rodoviários por onde passam, desde que equipados com sistemas que permitam essa deteção. No contexto abrangente das cidades inteligentes, é possível instalar sistemas automáticos nestes veículos que usem técnicas de visão computacional, para deteção, classificação e geolocalização de deficiências estruturais nos pavimentos rodoviários. Tirando partido desta capacidade instalada, estes sistemas serão capazes de notificar automaticamente as entidades decisoras sobre os problemas, contribuindo para a sua resolução atempada.

Neste âmbito, este artigo dá a continuidade a uma investigação conduzida pelos mesmos autores (Gonçalves et al., 2023), que apresentou uma revisão da literatura e identificou as técnicas de visão computacional mais promissoras, para a deteção de defeitos em pavimentos rodoviários. Esta deteção permitiria facilitar e agilizar os processos de classificação e geolocalização desses defeitos. Neste artigo, apresenta-se o desenho, implementação e validação de um protótipo funcional para demonstração do conceito. No desenvolvimento do protótipo foram utilizadas três redes neuronais convolucionais YOLOv4-tiny (Jiang et al., 2020), SSD MobileNet (Kumar et al., 2023) e RetinaNet (T. Y. Lin et al., 2017), para comparação e avaliação de desempenho, e um microcomputador Raspberry Pi 4 (Raspberry Pi 4 Model B, 2023).

Este artigo encontra-se estruturado da seguinte forma. A Secção 2 apresenta as técnicas de visão computacional consideradas, a descrição do *dataset* usado, o cenário de *benchmark* e a avaliação de desempenho. A Secção 3 apresenta o desenho e implementação do protótipo funcional para demonstração do conceito. Finalmente, na Secção 4, apresentam-se as conclusões e o trabalho futuro.

2. Técnicas de Visão Computacional

A visão computacional é uma subárea da inteligência artificial e *machine learning* que utiliza computadores para obter uma compreensão detalhada de dados visuais. Isto é, dá a capacidade de visão à máquina, semelhante à dos humanos. Esta capacidade permite que a máquina retire informações úteis através da interpretação dos recursos visualizados. Isto tudo é feito com o uso de câmaras e computadores, o que torna o processo automático, sem a necessidade de interação manual humana (Computer Vision Techniques, 2023; Ozgon, 2021; Xu et al., 2020). Nesse sentido, foram criados vários modelos com diversas funções, especialmente os modelos para a deteção de objetos.

A grande maioria dos modelos de visão computacional usa redes neurais convolucionais (CNNs), arquiteturas semelhantes ou baseadas nestas, para a deteção de objetos. As CNNs são compostas por uma camada de convolução, outra de *pooling* e outra totalmente ligada. A primeira extrai características das imagens que ainda não foram processadas (Awati, 2023; Goodfellow et al., 2016; Le, 2018; Smeda, 2019). A segunda camada consiste na redução do tamanho das imagens, número de parâmetros e cálculos na rede, enquanto preserva as características mais importantes, acelera a computação e controla a ocorrência de *overfitting* (O Que é Sobreajuste?, 2023). Para finalizar, a camada totalmente ligada usa uma função de ativação *softmax* ou *sigmoid*, que serve para estabelecer uma ligação entre o resultado das camadas convolução e *pooling*. Desta forma é possível classificar os objetos presentes na imagem, no leque de classes existentes (Awati, 2023; Islam & Sadi, 2019; Le, 2018; O Que é Sobreajuste?, 2023; What Is a CNN, 2023; Smeda, 2019; Zhao et al., 2018).

A próxima subsecção parte das conclusões apresentadas num artigo anterior dos mesmos autores (Gonçalves et al., 2023), que identificou as técnicas de visão computacional mais promissoras no contexto do problema a resolver, discutiu problemas e desafios em aberto.

2.1. Modelos de Deteção de Objetos

De seguida, descrevem-se as principais características das técnicas de visão computacional que serão usadas no protótipo proposto no contexto deste trabalho. A escolha recaiu em técnicas com a capacidade de deteção de defeitos em tempo real, passíveis de ser implementadas em dispositivos de tamanho reduzido como microcomputadores e telemóveis.

O YOLO (Redmon & Farhadi, 2018) é um dos modelos mais referenciados para a deteção de objetos. Apresenta uma arquitetura que possui 24 camadas convolucionais e 2 camadas totalmente ligadas, o que permite produzir alta precisão e velocidade de processamento (Zvornicanin, 2023). O YOLO divide a imagem numa grelha de células

$W \times W$, define as *bounding boxes* de cada célula da grelha e a taxa de confiança das mesmas. Esta taxa irá indicar a certeza da presença do objeto na *bounding box*, assim como o quanto o modelo acredita o quão precisa é a previsão da mesma. O modelo reconhece as *bounding boxes* através de 4 valores: o centro das mesmas, a largura e a altura (Huang et al., 2018; Redmon & Farhadi, 2018; Zvornicanin, 2023). O YOLO tem várias vantagens sobre sistemas baseados em classificação, pois visualiza a imagem como um todo durante o tempo de teste, o que faz com que as previsões sejam baseadas em todo o conteúdo presente na imagem (YOLO: Real-Time Object Detection, 2023).

O RetinaNet (T.-Y. Lin et al., 2016) é outro dos modelos mais usado na deteção de objetos. Este já provou ter bons resultados na deteção de objetos densos e de pequena escala (T.-Y. Lin et al., 2016). A arquitetura do RetinaNet é formada por 4 partes: *backbone network*, *Feature Pyramid Network* (FPN), sub-rede de classificação e sub-rede de regressão. A *backbone network* utilizada denomina-se de *ResNet* e é responsável por extrair características da imagem. A FPN realiza processamento adicional das características extraídas da *backbone network*. A sub-rede de classificação apresenta a probabilidade de um objeto estar presente em cada localização espacial para cada *anchor box*. Para finalizar, a sub-rede de regressão regride o desvio para as *bounding boxes*, a partir das *anchor boxes* para cada objeto de referência (T.-Y. Lin et al., 2016; Tian et al., 2020). O RetinaNet recorre ao *Focal Loss* (FL) que serve para resolver o problema de desequilíbrio de classes dos modelos de deteção de objetos de uma fase. Isto advém do facto de que, geralmente, estes têm de processar a maior quantidade possível de objetos. No modelo, podem estar presentes milhares de *anchor boxes* em cada camada e no caso da existência de objetos com alta probabilidade de deteção, estes podem afetar coletivamente o modelo. Daí ser necessário o uso de FL, que reduz a contribuição dessas deteções e aumenta a importância de corrigir deteções mal classificadas (T.-Y. Lin et al., 2016; Tian et al., 2020).

O Single Shot MultiBox Detector (SSD) (Liu et al., 2016) é um modelo baseado em redes neuronais simples, onde a informação apenas segue uma direção (*feed-forward*), ao contrário das redes onde os seus nós formam ciclos. As redes convolucionais *feed-forward* produzem coleções de *bounding boxes* de tamanho fixo e apresentam uma pontuação de presença de um objeto dentro destas. De seguida, é realizada uma etapa de compressão não máxima onde a *bounding box*, com a sobreposição máxima, obtém a sua pontuação máxima e deteta o objeto. Este modelo é semelhante ao YOLO visto que ambos dividem a imagem em grelhas de tamanhos iguais (Liu et al., 2016). O SSD pode ser combinado com uma arquitetura MobileNet, onde esta é usada como *backbone* para o SSD (Brownlee, 2021; Mehta, 2021). A arquitetura MobileNet foi desenvolvida para o uso em aplicações *mobile* e é baseada em *depthwise separable convolutions*, o que permite reduzir o número de parâmetros, comparado a uma rede convolucional comum, tornando a rede mais leve (Howard et al., 2017).

2.2. Avaliação de Desempenho

Esta subseção centra-se na avaliação de desempenho dos modelos CNN descritos, para deteção de defeitos em pavimentos rodoviários. Para este fim, apresenta-se o *dataset* usado no âmbito deste trabalho, o cenário de *benchmark*, as métricas de desempenho consideradas, e finalmente discutem-se os resultados observados.

Descrição do Dataset

O *dataset* utilizado para o treino dos modelos, foi criado a partir de imagens retiradas de um outro *dataset* (MAEDA, 2023), que já têm anotações dos defeitos estruturais subdivididas em 7 classes. Para não aumentar a complexidade do modelo, foram escolhidas apenas 4 classes de defeitos estruturais considerados mais comuns. Assim, as anotações do *dataset* criado estão divididas em 4 classes: D00, D10, D20, D40, dependendo do tipo de defeito. A classe D00 representa as fissuras longitudinais, D10 as fissuras transversais, D20 as fissuras de jacaré e D40 os buracos. A Figura 1 apresenta exemplos de imagens classificadas em função do tipo de defeito estrutural.

Todos os processos realizados ao *dataset*, como a adição de imagens, separação em conjuntos de imagens de treino, teste e validação e pré-processamento das imagens, foram realizados no Roboflow (Roboflow, 2023). O Roboflow é um site que fornece ferramentas para criar um modelo de visão computacional. Permite criar *datasets* desde o *upload* da imagem até à anotação dos objetos, o que permite simplificar todo o processo de preparação do *dataset* (Roboflow Docs, 2023).

Após todas as interações, o *dataset* final ficou com 3420 imagens e 7336 anotações de defeitos estruturais. Destas anotações, 2010 são da classe D00, 1373 da classe D10, 2010 da classe D20 e 1943 da classe D40. As imagens foram divididas da seguinte forma, 66%, que equivale a 2300 do total das imagens, para o conjunto de treino e 34%, que equivale a 1100 do total de imagens, para o conjunto de validação.

Foram usados dois passos de pré-processamento, *auto-orient* e *resize*. O primeiro é usado para descartar rotações *EXIF* e para padronizar a orientação dos pixéis. O segundo, redimensiona as imagens para que estas tenham todas o mesmo tamanho (600 x 600).



Figura 1 – Diferentes classes de defeitos estruturais em pavimentos rodoviários.

Foi criada uma segunda versão do *dataset*, no Roboflow, para o treino do SSD MobileNet. Esta nova versão tem as mesmas imagens e passos de pré-processamento. No entanto, foram usadas 2700 imagens para o conjunto de treino, 336 imagens para o conjunto de validação e outras 336 imagens para o conjunto de teste. Para o treino do YOLO também foi criada uma terceira versão do *dataset*, contendo as mesmas imagens e passos de pré-processamento, mas em que 2800 imagens foram utilizadas para o conjunto de treino, 500 imagens para o conjunto de validação e 49 imagens para o conjunto de teste.

Cenário de Benchmark

A versão escolhida para o YOLO foi a versão 4, devido à sua precisão (Nepal & Eslamiat, 2022) em relação às restantes (Supeshala, 2020). Para além disso, versões posteriores já não fazem o uso de *darknet* (GitHub-AlexeyAB/Darknet, 2023). O *darknet* é uma *framework* de rede neuronal *open source* do YOLO escrita em C e CUDA. O *darknet* é bastante rápido e fácil de instalar e suporta computação em CPU e GPU (What Is Darknet and Why Is It Needed?, 2023). Como é bastante rápido, pode ser muito útil para tarefas que envolvem tempo real (Kundu, 2023), como o problema que se pretende resolver neste artigo.

Devido à complexidade do modelo YOLOv4, foi utilizada uma configuração YOLOv4-tiny. O YOLOv4-tiny é uma versão comprimida do YOLOv4 que foi criada com o propósito de tornar a estrutura da rede mais simples e reduzir o número de parâmetros. Esta compressão permite que se torne viável para ser utilizado em dispositivos móveis ou embebidos. Esta versão permite que o treino do modelo se torne mais rápido e que realize uma deteção mais rápida, tendo 8 vezes mais *frames* por segundo (FPS) que o modelo base. No entanto, a precisão é afetada, conseguindo apenas aproximadamente 2/3 da precisão do modelo base (Techzizou, 2021; YOLOv4 Tiny, 2023).

Para que a máquina alvo pudesse realizar o treino das redes neuronais foram instaladas ferramentas como *CUDA* (CUDA Toolkit, 2023) e *cuDNN* (CUDA CuDNN, 2023). O *CUDA* é uma plataforma desenvolvida pela *NVIDIA* que permite o processamento paralelo através da utilização de uma *GPU*. A *cuDNN* é uma biblioteca de primitivas para *deep neural networks*, responsável por fornecer rotinas altamente otimizadas para ações relacionadas com a utilização de redes neuronais. Isto permite uma utilização eficiente das redes neuronais, tirando partido do poder de processamento paralelo da *GPU*.

O modelo RetinaNet utilizado foi o *Keras-RetinaNet* (GitHub-Fizyr/Keras-Retinanet, 2023; GitHub-Tensorflow/Models, 2023) que disponibiliza 3 *backbones* diferentes: *ResNet50*, *ResNet101* e *ResNet152*. Para o treino foi escolhido o *backbone ResNet152*, capaz de realizar a menor quantidade de erros na deteção (He et al., 2015).

Para o modelo SSD MobileNet foi usado o modelo pré treinado presente no repositório GitHub “Model Garden for TensorFlow” (GitHub-Tensorflow/Models, 2023). Neste modelo foi realizado um fine-tune (Fine-Tune a Pretrained Model, 2023) para detetar defeitos estruturais nos pavimentos rodoviários, no sentido de que um modelo pré-treinado foi instalado. Esse modelo foi o SSD MobileNet v2 com *Feature Pyramid Network Lite* (FPN Lite), que realiza processamento adicional das características extraídas numa resolução de 640x640 (Hongkun Yu et al., 2020).

Os modelos foram treinados em diferentes dispositivos. Para treinar e testar o YOLOv4-tiny, o ambiente de treino utilizado foi hospedado numa máquina com as seguintes especificações: processador AMD Ryzen 5 3600, placa gráfica NVIDIA GeForce GTX 1060, 16 GB de memória principal e Windows 10. Para treinar e testar o RetinaNet, o ambiente de treino utilizado foi hospedado numa outra máquina com as seguintes especificações: processador AMD Ryzen 5 3500X, placa gráfica NVIDIA GeForce GTX 1660 SUPER, 16 GB de memória principal e Windows 10. Para finalizar, para treinar e testar o SSD MobileNet foi utilizada uma máquina do Google Colab (Google Colab, 2023), com uma placa gráfica NVIDIA Tesla T4.

Métricas de Desempenho

O desempenho dos modelos será avaliado com base nas métricas *mAP* e *Loss*. O *mAP* permite avaliar o desempenho na deteção de objetos. Esta métrica está diretamente ligada com a *average precision* (AP) que compara as *ground-truth bounding boxes* com as *bounding boxes* detetadas e retorna uma pontuação de precisão. O AP apenas retorna a pontuação de uma classe. Se, tecnicamente existem 4 classes no contexto do problema a resolver, então serão retornados 4 valores de AP, um para cada classe. O *mAP* é a média das pontuações de AP de todas as classes, que irá indicar o desempenho do modelo (Ahmed, 2023).

A *Loss* resulta da soma da *classification loss*, *localization loss* e *confidence loss*. Se um objeto for detetado, a *classification loss* em cada célula da grelha de deteção do YOLO é o erro quadrático das probabilidades condicionadas para cada classe. A *localization loss* permite medir os erros nas localizações e nos tamanhos previstos das *bounding boxes*. A *confidence loss* refere-se à perda de confiança do modelo quando um objeto é ou não detetado numa *bounding box* (Hui, 2018). A *Loss* permite então quantificar o erro produzido pelo modelo. Ou seja, valores altos de *Loss* indicam que o modelo está a produzir resultados incorretos e valores baixos indicam que há menos erros no modelo (baedlung, 2023). A análise do gráfico de *Loss* pode ser bastante útil na descoberta de erros de treino do modelo (Como Interpretar Curvas de Perda, 2022).

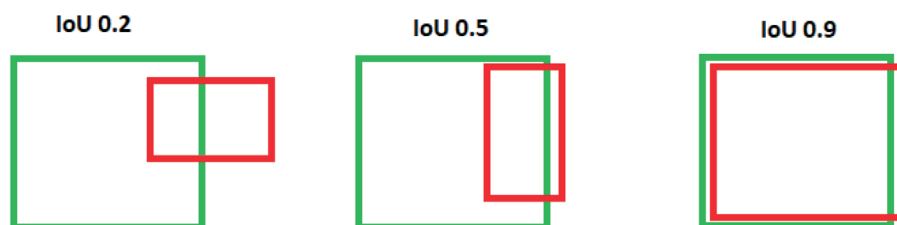


Figura 2 – Comparação de desempenho de *IoU*.

Importa também discutir o *IoU threshold*, pois o *mAP* está associado a esta métrica. O *IoU* mede o quanto as coordenadas das *bounding boxes* previstas se sobrepõem às coordenadas das *bounding boxes* de *ground-truth* (Ahmed, 2023). Neste sentido o *IoU* é a razão entre a intersecção das áreas das duas *bounding boxes*, a *bounding box* de

ground-truth e a *bounding box* detetada pelo modelo, e as suas áreas combinadas. É possível escolher uma pontuação mínima de *IoU* necessária para que a *bounding box* prevista seja considerada uma deteção positiva precisa. Isto permite usar o *IoU* como um limite para a deteção de objetos, daí se denominar de *IoU threshold*. Não existe um limite recomendado que sirva para todos os *IoU*, pois depende bastante da tarefa que se pretende realizar (Shah, 2023). No entanto, o *IoU threshold* mais comum é 0.5 pelo que a *bounding box* tem de ter um *IoU* score de pelo menos 0.5. A Figura 2 mostra um exemplo da comparação de desempenho de *IoU*.

Resultados e Discussão

O treino do RetinaNet demorou cerca de 10 horas, com o *backbone* *ResNet152*, realizando um total de 50000 passos, com verificação no final de cada época. A Figura 3 apresenta graficamente as métricas *mAP* (a) e *Loss* (b) do modelo treinado. Segundo o primeiro gráfico (a), após completar 8000 passos, o resultado obtido foi aproximadamente de 28% de precisão média, isto para 0.50 de *IoU threshold*. O segundo gráfico (b) da Figura 3 demonstra que a curva de *Loss* está a diminuir o que indica que o modelo foi treinado com sucesso.

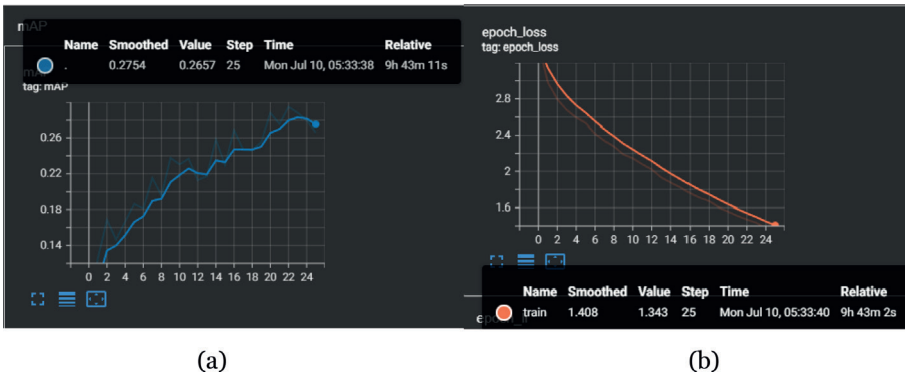


Figura 3 – *mAP* (a) e *Loss* (b) obtidos após treino do RetinaNet.

O treino do SSD MobileNet demorou cerca de 4 horas, realizando um total de 8000 passos. A Figura 4 apresenta o *mAP* e o gráfico de *Loss*. O *mAP* foi calculado através de uma ferramenta denominada de Cartucho (Cartucho J, 2018). Esta ferramenta avalia a rede neuronal usando o critério *mAP* definido em (VOC 2012, 2012). O *mAP* foi calculado em vários *IoU thresholds* entre 0.5 e 0.95 e a média de todos estes *IoU threshold* irá dar o *mAP* final, identificado na Figura 4 como *Overall*.

Pode observar-se na Figura 4 que, após completar 8000 passos, o resultado obtido de *mAP* foi de aproximadamente 5%, sendo este um resultado bastante baixo. A razão para esse resultado pode estar relacionada com a realização do *fine-tune* num modelo pré-treinado. O modelo deixa de ter precisão na deteção de defeitos a partir de 0.80 *IoU threshold*. Com 0,50 de *IoU threshold* foi obtido aproximadamente 17% de *mAP*. A Figura 4 mostra ainda graficamente a curva de *Loss* após o treino do SSD

MobileNet. Como se verifica, a curva está a diminuir o que indica que o modelo foi treinado com sucesso.

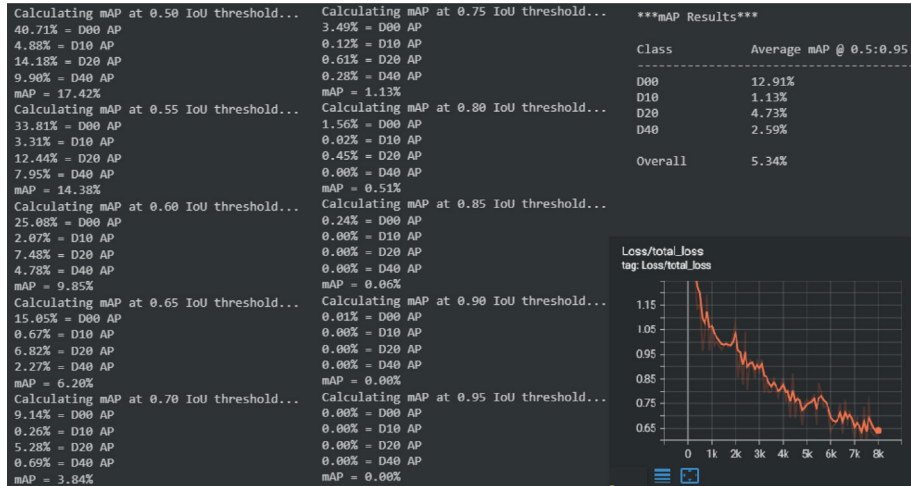


Figura 4 – Cálculo do *mAP* do SSD MobileNet em diferentes *IoU* thresholds e gráfico de *Loss*.

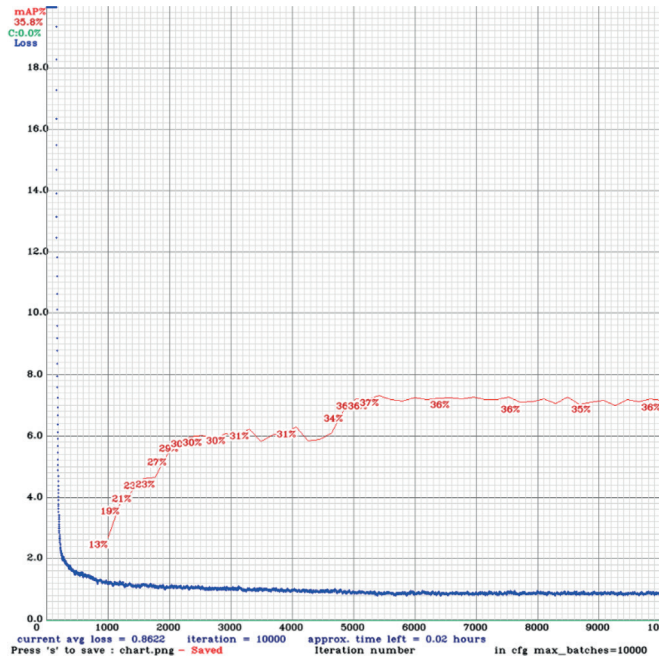


Figura 5 – *mAP* e *Loss* do YOLOv4-tiny.

O treino do YOLOv4-tiny demorou cerca de 8 horas, realizando um total de 10000 passos. A Figura 5 apresenta o gráfico da curva de *mAP* e da curva de *Loss* deste modelo. Segundo o gráfico, após completar os 10000 passos, o resultado obtido foi de aproximadamente 36% de precisão média com um *IoU threshold* de 0.50. O gráfico mostra ainda a curva de *Loss* a diminuir. Tal como nos casos anteriores, revela que o modelo foi treinado com sucesso.

Na análise global, se for considerado 0.50 de *IoU threshold*, é possível verificar que o *mAP* do YOLOv4-tiny é superior ao *mAP* do RetinaNet e que este é ainda superior ao *mAP* do SSD MobileNet. Face a estes resultados, conclui-se que o modelo mais adequado para ser utilizado na implementação do protótipo é o YOLOv4-tiny, pois foi o que obteve a maior precisão entre os três modelos avaliados. Assim, na próxima secção descreve-se o desenho e implementação de um protótipo funcional para demonstração do conceito, com base no modelo YOLOv4-tiny.

3. Protótipo

Esta secção descreve as etapas de proposta, implementação, teste e validação do protótipo funcional para detecção, classificação e sinalização de defeitos em pavimentos rodoviários com base em técnicas de visão computacional. O protótipo contém cinco componentes essenciais para o seu funcionamento: a rede neuronal convolucional YOLOv4-tiny, o computador que realiza o treino da rede neuronal, o Raspberry Pi 4 e um servidor de base de dados.

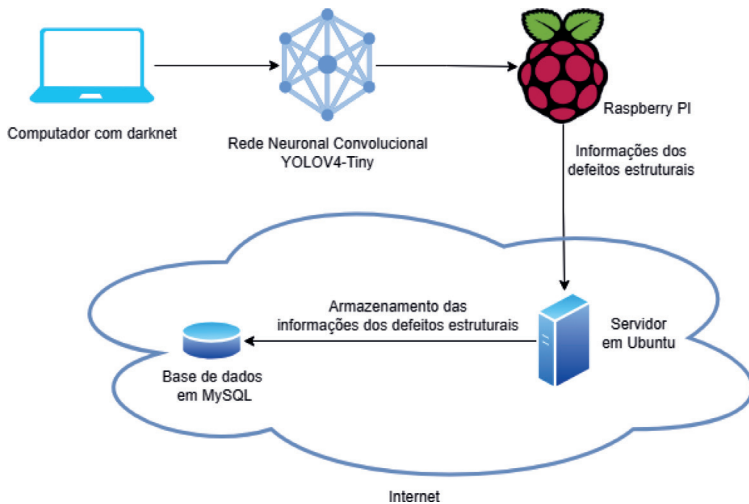


Figura 6 – Arquitetura do protótipo.

A arquitetura do protótipo é apresentada na Figura 6. O computador com *darknet* desempenha a função de treinar a rede neuronal e prepará-la para a integração no Raspberry Pi, sendo este o componente responsável pela detecção dos defeitos estruturais.

Este componente deve estar equipado com uma câmara e um sistema de geolocalização. No cenário proposto, este equipamento é montado num veículo, num local que permita captar vídeo em tempo real dos pavimentos rodoviários, enquanto o veículo circula. O servidor é responsável por servir de ponto de comunicação entre o Raspberry Pi e a base de dados. Na base de dados são guardados todos os dados que caracterizam os defeitos transmitidos pelo Raspberry Pi, nomeadamente, a sua geolocalização, o tipo de defeito estrutural e a data da deteção.

Nas subsecções seguintes, apresenta-se o hardware utilizado para construção do protótipo, a implementação de software e as configurações necessárias para a operação deste, e finalmente o teste e validação do protótipo.

Componente de Hardware

O protótipo apresentado tem por base um Raspberry Pi 4, que é responsável por executar a rede neuronal para realizar a deteção dos defeitos nos pavimentos rodoviários, bem como os processos secundários de comunicação com o servidor. Em questões de software, o Raspberry Pi está a operar com o sistema operativo Raspbian, sendo este baseado em Debian Linux utilizado geralmente por todos os microcomputadores da família Raspberry Pi (Raspberry Pi OS, 2023). Este sistema operativo contém todas as *drivers* necessárias para utilizar o Raspberry Pi e vem com ambiente de trabalho. Para a sua instalação foi usado um cartão microSD ligado ao Raspberry Pi e que exerce também a função do armazenamento.

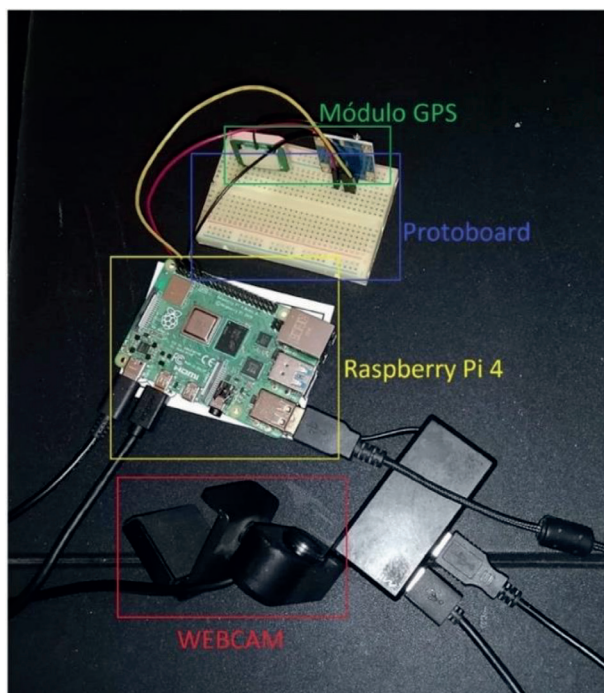


Figura 7 – Hardware do protótipo.

A alimentação deste equipamento é feita através de uma interface USB-C utilizando um transformador com saída DC de 5,1V e 3,0A. Para a aquisição de vídeo, foi utilizada uma *Webcam* USB com vídeo a 1080p *FULL HD* (Webcam Full HD, 2023), conectada a uma entrada USB do Raspberry Pi. Para aquisição da geolocalização foi usado um módulo GPS NEO-6M (Santos & Santos, 2023), compatível com Raspberry Pi, que permite obter as coordenadas de maneira precisa e rápida. O módulo de GPS está ligado ao Raspberry Pi 4 através de uma placa de *protoboard*. Para tal foram realizadas ligações onde o pino VCC do NEO 6M tem de estar conectado com o pino de 5v do Raspberry Pi, o pino GND do NEO 6M está ligado ao GND do Raspberry Pi e para finalizar o pino TX do NEO 6M tem de estar conectado ao pino RX do Raspberry Pi. Os três pinos do NEO 6M foram ligados à *protoboard* na “área de trabalho”, onde existem barramentos verticais. Estes barramentos conduzem a corrente elétrica, pelo que foram ligados cabos feminino-masculino, um em cada barramento vertical do *protoboard*. De seguida, cada cabo foi conectado ao pino correspondente do Raspberry Pi, como se verifica na Figura 7. Para a implementação e configuração do módulo de GPS foram seguidos os passos presentes em (Das, 2019). O protótipo funcional pode ser visto na Figura 7.

Para o servidor, considerou-se uma máquina acedida remotamente com um sistema operativo baseado em Ubuntu. Esta é uma máquina virtual hospedada por uma outra máquina com as seguintes especificações: processador AMD Ryzen 5 3600, placa gráfica NVIDIA GeForce GTX 1060, 16 GB de memória principal e Windows 10. Na máquina virtual foi realizada a instalação e configuração de um servidor MySQL. Para configurar o servidor MySQL, foram seguidos os passos descritos em (Fadheli, 2022). Para permitir o acesso remoto a esta, foi aberto o porto 3306 da rede onde a máquina está hospedada, que possibilitou a comunicação entre o protótipo e o servidor através de um endereço IP.

Componente de Software

Para realizar a deteção e classificação de defeitos em pavimentos rodoviários, foi utilizado o software *darknet* (GitHub-AlexeyAB/Darknet, 2023). Para tal, foi necessário instalar as bibliotecas OpenCV 4.5.3 (OpenCV, 2023) e *libopencv-dev*. A *darknet* é utilizada tanto pelo computador para realizar o treino de novas iterações da rede neuronal, como pelo Raspberry Pi para efetuar as deteções dos defeitos nos pavimentos rodoviários.

O *darknet* permite realizar a deteção através de imagens, vídeos, ou do input de uma câmara. Durante o processo de deteção, é produzido um *output* de texto, que contém os FPS e a informação sobre os defeitos estruturais detetados, incluindo a precisão e a classe associada. A Figura 8 demonstra o output produzido pela *darknet*.

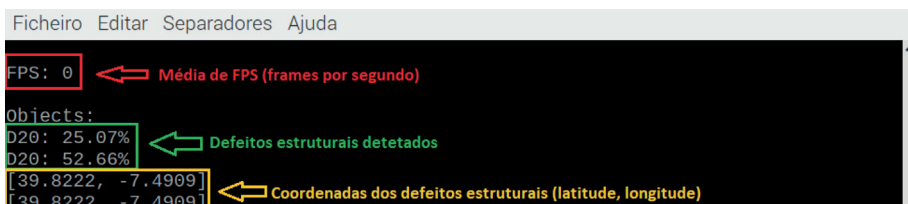


Figura 8 – Output produzido pela *darknet* na deteção de defeitos.

A Figura 9 mostra o comando que deve ser utilizado no *darknet*, para realizar detecções a partir de um vídeo de *input*. Também é possível utilizar o script *darknet_video.py*, que pode ser modificado conforme necessário. No contexto deste trabalho, o *script* foi alterado para criar uma ligação entre o Raspberry Pi e a base de dados MySQL, que se encontra hospedada na máquina remota, e para a aquisição das coordenadas do local de detecção do defeito no pavimento, através do módulo de GPS. O Raspberry Pi comunica com a base de dados alojada remotamente, através do endereço IP da máquina remota, e comunica com o módulo GPS através de uma porta *serial*. Sempre que o Raspberry Pi realiza uma detecção, recebe as coordenadas do local, enviadas a partir do módulo de GPS para a sua porta *serial*. Todas as informações do defeito estrutural, incluindo as coordenadas GPS, são depois enviadas para a base de dados a partir do Raspberry Pi. Para criar a conexão entre o Raspberry Pi e a base de dados, foi importada a biblioteca *mysql.connector* e desenvolvido código Python, para que seja possível o envio da informação.

```
./darknet detector demo data/obj.data yolov4tiny.cfg yolov4.weights -ext_output test.mp4_
```

Figura 9 – Comando para iniciar um teste de detecção e classificação no modelo YOLOv4-tiny.

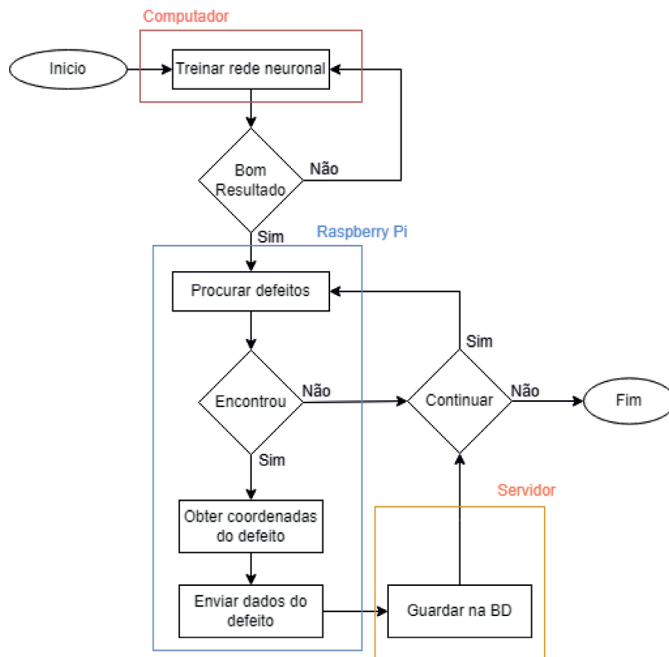


Figura 10 – Fluxograma do funcionamento do protótipo.

Para obter as coordenadas do módulo de GPS, foi importada a biblioteca *pynmea2* e desenvolvido código Python para aceder à porta *serial ttyAMA0*. As coordenadas são

obtidas de meio em meio segundo. Quando é realizada uma deteção de um defeito no pavimento, as informações vindas do módulo de GPS são lidas e formatadas para obter as coordenadas de latitude e longitude.

Após este processo, são enviadas para a base de dados, as informações relativas à deteção, tais como as coordenadas de latitude e longitude, o tipo de defeito estrutural. Quando a base de dados recebe estes dados, associa-lhes um *Id*, para identificar unicamente cada defeito, e a data e hora em que ocorreu a deteção, o que permite poupar processamento ao Raspberry Pi. A Figura 10 apresenta o fluxograma do conceito de funcionamento do protótipo.

A interação com o protótipo é feita sem periféricos, de forma remota, através de *headless* SSH (*Secure Shell*) (Ylonen, 2023). A conexão com o Raspberry Pi por SSH é realizada através da aplicação PuTTY (PuTTY, 2023) a partir do computador. O computador realiza novas iterações ao modelo de modo a aumentar a sua precisão. Posteriormente são enviadas estas iterações para o Raspberry Pi através da aplicação WinSCP (WinSCP, 2023) que permite copiar ficheiros de uma máquina local para uma máquina remota.

Teste e Validação

Para iniciar os testes e validação, e de modo a obter resultados num ambiente controlado, foi utilizado um vídeo com resolução 1920*1080 com extensão *mp4*, no qual existem trechos de pavimentos que contêm vários defeitos estruturais. A resolução de *input* vídeo não é relevante, já que o YOLOv4-tiny redimensiona a imagem durante o processo de deteção. Este redimensionamento é de 416*416, como está descrito no ficheiro de configuração do modelo. Tal resulta em que, independentemente da resolução dos *frames*, o tempo de processamento é o mesmo.

Face ao fraco poder computacional do dispositivo Raspberry Pi, decidiu-se eliminar o *output* que resulta do processamento do vídeo em tempo real. Para tal, foi alterada a *flag* “*-dont_show*” do *darknet*. Assim, o único *output* fornecido seria o texto apresentado na Figura 8. Para verificar o correto funcionamento do modelo, foi configurada a *flag* “*-out_filename*” que permite, no final do processo de deteção, obter um vídeo resultante do modelo em funcionamento, mostrando a deteção e classificação. A Figura 11 mostra um *frame* exemplo do vídeo, obtido após execução do modelo.

Usando estas configurações, foi possível garantir que o Raspberry Pi consegue realizar todas as tarefas necessárias, desde a deteção e classificação dos defeitos estruturais, até ao envio dos dados para a base de dados. No entanto, face às limitações computacionais do Raspberry Pi e à capacidade de processamento necessária para utilizar a rede neuronal, a deteção e classificação acontece a uma média abaixo do 1 FPS, como é mostrado na Figura 8 a vermelho.

O processo de deteção e classificação de defeitos em pavimentos numa simples imagem, permite concluir quanto tempo demora um *frame* a ser processado. Neste sentido, foi realizado um teste numa imagem e foi observado um *output* de aproximadamente 3,6 segundos, como demonstra a Figura 12 a vermelho. Este tempo é elevado e consequentemente não satisfaz o requisito de tempo real. Não garante a deteção dos defeitos nos pavimentos, durante o movimento normal do veículo.



Figura 11 – *Frame* de vídeo após processamento pelo YOLOv4-tiny, mostrando a deteção e classificação.

Esta limitação poderia ser resolvida usando recursos de computação adicionais, como por exemplo o Intel Neural Compute Stick, que se trata de um dispositivo USB (Intel Movidius, 2023) para inferências de *deep learning* o mais próximo possível da fonte de dados. Este dispositivo apresenta 12 núcleos programáveis que permitem acelerar a rede neuronal (Intel Movidius Support, 2023). Com este recurso, o modelo iria ser processado com maior eficiência, perspetivando-se que poderia ser uma solução viável para assegurar a deteção e classificação em tempo real.

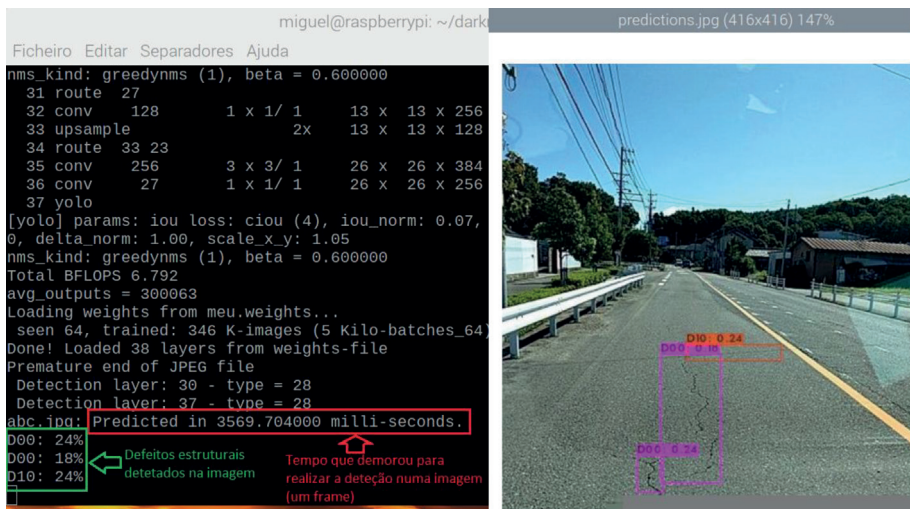


Figura 12 – Teste de deteção e classificação numa imagem.

4. Conclusões e Trabalho Futuro

O trabalho apresentando neste artigo é a segunda etapa de um projeto de investigação em curso, que tem por objetivo propor e avaliar uma solução para automatizar o processo de deteção de defeitos em pavimentos rodoviários, num contexto de cidade inteligente. O estudo incidiu sobre o uso de técnicas de visão computacional, como uma ferramenta tecnológica adequada para este fim.

Partindo de um estudo anterior dos mesmos autores, foi avaliado o desempenho das técnicas de visão computacional consideradas mais adequadas para este cenário: YOLOv4-tiny, RetinaNet e SSD MobileNet. Observou-se que o YOLOv4-tiny apresentou uma precisão superior. Partindo desta conclusão, foi descrita a proposta e processo de implementação de um protótipo funcional, onde foi implementado o YOLOv4-tiny num microcomputador Raspberry Pi, para a deteção e classificação de defeitos nos pavimentos rodoviários. Além de uma câmara responsável pela captação de vídeo, foi utilizado um módulo de GPS, que permite fornecer as coordenadas do microcomputador, instalado no veículo. Cada vez que é detetado um defeito no pavimento, a informação com a classificação do tipo de defeito, a geolocalização, o dia e hora da deteção, é enviada para uma máquina remota para ser armazenada numa base de dados em MySQL. O protótipo foi testado e validado com sucesso.

Permanecem em aberto diversos pontos para trabalho futuro, dos quais se destacam: 1) melhorar o *dataset* de treino e validação; 2) testar e avaliar o desempenho de outros modelos de redes neurais convolucionais; 3) testar alternativas que possam melhorar o desempenho computacional do Raspberry Pi, tais como Intel Neural Compute Stick ou Jetson Nano; 4) implementação de mecanismos de segurança nas trocas de informação entre os vários componentes da solução.

Contribuições dos Autores

Miguel Gonçalves, Tomás Marques: investigação, metodologia, análise formal, validação, preparação de redação-original.

Pedro D. Gaspar: revisão-escrita e edição, supervisão.

Vasco N. G. J. Soares, João M. L. P. Caldeira: análise formal, validação, revisão-escrita e edição, supervisão.

Referências

- Ahmed, N. (2023). mAP Guide. <https://kili-technology.com/data-labeling/machine-learning/mean-average-precision-map-a-complete-guide>
- Amazon (2023). O que é sobreajuste? <https://aws.amazon.com/pt/what-is/overfitting/>
- ARM (2023). What is a CNN. <https://www.arm.com/glossary/convolutional-neural-network>
- Awati, R. (2023). What are Convolutional Neural Networks? <https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network>

- Baeldung. (2023). Training and Validation Loss. <https://www.baeldung.com/cs/training-validation-loss-deep-learning>
- Bochkovskiy, A. (2023). GitHub-AlexeyAB/darknet. <https://github.com/AlexeyAB/darknet>
- Brownlee, J. (2021). Object Recognition With Deep Learning. <https://machinelearningmastery.com/object-recognition-with-deep-learning/>
- Cartucho J. (2018). GitHub-Cartucho/mAP. <https://github.com/Cartucho/mAP>
- Das, A. (2019). Use GPS Module with Raspberry Pi. <https://sparklers-the-makers.github.io/blog/robotics/use-neo-6m-module-with-raspberry-pi/>
- Data Science (2023). what is darknet and why is it needed?. <https://datascience.stackexchange.com/questions/65945/what-is-darknet-and-why-is-it-needed-for-yolo-object-detection>
- Developer NVIDIA (2023). CUDA cuDNN. <https://developer.nvidia.com/cudnn>
- Developer NVIDIA (2023). CUDA Toolkit. <https://developer.nvidia.com/cuda-toolkit>
- Developers google (2022) Como interpretar curvas de perda. <https://developers.google.com/machine-learning/testing-debugging/metrics/interpreting?hl=pt-br>
- Fadheli, A. (2022). Connect to a Remote MySQL Database in Python. <https://thepythoncode.com/article/connect-to-a-remote-mysql-server-in-python>
- Fizyr (2023). GitHub-fizyr/keras-retinanet. <https://github.com/fizyr/keras-retinanet>
- GOEIK (2023). Webcam Full HD. https://www.goeik.com/product.php?prod_id=1966
- Gonçalves, M., Marques, T., Gaspar, P. D., Soares, V. N. G. J., & Caldeira, J. M. L. P. (2023). Road Pavement Damage Detection using Computer Vision Techniques: Approaches, Challenges and Opportunities. *Revista de Informática Teórica e Aplicada*, 30(2), 22–35. <https://doi.org/10.22456/2175-2745.129787>
- Goodfellow, I., Courville, A., & Bengio, Y. (2016). Deep Learning Book. <https://www.deeplearningbook.org/>
- Google Colab. (2023). Colaboratory. <https://colab.research.google.com/>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. <https://arxiv.org/abs/1704.04861v1>
- Huang, R., Pedoeem, J., & Chen, C. (2018). YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers. In: *Proceedings - 2018 IEEE International Conference on Big Data*, Big Data 2018, 2503–2510. <https://doi.org/10.1109/BigData.2018.8621865>

- Huggingface (2023). Fine-tune a pretrained mode. <https://huggingface.co/docs/transformers/training>
- Hui, J. (2018). Real-time Object Detection YOLO, YOLOv2 and YOLOv3. <https://jonathan-hui.medium.com/real-time-object-detection-with-yolo-yolov2-28b1b93e2088>
- Infraestruturas de Portugal. (2023). <https://www.infraestruturasdeportugal.pt/>
- Intel (2023). Intel Movidius Support. <https://www.intel.com.br/content/www/br/pt/support/articles/000033354/boards-and-kits/neural-compute-sticks.html>
- Intel (2023). Intel Movidius. <https://www.intel.com/content/www/us/en/products/sku/125743/intel-movidius-neural-compute-stick/specifications.html>
- Islam, M. M., & Sadi, M. S. (2019). Path hole detection to assist the visually impaired people in navigation. 4th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2018, 268–273. <https://doi.org/10.1109/CEEICT.2018.8628134>
- Javapoint. (2023). Computer Vision Techniques. <https://www.javatpoint.com/computer-vision-techniques>
- Jiang, Z., Zhao, L., Li, S., Jia, Y., & Liquan, Z. (2020). Real-time object detection method based on improved YOLOv4-tiny. <https://arxiv.org/abs/2011.04244v2>
- Kumar, S., Kumar, R. (2023). Real-Time Detection of Road-Based Objects using SSD MobileNet-v2 FPNlite with a new Benchmark Dataset. 2023 4th International Conference on Computing, Mathematics and Engineering Technologies: Sustainable Technologies for Socio-Economic Development, ICoMET 2023. <https://doi.org/10.1109/ICOMET57998.2023.10099364>
- Kundu, R. (2023). YOLO Explained. <https://www.v7labs.com/blog/yolo-object-detection>
- Le, J. (2018). Convolutional Neural Networks: The Biologically-Inspired Model. https://www.codementor.io/@james_aka_yale/convolutional-neural-networks-the-biologically-inspired-model-iq6s48zms
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(2), 318–327. <https://doi.org/10.1109/TPAMI.2018.2858826>
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2016). Feature Pyramid Networks for Object Detection. <https://arxiv.org/abs/1612.03144v2>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9905 LNCS, 21–37. https://doi.org/10.1007/978-3-319-46448-0_2
- Maeda, H. (2023). RDD Dataset. <https://github.com/sekilab/RoadDamageDetector>

- Mehta, V. (2021). Object Detection using SSD Mobilenet V2. <https://vidishmehta204.medium.com/object-detection-using-ssd-mobilenet-v2-7ff3543d738d>
- Nepal, U., & Eslamiat, H. (2022). Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. *Sensors*, 22(2), 464. <https://doi.org/10.3390/s22020464>
- OpenCV. (2023). <https://opencv.org/releases/>
- Ozgon, D. (2021). Top 10 Computer Vision Techniques with Deep Learning. <https://becominghuman.ai/top-10-computer-vision-techniques-with-deep-learning-124fcbd3c20>
- PuTTY. (2023). <https://www.putty.org/>
- Raspberry Pi. (2023) Raspberry Pi 4 Model B. <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>
- Raspberry Pi. (2023). Raspberry Pi OS. <https://www.raspberrypi.com/software/>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. <https://arxiv.org/abs/1804.02767v1>
- Redmon, J.C. (2023). YOLO: Real-Time Object Detection. <https://pjreddie.com/darknet/yolo/>
- Roboflow (2023). YOLOv4 Tiny. <https://roboflow.com/model/yolov4-tiny>
- Roboflow. (2023). <https://roboflow.com/>
- Roboflow. (2023). Roboflow Docs. <https://docs.roboflow.com/>
- Santos, R., & Santos, S. (2023). Guide to GPS Module Arduino. <https://randomnerdtutorials.com/guide-to-neo-6m-gps-module-with-arduino/>
- Shah, D. (2023). Intersection over Union (IoU): Definition, Calculation, Code. <https://www.v7labs.com/blog/intersection-over-union-guide>
- Smeda, K. (2019). Understand the architecture of CNN. <https://towardsdatascience.com/understand-the-architecture-of-cnn-90a25e244c7>
- Supeshala, C. (2020). YOLO v4 or YOLO v5 or PP-YOLO? <https://towardsdatascience.com/yolo-v4-or-yolo-v5-or-pp-yolo-dad8e40f7109>
- Techzizou. (2021). YOLOv4 vs YOLOv4-tiny. Analytics Vidhya. <https://medium.com/analytics-vidhya/yolov4-vs-yolov4-tiny-97932b6ec8ec>
- Tensorflow (2023). GitHub-tensorflow/models. <https://github.com/tensorflow/models/tree/master>
- Tian, H., Zheng, Y., & Jin, Z. (2020). Improved RetinaNet model for the application of small target detection in the aerial images. *IOP Conference Series: Earth and Environmental Science*, 585(1). <https://doi.org/10.1088/1755-1315/585/1/012142>
- VOC 2012. (2012). <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>

- WinSCP. (2023). <https://winscp.net/eng/index.php>
- Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A. M., & Wang, X. (2020). Computer Vision Techniques in Construction: A Critical Review. *Archives of Computational Methods in Engineering*, 28(5),3383–3397. <https://doi.org/10.1007/S11831-020-09504-3>
- Yin, C. T., Xiong, Z., Chen, H., Wang, J. Y., Cooper, D., & David, B. (2015). A literature survey on smart cities. *Science China Information Sciences*, 58(10),1–18. <https://doi.org/10.1007/S11432-015-5397-4>
- Ylonen, T. (2023). What is SSH? <https://www.ssh.com/academy/ssh>
- Yu, H., Chen, C., Du, X., Li, Y., Rashwan, A., Hou, L., Jin, P., Yang, F., Liu, F., Kim, J., & Li, J. (2020). ssdmobilenetv2 Config. https://github.com/tensorflow/models/blob/master/research/object_detection/configs/tf2/ssd_mobilenet_v2_fpnlite_640x640_coco17_tpu-8.config
- Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2018). Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>
- Zvornicanin, E. (2023). What Is YOLO Algorithm? <https://www.baeldung.com/cs/yolo-algorithm>

A relação entre a inovação tecnológica e o desempenho nos meios de hospedagem no contexto da pandemia da Covid-19

Mariana Pires Vidal López¹, Patrícia Oliveira de Freitas², Sandra Martins Lohn Vargas³

marividal@gmail.com; p.defreitas@hotmail.com; sandrabiblio@hotmail.com

¹ Professora Doutora da Universidade Rural do Rio de Janeiro (UFRRJ), 23890-000, Seropédica, Rio de Janeiro, Brasil,

² Professora Doutora da Universidade Rural do Rio de Janeiro (UFRRJ), 23890-000, Seropédica, Rio de Janeiro, Brasil,

³ Professora Doutora do Instituto Federal de Santa Catarina (IFSC), 88075-010, Florianópolis, Santa Catarina, Brasil.

DOI: 10.17013/risti.52.45-60

Resumo: O presente estudo tem como objetivo discutir a relação entre a inovação tecnológica e o desempenho nos Meios de Hospedagens das regiões turísticas do Estado do Rio de Janeiro no contexto da pandemia da Covid-19. A metodologia proposta caracteriza-se como um estudo quantitativo, utilizando-se como instrumento para coleta de dados o questionário, com caráter exploratório e descritivo. A aplicação das pesquisas ocorreu entre 11 de janeiro a 06 de fevereiro de 2023, totalizando 300 entrevistados, baseado no cálculo amostral, optando-se assim pela Análise Fatorial Exploratória. O estudo revelou que os recursos tecnológicos podem influenciar positivamente o desempenho financeiro dos hotéis, fornecendo informações para que os gestores tomem melhores decisões. Compreendeu-se que a pandemia de Covid-19 impactou de forma significativa os Meios de Hospedagens do Estado do Rio de Janeiro tendo sido de extrema importância a adoção de novas tecnologias, principalmente aquelas voltadas aos processos comunicacionais, operacionais e sanitários.

Palavras-chave: Inovação Tecnológica; Desempenho; Hotelaria; Covid-19.

The relationship between technological innovation and performance in accommodation facilities in the context of the Covid-19 pandemic

Abstract: The present study aims to discuss the relationship between technological innovation and performance in accommodation facilities in tourist regions in the State of Rio de Janeiro in the context of the Covid-19 pandemic. The proposed methodology is characterized as a quantitative study, using the questionnaire as an instrument for data collection, with an exploratory and descriptive character. The surveys were applied between January 11th and February 6th, 2023, totaling 300

interviewees, based on the sample calculation, thus opting for Exploratory Factor Analysis. The study revealed that technological resources can positively influence the financial performance of hotels, providing information so that managers can make better decisions. It was understood that the Covid-19 pandemic had a significant impact on accommodation facilities in the State of Rio de Janeiro, with the adoption of new technologies being extremely important, especially those focused on communication, operational and health processes.

Keywords: Technological Innovation; Performance; Hospitality; Covid-19.

1. Introdução

A pandemia ocasionada pelo Covid-19 apresenta quatro diferenças fundamentais em relação a outros eventos catastróficos ocorridos nos últimos anos que a tornam mais grave: intensidade, abrangência geográfica, duração e grau de incerteza, visto que o setor hoteleiro foi extremamente impactado pela pandemia (Hidalgo, Martín-Barroso, Nunez-Serrano, Turrión & Velazquez, 2022). Sabe-se que o setor do turismo ficou sem receita por quase um ano e com hotéis funcionando de forma quase programada nos poucos períodos em que o vírus se mostrou mais ameno, o setor teve queda nas vendas de cerca de 50%, de acordo com os dados da Associação Brasileira das Operadoras de Turismo (Braztoa, 2020).

Nesse sentido, conforme afirma Ivanov, Webster, Stoilova & Slobodskoy (2020), para gerenciar uma crise torna-se necessário o fornecimento de tecnologia de automação às empresas relacionadas ao turismo a fim de mitigar os impactos negativos das ameaças de biossegurança em seu desempenho econômico, sobretudo no gerenciamento de receitas. Nesse contexto, identifica-se que a inovação poderá implementar decisões deliberadas feitas para melhorar a estratégia da empresa, explorando oportunidades de mercado e respondendo aos desafios do ambiente de negócios, o que poderá ajudar o setor econômico no período de pandemia e pós-pandemia quanto aos Meios de Hospedagem.

Nesse sentido, a capacidade de inovação poderá desempenhar um papel fundamental na recuperação de crises e tornar-se uma questão altamente atual à medida que as empresas se esforcem para lidar com a crise da saúde vivenciada em todos os setores, mas especialmente no setor hoteleiro (Breier, Kallmuenzer, Clauss, Gast, Kraus & Tiberius, 2021). Complementando com a ideia de Wooder e Baker (2012) e Ramos, (2021), a inovação tecnológica poderá ser uma combinação de inovações associada ao desenvolvimento de tecnologia com o objetivo de melhorar o produto e serviços existentes de forma incremental ou criando novos.

Assim, entende-se que o presente estudo é de interesse social por contribuir com investigações, discussões e ações referentes ao uso de novas tecnologias e estratégias no referido setor a fim de avaliar como estas influenciam o desempenho nos meios de hospedagem. Nesse sentido, tem-se por objetivo neste estudo discutir a relação entre a inovação tecnológica e o desempenho nos meios de hospedagem das regiões turísticas do Estado do Rio de Janeiro no contexto da pandemia da Covid-19.

2. Revisão de Literatura

2.1. O uso da inovação tecnológica como estratégia organizacional

A organização que apresenta a capacidade de inovar torna-se capaz de responder aos desafios de maneira mais rápida e explorar novos produtos e oportunidades de mercado de forma mais eficaz do que aquelas organizações consideradas não inovadoras. A capacidade de inovar em uma organização significa a capacidade de descobrir e introduzir novos produtos no mercado de maneira que se possa combinar estratégias com comportamentos e processos inovadores e atuar no mercado. Pode-se compreender a capacidade de inovar como um alinhamento das práticas de inovação com as práticas organizacionais, o que consequentemente transforma o conhecimento e aprendizagem em inovações radicais e incrementais (Welter, Sausen & Cappellari, 2019; Froehlich & Nodari, 2021).

A capacidade de inovação significa a capacidade de transformar continuamente conhecimento e ideias em novos produtos, processos e sistemas para o benefício da empresa e de seus stakeholders (Silva, 2021).

Compreende-se que a organização com capacidade de inovar baseia-se também em uma perspectiva de aquisição de novos conhecimentos desenvolvimentos de novas ideias e experiências dentro da organização e junto às suas equipes, ou seja, existem diferentes combinações de capacidades ou diferentes conjuntos específicos de conhecimentos, habilidades e rotinas que permitem que as organizações tornem-se capazes de desenvolver, transformar, adaptar e processar recursos, cujo objetivo é atingir metas de competitividade e de desempenho (Silva, 2021).

Baseado na ideia de Silva (2021), a capacidade de uma organização em inovar baseia-se em alguns fatores essenciais como a liderança transformadora, a intenção estratégica para inovar, o gerenciamento para a inovação, o conhecimento do cliente e do mercado, a gestão estratégica para o uso da tecnologia, a estrutura organizacional, o gerenciamento de projetos e o desempenho baseado na inovação, de forma que tais fatores possam contribuir para a geração de resultados para a organização e seus negócios (Silva, 2021).

A inovação tecnológica está relacionada a novas estratégias implementadas pelas organizações, o que requer novas estruturas organizacionais e mudanças para todos os membros envolvidos na equipe. Nesse sentido, para que ocorra a inovação tecnológica e um melhor desempenho da organização no mercado necessita de uma estrutura organizacional prévia e a capacidade de transformação frente a novos paradigmas (Oliveira & Avellar, 2020; Ramos, 2021).

Nesse sentido, a busca pelo melhor desempenho frente ao mercado e ao cliente poderá aumentar a capacidade da organização em inovar, de forma a aprimorar o processo de produção e permitir o desenvolvimento de novos produtos e serviços, práticas organizacionais e promover novos conhecimentos (OECD, 2018). Destaca-se aqui a inovação tecnológica, sendo esta uma combinação de inovações associada ao desenvolvimento de tecnologia, cujo objetivo é melhorar os produtos e os serviços existentes na organização de forma incremental ou criando novos produtos ou serviços para o mercado (Wooder & Baker, 2012).

2.2.A prática da inovação tecnológica nos meios de hospedagens

Com a revolução digital houve uma grande transformação na operação e na gestão hoteleira, onde as tecnologias digitais têm sido reconhecidas como as principais fontes de eficiência e vantagem competitiva no hotel. Nos últimos anos o setor hoteleiro obteve uma explosão de novas tecnologias e testemunharam como as tecnologias inovadoras remodelam a indústria hoteleira. Tecnologias digitais, como tecnologias de autoatendimento, robôs, *blockchain*, análise de big data, aplicativos móveis, mídia social e sistemas de informação, podem alterar a forma como os hotéis gerenciam suas gestões/operações como também seus *stakeholders* (Choi, Choi, Oh & Kim, 2020; Mingotto, Montaguti & Tamma, 2021; Ramos, 2021).

Os hotéis podem usar essas tecnologias para gerenciar serviços, pedidos e processos, porém demanda diferentes capacidades e recursos, relacionamento com o cliente e o fornecedor, alcançando um melhor desempenho em termos financeiros, vantagem competitiva, qualidade nos serviços, melhor utilização de recursos, flexibilidade e inovação em seus negócios (Ivanov *et al.*, 2020; Lenuwat & Boon-itt, 2021).

Orfila-Sintes e Mattsson (2009) salientam que empresas hoteleiras mostram uma maior propensão em inovar através da introdução de Pesquisa e Desenvolvimento (P&D) incorporada em tecnologia (novos equipamentos e/ou *software*) e uma menor predisposição para realizar atividades internas de P&D. Baseado no estudo de Guisado-González, Guisado-Tato e Sandoval-Pérez (2013), perceberam que a inovação mais comum na hotelaria como estratégia é a aquisição de novas tecnologias através da compra de máquinas e equipamentos, seguido pela formação do pessoal envolvido na execução das atividades, a compra de tecnologia intangível e atividades de marketing relacionadas com a introdução de novos serviços.

Em suma, a inovação no setor hoteleiro é ilustrada na adoção de novas tecnologias e se traduz em vantagens competitivas, seja por eficiência produtiva, entendida como a redução de custos, ou uma maior capacidade de diferenciação através de melhorias nos serviços e sua maior adaptação às necessidades do cliente. Especificamente, a inovação neste setor está amplamente associada a avanços em tecnologias da informação (Nieves, 2017).

2.3.A influência do desempenho na estratégia organizacional

Desempenho ou performance organizacional é um conceito amplo e complexo em razão de diferentes empresas poderem ter indicadores de desempenho distintos, dependendo da natureza de suas estratégias. Deve-se considerar ainda que diferenças estruturais entre os vários mercados de atuação de uma empresa, seja local, regional ou internacional, podem conduzir a diferentes variáveis que podem determinar sua performance (Oliveira & Avellar, 2020; Ramos, 2021).

A estratégia pode ser definida como a forma para se pensar o futuro integrada ao processo decisório das organizações, assim como baseada em procedimentos formalizados e articulados pelos gestores (Eisenhardt & Santos, 2000).

Para que as organizações possam criar um ambiente organizacional inovador, espera-se que tenha um conjunto mínimo de capacidades criativas, de maneira que consiga

executar suas atividades e gerar novas ideias de forma integrada com suas equipes (Silva, Arvela & Ramos, 2021).

Nesse contexto, o uso da tecnologia torna-se um recurso necessário para gerar e gerenciar atividades inovadoras em produtos, processos e organização da produção, sistemas organizacionais, equipamentos e engenharia de projetos. Os recursos estão incorporados nas habilidades, experiências e qualificações formais dos profissionais pertencentes às equipes, assim como, no sistema organizacional, rotinas e procedimentos da empresa (Silva, Arvela & Ramos, 2021).

Compreende-se que a utilização de recursos tecnológicos é “um critério de desempenho que avalia a eficiência com que os recursos são utilizados na prestação de serviços” (Cho *et al.*, 2012, p. 807). Nos estudos de Jeong, Lee & Nagesvaran (2016) foram testados empiricamente os impactos do uso de dispositivos móveis na equipe de hotéis de luxo nos Estados Unidos pode-se encontrar uma associação positiva entre o uso de dispositivos móveis e o desempenho da equipe do hotel. Os autores argumentam que os dispositivos móveis permitem que a equipe do hotel execute tarefas relacionadas ao trabalho em qualquer local com eficiência.

Seguindo os estudos de Cho *et al.* (2012), as dimensões de desempenho do hotel podem ser categorizadas como desempenho financeiro, competitividade, qualidade do serviço, utilização de recursos, flexibilidade e inovação. Em estudos realizados na Malásia, Ahmad e Scott (2019) constataram que, além da redução de custos, os recursos tecnológicos podem influenciar positivamente o desempenho financeiro dos hotéis, fornecendo informações para que os gestores tomem melhores decisões e proporcionando aos colaboradores mais tempo para focar em funções estratégicas e ser mais proativo no planejamento.

Baseado em estudos de Garrido-Moreno, García-Morales, King & Lockett (2020) realizados em hotéis na Espanha, estes descobriram que o uso de mídias sociais tem uma influência positiva no desempenho do hotel, aprimorando os recursos de engajamento do cliente e o gerenciamento de relacionamento com o cliente social, citando aqui como ferramenta o uso do CRM (*Customer Relationship Management*) como forma de melhorar o seu relacionamento com o cliente. Em outro estudo citado por Palácios-Marqués, Merigó & Soto-Acosta (2015) sobre hotéis quatro e cinco estrelas na Espanha, apresentam também que o uso de mídias sociais gera maior desempenho e maior capacidade de inovação no setor hoteleiro.

Além disso, Garrido-Moreno *et al.*, (2020) citam ainda em seu estudo que os hotéis podem sim reduzir os custos das atividades de marketing e se comunicar diretamente com os clientes utilizando plataformas de mídia social que podem reduzir custos, gerar aumento na imagem da marca e crescimento de vendas, ou seja, o uso de mídias sociais poderá gerar o aumento da capacidade de inovação dos hotéis, o que também permite entender e antecipar os desejos e necessidades dos clientes.

De acordo com Ivanov *et al.* (2020), os custos operacionais podem ser reduzidos de maneira que permita que os profissionais atuantes nos hotéis possam ter um desempenho mais eficiente e eficaz. O uso dos robôs, por exemplo, pode gerar receita adicional ao fornecer serviços adicionais, permitindo que os hotéis ofereçam aos clientes experiências robóticas e diferenciadas, assim como possam liberar os colaboradores para executar demais atividades de valor agregado. Além disso, o uso das tecnologias de

autoatendimento também pode reduzir os custos de mão de obra. No entanto, embora os robôs e as tecnologias de autoatendimento possam reduzir os custos operacionais dos hotéis, eles devem ser usados com cautela devido à importância do atendimento personalizado, interação e emoção no setor hoteleiro.

Salienta-se que os comentários online dos clientes são importantes para a gestão da qualidade e os hotéis podem melhorar a qualidade dos serviços com base nas avaliações publicadas, visto que sistemas inteligentes, aplicativos móveis, robôs e tecnologias de autoatendimento tendem a aprimorar a experiência do cliente (Leung, 2019; Silva, Arvela & Ramos, 2021). Assim como, os autores apontam que o uso das ferramentas tecnológicas facilita as reservas de serviços, como spa, restaurante e reservas de quartos, oferece controle de tecnologia no quarto (luzes, TV, termostatos), fornece serviços pessoais em tempo real, como coleta de roupas e serviços de limpeza, ou outros pedidos especiais que os clientes possam solicitar (Lei, Wang & Law, 2019).

Sabe-se que no período de pandemia da COVID-19 as tecnologias demonstraram desempenhar um papel importante na superação dos desafios trazidos à tona pela pandemia, por exemplo, devido ao COVID-19, as pessoas preferiram experiências menos arriscadas e a realidade virtual permitiu que os clientes avaliassem os riscos antes de realizarem suas reservas (Schiopu, Hornoiu, Padurean & Nica, 2021). No entanto, a análise das estratégias de gestão hoteleira e suas implicações no desempenho tornou-se uma questão urgente à luz da crise global derivada da pandemia do COVID-19 (Marco-Lajara *et al*, 2021).

Entende-se que quando se busca melhorar o desempenho da empresa, aumentará a capacidade de inovar, assim como estará aprimorando o processo de produção e permitindo o desenvolvimento de novos produtos e práticas organizacionais, o que consequentemente também fará aumentar a capacidade da empresa em adquirir e criar novos conhecimentos (OECD, 2018).

3. Metodologia

A metodologia proposta nesta pesquisa caracteriza-se como um estudo quantitativo, utilizando-se como instrumento para coleta de dados o questionário, permitindo trazer para o estudo um caráter exploratório e descritivo. Na primeira etapa da pesquisa, se estendendo por todo o período de sua realização, foram feitos levantamentos bibliográficos que tratam da temática da inovação tecnológica e desempenho. Utilizou-se a pesquisa bibliográfica para estruturar a fundamentação teórica, a fim de adotar artigos nacionais e internacionais, pesquisados em Bases de Dados como EBSCO, *Web Of Science*, SCIELO, Portal da Capes e *Scientific Periodicals Eletronic Library* (SPELL).

A abrangência proposta para a investigação teve como universo os Meios de Hospedagem das regiões turísticas do Estado do Rio de Janeiro, como forma de analisar a inovação tecnológica adotada no período de pandemia da Covid-19. A segunda etapa da pesquisa envolveu a elaboração do material de coleta de dados, a realização de pré-teste e aplicação do questionário aos gestores dos Meios de Hospedagem, sendo hotéis, pousadas, hostel/albergues, resorts, entre outros do Estado do Rio de Janeiro, conforme tabela 1 representando as regiões pesquisadas.

Regiões	Universo	%	Entrevistas
Metropolitana	281	14,4	43
Agulhas Negras	56	2,9	9
Baixada Fluminense	16	0,8	2
Caminhos da Mata	8	0,4	1
Costa Doce	14	0,7	2
Costa do Sol – Região dos Lagos	865	44,4	134
Costa Verde	443	22,7	68
Noroeste das Águas	16	0,8	2
Serra Norte	19	1,0	3
Serra Verde Imperial	153	7,8	24
Vale do Café	79	4,1	12
TOTAL	1.950	100,0	300

Fonte: Dados primários (2023).

Tabela 1 – Regiões pesquisadas

A população alvo deste estudo foi constituída por gestores localizados nas diferentes regiões do Estado do Rio de Janeiro, conforme descrito no quadro. Inicialmente foi feito um mapeamento do total de Meios de Hospedagem no CADASTUR-RJ que é o Cadastro de Prestadores de Serviços Turísticos, do Ministério do Turismo (MT), sendo definido, no site do MT, como cadastro de pessoas físicas e jurídicas que atuam no setor turístico, sendo obrigatório para Meios de Hospedagem, Agências de Turismo, Transportadoras Turísticas, Organizadoras de Eventos, Parques Temáticos, Acampamentos Turísticos e Guias de Turismo-MEI.

A opção por investigar o Estado como um todo tem o objetivo de perceber como há predomínio da adoção de determinado modelo de inovação tecnológica em alguma região e se essas ferramentas podem ser aplicadas por regiões que não as adotem.

Quanto ao plano amostral, o universo correspondia a 1.950 participantes, sendo o número total de entrevistas 300 gestores, com margem de erro para cada público de $\pm 5,2\%$ e coeficiente de confiança de $\pm 95\%$. As entrevistas ocorreram por telefone e foram realizadas entre 11 de janeiro a 06 de fevereiro de 2023, por agentes pesquisadores com treinamento específico para esta tarefa. Salienta-se que foram realizados os pré-testes com 30 Meios de Hospedagens e não havendo alteração no instrumento foi dada continuidade.

Por meio dos questionários respondidos pôde-se dar início ao tratamento e análise dos dados, utilizando-se primeiramente planilhas eletrônicas para organizar os dados coletados. Nesse caso, o software Excel pôde auxiliar na organização dos dados coletados e na sequência houve a importação dos dados para o software estatístico *SPSS Statistics* versão 21.

Nesta pesquisa optou-se pela Análise Fatorial Exploratória para a análise estatística dos dados, buscando-se a redução de um determinado número de variáveis. Entende-se que a Análise Fatorial Exploratória (AFE) gera fatores e a Análise de Componentes Principais (ACP) gera componentes, sendo que a diferença entre fatores e componentes está relacionada à forma como os itens são retidos. As duas técnicas de redução de dados, partem do pressuposto de que cada item do questionário tem 3 tipos de variâncias, sendo elas: a específica, a comum e a de erro.

No caso desta análise, primeiramente foi feita uma Análise Fatorial Exploratória, partindo de como foi elaborado e aplicado o questionário de 18 perguntas distribuídas em 2 blocos, sendo eles a Capacidade de Inovação (Bloco 1) e o Desempenho (Bloco 2), além de 4 perguntas referentes à caracterização dos Meios de Hospedagens. Esta é uma análise utilizada como uma etapa na criação e validação dos dados, sendo ela muito mais teórica que matemática. Salienta-se que analisando a estatística descritiva, todas as questões foram respondidas por todos os entrevistados (300) e não houve omissão nas respostas.

O instrumento para mensurar capacidade de inovação foi baseado nos modelos de Gallouj (2002) e Jimenez-Jimenez & Sanz-Vale (2011), baseado em seus artigos e posterior tradução dos respectivos instrumentos, e para o desempenho, foi baseado nos modelos de Zou, Taylor & Osland (1998), conforme quadro 1.

Foi realizado o teste KMO para avaliar a adequação do tamanho da amostra, sendo que a variação desse índice vai de 0 a 1, sendo que zero é inadequado para a análise fatorial, sendo aceitável se for maior que 0,5 e recomendado acima de 0,8. Foi realizado o Teste de esfericidade sendo este um teste estatístico para verificação da presença de correlação entre as variáveis, sendo que irá medir a adequação da análise, devendo-se assim encontrar um p-valor (Sig) abaixo de 0,05.

As comunalidades ou Cos2, são quantidades das variâncias (correlações) de cada variável explicada pelos fatores, visto que quanto maior a comunalidade, maior será o poder de explicação daquela variável pelo fator. Cos2 é uma medida que indica a qualidade da representação para variáveis. A comunalidade é a soma dos quadrados das correlações entre cada variável e a componente principal. As comunalidades são quantidades das variâncias (correlações) de cada variável explicada pelos fatores. Salienta-se que quanto maior a comunalidade, maior será o poder de explicação daquela variável pelo fator. Nesse caso, deseja-se comunalidades superiores a 0,5, ou seja, superiores a 50%.

4. Resultados e Discussão

Nesse momento serão descritos os dados dos Meios de Hospedagens pesquisados, sendo o total de respondentes 300 empresas em regiões do Estado do Rio de Janeiro, no período entre 11 de janeiro a 06 de fevereiro de 2023, realizadas por agentes pesquisadores com treinamento específico e por telefone.

Dos Meios de Hospedagens pesquisados, 134 (44,7%) estão localizadas na Costa do Sol - Região dos Lagos, 68 (22,7%) na Costa Verde, 43 (14,3%) na Região Metropolitana,

24 (8%) na Região da Serra Verde Imperial, 12 (4%) na Região do Vale do Café, 9 (3%) na Região de Agulhas Negras, 3 (1%) na Região da Serra Norte, 2 (0,7%) na Região da Baixada Fluminense, 2 (0,7%) na Região da Costa Doce, 2 (0,7%) na Região do Noroeste das Águas, e 1 (0,3%) na Região de Caminhos da Mata.

Constructos	Variáveis	Autores
Capacidade de Inovação	Q5. O Meio de Hospedagem adotou um protocolo com medidas de prevenção e controle da infecção por vírus (Covid-19) baseado nos Órgãos Públicos Sanitários.	Gallouj (2002) e Jimenez-Jimenez e Sanz-Vale (2011)
	Q6. Antes da pandemia, recursos/ferramentas tecnológicas eram utilizados no Meio de Hospedagem.	
	Q7. O contexto gerado pela pandemia impulsionou a implementação de inovações tecnológicas no Meio de Hospedagem.	
	Q8. A implementação de novas tecnologias (ferramentas e recursos) foi eficiente para a operacionalização do Meio de Hospedagem durante a pandemia.	
	Q9. O comportamento da maioria dos hóspedes demonstra aceitação do uso gradativo de tecnologias em substituição aos serviços prestados pelos colaboradores do Meio de Hospedagem.	
	Q10. Quanto ao uso de novas ferramentas tecnológicas, o Meio de Hospedagem oferece programas formais de treinamento e qualificação para seus colaboradores.	
	Q11. O Meio de Hospedagem estimula o comprometimento dos colaboradores para a criatividade e a inovação quanto ao uso de novas ferramentas tecnológicas.	
	Q12. A pandemia causou impactos na força de trabalho do Meio de Hospedagem quanto ao uso de novas ferramentas tecnológicas.	
	Q13. O Meio de Hospedagem desenvolve habilidades para adotar novas ideias, tecnologias ou processos para oferecer produtos e serviços antes dos concorrentes.	
	Q14. O Meio de Hospedagem nos últimos 2 anos tem valorizado a inovação em vez de focar na comercialização de produtos/serviços já existentes.	
	Q15. O Meio de Hospedagem tem ou teve receita disponível para implementar inovações tecnológicas no momento da pandemia.	
	Q16. Para o Meio de Hospedagem é importante o uso das inovações tecnológicas visando medidas para o futuro do negócio.	
Desempenho	Q17. O Meio de Hospedagem nos últimos 2 anos aumentou as vendas para novos clientes com o uso de OTA's (plataformas como: Booking, Tripadvisor, hoteis.com).	Zou, Taylor e Osland (1998)
	Q18. O Meio de Hospedagem nos últimos 2 anos aumentou as vendas para os próprios clientes com o uso de novas tecnologias.	
	Q19. O Meio de Hospedagem nos últimos 2 anos aumentou sua produtividade com o uso de novas tecnologias.	
	Q20. O Meio de Hospedagem nos últimos 2 anos melhorou sua competitividade com o uso de novas tecnologias.	
	Q21. O Meio de Hospedagem nos últimos 2 anos melhorou seu posicionamento estratégico no mercado com o uso de novas tecnologias.	
	Q22. O Meio de Hospedagem nos últimos 2 anos aumentou sua participação de mercado com o uso de novas tecnologias.	

Fonte: Dados da pesquisa (2023).

Quadro 1 – Dimensões e variáveis investigadas para capacidade de inovação e desempenho

Quanto à classificação dos Meios de Hospedagens pesquisados, identifica-se que 105 (35%) são classificados como Hotel, 159 (53%) são pousadas, 6 (2%) são classificados como Hotel Fazenda, 2 (0,7%) são Resorts e 28 (9,3%) são classificados como Hostel ou Albergue. Salienta-se que 33% dos Meios de Hospedagens fazem parte de uma rede hoteleira e 67% não pertencem à rede hoteleira.

Quanto ao porte dos Meios de Hospedagens, identifica-se que em empresas consideradas de pequeno porte (até 100 unidades habitacionais) predominam as Pousadas (58,7%), de médio porte (de 101 até 399 unidades habitacionais) e de grande porte (mais de 400 unidades habitacionais), predominam os hotéis, 87,5% e 75%, respectivamente.

Na sequência, a análise fatorial exploratória foi realizada para identificar se as discussões e ações referentes ao uso das novas tecnologias e estratégias nos meios de hospedagem influenciam o desempenho nos meios de hospedagem. Os itens que compõem as variáveis se agrupam nos aspectos abordados e se necessário pode-se reduzir o número de itens que compõem cada aspecto. Seguindo o critério indicado por Hair *et al.* (2005) foram considerados em cada fator os itens com coeficientes de carga fatorial acima de 0,5, obedecendo também os critérios de variância explicada já descritos na metodologia.

Observa-se que no teste de KMO foi encontrado o valor 0,86, sendo um indicador que demonstra que a amostra está adequada para a realização da análise fatorial. Caso esse índice fosse abaixo de 0,5 o estudo não seria feito, sendo que pelo teste de esfericidade de Bartlett, a matriz possui esfericidade se a Significância for abaixo de 0,05.

Foi gerada a Matriz de Correlações, sendo que a matriz gerada possui Determinante = 0,001. Na sua diagonal todos os valores obtidos são iguais a um, verificando-se que acima do eixo diagonal que contém seus valores espelhados com os valores abaixo da diagonal central. Conforme descrito na metodologia o questionário elaborado e aplicado foi distribuído em dois blocos sendo eles a Capacidade de Inovação (Bloco 1) e o Desempenho Geral (Bloco 2). Os resultados da análise fatorial destes serão discutidos a seguir.

4.1. Análise fatorial da Inovação Tecnológica

No bloco da Inovação Tecnológica houve alterações das variáveis das questões sinalizadas na construção do questionário. As questões 20, 21 e 22 que anteriormente eram do bloco do desempenho vieram para o bloco da inovação tecnológica. Também se destaca que as questões 5 e 6 foram descartadas pois não tiveram significância. A Tabela 2, é composta pelas comunalidades e as cargas fatoriais dos itens que compõem o fator extraído, o total de variância explicada pelos itens componentes de cada fator, bem como o valor de KMO.

Questões do instrumento	Comunalidade	Carga Fatorial
Questão 13	0,543	0,784
Questão 11	0,524	0,752
Questão 10	0,500	0,725
Questão 14	0,422	0,665
Questão 21	0,738	0,599

Questões do instrumento	Comunalidade	Carga Fatorial
Questão 22	0,687	0,578
Questão 20	0,559	0,556
Variância Explicada	32,62%	
KMO	0,861	

Fonte: Dados da pesquisa (2023).

Tabela 2 – Valores da Comunalidade, Carga Fatorial, Variância Explicada e KMO (Inovação)

Verifica-se que os resultados alcançados pela comunalidade, carga fatorial, variância explicada e KMO demonstraram que todas as variáveis obtiveram carga fatorial adequada. As comunalidades das variáveis obtiveram seus resultados acima de 0,5, exceto na questão 14, porém a carga fatorial da questão está alta, o que sinaliza uma maior correlação.

As cargas fatoriais foram extraídas acima de 0,5, porém houve questões que não se agruparam com as outras variáveis, o que sinaliza a falta de importância que os entrevistados dão a elas, sendo uma variável que poderia ser mais bem trabalhada na hotelaria. A variância total explicada demonstra que este constructo está sendo explicado em 32,62%, atendendo ao critério de Hair *et al.* (2005), e o teste de KMO foi de 0,86, o que demonstra a adequação da amostra.

Observa-se que o fator da capacidade de inovação corrobora com os estudos de Guisado-González, Guisado-Tato & Sandoval-Pérez (2013), que perceberam que a inovação mais comum na hotelaria como estratégia é a aquisição de novas tecnologias através da compra de máquinas e equipamentos, seguido pela formação do pessoal envolvido na execução das atividades, a compra de tecnologia intangível e atividades de marketing relacionadas com a introdução de novos serviços. Verifica-se que a inovação neste setor está amplamente associada a avanços em tecnologias da informação (Orfila-Sintes & Mattsson, 2009; Fernández, Cala & Domecq, 2011; Nieves, 2017).

Como já citado anteriormente, as tecnologias digitais, como as tecnologias de autoatendimento, robôs, *blockchain*, análise de big data, aplicativos móveis, mídia social e sistemas de informação, surgem para alterar a maneira como os hotéis gerenciam suas operações. E para que ocorra a inovação tecnológica e consequentemente um melhor desempenho das organizações no mercado demanda diferentes capacidades e recursos, relacionamento com o cliente e o fornecedor. Dessa forma, torna-se possível alcançar um melhor desempenho em termos financeiros, vantagem competitiva, qualidade nos serviços, melhor utilização de recursos, flexibilidade e inovação em seus negócios (Cho *et al.*, 2012; Ivanov *et al.*, 2020; Oliveira & Avellar, 2020, Anser *et al.*, 2020; Choi *et al.*, 2020; Mingotto, Montaguti, & Tamma, 2021; Lenuwat & Boon-itt, 2021).

4.2. Análise fatorial do desempenho

Quanto ao bloco de desempenho, também houve alterações das variáveis das questões sinalizadas na construção do questionário. As questões 7, 8, 15 e 12, que anteriormente eram do bloco da inovação tecnológica, vieram para o bloco do desempenho. A tabela

3, é composta pelas comunalidades, as cargas fatoriais dos itens que compõem o fator extraído, o total de variância explicada pelos itens componentes de cada fator, bem como o valor de KMO.

Questões do instrumento	Comunalidade	Carga Fatorial
Questão 7	0,449	0,673
Questão 8	0,446	0,643
Questão 19	0,589	0,603
Questão 17	0,414	0,543
Questão 15	0,312	0,533
Questão 12	0,244	0,517
Variância Explicada	9,20%	
KMO	0,861	

Fonte: Dados da pesquisa (2023).

Tabela 3 – Valores da Comunalidade, Carga Fatorial, Variância Explicada e KMO (Desempenho)

Observa-se que os resultados alcançados pela comunalidade, carga fatorial, variância explicada e KMO demonstraram que todas as variáveis obtiveram carga fatorial adequada. Nem todas as comunalidades das variáveis obtiveram seus resultados acima de 0,5, no entanto, Stevens (1992) sinaliza que com um número menor do que 20 variáveis podem-se ter comunalidades baixas (< 0,4).

As cargas fatoriais foram extraídas acima de 0,5, o que sinaliza que os entrevistados dão importância a essas variáveis. A variância total explicada demonstra que este constructo está sendo explicado em 9,20%, atendendo ao critério de Hair *et al.* (2005), e o teste de KMO foi de 0,86, o que demonstra a adequação da amostra.

Verifica-se que o fator de desempenho corrobora com o estudo de Ivanov *et al.* (2020), que sinaliza que os custos operacionais podem ser reduzidos de maneira que permita que os profissionais atuantes nos hotéis possam ter um desempenho mais eficiente e eficaz. Já no estudo de Ahmad e Scott (2019), estes sinalizam que os recursos tecnológicos podem influenciar positivamente o desempenho financeiro dos hotéis, fornecendo informações para que os gestores tomem melhores decisões e proporcionando aos colaboradores mais tempo para focar em funções estratégicas e ser mais proativo no planejamento.

Já foi citado neste estudo que o uso das tecnologias se torna também um critério de desempenho capaz de avaliar a eficiência com que os recursos são utilizados na prestação dos serviços nos hotéis. Pode-se citar como exemplo o impacto sobre o uso de dispositivos móveis nos hotéis, visto que os dispositivos móveis permitem que a equipe do hotel execute tarefas relacionadas ao trabalho em qualquer local com maior eficiência. Dessa forma, as dimensões de desempenho em um hotel poderão ser categorizadas como desempenho financeiro, competitividade, qualidade do serviço, utilização de recursos, flexibilidade e inovação (Cho *et al.*, 2012; Jeong, Lee & Nagesvaran, 2016).

Nesse contexto, compreende-se que quando se busca melhorar o desempenho dos hotéis, estará aumentando a capacidade de inovar, aprimorando o processo de produção e permitindo o desenvolvimento de novos produtos e práticas organizacionais (OECD, 2018).

5. Conclusões

Este artigo teve como objetivo discutir a relação entre a inovação tecnológica e o desempenho nos Meios de Hospedagens das regiões turísticas do Estado do Rio de Janeiro no contexto da pandemia da Covid-19.

Com base nos resultados foi possível perceber, na perspectiva dos respondentes, que a adoção de novas ideias, tecnologias ou processos melhorou o posicionamento estratégico e a competitividade no mercado, bem como, aumentou a participação de mercado, a produtividade, vendas para novos clientes com o uso de OTA's (Agências de Viagens Online). Vale ressaltar ainda que houve um reconhecimento de que o contexto gerado pela pandemia impulsionou a implementação de inovações tecnológicas.

Foram apontados que são desenvolvidas habilidades para adotar novas ideias, tecnologias ou processos para oferecer produtos e serviços antes dos concorrentes, sendo oferecidos programas formais de treinamento e qualificação para seus colaboradores. Quanto ao uso de novas ferramentas tecnológicas, os colaboradores são estimulados para a criatividade e a inovação quanto ao uso de novas ferramentas tecnológicas.

Foram citadas como principais ferramentas tecnológicas adotadas nos últimos dois anos, respectivamente: *WhatsApp, Booking, Instagram, Facebook*, Site da organização, Sistema/Programas/Motor de reservas, *Check-in/out online*, Email, *TripAdvisor*, *Hoteis.com*, *QR Code*, *Airbnb*, *Expedia*, Decolar, PIX e 123 milhas, que não eram utilizadas de forma significativa antes da pandemia.

Vale ressaltar que há diferenças significativas entre as regiões turísticas do Estado do Rio de Janeiro, onde em determinadas regiões como a Metropolitana predominam hotéis e em outras regiões como a Costa Verde e a Costa do Sol – Região dos Lagos, onde predominam as pousadas, por se tratar de regiões voltadas ao turismo de praia e sol onde há uma forte presença de pessoas que residem no próprio Estado e que muitas vezes buscam acomodações com custo mais baixo. Baseado nessa percepção, sugere-se para estudos futuros uma análise mais aprofundada por regiões a partir de um olhar mais criterioso e individualizado.

A pandemia de Covid-19 impactou de forma significativa os Meios de Hospedagens do Estado do Rio de Janeiro tendo sido de extrema importância a adoção de novas tecnologias, principalmente aquelas voltadas aos processos comunicacionais, operacionais e sanitários. Como já demonstrado em estudos anteriores, algumas destas inovações cumpriram seu papel em um determinado momento, mas outras vieram para ficar, pois permitem aos Meios de Hospedagens se conectar com novas realidades, dinâmicas e entender melhor novos hábitos, práticas e visões dos hóspedes pós pandemia. Porém, recomenda-se que os Meios de Hospedagens ao fortalecer e ampliar o uso das tecnologias preservem a humanização da relação com seus hóspedes, pois ainda

que a inovação tecnológica tenha facilitado a comunicação e a gestão dos processos, os hóspedes ainda valorizam o contato pessoal.

Referências

- Ahmad, R. & Scott, N. (2019). Technology innovations towards reducing hospitality human resource costs in Langkawi. *Malays. Tourism Review*, 74, 547–562. <https://doi.org/10.1108/TR-03-2018-0038>
- Anser, M. K., Yousaf, Z., Usman, M. & Yousaf, S. (2020). Towards strategic business performance of the hospitality sector: nexus of ICT, e-marketing and organizational readiness. *Sustainability*, 12, 1346. <https://doi.org/10.3390/su12041346>
- Braztoa. (2020). Associação Brasileira das Operadoras de Turismo. Setor hoteleiro vê queda de ocupação no carnaval e faz projeção para 2021. Recuperado de <https://www.braztoa.com.br/>.
- Breier, M., Kallmuenzer, A., Clauss, T., Gast, J., Kraus, S. & Tiberius, V. (2021). The role of business model innovation in the hospitality industry during the covid-19 crisis. *International Journal of Hospitality Management*, 92, 102723. <https://doi.org/10.1016/j.ijhm.2020.102723>
- Choi, Y., Choi, M., Oh, M. & Kim, S. (2020). Service robots in hotels: understanding the service quality perceptions of human-robot interaction. *Journal of Hospitality Marketing & Management*, 29, 613–635. <https://doi.org/10.1080/19368623.2020.1703871>
- Eisenhardt, K. & Santos, F. (2000). Knowledge-Based View: A New Theory of Strategy? In: Pettigrew, A., Thomas, H., Whittington, R. (Eds.). *Handbook of Strategy and Management*. London: Sage Publications. <https://doi.org/10.4135/9781848608313>.
- Fernández, J. I. P., Cala, A. S. & Domecq, C. F. (2011). Critical external factors behind hotels' investments in innovation and technology in emerging urban destinations. *Tourism Economics*, 17(2), 339–357. <https://doi.org/10.5367/te.2011.0033>
- Froehlich, C. & Nodari, C. H. (2021). Dynamic Capabilities and innovation in health services. *Revista Eletrônica e Estratégia & Negócios*, 14 (2), 115-135. <https://doi.org/10.19177/reen.v14e22021115-135>
- Gallouj, F. (2002). Innovation in the service economy: the new wealth of nations. Cheltenham: Edward Elgar. <https://doi.org/10.4337/9781843765370>
- Garrido-Moreno, A., García-Morales, V., King, S. & Lockett, N. (2020). Social media use and value creation in the digital landscape: a dynamic-capabilities perspective. *Journal of Service Management*, 31, 313–343. <https://doi.org/10.1108/JOSM-09-2018-0286>
- Guisado-González, M., Guisado-Tato, M. & Sandoval-Pérez, Á. (2013). Determinants of innovation performance in Spanish hospitality companies: Analysis of the coexistence of innovation strategies. *The Service Industries Journal*, 33(6), 580–593. <https://doi.org/10.1080/02642069.2011.614343>

- Hidalgo, A., Martín-Barroso, D., Nunez-Serrano, J., Turrion, J. & Velazquez, F. (2022). Does hotel management matter to overcoming the COVID-19 crisis? The Spanish case. *Tourism Management*, 88. <https://doi.org/10.1016/j.tourman.2021.104395>
- Ivanov, S., Webster, C., Stoilova, E. & Slobodskoy, D. (2020). Biosecurity, crisis management, automation technologies and economic performance of travel, tourism and hospitality companies – A conceptual framework. *Tourism Economics*, 28(1). <https://doi.org/10.1177/1354816620946541>
- Jeong, M., Lee, M. & Nagesvaran, B. (2016). Employees' use of mobile devices and their perceived outcomes in the workplace: a case of luXury hotel. *International Journal of Hospitality Management*, 57, 40–51. <https://doi.org/10.1016/j.ijhm.2016.05.003>
- Jimenez-Jimenez, D. & Sanz-Valle, R. (2011). Innovation, organizational learning, and performance. *Journal of Business Research*, 64, 408–417. <https://doi.org/10.1016/j.jbusres.2010.09.010>
- Lei, S.I., Wang, D. & Law, R. (2019). Hoteliers' service design for mobile-based value co-creation. *International Journal of Contemporary Hospitality Management*, 31 (11), 4338–4356. <https://doi.org/10.1108/IJCHM-03-2018-0249>
- Lenuwat, P. & Boon-itt, S. (2021). Information technology management and service performance management capabilities: an empirical study of the service supply chain management process. *Journal of Advances in Management Research*, 19, 55–77. <https://doi.org/10.1108/JAMR-01-2021-0039>
- Leung, R. (2019). Smart hospitality: Taiwan hotel stakeholder perspectives. *Tourism Review*, 74, 50–62. <https://doi.org/10.1108/TR-09-2017-0149>
- Marco-Lajara, B., Ruiz-Fernández, L., Seva-Larrosa, P. & Sánchez-García, E. (2021). Hotel strategies in times of COVID-19: a dynamic capabilities approach. *Anatolia*, 1(12), 525–536. <https://doi.org/10.1080/13032917.2021.1969960>
- Mingotto, E., Montaguti, F. & Tamma, M. (2021). Challenges in re-designing operations and jobs to embody AI and robotics in services. Findings from a case in the hospitality industry. *Electronic Markets*, 31(3), 493–510. <https://doi.org/10.1007/s12525-020-00439-y>
- Nieves, J. & Osório J. (2017). Using information technology to achieve management innovation. *Academia Revista Latinoamericana de Administración*, 32(1), 20–39. <https://doi.org/10.1108/ARLA-02-2016-0037>
- OECD. (2018). Oslo Manual 2018. 4th ed. Paris: OECD Publishing. <https://doi.org/10.1787/9789264304604-en>.
- Oliveira, C. & Avellar, A. (2021). Evidências da Relação entre Inovação Organizacional e Inovação Tecnológica na Indústria Brasileira. *RASI, Volta Redonda*, 7 (3), 09–29. <https://doi.org/10.20401/rasi.7.3.463>
- Orfila-Sintes, F. & Mattsson, J. (2009). Innovation behaviour in the hotel industry. *Omega*, 37(2), 380–394. <https://doi.org/10.1016/j.omega.2007.04.002>

- Ramos, C. M. Q. (2021). Transformação digital: Efeitos na educação, comércio e sustentabilidade ambiental. *RISTI*, 44, 1-4, dez. <https://doi.org/10.17013/risti.44.1-4>
- Schiopu, A. F., Hornoiu, R. I., Padurean, M. A., & Nica, A. (2021). Virus tinged? Exploring the facets of virtual reality use in tourism as a result of the COVID-19 pandemic. *Telematics and Informatics*, 60(2), 101575. <https://doi.org/10.1016/j.tele.2021.101575>
- Silva, G. B., Arvela, M. A. & Ramos, C. M. Q. (2021). A influência das *Social Media* nas compras online. *RISTI*, 44, 50-66. <https://doi.org/10.17013/risti.44.50-66>
- Silva, S. B. (2021). Improving the firm innovation capacity through the adoption of standardized innovation management systems: A comparative analysis of the ISO 56002:2019 with the literature on firm innovation capacity. *International Journal of Innovation*, 9(2), 389-413. <https://doi.org/10.5585/iji.v9i2.19273>
- Stevens, J. P. (1992). *Applied multivariate statistics for the social sciences*. 2th ed. Hillsdale: Erlbaum. <https://doi.org/10.4324/9781410604491>.
- Welter, C. V. N., Sausen, J. O. & Cappellari, G. (2019). Tipologias de inovação: Um estudo em organizações graduadas de incubadoras de base tecnológica. *Revista Iberoamericana de Estratégia – RIAE*, 18(4), 576-597. <https://doi.org/10.5585/riae.v18i4.16199>.
- Wooder, S. & Baker, S. (2012). Extracting Key Lessons in Service Innovation. *Journal of Product Innovation Management*, 29(1), 13-20. <https://doi.org/10.1111/j.1540-5885.2011.00875.x>.
- Zou, S., Taylor, C. R., & Osland, G. E. (1998). The EXPERF scale: a cross-national generalized export performance measure. *Journal of International Marketing*, 6(3), 37-58. <https://doi.org/10.1177/1069031X9800600304>.

Web scraping: Uso de plataformas de extracción de datos aplicadas a un sitio web sobre profesiones en México

Felipe Cocón¹, Dámaris Pérez-Cruz¹, José Ángel Pérez-Rejón¹,
Patricia Zavaleta-Carrillo¹, Ulises Barradas-Arenas¹,
Rubí Gómez-Ramón¹, José Alonso Pérez Cruz¹.

{jcocon, dperez, japerez; pzavaleta; ubarradas; rgramon; japacruz}@pampano.unacar.mx

¹ Facultad de Ciencias de la Información, Universidad Autónoma del Carmen, 24180, Campeche, México.

DOI: 10.17013/risti.52.61–73

Resumen: Este artículo realiza una revisión exhaustiva de las principales herramientas de *web scraping* disponibles en el mercado y compara sus características y funcionalidades. Se selecciona una herramienta específica para demostrar su uso en la obtención de datos sobre porcentajes de egresados en diversas carreras en México, así como la distribución relacionada con género y salarios en varios estados del país. El artículo tiene como objetivo principal ilustrar cómo se pueden recopilar datos mediante una herramienta de extracción. Además, se destaca la importancia de acceder a fuentes de datos confiables y se proporciona una descripción detallada del proceso de extracción de datos utilizando la herramienta WebHarvy. En última instancia, se destaca la importancia del *web scraping* como una técnica poderosa y ética profesional para recopilar datos valiosos de la web de manera eficaz y responsable.

Palabras-clave: Educación; extracción; laboral; mercado; scraping.

Web scraping: Use of data extraction platforms applied to a website about professions in Mexico

Abstract: This article provides a thorough review of the main web scraping tools available on the market and it is comparing their features and functionalities. A specific tool is selected to demonstrate its use in obtaining data on percentages of graduates in various careers in Mexico, as well as the distribution related to gender and salaries in several states of the country. The main objective of the article is to illustrate how data can be collected using a data extraction tool. Additionally, the importance of accessing reliable data sources is highlighted and a detailed description of the data extraction process using the WebHarvy tool is provided. Ultimately, it is highlighting the importance of web scraping as a powerful technique and professional ethical to collect valuable data from the web to effectively and responsibly.

Keywords: Education; extraction; employment; scraping; professions.

1. Introducción

Realizar una búsqueda de información en la *World Wide Web* se ha vuelto cada vez más problemático, esto debido a la creciente disponibilidad y diversidad de documentos que proporciona, ya el proceso de revisión de la información, el análisis y la extracción de manera manual demanda tiempo de modo que, se vuelve un desafío para los usuarios de la web.

El *web scraping* es una técnica que consiste en extraer o raspar datos de sitios web de forma automática mediante scripts o programas (Sinche & Torres, 2021, Mitchel, 2018). Aplicar esta técnica ofrece múltiples aplicaciones y beneficios, como la recopilación de información relevante, la automatización de tareas repetitivas, el ahorro de tiempo y dinero, y la generación de contenidos de calidad (Ribeiro et al., 2022)

Existen diversas herramientas de web scraping que facilitan la extracción de datos de la web sin requerir conocimientos técnicos. Estas herramientas suelen ofrecer una interfaz gráfica que permite seleccionar y nombrar los elementos deseados, así como opciones para guardar y exportar los datos en diferentes formatos.

En este sentido, el *Web Scraping* (raspado) permite extraer (*scrapear*) y recopilar información de páginas web y PDF de forma automatizada. Este procedimiento funciona a través del uso de programas o scripts, también conocidos como *scrapers*, capaces de “navegar por múltiples sitios web” y así “identificar y extraer información relevante de acuerdo con criterios preestablecidos” (Kinsta, 2022). Como se aprecia en la Figura 1 la información que se extrae se recopila y exporta a un formato que sea útil para el usuario.

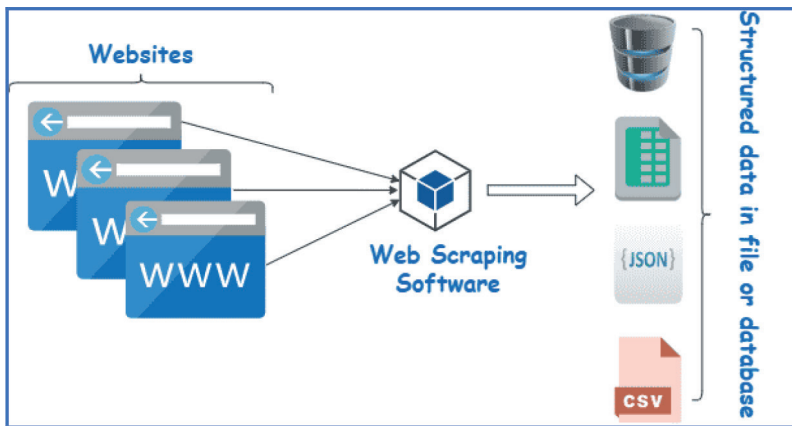


Figura 1 – Representación de Web scraping, fuente: Obtenido de (Kinsta, 2022).

El campo de aplicación para esta técnica es amplio, entre los usos comunes se mencionan los siguientes:

- Investigación de mercado para recopilar datos de precios, características de productos y reseñas de productos en línea, para realizar un análisis más detallado de sus competidores.

- Extraer opiniones y comentarios de usuarios en redes sociales, foros y blogs para evaluar la percepción pública de una marca, producto o servicio.
- Recopilar información de contacto de posibles clientes en directorios o sitios web de empresas, para estudiar tendencias del mercado, demanda y competidores mediante la recopilación y análisis de información en línea.
- Noticias y contenido.
- También se emplea para crear bases de datos para inteligencia artificial y aprendizaje automático.
- Extracción de datos de manera ilegal o no ética, para obtener información personal o confidencial de los usuarios de un sitio web sin su consentimiento.

Para agilizar el trabajo de los *scrapers*, los *Crawlers* o arañas navegan por la web buscando e indexando (Rama-Rico, 2022; Gilabert, 2021). De igual modo, los motores de búsqueda como Google utilizan rastreadores para actualizar los índices y las clasificaciones de los sitios web.

En este artículo se presenta una revisión de las principales herramientas de *web scraping* disponibles en el mercado, se comparan algunas de sus características y funcionalidades, con la intención de seleccionar una herramienta scraping para buscar y extraer datos sobre porcentajes de egresados de las diferentes carreras ofertadas en México, así como, la distribución con respecto a género y salario acorde con las profesiones en distintos estados de la República Mexicana.

2. Problemática

Para presentar la utilidad de las técnicas extracción de datos, se planteó obtener datos sobre la empleabilidad de los egresados en México. La educación superior es un factor clave para el desarrollo económico y social de un país, ya que contribuye a la formación de capital humano calificado, a la generación de conocimiento y a la innovación (Medina et al, 2021). Sin embargo, existen diversos desafíos y desigualdades que afectan al sistema educativo y al mercado laboral en México, tales como la baja cobertura, la calidad y pertinencia de la oferta educativa, la inserción laboral de los egresados, las condiciones de trabajo y los salarios de los profesionistas, y las brechas de género y territoriales (Fernández-Fassnach, 2017).

Para estudiar estos aspectos, es necesario contar con información estadística confiable y actualizada que permita conocer el perfil de los egresados de nivel superior por carreras, la cantidad de mujeres y hombres empleados, las carreras mejor y peor pagadas, las carreras con mayor porcentaje de empleos, la diferencia salarial de empleos por estados, las áreas profesionales con mayor y menor número de profesionistas, entre otros indicadores. Estos datos ayudan a identificar las fortalezas y debilidades del sistema educativo superior y del mercado laboral, así como las oportunidades y necesidades de los egresados y los empleadores.

Para obtener o recuperar la información, es necesario realizar un proceso de extracción de datos en los sitios web que contienen los datos del comportamiento de las profesiones en México. Por tanto, los datos recuperados pueden responder a interrogantes como las que se mencionan a continuación:

- ¿Qué carreras tienen mayor demanda y oferta laboral en el país y en cada entidad federativa?
- ¿Qué carreras tienen mayor retorno económico y social para los egresados?
- ¿Qué carreras presentan mayor equidad de género en términos ocupación y salario?

Por lo anterior, es preciso obtener información que permitan la generación de datos analizables. Para lo cual, en primer lugar, se necesita encontrar la fuente de información que satisface el estudio requerido. En segundo lugar, revisar las posibles herramientas *Web Scrapers*, finalmente, realizar la extracción de los datos con una herramienta seleccionada.

3. Fuentes para la obtención de datos

Para lograr la extracción de datos planteada, primero, se requiere la recopilación de información de empleabilidad de los egresados de las distintas carreras que se ofrecen en México, por lo tanto, es necesario realizar una búsqueda de sitios oficiales que administran información de este tipo. Dependiendo del sitio *web*, la información disponible se encuentra en publicaciones periódicas o en anuarios estadísticos, para lo cual se realiza un primer filtrado de datos para consensuar si la información satisface los requisitos planteados.

A continuación, se mencionan algunas fuentes oficiales más destacadas que proporcionan información sobre egresados en México:

- **Sistema Nacional de Información Estadística Geografía (INEGI).** Es un organismo público, autónomo responsable de normar y coordinar el Sistema Nacional de Información Estadística y Geográfica, así como de captar y difundir información de México en cuanto al territorio, los recursos, la población y economía, que permita dar a conocer las características de nuestro país y ayudar a la toma de decisiones. El INEGI a través del levantamiento de datos en la Encuesta Nacional de Ocupación y Empleo (ENOE) recopila y publica información referente a: empleo y ocupación, hogares y vivienda, población, salud y seguridad social, tecnologías de la información y comunicaciones. En el caso de las características educativas de la población mexicana, proporciona información sobre la población que asiste a la escuela, así como alfabetismo y nivel de escolaridad. También, incluye información sobre el número de alumnos inscritos en los diferentes niveles educativos, indicadores de eficiencia y sobre los recursos humanos del sistema educativo nacional (Instituto Nacional de Estadística y Geografía, 2023).
- **Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES).** La Asociación Nacional de Universidades e Instituciones de Educación Superior, en constante esfuerzo por mejorar el sistema de información de la educación superior en México, presenta el Anuario Estadístico en su versión digital, el cual contiene información de la población escolar y del personal docente de los tipos de educación media superior y educación superior en los niveles técnico superior universitario, licenciatura universitaria y

tecnológica, licenciatura en educación normal y posgrado. La ANUIES, plantea la necesidad de contar con un *sistema consolidado de seguimiento de egresados universitarios que oriente sobre la demanda laboral*, que permita obtener información confiable que oriente a las instituciones de educación superior sobre los profesionales que demanda el mundo laboral (Asociación Nacional de Universidades e Instituciones de Educación Superior, 2023).

- **Observatorio Laboral (OLA).** Es un servicio público de información confiable y gratuito que la Secretaría del Trabajo y Previsión Social (STPS), a través del Servicio Nacional de Empleo (SNE) ofrece, sobre las principales carreras profesionales del país, con la finalidad de que los jóvenes, los estudiantes, y los padres de familia, cuenten con información confiable y veraz que les permita tomar decisiones sobre qué carrera elegir y como insertarse en el mundo del trabajo (Observatorio Laboral, 2023).

Existen otras fuentes oficiales de información desde el portal del Gobierno de México (2018), pero también se basan en datos de la INEGI y en los anuarios de la ANUIES. Consecuentemente, para el caso de estudio propuesto se determina que la fuente de información disponible en el portal del Observatorio Laboral es factible y suficiente para realizar un *scraping* y generar la información requerida.

3.1. El futuro de las profesiones según OLA

Sobre el futuro de las carreras en las universidades, el Observatorio Laboral (OLA) menciona:

- Carreras que tienen mayor captación de empleo son las relacionadas con:
 - Nuevas tecnologías, Internet y el sector digital.
- Las carreras más prometedoras tienen que ver con Tecnologías:
 - La informática, la telemática, la telefonía celular, la ingeniería genética, la biotecnología, la biónica, la realidad virtual, la información multimedia y los nuevos materiales cerámicos.
- Serán prometedoras las profesiones relacionadas con:
 - Los cuidados a distancia para la tercera edad y la infancia, la teleasistencia sanitaria, los cultivos acuáticos, la robótica, la domótica, los sistemas de seguridad pública y la inteligencia artificial.
- La tendencia internacional (según la OCDE):
 - Ingeniería Molecular, Nanotecnología, Biomedicina, Investigación Espacial, Cibernética, Mecatrónica, Ciencias de la Tierra.
- Deja fuera las carreras relacionadas con las Ciencias Sociales, Humanidades y Artes.

3.2. Herramientas disponibles para web scraping

El web *scraping* es el proceso de automatizar la extracción de datos de sitios web de manera estructurada. Permite recopilar grandes cantidades de información de forma eficiente y automatizada, lo que ahorra tiempo y esfuerzo considerable. Las herramientas de *scraping* desempeñan un papel fundamental en este proceso al proporcionar una interfaz o biblioteca que simplifica la extracción de datos.

Estas herramientas utilizan diferentes métodos y técnicas para extraer datos de sitios web. Algunas herramientas se basan en técnicas de web *scraping* tradicionales, como el análisis de HTML y el uso de expresiones regulares para extraer información específica de las páginas web. Otras herramientas más avanzadas utilizan tecnologías como la automatización de navegadores o la inteligencia artificial para realizar la extracción (Jezreel & Ramirez, 2016).

Existen diversas herramientas de *scraping* disponibles, desde bibliotecas de programación, hasta aplicaciones con interfaz gráfica. Algunas de las herramientas más utilizadas incluyen: BeautifulSoup, Scrapy, Selenium, Octoparse, Import.io, dexi.io, WebHarvy, ParseHub, Web Scraper Chrome Extension, Apify, OutWit Hub, visual scraper, scraping hub, UiPath entre otras. Cada herramienta tiene sus propias características, ventajas y limitaciones, por lo que es importante elegir la herramienta adecuada según las necesidades y requerimientos del problema o proyecto.

Para seleccionar una herramienta de *scraping*, es importante considerar criterios como el modelo de negocio (de pago o libre), los formatos de documentos a scrapear (ej. html o pdf), los formatos de exportación de datos (ej. CSV, XLS o JSON), la cantidad de *scrapers* permitidos, la compatibilidad con sistemas operativos y otras características adicionales ofrecidas por cada herramienta. En la Tabla 1 se muestra el resultado de la revisión de 11 herramientas de *scraping*, teniendo en cuenta los criterios antes mencionados y la compatibilidad con sistemas operativos (SO).

Herramienta	Modelo de negocio	Fuente	Exportación	Scrapers permitidos	Compatibilidad con SO
<i>BeautifulSoup</i>	Libre y de código abierto	HTML y XML	Combinación con Phyton para exportar CSV, JSON, XML	Sin límite	Windows, macOS y Linux
<i>Scrapy</i>	Libre y de código abierto	HTML y XML	JSON, CSV y XML	Sin límite	Windows, macOS y Linux
<i>Selenium</i>	Libre y de código abierto	HTML	Combinación con Phyton para exportar CSV, JSON	Sin límite	Windows, macOS y Linux
<i>Octoparse</i>	Gratuitas y de pago	HTML y PDF	CSV, Excel, JSON y base de datos SQL	Gratuita: hasta 10 scrapers Pago: sin límite	Windows
<i>Import.io</i>	Gratuitas y de pago	HTML	CSV, Excel, JSON y base de datos SQL	Gratuita: 1 scraper Pago: sin límite	Windows, macOS y Linux
<i>ParseHub</i>	Gratuitas y de pago	HTML	CSV, Excel, JSON y base de datos SQL	Gratuita: hasta 5 proyectos. Pago: sin límite.	Windows, macOS y Linux
<i>WebHarvy</i>	Gratuitas y de pago	HTML	CSV, Excel, JSON y base de datos SQL	Sin límite	Windows

Herramienta	Modelo de negocio	Fuente	Exportación	Scrapers permitidos	Compatibilidad con SO
Apify	Gratuitas y de pago	HTML	JSON, CSV y Excel	Gratuita: hasta 10 actores Pago: sin límite	Windows, macOS y Linux
OutWit Hub	Gratuitas y de pago	HTML	CSV, Excel y XML	Sin límite	Windows y macOS
Data Miner	Gratuitas y de pago	HTML	CSV, Excel y Google Sheets	No limitada	Compatible con Google Chrome como una extensión
Web Scraper Chrome Extension	Gratuita	HTML	CSV y JSON	Sin límite	Compatible con Chrome en Windows, macOS y Linux

Tabla 1 – Comparación de herramientas *scrapers*. Elaboración propia.

Como se muestra en la Figura 2, para efectos de este trabajo, se realizó una prueba de conocimiento con tres herramientas que se ajustaron a las condicionantes de la investigación: acceso gratuito, fácil instalación, facilidad de uso, varios sitios para procesar y el formato de exportación en csv (<https://www.octoparse.com/>, <https://www.webharvy.com/index.html>, <https://webscraper.io/>). En la Figura 2, la ventana superior es una vista de escrapeado del sitio OctoParse, la ventana a la izquierda corresponde a la ejecución en Web Harvy y la ventana derecha corresponde al sitio web scraper.

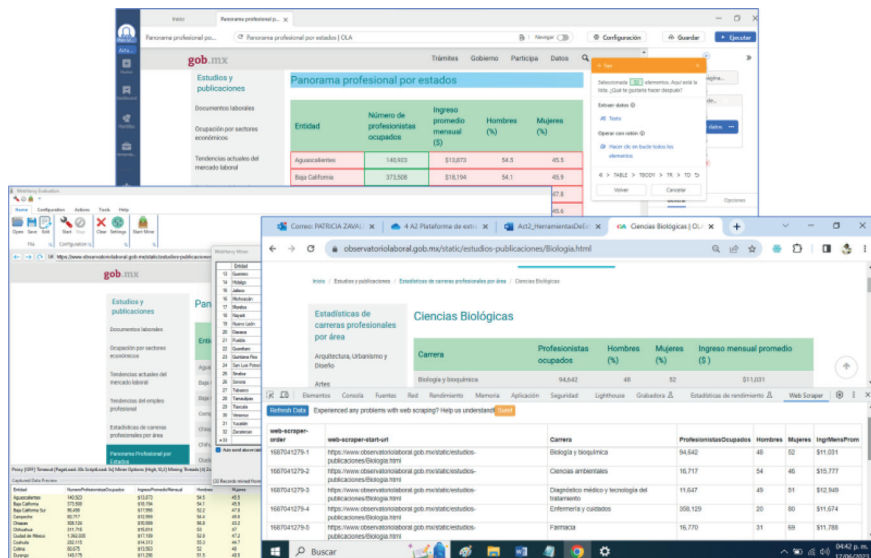


Figura 2 – Escrapeado con octoparse, web escrapper y webharvy

En la Figura 3, se muestra una arquitectura de *entrada-proceso-salida* para representar un *web scraping* exitoso siguiendo la siguiente secuencia de pasos:

1. Especificar las URLs de los sitios web y las páginas que se desean scrapear.
2. Acceder a la página web mediante una solicitud HTTP.
3. Analizar el contenido HTML de la página para identificar la información deseada. Utilizar localizadores como expresiones regulares para extraer la información.
4. Extraer la información y almacenarla en una base de datos o en un formato estructurado, como CSV, JSON u otro formato disponible.

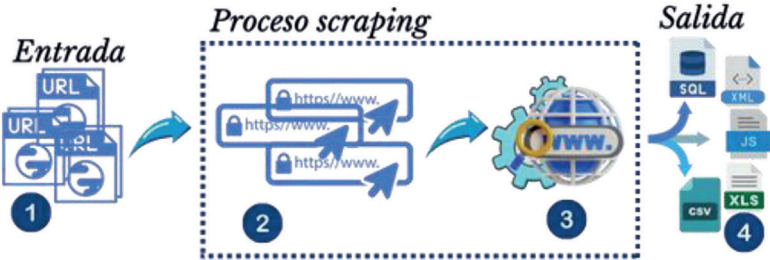


Figura 3 – Arquitectura web scraping. Elaboración propia.

3.3.Extracción empleabilidad de egresados con la herramienta WebHarvy

Para el caso de la obtención de datos de egresados planteado como problemática para extracción, se eligió la herramienta de WebHarvy, ya que, durante las pruebas de raspado las tres herramientas (Octoparse, WebHarvy y Web scraper), WebHarvy arrojó mejores resultados, tanto en la extracción como en la exportación de los datos. A continuación, se describe el proceso seguido.

3.3.1. Paso 1. URL del portal web para la fuente de información

En este paso se proporciona el sitio que contiene los datos, es el más sencillo, y corresponde a proporcionar el enlace (*link*) o *URL*, para iniciar el raspado, como se puede observar en la Figura 4.



Figura 4 –URL de la página para su raspado.

3.3.2. Paso 2. Proceso de scrapeado (raspado)

Posteriormente de ingresar el URL, se oprime el ícono de “Start”. Como se puede observar en la Figura 5, es necesario seleccionar un campo de la columna deseada, posteriormente oprimir “Capture Text”, en la ventana emergente colocar el nombre que se le dará a la columna del dato indicado, finalmente presionar el botón “OK”.

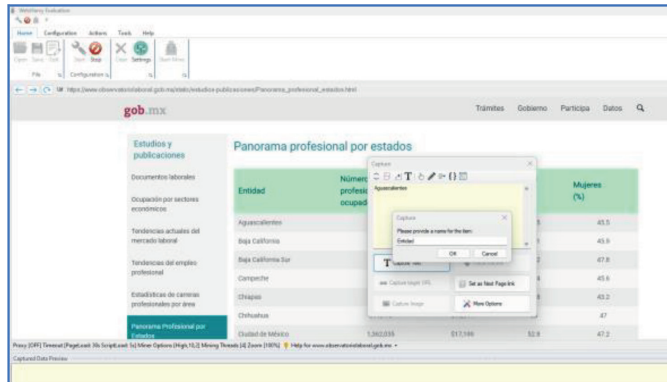


Figura 5 – Ejemplo de columnas para el raspado.

En la Figura 6 se puede observar que en la parte inferior se recuperan los datos de la columna seleccionada, identificada con el nombre que se escribió como encabezado; se repetirán estos pasos para completar todas y cada una de las columnas deseadas. Al terminar de raspar (scrapear) toda la tabla, se oprime el botón “Stop” de la parte superior, para detener el proceso.

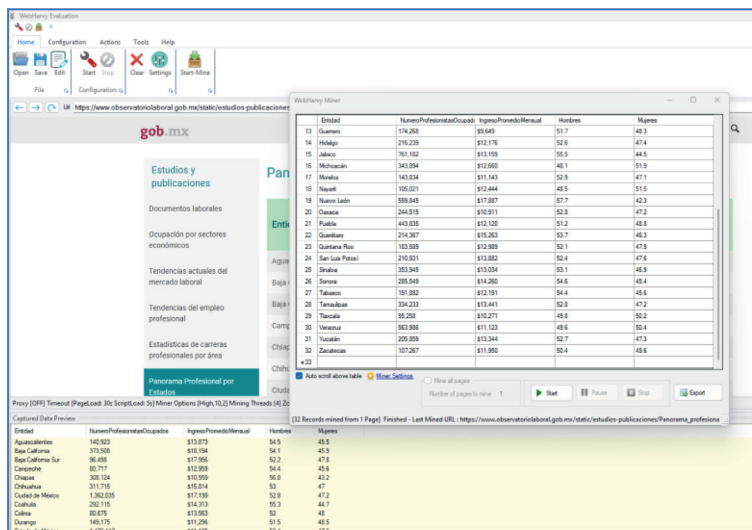


Figura 6 – Raspado de datos en WebHarvy.

Este proceso finaliza oprimiendo el icono “*Start-Mine*”, posteriormente presionar el botón “*Start*” para proceder a generar todos los datos antes seleccionados y almacenarlos en una tabla, para su exportación.

En este caso los datos se guardaron en un archivo de hoja de cálculo permitiendo elegir el destino y el nombre del archivo. De igual modo, esta opción permite seguir guardando en el archivo, sobrescribirlo (*Overwrite*), agregar registros (*Append*) o actualizar el documento (*Update*).

3.3.3. Paso 3. Resultados: datos obtenidos

Una vez terminado el proceso anterior, abrimos el documento de hoja de cálculo y aparecen los datos de todas las páginas raspadas, acomodados en columnas, como se observa en la Figura 7. Para efectos de la muestra del raspado, solo se muestran algunas filas de la información obtenida.

A	B	C	D	E	F	G
Área	Carrera	Profesionistas	Ocupados	Hombres	Mujeres	Ingresopromediomensual
Arquitectura, Urbanismo y Diseño	Arquitectura y urbanismo	240,510	71	29	\$15,031	
Artes	Artes, programas multidisciplinarios o generales	9,336	26	74	\$11,756	
Artes	Bellas artes	18,580	40	60	\$8,863	
Artes	Diseño	44,114	29	71	\$12,440	
Artes	Música y artes escénicas	31,205	63	37	\$9,420	
Artes	Técnicas audiovisuales y producción de medios	156,300	53	47	\$13,889	
Ciencias Biológicas	Biología y bioquímica	94,842	48	52	\$11,031	
Ciencias Biológicas	Ciencias ambientales	16,717	54	46	\$15,777	
Ciencias Biológicas	Diagnóstico médico y tecnología del tratamiento	11,647	49	51	\$12,949	
Ciencias Biológicas	Enfermería y cuidados	358,129	20	80	\$11,674	
Ciencias Biológicas	Farmacología	16,770	51	49	\$11,788	
Ciencias Biológicas	Medicina	324,198	54	46	\$18,569	
Ciencias Biológicas	Odontología	162,107	48	54	\$14,809	
Ciencias Biológicas	Psicología	357,044	25	75	\$11,395	
Ciencias Biológicas	Química	31,468	40	60	\$13,278	
Ciencias Biológicas	Salud pública	14,692	46	54	\$18,698	
Ciencias Biológicas	Terapia y rehabilitación	107,851	27	73	\$10,209	
Ciencias Biológicas	Veterinaria	77,658	75	25	\$12,378	
Ciencias Físico Matemáticas	Ciencias de la tierra y de la atmósfera	14,428	63	37	\$16,742	
Ciencias Físico Matemáticas	Estadística	8,557	62	38	\$17,232	
Ciencias Físico Matemáticas	Física	9,967	82	18	\$12,723	
Ciencias Físico Matemáticas	Matemáticas	36,313	61	39	\$13,664	
Ciencias Sociales	Bibliotecología	4,877	47	53	\$10,179	
Ciencias Sociales	Ciencias políticas	77,835	45	55	\$14,479	
Ciencias Sociales	Ciencias sociales y estudios del comportamiento	12,692	58	42	\$14,125	
Ciencias Sociales	Comunicación y periodismo	198,551	47	53	\$12,425	
Ciencias Sociales	Criminología	58,163	52	48	\$12,028	
Ciencias Sociales	Derecho	908,694	61	39	\$13,123	
Ciencias Sociales	Sociología y antropología	46,252	56	44	\$14,218	
Ciencias Sociales	Trabajo y atención social	99,296	10	90	\$10,018	
Económico Administrativas	Administración y gestión de empresas	1,091,842	49	51	\$13,137	
Económico Administrativas	Comercio	281,002	49	51	\$15,780	
Económico Administrativas	Contabilidad y fiscalización	803,231	51	49	\$13,418	
Económico Administrativas	Economía	81,937	58	42	\$15,256	

Figura 7 – Documento en hoja de cálculo con datos raspados.

3.4. Recomendaciones sobre la extracción obtenida

Cabe mencionar que para realizar el raspado de varias páginas web, cómo fue el caso planteado, se debe realizar página por página, de manera individual, es decir, que se realiza el raspado de la primera página, se genera un documento en Excel y se guarda para exportar el archivo. Posteriormente con la opción *Append* se repite el procedimiento para la página siguiente y anexarlo al final del último registro.

El proceso anterior es iterativo para todas las páginas; se debe considerar que las otras páginas pueden tener un orden diferente a las columnas de información extraídas, por ello, es importante que cada vez que se realiza el raspado, se verifique el orden con respecto al nombre de las columnas para garantizar que los registros coincidan con el encabezado previamente seleccionado.

Una utilidad que se resalta de WebHarvy es la facilidad para seleccionar las columnas deseadas y el orden en que se requieren, permitiendo la asignación del identificador (nombre) para cada columna.

En los inconvenientes se mencionan que, no se puede scrapear otra URL de manera continua. Otro inconveniente es que la licencia tiene límite de tiempo de 16 días.

La técnica *Scraping* es de gran utilidad, sin embargo, acceder a la información conlleva a un proceso de ética y responsabilidad sobre el acceso y protección de datos, sobre todo aquellos que no están disponibles para uso de dominio público o aquellos que son utilizados en perjuicio de terceros.

4. Conclusiones

Este trabajo presentó a las herramientas *Scrapers* para la obtención de datos de sitios *web*, las cuales otorgan facilidades cuando se requiere descargar y tratar considerables cantidades de información, que pueden apoyar en la toma de decisiones.

La selección de una herramienta *scraping* requiere de un proceso de selección en el que el usuario debe dar importancia a la determinación de criterios que permitan elegir una herramienta que mejor se adapte tanto a las necesidades de obtención y procesamiento de datos, al modelo de negocio esperado (de pago o libre), así como también a los requisitos de compatibilidad con sistemas operativos entre otras características adicionales ofrecidas por cada herramienta.

La herramienta *WebHarvy* utilizada para el proceso de raspado del sitio “Observatorio Laboral”, ofrece grandes ventajas a la hora de raspar los datos, con la opción de *Append*, para anexas los datos al final del archivo y la posibilidad de omitir columnas innecesarias. Así también presentó algunos inconvenientes, como lo fueron: el orden diferente de las columnas en las otras páginas, la necesidad de cerrar las ventanas abiertas para cambiar de URL y la duración limitada de la licencia. Al realizar el proceso de manera repetida, la herramienta se vuelve mecánica y sencilla de trabajar.

Como continuación a este trabajo, se propone contrastar los datos presentados por el Observatorio Laboral, con los resultados escrapeados llevados a un segundo nivel, utilizando una herramienta de inteligencia de negocios.

Por último, el *scraping* es una técnica que permite recoger datos de la *web* de forma rápida y automatizada. Con estos datos se puede obtener información útil para diversos fines, como estudios de mercado, comparación con la competencia, captación de clientes y más. Si se respeta la ética y los términos de servicio de los

sitios *web*, el *scraping* puede ser una herramienta eficaz para extraer conocimientos valiosos de la gran cantidad de datos que hay en la web.

Referencias

- Asociación Nacional de Universidades e Instituciones de Educación Superior. (2023). Anuarios Estadísticos de Educación Superior. <http://www.anuies.mx/informacion-y-servicios/informacion-estadistica-de-educacion-superior/anuario-estadistico-de-educacion-superior>
- Dirección General de Educación Superior (DGESUI). (2023). Consultas. <https://dgesui.ses.sep.gob.mx/>
- El comercio Perú. (2023). Web scraping: qué es y cómo funciona la herramienta que extrae información de los sitios web. (M. M. Saravia, Editor) <https://elcomercio.pe/tecnologia/ciberseguridad/web-scraping-que-es-y-como-funciona-la-herramienta-que-extrae-informacion-de-los-sitios-web-espana-mexico-usa-noticia/>
- Escuela de datos. (2023). Introducción a la extracción de datos de sitios web: scraping. <https://escueladedatos.online/introduccion-a-la-extraccion-de-datos-de-sitios-web-scraping/>
- Fernández-Fassnach, E. (2017). Una mirada a los desafíos de la educación superior en México. *Innovación Educativa*, 17(74), 183-207. Recuperado el 20 de octubre 2023, de <https://www.scielo.org.mx/pdf/ie/v17n74/1665-2673-ie-17-74-00183.pdf>
- Gilabert-Perea, X. (2021). Diseño e implementación de un extractor de noticias automatizado. Trabajo de Fin de Grado en Ingeniería Informática. Universitat Politècnica de València, España. Recuperado de <https://riunet.upv.es/bitstream/handle/10251/174252/Gilabert%20-%20Diseno%20e%20implementacion%20de%20un%20extractor%20de%20noticias%20automatizado.pdf>
- Gobierno de México. (2018). Educación por niveles. <https://www.gob.mx/sep/acciones-y-programas/educacion-por-niveles?state=published>
- Instituto Nacional de Estadística y Geografía. (2023). México en Cifras. <https://www.inegi.org.mx/app/areasgeograficas/#collapse-Resumen>
- Jezreel, M., & Ramirez, H. (2016). Estableciendo controles y perímetro de seguridad para una página web de un CSIRT. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, 17, 1-14. <https://www.risti.xyz/issues/risti17.pdf>
- Kinsta. (2022). ¿Qué Es el Web Scraping? Cómo Extraer Legalmente el Contenido de la Web. <https://kinsta.com/es/base-de-conocimiento/que-es-web-scraping/>
- Medina Delgado, B., Palacios Alvarado, W. Camargo Ariza, L. L. (2021). Economía del conocimiento en la educación superior: factor clave en la calidad educativa. *REDIPE* 10(7). <https://doi.org/10.36260/rbr.v10i7.1347>

- Mitchell, R. (2018). Web scraping with Python: Collecting more data from the modern web. .O'Reilly Media, Inc..
- Observatorio Laboral (2023). Estadísticas de carreras profesionales por área. https://www.observatoriolaboral.gob.mx/static/estudios-publicaciones/Ola_indice_estadisticas_area.html
- Papeles de Inteligencia. (2018). 10 herramientas de web scraping para extraer datos online de forma automática. <https://papelesdeinteligencia.com/herramientas-de-web-scraping/#:~:text=Las%20herramientas%20de%20web%20scraping%20est%C3%A1n%20especialmente%20dise%C3%B1adas%20para%20extraer,datos%20de%20una%20p%C3%A1gina%20web.>
- Rama Rico, J. A. (2022). Desarrollo de un software para la búsqueda de apuestas seguras. Trabajo Fin de Grado. Grado en Ingeniería en Tecnologías Industriales. Universidad de Sevilla, España. Recuperado de: https://idus.us.es/bitstream/handle/11441/140647/TFG4331_Rama%20Rico.pdf?sequence=1&isAllowed=y
- Ribeiro, A. J., Mendes, R., & Duarte, M. do C. (2022). Requisitos para a ciência de dados: Analisando anúncios de vagas de emprego com mineração de texto. RISTI - Revista Ibérica de Sistemas y Tecnologias de Información, 46, 54-70. <https://www.risti.xyz/issues/risti46.pdf>
- Secretaría del Trabajo y Previsión Social. (2023). 90 Datos en datos.gob.mx. Catálogo de Datos Abiertos del Gobierno de la República: <https://datos.gob.mx/busca/organization/stps>
- Sinche, J. U., & Torres, J. C. (2021). Análisis de sentimientos en los mensajes recibidos en el entorno virtual de aprendizaje de la modalidad abierta y a distancia de la UTPL. RISTI - Revista Ibérica de Sistemas y Tecnologías de la Información, 41, 98-113. <http://www.risti.xyz>.
- Sitelabs. (2016). Qué es el Web scraping? Introducción y herramientas. (M. Marq, Editor) <https://sitelabs.es/web-scraping-introduccion-y-herramientas/>

Perspectivas latinoamericanas del uso de las TIC en estudiantado universitario

María Guadalupe Veytia Bucheli¹, Armando Sánchez Macías²,
Maritza Librada Cáceres Mesa¹, José Frank Vázquez Horta³

maria_veytia@uaeh.edu.mx; armando.sanchez@uaslp.mx;
maritza_caceres3337@uaeh.edu.mx; fvhorta@gmail.com

¹ Instituto de Ciencias Sociales y Humanidades (ICSHu) Universidad Autónoma del Estado de Hidalgo (UAEH). Carretera Pachuca – Actopan km 4/5 s/n. San Cayetano el Bordo, CP 42084, Pachuca de Soto, Hidalgo, México,

² Coordinación Académica. Región Altiplano Oeste. (CARAO). Universidad Autónoma de San Luis Potosí (UASLP). Santo Domingo Salinas, 200. CP 78600, Salinas Hidalgo, S.LP., México,

³ Edunetic, La Paz, Bolivia.

DOI: 10.17013/risti.52.74–96

Resumen: El empleo de las Tecnologías de la Información y la Comunicación (TIC) está cada vez más presente en los distintos ámbitos de la vida del ser humano, y el ámbito educativo no es la excepción, las cuales favorecen la realización de actividades y tareas, y su empleo también contribuye al desarrollo de habilidades digitales y de pensamiento. Este artículo tiene como objetivo describir la percepción que tiene el estudiantado universitario del empleo de las TIC en el ámbito educativo. Se desarrolla en el primer semestre del 2023 con 182 universitarios de cuatro países latinoamericanos (México, Guatemala, Ecuador y Bolivia). Se aborda desde el enfoque empírico-analítico, desde un estudio cuantitativo, de alcance descriptivo-transversal y no experimental. Se utiliza un cuestionario, aplicado en *Google Forms*. Los resultados reflejan diferencias estadísticamente significativas, en el uso de la TIC en el ámbito educativo, en el manejo de programas y en la alfabetización digital.

Palabras-clave: educación; estudiantado universitario; TIC; universidad.

Latin American perspectives on the use of ICT in university students

Abstract: The use of Information and Communication Technologies (ICT) is increasingly present in different areas of human life, and the educational field is no exception, as they favour the performance of activities and tasks, and their use also contributes to the development of digital and thinking skills. The aim of this article is to describe how university students perceive the use of ICT in education. It was carried out in the first semester of 2023 with 182 university students from four Latin American countries (Mexico, Guatemala, Ecuador and Bolivia). It is approached from an empirical-analytical, quantitative, descriptive-transversal and non-experimental approach. A questionnaire was used, applied using *Google Forms*. The results reflect statistically significant differences in the use of ICT in education, in the use of programmes and in digital literacy.

Keywords: education; university students; ICT; university.

1. Introducción

La Sociedad del Siglo XXI se caracteriza por el fenómeno de la globalización, cambios constantes, nuevos paradigmas de la comunicación, así como el empleo de las Tecnologías de la Información y la Comunicación (TIC) (Vidal y Pinargote, 2019). Esto ha permitido en ocasiones reducir los costos laborales, facilitar el acceso a la información que se presenta en distintos formatos, así como favorecer la conexión entre las personas tanto de manera sincrónica como asincrónica.

El manejo de dispositivos digitales e internet es una tendencia creciente por personas de distintas edades a nivel mundial (Murillo, et al., 2021). Se observa un avance significativo en la realización de actividades en ámbitos como el económico, social, político y educativo (Soomro, et al., 2020). Esto ha modificado la manera en cómo se lee, construye, procesa la información y se comparte a los demás (Maldonado, et al., 2019).

El desarrollo de las TIC en el ámbito educativo está presente en contextos formales (Instituciones Educativas), no formales (instituciones que no pertenecen al sistema escolar) e informales (los espacios en los cuales se desarrolla la persona). El acceso a la información no se reduce a material físico o proporcionado por el docente, lo que permite al estudiante profundizar sobre un tema de interés de manera independiente y generar sus Entornos Personales de Aprendizaje (Tirado y Roque, 2019).

Otro de los aspectos relevantes a considerar es la emergencia sanitaria que se vivió a nivel mundial provocada por la pandemia por COVID-19. En esta se transitó de espacios presenciales a espacios remotos, en la mayoría de los casos sin una preparación previa por parte de los docentes. Se emplearon distintas herramientas tecnológicas ahora desde una perspectiva didáctica, entre las más recurrentes se encuentra el uso de *WhatsApp*, el *Facebook*, así como *Zoom* y *Meet* (Huamán, et al., 2021).

Desde una perspectiva internacional, uno de los objetivos clave de los estados es la fortalecer la Educación de calidad que establece el cuarto Objetivo de Desarrollo Sostenible (ODS) de la Agenda 2030 (ONU, 2018). Se destaca la inclusión, la equidad y la promoción de oportunidades de aprendizaje para todos de manera permanente. Se asumen como indicadores el acceso a internet con fines pedagógicos orientados a enriquecer los procesos de enseñanza – aprendizaje en todos los niveles educativos.

Al referirse de manera particular del nivel de Educación Superior, la UNESCO (2021) aborda temas clave, como la importancia de la educación para todos, el aprendizaje centrado en el estudiante, la organización didáctica de los conocimientos desde una perspectiva constructivista, en la cual se generen mediaciones que favorezcan el trabajo individual y colaborativo. La educación formal que se desarrolla en las Instituciones Educativas requiere vincularse a distintos niveles y contextos, y más aún cuando se trata del nivel superior. Por ejemplo, el vínculo que se establece entre la institución educativa, el contexto y las comunidades, lo que se reflejará en proyectos de incidencia social.

Se señala además que el uso de herramientas tecnológicas facilita las acciones para ampliar el acceso a la Educación Superior, la digitalización de materiales, el consumo y/o

construcción de Recursos Educativos Abiertos en diferentes formatos que favorezcan los distintos estilos de aprendizaje de los estudiantes y generen aprendizajes con mayor significado, sentido y aplicación que permita la resolución de problemáticas en contextos reales (UNESCO, 2021).

Ante esta realidad, los sistemas educativos en sus distintos niveles y modalidades se enfrentan al reto de evolucionar los procesos de enseñanza y aprendizaje con la incorporación de las TIC (UNESCO, 2014), que no se limite a una perspectiva técnica, sino desde una concepción integral que permita la construcción de conocimientos y la resolución de problemáticas, que les permitan trascender a lo largo de toda su vida. Generar dinámicas para la redefinición de los roles del profesorado y el estudiantado es fundamental para lograr un cambio no de herramientas, sino de metodologías y mediaciones en la generación de ambientes de aprendizajes.

El estudiantado universitario se encuentra en un contexto en el que las TIC se utilizan de manera cotidiana para sus actividades escolares, para desarrollar tareas o como instrumento para apoyar los procesos de aprendizaje (Díaz-García, et al., 2020). El uso personal o educativo de las TIC por parte de los estudiantes, se relaciona con la construcción de su aprendizaje que puede ser desde un enfoque superficial o uno profundo.

Ponce, Zambrano y de la Peña (2023) llevan a cabo un estudio en el que se describen nuevas maneras de generar procesos de enseñanza – aprendizaje en el estudiantado universitario desde la perspectiva de los docentes, en donde se destaca un mayor empleo de plataforma educativas, una de las más empleadas es Moodle, en donde existe una satisfacción positiva por parte de los usuarios, las cuales facilitan la comunicación tanto sincrónica como asincrónica, la flexibilidad en el modelo, así como la diversidad de recursos didácticos.

De manera particular, el empleo de las TIC se enriquece de forma significativa durante la pandemia sanitaria del COVID-19, Arriaga, Bautista y Montenegro (2021) llevan a cabo una investigación con estudiantado universitario en la cual se describe el apoyo de las TIC en el proceso de enseñanza – aprendizaje cuando se transitó de espacios presenciales a espacios virtuales, entre los principales hallazgos se identifican principalmente cuatro dimensiones: la enseñanza, la práctica, la simulación y resolución de problemas, así como la generación de productos.

Otro estudio que aborda la percepción de la Educación Virtual en Instituciones de Educación Superior es el que presentan Chávez, Rivera y Jaro (2020) en donde analizan las principales problemáticas que tuvieron los estudiantes al acceder a las clases virtuales en la emergencia sanitaria, en donde se destaca como uno de los principales hallazgos la competencia de autogestión que desarrollaron los estudiantes durante el proceso de enseñanza – aprendizaje, y que hasta la actualidad constituye una de las más importantes, ya que permiten aprender no solo para la vida, sino durante toda la vida.

La Innovación Educativa está presente en la Educación Superior, y el empleo de las TIC ha contribuido en algunos casos, como lo exponen Deroncele-Acosta, et al (2021) en donde se recuperan la experiencia de los docentes de cuatro universidades latinoamericanas encontrando entre los principales hallazgos para el logro de las

innovaciones la participación de comunidades profesionales de aprendizaje con TIC, los procesos de capacitación y actualización como formación continua de los docentes, así como el equipamiento de los laboratorios y centros de cómputo.

La Movilidad Estudiantil es otro de los factores que se ha beneficiado entre los estudiantes universitarios gracias al empleo de las Tecnologías de la Información y la Comunicación, Santos-Rego, Mella-Núñez y Sotelino-Losada (2020) presentan una experiencia sobre esta temática desde una perspectiva global y tecnológica, en donde existe la posibilidad de realizar actividades tanto de manera presencial como virtual mediante una comunicación sincrónica y asincrónica, y de esta manera enriquecer la formación del estudiantado universitario.

Después de haber revisado algunas investigaciones en torno al empleo de las TIC desde distintas perspectiva en estudiantado universitario de diferentes países, el preesnte este estudio se desarrolla en el contexto del estudiantado que cursa programas de pregrado y posgrado en cuatro países latinoamericanos: Bolivia, Ecuador, Guatemala y México. El objetivo es describir la percepción que tiene el estudiantado universitario en el empleo de las TIC en el ámbito educativo.

La Introducción constituye el primer apartado del estudio, en un segundo momento se presentan los referentes teóricos que permiten definir a las Tecnologías de la Información y la Comunicación, así como su empleo en los ámbitos académicos, en esta caso de manera particular en la Educación Superior, el tercer apartado lo conforma la metodología empleada que fue desde un enfoque empírico – analítico, con una metodología cuantitativa de alcance descriptivo, después se presentan los resultados, y finalmente las principales conclusiones a las que se llega.

2. Referentes teóricos

Para Bernard (2006), caracterizar las TIC es referirse en sentido amplio a la variedad de usos que se pueden generar con su empleo desde distintos ámbitos; entre ellos se encuentra el económico, con un enfoque en la competitividad. También es pertinente precisar que el acceso a vasta cantidad de información no es sinónimo de adquirir conocimiento.

Las TIC están dotadas de una vida social doble, aseguran Loveless y Williamson (2017), ya que contribuyen a influir y dar forma tanto al pensamiento como a la acción de los seres humanos. Así mismo, tienen una incidencia significativa en la forma y estructura de la sociedad; este se refleja en los cambios que se han presentado en cuanto a dinámicas en distintos ámbitos sociales.

Para Coll y Monereo (2011), las TIC han sido desde sus diferentes niveles de desarrollo, instrumentos para la concreción de actividades orientadas a procesos de pensamiento, aprendizajes, conocimientos, representaciones y recursos para transmitir información a personas de la misma o de otras generaciones, desde una perspectiva analítica, reflexiva y activa, que favorezca la promoción de nuevos saberes.

En la educación de finales del siglo XX y principios del siglo XXI, las TIC, las propuestas promovidas por los organismos internacionales y el vínculo con el ámbito laboral,

confirman que las mediaciones tecnológicas ocupan un lugar cada vez más destacado en los procesos de enseñanza – aprendizaje. Se constituyen en un recurso para llegar a un mayor número de personas tanto en espacios presenciales como semipresenciales y en línea (Crovi, 2006).

La llamada Sociedad de la Información que busca transitar hacia la Sociedad del Conocimiento, según Cebrián (2014), se caracteriza principalmente por los siguientes rasgos: (a) la complejidad, la interdependencia y la imprevisibilidad que caracteriza las actividades y las interacción de los individuos en la actualidad; (b) la información, la sobre información y el ruido, es notable el incremento de la cantidad de información a la cual se puede tener acceso de manera gratuita, sin embargo, este fenómeno no garantiza que las personas estén más y mejores informadas, es necesario contar con criterios que permitan valorar la información confiable y contrastar su veracidad, (c) la rapidez con la que se generan los procesos, los cambios y transformaciones, donde la vigencia de la información es cada vez más reducida; (d) la escases de espacios y tiempos para realizar procesos de abstracción y de reflexión, la velocidad con la que se vive en la actualidad es contraria a los procesos de reflexión y; (e) la transformación de las coordenadas de los espacios temporales de la comunicación, y la diversidad de formatos con los cuales se puede llevar a cabo.

Es en este sentido, que la educación como base de la sociedad del siglo XXI requiere reconocer sus demandas y necesidades, para responder con pertinencia y oportunidad a las diferentes problemáticas que se presentan al estudiantado. Se considera entonces que la promoción del uso de las TIC constituye una herramienta fundamental para el acceso a la información, la comunicación sincrónica y asincrónica, la participación activa de estudiantado y profesorado mediante la generación de Comunidades de Aprendizaje, así como a través de experiencias innovadoras para la resolución de tareas empleando distintos formatos electrónicos (Mañas y Roig-Vila, 2019).

La presencia de las TIC en la Educación Superior es cada vez más frecuente (Poveda-Pineda y Cifuentes-Medina, 2020), no obstante, es necesario que el estudiantado universitario las utilice no únicamente de manera instrumental, sino como un recurso para fortalecer el desarrollo del pensamiento, la generación y construcción de conocimientos, tanto de manera individual como de forma colaborativa, que permitan transferir saberes conceptuales, procedimentales y actitudinales a contextos reales, desde donde emerjan nuevas formas de pensamiento que condicionen la gestión del conocimiento (Huanca y Portal, 2023). En ello el profesorado tiene una labor clave por desarrollar.

La UNESCO, según las aportaciones de Valencia-Molina, et al., (2017) establece los niveles de apropiación de las TIC por parte del profesorado y como consecuencia el manejo que tienen de ellas el estudiantado. Un primer nivel se refiere a la integración, es decir, el empleo de herramientas propiamente para economizar tiempo. Las principales tareas se relacionan con digitalizar contenidos, almacenarlos y realizar exposiciones utilizando distintos recursos y a la vez el apoyo en la resolución de problemas apoyado en estos recursos tecnológicos, condiciona la incorporación gradual de estas habilidades en los estudiantes (Paucar-Curasma, et al., 2022). En un segundo nivel se encuentra el empleo de las TIC para fortalecer los procesos de aprendizaje del estudiantado, orientándolas hacia las habilidades para comprender, analizar y reflexionar, mediante actividades como resumir, debatir, buscar información, jugar o hacer anotaciones

digitales. Finalmente, en el tercer nivel se busca que el estudiantado no sea únicamente consumidor de las TIC, sino que genere escenarios virtuales y recursos orientados al desarrollo de un pensamiento crítico y la resolución de problemas, como pueden ser la creación de videos, de podcast, exámenes en línea, por mencionar algunos.

Es por ello, que la importancia de incorporar las TIC en el nivel superior no radica en la herramienta a emplear, sino en la transformación del proceso de enseñanza – aprendizaje y su articulación con la formación profesional (Zambrano y Zambrano, 2019). Se espera que el estudiantado vincule los conocimientos adquiridos en la Universidad con los contextos profesionales como son empresas, organizaciones civiles, instituciones educativas, por mencionar algunos. Así como generar proyectos para la solución de problemáticas y la presentación de propuestas innovadoras.

El empleo de las TIC propicia el desarrollo de los procesos de enseñanza y aprendizaje no solo en espacios presenciales, sino semipresenciales y en línea, lo que favorece la incorporación de nuevos paradigmas educativos que se centran en los ritmos y necesidades del estudiante, favoreciendo el aprendizaje profundo, la movilización de saberes, la selección de los materiales más significativos, así como la construcción de los Entornos Personales de Aprendizaje (Luna-Romero, et al., 2019), a través de los cuales se erigen nuevas formas de pensamiento que sustenta la construcción activa del saber, desde una perspectiva individual y colegiada.

El centro del proceso de enseñanza y aprendizaje en la actualidad se ha desplazado a la figura del estudiantado, lo que modifica las relaciones e interacciones que genera con sus compañeros, con el profesorado, con el contexto en el que se desarrolla, a partir de las TIC que emplea y la finalidad con las cuales las utiliza, destacando en este sentido una dinámica compleja y dialéctica (Crovi, 2006) en las formas de operar e interactuar con el objeto de conocimiento, como incentivo que ressignifica la construcción de nuevos conocimientos.

Desde este nuevo panorama, el estudiantado debe ser capaz de enfrentar incertidumbres, fortalecer el trabajo autónomo, el espíritu de búsqueda, la responsabilidad, la creatividad, la capacidad para expresar sus ideas desde distintos formatos virtuales y renovar de manera constante su deseo por aprender (Bernate y Vargas, 2020). Se presenta entonces un doble desafío, en lugar de dejarse conducir, debe administrar sus tiempos, manejar sus recursos, trabajar de manera individual, pero también aprender a construir de forma colaborativa y a gestionar de forma individual y equipos los conocimientos y saberes, que sustenten el quehacer académico.

Al hablar de los procesos y redes de colaboración que se establecen desde el empleo de las TIC en un diseño tecno-pedagógico y sus prácticas de uso, Coll et al., (2011) proponen una tipología que se conforma por cinco etapas: (a) las TIC como instrumentos mediadores de las relaciones que se establece entre el alumnado y los contenidos de aprendizaje; (b) las TIC como instrumentos mediadores entre las relaciones de los contenidos y el profesorado; (c) las TIC como instrumentos mediadores entre profesorado y alumnado; (d) las TIC como instrumentos mediadores de la actividad colaborativa entre el profesorado y el alumnado y; (e) las TIC como instrumentos que permiten la configuración de entornos de trabajo y aprendizaje; los cuales constituyen referentes para una adecuada organización de ambientes de aprendizajes, que de manera

gradual indiquen en la formación del estudiante universitario, como usuario autónomo y responsable en su aprendizaje a lo largo de toda la vida (López García et al., (2021).

Es también relevante abordar estas implicaciones de colaboración desde el rol que desempeñan las instituciones educativas como espacios formadores del estudiantado. Contribuir con la conformación del perfil digital del futuro ciudadano, así como reducir la brecha digital a partir de fortalecer las acciones para el desarrollo de saberes conceptuales, procedimentales y actitudinales (Torres, 2018), considerando la interacción del aprendizaje, la fluidez tecnológica y la responsabilidad social en el uso ético de los saberes, como insumos básicos que condicionan la conformación de una ciudadanía digital.

Después de hacer una revisión teórica de los conceptos que se relacionan con el uso de las TIC en el ámbito educativo, se recupera una serie de investigaciones que se han realizado con el objetivo de identificar el manejo y uso de las TIC en actividades que se relacionan de manera directa con su proceso de formación desde la perspectiva del estudiantado universitario de distintos países.

Quiroga et al., (2019) aseguran que se observa un cambio significativo en cuanto al proceso de enseñanza – aprendizaje, ya que con el empleo de las TIC se favorece el uso de metodologías activas como el aula invertida, la gamificación, el aprendizaje basado en problemas y el aprendizaje basado en retos. Sin embargo, también destacan algunos riesgos como la adicción al uso de redes sociales por altos periodos de tiempo.

Molinero y Chávez (2019) profundizan en torno a las TIC en el proceso de enseñanza aprendizaje en estudiantes de Educación Superior. En una muestra de 224 estudiantes tanto de licenciatura como de posgrado, encontraron que las herramientas ofimáticas son las más utilizadas, que ha sido poco el empleo de videoconferencias y que las plataformas en las que se trabaja con mayor frecuencia son Canvas y Socrative.

El empleo de distintas herramientas tecnológicas en Educación Superior ha permitido generar un proceso más dinámico entre estudiantado y profesorado. En el estudio realizado por Pardo-Cueva y otros (2020), se analizó la relación entre el empleo del Padlet de manera didáctica y el rendimiento académico. Se integró una muestra de 92 personas. Se observa una relación significativa entre los resultados del estudiantado y el uso del padlet, además de que el empleo de esta herramienta favorece la identificación de temas que son complejos en las asignaturas.

Otro estudio que aborda la temática de interés es el que realizan Amador y Velarde (2019) en una muestra de 310 integrantes del estudiantado en el Instituto Tecnológico José Mario Molina Pasquel y Henríquez (TECMM). Las competencias digitales que se manifestaron en niveles más altos son las relacionadas a la ciudadanía digital. Mientras que los más bajos se presentaron en las actividades de comunicación, colaboración, creatividad e innovación.

Sánchez-Otero et al., (2019) presentan el resultado de analizar cómo se han transformado las estrategias pedagógicas con la incorporación de las TIC en el ámbito educativo. Las principales conclusiones muestran que la difusión de las estrategias pedagógicas en los procesos de enseñanza – aprendizaje ha fortalecido de manera significativa la labor del profesorado.

Una de las herramientas que se emplean para acceder a las TIC son las aplicaciones móviles, fenómeno que estudian Rodríguez y otros (2019) quienes analizan las ventajas y desventajas de su uso. Además de realizar una clasificación de sus principales aplicaciones. También se destaca la importancia de las TIC en el ámbito educativo con una tendencia a incrementar el desempeño de los estudiantes.

Licona y Veytia (2019) aseguran que el empleo de las TIC en la Educación Superior generan relaciones de transformación e innovación en las Instituciones Educativas, posibilitan la integración de nuevos escenarios de enseñanza – aprendizaje por lo que aseguran es relevante incorporar su empleo de manera transversal en los distintos diseños curriculares de los programas educativos.

Por lo tanto, se asume que las TIC tienen un impacto significativo en la educación superior y la sociedad en general. Su versatilidad permite una variedad de usos, desde mejorar la competitividad económica hasta influir en el pensamiento y la acción de las personas. Sin embargo, es importante recordar que el acceso a una gran cantidad de información no garantiza la adquisición de conocimientos. Las TIC están transformando el proceso de enseñanza-aprendizaje, centrandose en el estudiantado, sus características particulares, estilos de aprendizaje y fomentando la colaboración. Las Instituciones Educativas desempeñan un papel crucial en la formación de ciudadanos digitales y en la reducción de la brecha digital. Se observan beneficios en el uso de las TIC, como el aumento del aprendizaje activo y la mejora del rendimiento académico. La evolución de las estrategias pedagógicas fortalece la labor del profesorado, y la incorporación de las TIC en la educación superior genera transformación e innovación en las instituciones educativas. Se destaca la importancia de integrar las TIC de manera transversal en los programas educativos para preparar a los estudiantes para los desafíos de la sociedad del conocimiento. Aunque las realidades Latinoamericanas tienen a ser similares, se considera de importancia generar diagnósticos que identifiquen las potenciales diferencias, caracterizarlas y con ello aportar a la toma de decisiones para el desarrollo de intervenciones y políticas educativas al interior de las Instituciones Educativas de acuerdo a las condiciones que su contexto les impone.

3. Metodología

El estudio se aborda desde el enfoque empírico-analítico (Ortiz, 2015), tipo de investigación cuantitativa, alcance descriptivo, que consiste en determinar las características de un evento de estudio dentro de un contexto en particular (Rojas, 2009); transversal, es decir que la obtención de los datos se realiza en un solo periodo de tiempo, y no experimental, en donde las variables no son manipuladas ni controladas (Hernández, et al., 2014).

Este estudio permite dar respuesta a la pregunta de investigación ¿De qué manera emplean las TIC el estudiantado universitario de cuatro países Latinoamericanos en el ámbito educativo?

La investigación se lleva a cabo en Instituciones de Educación Superior de cuatro países Latinoamericanos que comparten realidades y retos estructurales. El estudio se desarrolla en el primer semestre del 2023 con estudiantado de educación superior. La muestra fue elegida de manera no probabilística por conveniencia, y quedó representada

por 182 sujetos. La selección se realizó atendiendo a los siguientes criterios de inclusión: (a) ser estudiantado universitario; (b) tener disposición de participar en la investigación y, (c) tener acceso a un dispositivo electrónico. Se garantizó el anonimato y al inicio de la participación se firmó un consentimiento informado.

La técnica para la recolección de la información fue la encuesta, definida como el conjunto que procedimientos de investigación mediante los cuales se recoge y analizan una serie de datos de una muestra de casos representativa de una población (Casas, et al, 2003), y el instrumento fue el cuestionario. Este permitió obtener de manera organizada los indicadores de las variables establecidas en el objetivo de la encuesta.

Se diseñó un cuestionario *ad hoc* para medir la variable en formato en línea a través del empleo de Google Forms. Se confirmó la validez de contenido del cuestionario mediante la técnica de juicio de expertos. Participaron cinco expertos. Para verificar la consistencia interna de las dimensiones se aplicó la prueba de Alpha de Cronbach que obtuvo un valor de 0.94.

El procesamiento de análisis de los datos se llevó mediante el programa estadístico SPSS Versión 23. La definición operacional de la variable *uso de las TIC en el ámbito educativo* se muestra en la tabla 1.

Variable	Dimensiones
Uso de las TIC en el ámbito educativo	El uso de las TIC en el ámbito educativo para la búsqueda y selección de la información.
	Acciones para la realización de investigaciones y trabajos académicos con el uso de la tecnología.
	Seguridad digital en ámbitos educativos.
	La colaboración digital en el ámbito educativo.
	Las redes sociales como medio de aprendizaje.
	Empleo de software con fines académicos.

Fuente. Elaboración propia

Tabla 1 –

4. Resultados

Las edades de los sujetos objeto de estudio, oscilan entre los 18 y 55 años, la media es de 27 años (s=7.7). La distribución geográfica es como sigue: Bolivia (30.2%), Ecuador (28.6%), Guatemala (12.6%) y México (28.6%). En cuanto al género, 31.9% declararon ser hombres, y 67.6% mujeres.

El 18.1% de la muestra cuenta únicamente con teléfono móvil, el 17% únicamente con laptop personal o familiar, mientras que el restante 64.3% cuenta con ambos. En una escala que va de 1. Nunca; 2. Algunas veces y; 3.Siempre; se evidencia que la mayoría de la muestra cuenta con servicio de internet en casa o mediante su teléfono móvil. Sólo una minoría debe pagar para conectarse en un lugar externo.

Un 82.4% de la muestra afirma conocer el significado de la abreviatura TIC. En cuanto al tipo de cursos que han tomado para mejorar su desempeño en el uso de las TIC, se

encontró que la mayoría (53.8%) no ha tomado ningún curso. El 19.8% si cuenta con cursos de tecnología educativa y el 9.3% cursos de cómputo en general. Los niveles de capacitación coinciden con la autopercepción del nivel de dominio de las TIC y el Internet. Sólo el 26.4% considera que cuenta con nivel avanzado, mientras que el 73.6% considera que es básico. Ninguno de los encuestados se identificó como un usuario experto.

Uso de las TIC en el ámbito educativo

Existe casi un acuerdo generalizado (91.8%) de que la utilización de recursos tecnológicos como apoyo didáctico es necesario. Sólo un 8.2% lo considera opcional. En este sentido la gran mayoría (98.4%) afirman además que favorecen la adquisición de aprendizajes.

En la figura 1. se muestra la percepción que tiene el estudiantado del empleo de las herramientas para la búsqueda y selección de información con finalidades académicas. Los porcentajes más altos se presentan en el empleo de los buscadores como Google, Yahoo y otros, y buscadores especializados como Google Académico. Un segundo grupo utiliza revistas especializadas y solo algunos emplean conscientemente los operadores booleanos para conectar palabras de búsqueda que permitan ampliar o reducir las posibilidades de resultados.

La mayoría de la muestra utiliza directamente el internet para buscar la información, por lo que es relevante fortalecer los procesos de búsqueda especializada que les permita encontrar con mayor precisión los documentos relacionados con el tema de interés, además de que presenten un rigor científico.

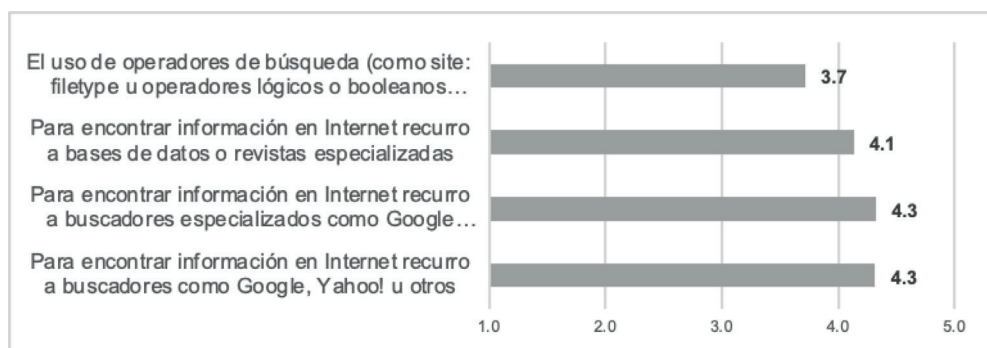


Figura 1 – Nivel de uso de herramientas de búsqueda de información
Fuente: Elaboración propia

En la figura 2. se recupera el proceso que llevan a cabo el estudiantado con la información obtenida, para realizar investigaciones y/o trabajos académicos. Los resultados más altos mencionan que contrastan la información encontrada con otras fuentes. Otro aspecto que destacan es la identificación de los argumentos, lo que les permite evaluarlos, analizarlos y obtener conclusiones. Finalmente, en un tercer nivel de importancia examinan la información e identifican si es confiable y veraz. También comparan distintos puntos de vista planteados por los autores, valoran semejanzas y diferencias entre ambos.

Este tipo de actividades que realizan los estudiantes con la búsqueda, selección y análisis de la información fortalece el desarrollo del pensamiento crítico, definido por Paul y Elder (2005, p. 7) como “el proceso de analizar y evaluar el pensamiento con el propósito de mejorarlo [...] presupone el conocimiento de las estructuras más básicas del pensamiento (los elementos del pensamiento), y los estándares intelectuales más básicos del pensamiento (estándares intelectuales universales)”.

Cangalaya (2020) destaca cuatro habilidades fundamentales para pensar de manera crítica: (a) la argumentación, es decir, la razón que permita sustentar y respaldar la conclusión final; (b) el análisis, que es la habilidad que permite separar el todo en sus partes, lo cual permite explicar el fenómeno desde sus características; (c) solución de problemas, que consiste en la ejecución de una serie de tareas para dar respuesta a una pregunta o una situación y; (d) la evaluación, que permite tomar decisiones ante las distintas situaciones que enfrenta el sujeto.

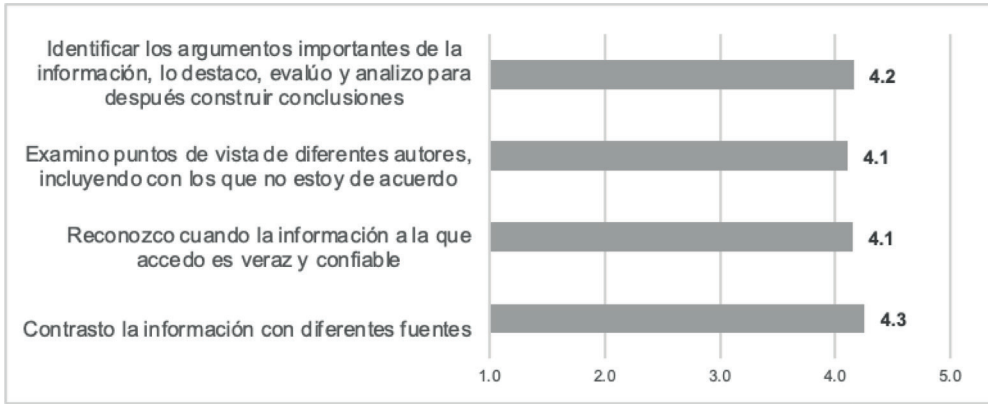


Figura 2 – Acciones para elaborar investigaciones o trabajos académicos
Fuente: Elaboración propia

Otro de los aspectos fundamentales que requiere considerar el estudiantado universitario al utilizar las TIC en el ámbito educativo es el desarrollo de hábitos y acciones que permitan realizar una práctica ética, responsable y segura. Esto se conoce actualmente con el término de ciudadanía digital, la cual, concibe al sujeto desde un rol activo que desarrolla su propia identidad. Es una nueva manera de definirse a sí mismo en ambientes virtuales (Lozano-Díaz y Fernández-Prados, 2019). Dado que la seguridad digital adquiere gran relevancia, se aborda en las preguntas que se presentan en la figura 3.

Dentro de los hallazgos más relevantes se encuentran que el 91% asegura que las contraseñas que utiliza están conformadas por números, letras y caracteres especiales, el 87% evita publicar información confidencial en internet como es su número telefónico, fotos y la localización geográfica, el 81% controla el acceso a sus perfiles, y organiza sus contactos en círculos y grupos, el 80% tienen instalado antivirus en su computadora. En general se observan la realización de acciones concretas que fortalecen la seguridad en la navegación en internet y en el desarrollo de actividades académicas.

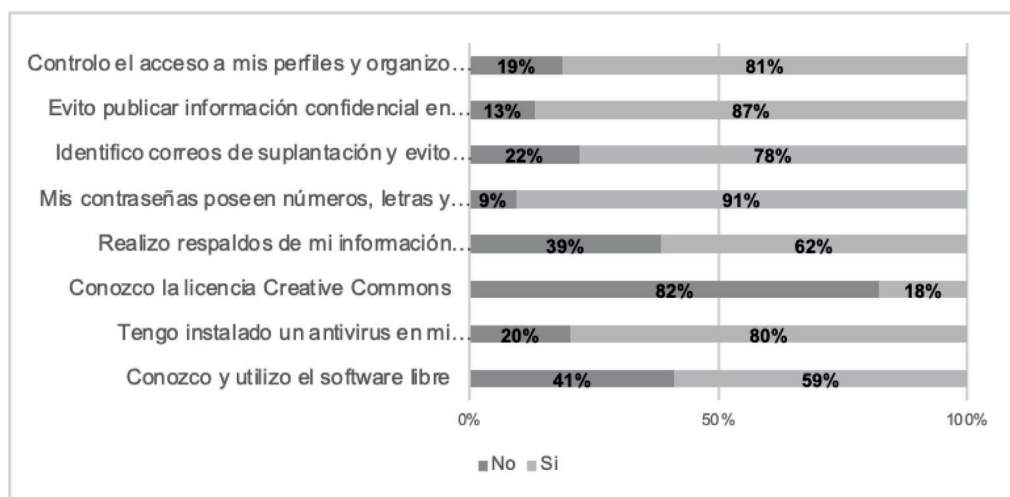


Figura 3 – Seguridad digital en ámbitos educativos
Fuente. Elaboración propia

El aprendizaje colaborativo constituye una de las competencias clave en la sociedad del siglo XXI. Adquiere cada vez más importancia en los procesos educativos en el nivel universitario, ya que la participación colaborativa fortalece la comunicación. Permite expresar y compartir ideas, puntos de vista, debatir, argumentar, favorecer el desarrollo de aprendizajes activos y significativos, tomar decisiones y solucionar problemáticas (Roig-Vila, 2019).

En la figura 4. Se mencionan algunas herramientas de comunicación que utiliza el estudiantado universitario con mayor frecuencia, entre los resultados se observa que lo más empleado son los chats en dispositivos móviles con aplicaciones como el *WhatsApp* y el *Telegram*. También el manejo de las redes sociales, que les ha permitido establecer comunicación con compañeros y generar grupos por temas de interés o actividades a realizar. Se continúa utilizando el correo electrónico, principalmente para cuestiones académicas. Con un menor porcentaje se utilizan las herramientas para llevar a cabo videoconferencias con aplicaciones como *Zoom*, *Meet* o *Skype*.

Las redes sociales son cada vez más empleadas por las personas, sin embargo, se ha observado que el estudiantado universitario hace uso de ellas no solo con una función social o de recreación, sino desde una perspectiva didáctica, y llevan a cabo la selección de la red social que consideran más adecuada para el tipo de actividades que van a realizar (Fernández, et al, 2019).

En la figura 5 se observa que se utiliza a las redes sociales como medio de aprendizaje al crear y compartir herramientas con un 72%, el uso de instagram o pinterest para compartir imágenes con colegas con un 50%. El manejo de grupos en *Facebook* para debatir alguna temática y generar trabajo colaborativo con un 41%. Se observa que el alumnado identifica las potencialidades de la red social y genera actividades para

fortalecer procesos de enseñanza – aprendizaje desde una manera significativa y con herramientas que ellos emplean de forma cotidiana.

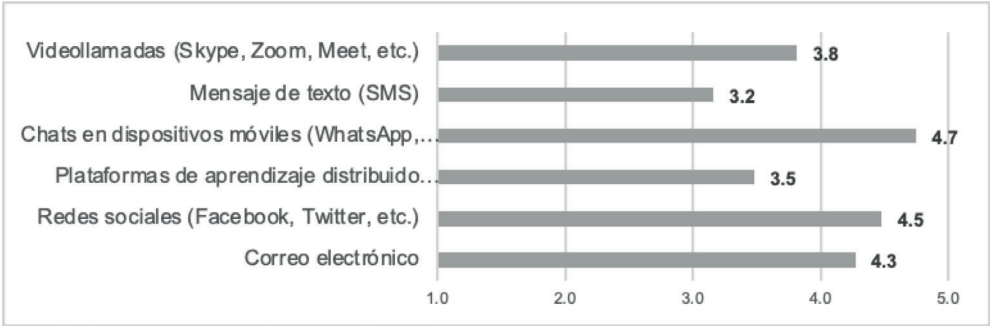


Figura 4 – Herramientas de comunicación.
Fuente. Elaboración propia



Figura 5 – Redes sociales como medio de aprendizaje.
Fuente: Elaboración propia.

El uso de software libre se define a los programas, procedimientos y documentos relacionados con el hardware, y es una herramienta para obtener resultados en el procesamiento de datos (Faría, 2020). Para el estudiantado universitario favorece los procesos de organización, análisis y evaluación de la información en las diferentes actividades y tareas académicas.

En la figura 6 se presentan las principales actividades que se llevan a cabo con fines académicos mediante el empleo del software libre. Destacan las funciones de revisión ortográfica y para dar formato al documento, el manejo de imágenes y videos para incorporar a presentaciones que utilizarán para desarrollar un tema, empleo de imágenes y videos para enriquecer un documento en *Word*. En un menor porcentaje se menciona el empleo de *diapositivas maestras* en un administrador de presentaciones y las funciones especiales como tablas de contenido e índices automatizados.

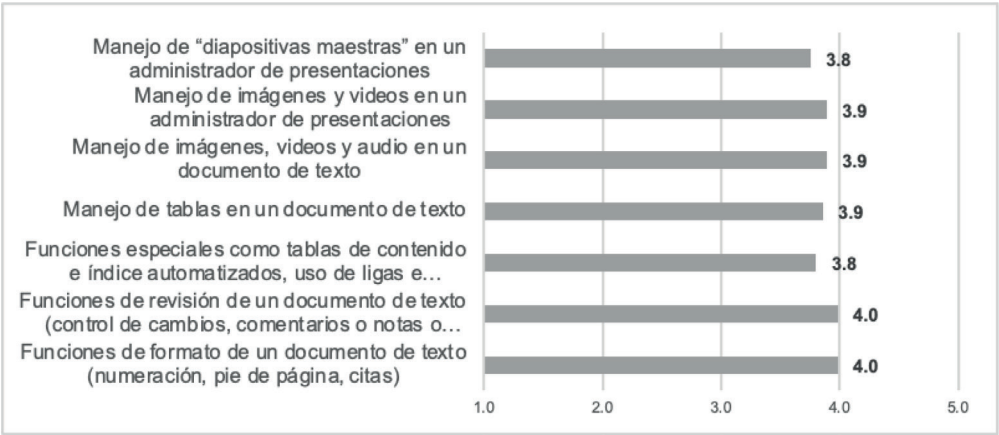


Figura 6 – Empleo de software con fines académicos.
Fuente. Elaboración propia

Un diferenciador de este estudio es la participación de estudiantado universitario de cuatro países latinoamericanos. En la figura 7 y la tabla 2 se presentan las frecuencias de las distintas dimensiones que constituyen a la variable uso de las TIC con fines académicos. Se destaca Ecuador por su participación en ciudadanía digital y uso de software, también México en ciudadanía digital. Un país que llama la atención por sus bajos índices es Guatemala, particularmente en o que se refiere al manejo de programas. Existen diferencias estadísticamente significativas en la variable *Uso de las TIC en el ámbito Educativo*, así como en las dimensiones *Manejo de Programas y Sistemas de Información relativos a su área de Conocimiento* y *Alfabetización digital* con valores p en la prueba de Kruskal-Wallis para muestras independientes de 0.026, 0.00 y 0.045 respectivamente.

	$\bar{x}$				Cuartiles		
	México	Guatemala	Ecuador	Bolivia	Q1	Q2	Q3
Uso de las TIC en el ámbito educativo	4.0	3.8	4.0	4.2	3.6	4.1	4.6
Uso digital	4.2	4.0	4.2	4.0	3.8	4.1	4.6
Ciudadanía digital	4.1	4.0	4.2	4.0	3.8	4.3	4.8
Comunicación, socialización y colaboración	4.0	4.1	4.2	4.0	3.8	4.1	4.5
Uso de Software	4.0	4.1	4.2	4.0	3.7	4.1	4.4
Manejo de programas y sistemas de información relativos a su área de conocimiento.	3.9	3.4	4.0	3.5	3.4	3.9	4.3
Alfabetización digital	4.0	3.9	4.1	4.9	3.8	4.0	4.4

Fuente: Elaboración propia.

Tabla 2 – Estadísticos descriptivos de la variable e indicadores

Las diferencias en el uso de las TIC en la educación entre países como México, Guatemala, Ecuador y Bolivia pueden ser atribuidas a una serie de factores que incluyen: (a) la disponibilidad y accesibilidad de dispositivos como computadoras, tabletas y acceso a Internet varían significativamente entre estos países (Jiménez-Pitre, et al., 2023); (b) los gobiernos y las instituciones educativas en estos países tienen diferentes niveles de inversión en infraestructura de TIC en las escuelas (Vázquez, et al., 2019); (c) la formación y capacitación de los docentes en el uso efectivo de las TIC en la enseñanza varía entre países. La falta de capacitación adecuada puede limitar la integración de la tecnología en el aula (Padilla y Ayala, 2021); (d) las políticas gubernamentales y los enfoques educativos pueden influir en el uso de las TIC en la educación (Ríos-Cabrera y Ríos-Bolívar, 2020); (e) la disponibilidad de fondos públicos y privados para la adquisición de tecnología educativa puede ser muy diferente entre estos países (Cabero-Almenara y Valencia-Ortiz, 2019); (f) las actitudes y percepciones sobre el uso de las TIC en la educación también varían, influyendo en su implementación (Ruíz-Aquino, et al., 2022); (g) la diversidad cultural y lingüística también puede influir en la forma en que se utilizan las TIC en la educación. La adaptación de la tecnología para atender a diferentes grupos étnicos y lingüísticos es un desafío en países latinoamericanos (Arispe, 2020).

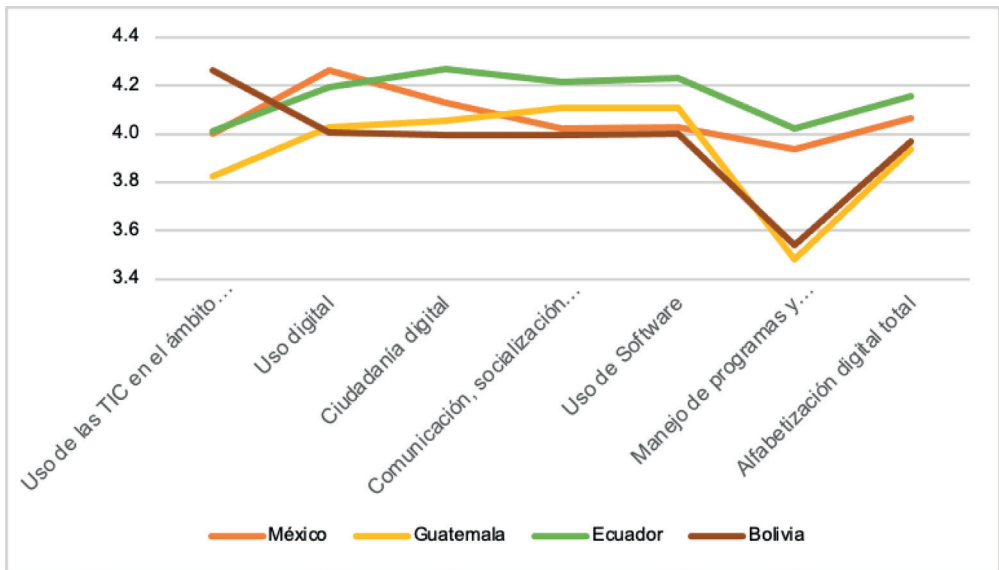


Figura 7 – Uso de las TIC para fines académicos y sus dimensiones por país
Fuente: Elaboración propia.

Mediante pruebas no paramétricas se encontraron relaciones estadísticamente significativas ($p < 0.05$) entre las características sociodemográficas de la muestra y la variable *Uso de las TIC en el ámbito educativo* y sus dimensiones, a saber:

Tener posgrado se asocia a mayores niveles en el uso de las TIC en el ámbito educativo. Esto puede deberse a que los programas de posgrado suelen fomentar la capacitación en tecnología y promover el uso de herramientas digitales en la enseñanza. Los estudiantes

de posgrado pueden tener una mayor disposición y habilidad para aprovechar las TIC en su práctica educativa, lo que puede conducir a un mayor uso de la tecnología en el aula. Estos resultados coinciden con los de Urcid (2022) en México y Ames (2019) en Perú.

Tener licenciatura se asocia a mayores niveles de ciudadanía digital y mayores niveles de manejo de programas y sistemas. Esto puede explicarse por la diferencia de edades de las muestras, (licenciatura 23.5 años $s=4.2$ y posgrado es de 36.5 años, $s=9.2$) y la implicación de pertenecer a los grupos de inmigrantes o nativos digitales. Según Claro y otros (2021), tener mayor edad, así como otros factores -género femenino, procedencia indígena y bajo nivel socioeconómico- influyen en el desarrollo de ciudadanía digital. Esto también se fundamenta en el hallazgo de que una menor edad se asocia a mayor manejo de programas y sistemas. Esto sugiere también, que los programas de licenciatura pueden incluir componentes educativos que promuevan el desarrollo de habilidades digitales y el conocimiento de programas y sistemas informáticos relevantes. Las personas con licenciatura pueden tener una mayor conciencia de las responsabilidades y los comportamientos éticos en línea, así como una mayor capacidad para utilizar diferentes programas y sistemas tecnológicos.

Reconocer qué significan las TIC se asocia a mayores niveles de: (a) comunicación, socialización y colaboración; (b) uso de software; (c) manejo de programas y sistemas. Esto es comprensible, ya que tener un conocimiento sólido sobre las TIC permite a las personas utilizar eficazmente diferentes herramientas y recursos digitales para comunicarse, colaborar y acceder a información. El reconocimiento y comprensión de las TIC son fundamentales para desarrollar habilidades digitales avanzadas y aprovechar al máximo el potencial de la tecnología (Recio, et al., 2020). En este sentido reconocer qué significan las TIC, se asocia a menores niveles de brecha digital académica en la muestra, así como considerar que los recursos tecnológicos favorecen la adquisición de aprendizajes se asocia a mayor uso de las TIC en el ámbito educativo.

Ser boliviano se asocia a un menor reconocimiento del concepto de las TIC. Esto podría estar relacionado con factores socioeconómicos, acceso limitado a recursos tecnológicos o diferencias en los sistemas educativos. Es importante abordar estas disparidades y trabajar hacia una mayor inclusión y acceso a la educación digital en todos los contextos. Sin embargo, Núñez-Ramírez, et al., (2021) afirman que una intervención que oriente hacia el reconocimiento de las TIC se asocia con su uso para fines educativos en esta población.

Finalmente, se encontró que estar en el rango de entre 30 y 40 años se asocia a un menor reconocimiento del concepto de las TIC. En este punto cabe destacar que el 92.6% de los encuestados tienen hasta 40 años. Esto puede deberse a que las personas de este grupo de edad pueden no haber crecido con la misma exposición a las tecnologías digitales como las generaciones más jóvenes. No obstante, es importante destacar que la alfabetización digital y el reconocimiento de las TIC son habilidades que se pueden desarrollar y mejorar en cualquier etapa de la vida (Valenzuela, et al., 2022).

5. Conclusiones

Es necesario promover la formación y acceso a programas de posgrado en el ámbito educativo. Dado que tener un posgrado se asocia con mayores niveles de uso de las TIC en la educación, y en particular para los profesionales de la educación, como sustento para

la mejora de su competencia digital y a utilizar de manera más efectiva las tecnologías en el aula y en los procesos de aprendizajes.

Fortalecer procesos formativos que incorporen el uso de las TIC, condiciona la alfabetización y promoción de una ciudadanía digital en programas de licenciatura, en función de generar un mayor nivel de dominio del uso diferentes herramientas tecnológicas, lo cual se asocia con mayores niveles de ciudadanía digital, por lo que se puede incidir en un proceso de diagnóstico, que de cuenta de las necesidades y demandas formativas, desde donde se genere una diversidad cursos y asignaturas que aborden específicamente temáticas que promuevan la resignificación y uso de las TIC en los programas de licenciatura. Esto permitiría que los futuros educadores adquirieran las habilidades y conocimientos necesarios para ser ciudadanos digitales responsables y éticos en el ejercicio de su práctica.

Brindar capacitación en el uso de programas y sistemas: dado que tanto tener el nivel de Licenciatura como reconocer qué significan las TIC, se asocian con mayores niveles de manejo de programas y sistemas, por lo que sería importante proporcionar capacitación y formación continua a los docentes y profesionales en general. Esto les permitiría adquirir las habilidades técnicas necesarias para utilizar eficientemente los programas y sistemas tecnológicos en su trabajo.

Fomentar el reconocimiento del concepto de las TIC, porque reconocer qué significan las TIC se asocia con diversos beneficios, como mayor comunicación, socialización, colaboración y uso de software, es esencial promover una mayor comprensión y conciencia de lo que abarca el término TIC. Esto se puede lograr a través de campañas de sensibilización, programas de educación y recursos informativos dirigidos tanto a docentes como a la comunidad en general.

Investigar las causas del menor reconocimiento del concepto de las TIC en Bolivia y entre personas de 30 a 40 años. Sería de interés llevar a cabo investigaciones para comprender las razones detrás de estos patrones observados. Esto podría incluir análisis de factores socioeconómicos, diferencias en los sistemas educativos, acceso a la tecnología y características culturales que podrían influir en la comprensión y adopción de las TIC. Estos hallazgos podrían servir de base para desarrollar estrategias específicas destinadas a cerrar la brecha digital en esos grupos.

En resumen, se considera necesario fomentar la formación académica, fortalecer la educación en ciudadanía digital, proporcionar capacitación en el uso de programas y sistemas, promover el reconocimiento del concepto de las TIC y llevar a cabo investigaciones para comprender las disparidades observadas. Estas acciones pueden contribuir a mejorar el uso y aprovechamiento de las TIC en el ámbito educativo y promover la alfabetización digital en general.

Referencias

Amador, C. M., Velarde, L. (2019). Competencias para el uso de las TIC en estudiantes de educación superior: un estudio de caso. *Revista Iberoamericana para la investigación y el Desarrollo Educativo*. 10, 1-27, <https://doi.org/10.23913/ride.v10i19.515>

- Ames, P. (2019). El uso de materiales audiovisuales y recursos digitales en la docencia universitaria: una experiencia de innovación a nivel de posgrado en Perú. *REDU Revista de Docencia Universitaria*. 17(1),167-182. <https://doi.org/10.4995/redu.2019.9894>
- Arriaga, W., Bautista, J.K., Montenegro, L. (2021). Las TIC y su apoyo en la educación universitaria en tiempo de pandemia: Una fundamentación facta teórica. *Conrado*. 17(78), 201-206. <http://scielo.sld.cu/pdf/rc/v17n78/1990-8644-rc-17-78-201.pdf>
- Arispe, V. (2020). Educación intercultural. La perspectiva de los pueblos indígenas de Bolivia. *Caracol*. (20), 166-187. <https://dialnet.unirioja.es/servlet/articulo?codigo=7888745>
- Bernard, M. (2006). *Formación, distancias y tecnología*. Ediciones Pomares.
- Bernate, J. A., Vargas, J. A. (2020). Desafíos y tendencias del siglo XXI en la Educación Superior. *Revista de Ciencias Sociales (RCS) FCES-LUZ*. 26(2), 141-154. <https://dialnet.unirioja.es/servlet/articulo?codigo=7599937>
- Cabero-Almenara, J., Valencia-Ortiz, R. (2019). TIC para la inclusión: una mirada desde Latinoamérica. *Aula Abierta*, 48(2), 139–146. <https://doi.org/10.17811/rifie.48.2.2019.139-146>
- Cangalaya, L. M. (2020). Habilidades del pensamiento crítico en estudiantes universitarios a través de la investigación. *Desde el sur*, 12(1), 141-153. <http://www.scielo.org.pe/pdf/des/v12n1/2415-0959-des-12-01-141.pdf>
- Casas, J., Repullo, J. R., Donado, J. M. (2003). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos. *Atención primaria*. 31(8), 527-538, [https://doi.org/10.1016/S0212-6567\(03\)70728-8](https://doi.org/10.1016/S0212-6567(03)70728-8)
- Cebrián, J. L. (2014). *La Red. Cómo cambiarán nuestras vidas los nuevos medios de comunicación*. Taurus.
- Chávez, M., Rivera, V., Haro, G. (2021). Percepción de la Educación Virtual en Instituciones de Educación Superior 2020-2020. *Revista de Investigación Enlace Universitario*. 20(1), 8-21. <http://doi.org/10.33789/enlace.20.1.81>
- Claro, M., Santana, L. E., Alfaro, A., Franco, R. (2021). *Ciudadanía digital en América Latina: revisión conceptual de iniciativas*, serie Políticas Sociales, N° 239 (LC/TS.2021/125), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL).
- Coll, C., Monereo, C. (2011). *Psicología de la educación virtual*. Morata.
- Coll, C., Mauri, T., Onrubia, J. (2011). La utilización de las tecnologías de la información y la comunicación en la educación: Del diseño tecno-pedagógico a las prácticas de uso. En: Coll, César, Monereo, Carlos. *Psicología de la Educación Virtual*. 74-103.
- Crovi, D. (2006). *Educación en la era de las redes*. UNAM.

- Deroncele-Acosta, A., et al (2021). Innovación Educativa con TIC en Universidades Latinoamericanas: Estudio Multi-País. REICE Revista Iberoamericana sobre Calidad, Eficacia y Cambio Social. 19(4), 145-161. <https://doi.org/10.15366/reice2021.19.4.009>
- Díaz-García, I., Almerich, G., Suárez-Rodríguez, J., Orellana, N. (2020). La relación entre las competencias TIC, el uso de las TIC y los enfoques de aprendizaje en alumnado universitario de educación. *Revista de Investigación Educativa*, 38(2), 549-566. <http://dx.doi.org/10.6018/rie.409371>
- Faría, J. A. (2020). Impacto del Software Libre en el Aprendizaje de los Estudiantes Universitarios. *Revista de Ciencias Sociales*, 1(4), 43-52. <http://eprints.rclis.org/43134/1/5.%20Impacto%20del%20software%20libre.pdf>
- Fernández, J. C., Otero, L., Fernández-Morante, C., Cebreiro, B. (2019). Actitudes y uso de redes sociales en estudiantes universitarios/as de Galicia: Implicaciones personales y sociales. *Revista Prisma Social*, 28, 145-160. <https://revistaprismasocial.es/article/view/3372/4085>
- Hernández, R., Fernández, C., Baptista, M. P. (2014). *Metodología de la Investigación*. McGraw-Hill.
- Huamán, L., Torres, L. A., Amancio, A. M., Sánchez, S. (2021). Educación remota y desempeño docente en las instituciones educativas de Huancavelica en tiempos de COVID-19. *Apuntes Universitarios*, 11(3), 45-59. <https://doi.org/10.17162/au.v11i3.692>
- Huana, J. C., Portal, J. A. (2023). Análisis de contenido cuantitativo sobre gestión del conocimiento en instituciones de educación superior latinoamericanas. *Revista Actualidades Investigativas en Educación*, 23(1), 1-30. <https://doi.org/10.15517/aie.v23i1.51513>
- Jiménez-Pitre, I., Molina-Bolívar, G., Gámez Pitre, R. (2023). Visión sistémica del contexto educativo tecnológico en Latinoamérica. *Región Científica*, 2(1), 202358. <https://doi.org/10.58763/rc202358>
- Licon, K., Veytia, M. G. (2019). El empleo de las TIC en la educación superior. *Educando para educar*, 37, 90-99. <https://beceneslp.edu.mx/ojs2/index.php/epe/article/view/47/46>
- López García C., Sánchez Gómez M. C., García-Valcárcel Muñoz-Repiso A. (2021). Desarrollo de la Competencia Digital en estudiantes de primaria y secundaria en tres dimensiones: fluidez, aprendizaje-conocimiento y ciudadanía digital. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, 44(12). 5-20 <http://dx.doi.org/10.17013/risti.44.5-20>
- Loveless, A., Williamson, B. (2017). *Nuevas identidades de aprendizaje en la era digital. Creatividad – Educación – Tecnología – Sociedad*. Nárcea Ediciones.

- Lozano-Díaz, A., Fernández-Prados, J. S. (2019). Hacia una educación para la ciudadanía crítica y activa en la universidad. *Revista Latinoamericana de Tecnología Educativa*, 18(1), 175-187, <https://dehesa.unex.es:8443/handle/10662/9796>
- Luna-Romero, A. E., Vega, F. Y., Luna, M. E. (2019). Las tecnologías de la información y la comunicación en la consolidación de la Educación Superior a distancia en América Latina. *Conrado*, 15(67), 32-37. <https://conrado.ucf.edu.cu/index.php/conrado/article/view/920>
- Maldonado, G. A., García, J., Sampedro-Requena, B. E. (2019). El efecto de las TIC y redes sociales en estudiantes universitarios. *RIED Revista Iberoamericana de Educación a Distancia*, 22(2), 153-176, <https://doi.org/10.5944/ried.22.2.23178>
- Mañas, A., Roig-Vila, R. (2019). Las Tecnologías de la Información y la Comunicación en el ámbito educativo: Un tándem necesario en el contexto de la sociedad actual. *Revista Internacional d'Hummanitas* 45, 75-86. <https://rua.ua.es/dspace/handle/10045/82089>
- Molinero, M. C., Chávez, U. (2020). Herramientas tecnológicas en el proceso de enseñanza – aprendizaje en estudiantes de Educación Superior. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*. 10(19), 1-31. <https://doi.org/10.23913/ride.v10i19.494>
- Murillo, G., Martínez, G., Jiménez, F. J., Silva, F. (2021). Estrategia pública para reducir la brecha digital en el sector educativo y salud en el Estado de Tabasco antes de la Pandemia. *Revista de Investigaciones Universidad del Quindío*, 33(2), 138-142. <https://doi.org/10.33975/riuv.vol33n2.747>
- Núñez-Ramírez, M. A., Atila-Lijerón, J. D., Banegas-Rivero, R. A., Esparza-García, I. G. (2021). Predictores de la intención hacia el uso de tecnologías de la información y la comunicación (TIC) por profesores universitarios en Bolivia durante la pandemia por COVID-19. *Formación Universitaria*, 14(6), 109-118. <http://dx.doi.org/10.4067/S0718-50062021000600109>
- ONU (2018). *La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe*. Organización de las Naciones Unidas.
- Ortiz, A. (2015). *Enfoques y métodos de investigación en las Ciencias Sociales*. Ediciones de la U.
- Padilla, J. C., Ayala, G. G. (2021). Competencias digitales en profesores de educación superior de Iberoamérica: una revisión sistemática. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 12(23). <https://doi.org/10.23913/ride.v12i23.1096>
- Pardo-Cueva, M., Chamba-Rueda, L. M., Gómez, A. H., Jaramillo-Campoverde, B. G. (2020). Las TIC y rendimiento académico en la educación superior: Una relación potenciada por el uso del Padlet. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, E28, 934-944, <https://www.proquest.com/docview/2388305711?pq-origsite=gscholar&fromopenview=true>

- Paul, R., Elder, L. (2005). *Miniguía del pensamiento crítico*. Fundación para el pensamiento crítico. <https://www.criticalthinking.org/resources/PDF/SP-ConceptsandTools.pdf>
- Ponce, F.F., Zambrano, J.M., De la Peña, G. (2023). El desarrollo de aulas virtuales en contextos universitarios: análisis de las percepciones desde la perspectiva de los docentes. *Revista Cubana de Educación Superior*. 42(2), 1-10. <https://cutt.ly/FwFkmKOU>
- Poveda-Pineda, D., Cifuentes-Medina, J. (2020). Incorporación de las tecnologías de información y comunicación (TIC) durante el proceso de aprendizaje en la educación superior. *Formación Universitaria*. 13(6), 95-104. <http://dx.doi.org/10.4067/S0718-50062020000600095>
- Quiroga, L. P., Vanegas, O. L., Pardo, S. (2019). Ventajas y desventajas de las TIC en educación desde la primera infancia hasta la educación superior. *Revista de Educación y Pensamiento*. 76-85. <http://www.educacionypensamiento.colegiohispano.edu.co/index.php/revistaeypp/article/view/103/92>
- Recio, F., Silvia, J. E., Abricot, N. (2020). Análisis de la competencia digital en la formación inicial de estudiantes universitarios: un estudio de meta-análisis en la Web of Science. *Pixel-Bit*. 59,125-146, <https://recyt.fecyt.es/index.php/pixel/article/view/77759/61040>
- Ríos-Cabrera, P., Ruiz-Bolívar, C. (2020). La innovación educativa en América Latina: lineamientos para la formulación de políticas públicas. *Innovaciones Educativas*, 22(32), <http://dx.doi.org/10.22458/ie.v22i32.2828>
- Rodríguez, A., Rocío, E., Zambrano-Cedeño, E., Zambrano, V., Rodríguez, G. (2019). TICS y aplicaciones móviles en la Educación Superior, del dicho al reto. *Revista Atlante Cuadernos de Educación y Desarrollo*. <https://www.eumed.net/rev/atlante/2019/01/tics-educacion-superior.html>
- Rojas, R. (2009). *Guía para realizar investigaciones sociales*. Plaza y Valdés.
- Roig-Vila, R. (2019). *Investigación e innovación en la Enseñanza Superior. Nuevos contextos, nuevas ideas*. Octaedro.
- Paucar-Curasma, R., Villalba-Condori, K.O., Fernando Viterbo, S.C., Nolan, J.J., Unsihuay-Tovar, R.F., Rondon, D. (2022). Fomento del pensamiento computacional a través de la resolución de problemas en estudiantes de ingeniería de reciente ingreso en una universidad pública de la región andina del Perú. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, 48(12),23-39. <https://doi.org/10.17013/risti.48.23-40>
- Ruíz-Aquino, M., Borneo, E., Alania-Contreras, R. B., Garcia, E. S., Zevallos, U. (2022). Actitudes hacia las TIC y uso de los entornos virtuales en docentes universitarios en tiempos de pandemia de la COVID-19. *Publicaciones*, 52(3), 107–120, <https://doi.org/10.30827/publicaciones.v52i3.22270>

- Sánchez-Otero, M., García-Guiliany, J., Steffens-Sanabria, E., Hernández-Palma, H. (2019). Estrategias Pedagógicas en Procesos de Enseñanza y Aprendizaje en la Educación Superior incluyendo Tecnologías de la Información y la Comunicación. *Información Tecnológica*. 30(3), 277-286. <http://dx.doi.org/10.4067/S0718-07642019000300277>
- Santos-Rego, M. A., Mella-Núñez, I., Sotelino-Losada, A. (2020). Movilidad y TIC en aprendizaje-servicio: perspectivas para una sociedad global y tecnológica. *RIED Revista Iberoamericana de Educación a Distancia*, 23(1), 67-84. <http://dx.doi.org/10.5944/ried.23.1.24180>
- Soomro, A. K., Kale, U., Curtis, R., Akcaoglu, M., Bernstein, M. (2020). Digital divide among higher education faculty. *International Journal of Educational Technology in Higher Education*. 17(21), 1-16, <https://doi.org/10.1186/s41239-020-00191-5>
- Tirado, P. J., Roque, M. P. (2019). TIC y contextos educativos: frecuencia de uso y función por universitarios. *EDUTEC. Revista Electrónica de Tecnología Educativa*. 67, 31-47. <https://doi.org/10.21556/edutec.2019.67.1135>
- Torres, C. A. (2018). Formas de participación en línea en estudiantes de la Facultad de Administración de la Universidad Veracruzana en México. *Revista Actualidades Investigativas en Educación*, 18(2), 1-29. <https://doi.org/10.15517/aie.v18i2.33131>
- UNESCO (2014). Enfoques estratégicos sobre las TICS en educación en América Latina y el Caribe. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000223251>
- UNESCO (2021). Caminos hacia 2050 y más allá. Resultados de una consulta pública sobre los futuros de la Educación Superior. UNESCO. <https://www.iesalc.unesco.org/los-futuros-de-la-educacion-superior/caminos-hacia-2050-y-mas-alla/>
- Urcid, R. (2022). Autoaprendizaje mediado por las TIC. Estudio de caso: alumnado de la maestría en educación. *EDUTEC. Revista Electrónica de Tecnología Educativa*. 79, (marzo), <https://hdl.handle.net/11162/245944>
- Valencia-Molina, T., Serna-Collazos, A., Ochoa-Angrino, S., Caicedo-Tamayo, A. M., Montes-González, J. A., Chávez-Vescance, J. D. (2016). *Competencias y estándares TIC desde la dimensión pedagógica. Una perspectiva desde los niveles de apropiación de las TIC en la práctica educativa docente*. UNESCO.
- Valenzuela, C., Rodríguez, F., Oliveros, S. (2022). Gobernanza electrónica e inclusión digital de personas mayores mediante estrategias de alfabetización digital e informacional en la localidad de Placilla, Valparaíso, Chile. *Palabra clave*, 12(1), <http://dx.doi.org/https://doi.org/10.24215/18539912e168>
- Vázquez, E. A., Bottamedi, J., Brizuela, M. L. (2019). Pensamiento computacional en el aula: el desafío en los sistemas educativos de Latinoamérica. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, (7). <https://doi.org/10.6018/riite.397901>

- Vidal, P., Pinargote, K. G. (2019). El impacto de la brecha digital en los procesos de enseñanza y aprendizaje en la Universidad Laica Eloi Alfaro de Manabí. *Revista Electrónica Formación y Calidad Educativa* 7(1), 1-14. <https://refcale.uleam.edu.ec/index.php/refcale/article/view/2970>
- Zambrano, D. L. y Zambrano, M. A. (2019). Las Tecnologías de la Información y las Comunicaciones /TICs) en la Educación Superior: Consideraciones teóricas. *Revista Electrónica Formación y Calidad Educativa*, 7(1), 213-218. <http://refcale.uleam.edu.ec/index.php/refcale/article/view/2750>

Impacto de ChatGPT en la enseñanza: Un enfoque de aula invertida para fundamentos de programación

Michael Apaza Calsin¹, Marco Aedo², Eveling Castro³

mapazacals@unsa.edu.pe; maedol@unsa.edu.pe; ecastro@unsa.edu.pe

^{1, 2, 3} Escuela Profesional de Ingeniería de Sistemas, Universidad Nacional de San Agustín, Santa Catalina 117, C.P. 04001, Arequipa, Perú.

DOI: 10.17013/risti.52.97-112

Resumen: Actualmente las herramientas de Inteligencia Artificial (IA) han surgido de forma disruptiva afectando a campos como el de las tecnologías de la información, específicamente al desarrollo de software, para el cual herramientas como ChatGPT abren nuevas oportunidades y desafíos en la enseñanza y el aprendizaje de conceptos de programación. Actualmente, esta se basa en métodos pedagógicos convencionales y consideramos la adaptación esencial pues estas herramientas de IA cuentan con un potencial que aún está en exploración. Integrar las herramientas de IA en la educación requiere una comprensión profunda de sus capacidades y limitaciones, por eso, en este artículo describimos nuestra experiencia en la adopción de un enfoque de aula invertida para la enseñanza de tópicos básicos de programación aplicada sobre estudiantes de primer año de universidad.

Palabras-clave: Inteligencia Artificial, ChatGPT, Educación, Programación

Impact of ChatGPT on Teaching: A Flipped Classroom Approach for Programming Fundamentals

Abstract: Currently, Artificial Intelligence (AI) tools have emerged disruptively, affecting fields such as information technology, specifically software development. Tools like ChatGPT, for instance, open new opportunities and challenges in teaching and learning programming concepts. Currently, education relies on conventional pedagogical methods, but we consider adaptation essential because these AI tools have untapped potential. Integrating AI tools into education requires a profound understanding of their capabilities and limitations. In this article, we describe our experience in adopting a flipped classroom approach for teaching basic programming topics to first-year university students..

Keywords: Artificial Intelligence, ChatGPT, Education, Programming

1. Introducción

En los recientes años se han dado varias mejoras en el campo tecnológico, especialmente en aquellas herramientas disruptivas basadas en inteligencia artificial, como lo es

ChatGPT, un gran modelo de lenguaje (LLM) desarrollado por OpenAI, un ChatBot soportado bajo el modelo GPT-3.5 (Chat Generative Pre-Trained Transformer, por sus siglas en inglés) enfocado en la conversación (Tian, Lu, On, Tang, Cheung, Klein & Bissyandé, 2023), que ha alcanzado un amplio rango de usuarios desde su lanzamiento en 2022, siendo una de las principales herramientas que más han influido y de las que más interés ha causado en los últimos años desde su apertura al público.

La gran aceptación de esta herramienta y la motivación e interés que supuso para las grandes tecnológicas estimuló la inversión de recursos en la elaboración de sus propios modelos, como Bard de Google o LLaMA de Meta. Este avance provocó una ola de preocupación en las personas sobre el impacto que podrían llegar a tener o cómo es que estas herramientas les afectaría en un futuro (Vorobeva et al., 2022), y según (Baidoo-anu & Owusu, 2023) una sensación dividida entre educadores, adoptando este tipo de herramientas como el futuro o tomando una posición escéptica al indicar que podría apaciguar o eliminar las habilidades analíticas de los estudiantes. A pesar de esto, estudios como el de (Zohair, 2018) señalan que estos estudios no tienden a revelar información fiable, puesto que existen barreras para predecir el futuro (trabajo y empleabilidad), proponiendo optar por un enfoque adaptativo sobre uno preventivo, entre estas barreras destacan ideas cómo, dónde deberían de aplicarse, como se realizaría, atender a las necesidades de un mercado en constante cambio y la realidad social y regional (Morgan et al. , 2019).

Con esta idea se han propuesto varias aplicaciones sobre cómo es que pueden afectar las herramientas actuales en distintas áreas como la investigación, el campo social, la educación, etc. Uno de los primeros sectores afectados fue el sector tecnológico, específicamente el del desarrollo de software. Y por supuesto, que en un principio también se comentó que los desarrolladores podrían llegar a ser reemplazados, según (J. Bommarito, M. Bommarito, D. Katz, & J. Katz 2023) estas herramientas podrían incentivar la mejora de las habilidades y conocimiento de los profesionales de la industria. La aparición de herramientas como GitHub Copilot en el 2021 fue uno de los principales factores de interés en la industria, un modelo de inteligencia artificial basado en Codex de OpenAI que se enfoca en el autocompletado de código fuente, capaz de mejorar la productividad de un desarrollador realizando las mismas actividades hasta un 50% más rápido en comparación con un desarrollo sin el apoyo de esta herramienta (Peng, Kalliamvakou, Cihon, & Demirer 2023) y por supuesto el uso de herramientas alternativas como Tabnine, que pueden llegar a completar alrededor de la mitad del código fuente escrito por el desarrollador y permitir que este se enfoque en la lógica del problema a solucionar en vez de detalles como la sintaxis del código (Reini, 2022).

La aparición de estas herramientas ha permitido un gran avance en la industria y también ha revelado la flexibilidad de los involucrados al aceptar estos cambios, evidencia de esto nos la brinda la Stack Overflow Survey de 2023 (StackOverflow 2023), que indica que las herramientas de IA tienen una adopción de aproximadamente la mitad de desarrolladores, siendo utilizada no solamente para el desarrollo, sino que también para el aprendizaje, además que muy probablemente se siga incrementando la cantidad de personas que las utilicen en los años posteriores.

Es por el gran impacto que pueden llegar a tener estas herramientas en un futuro y especialmente en el futuro de los estudiantes de las distintas áreas de TI que consideramos

necesario incluir el uso de estas herramientas en la enseñanza, integrando modelos de IA como ChatGPT como fuentes de información sobre tópicos relacionados al área, además de aprovechar las distintas ventajas que proveen con respecto a herramientas de enseñanza tradicionales. Sin embargo, también es necesario considerar las posibles desventajas de su uso, puesto que, como se señala en (King, 2023) los estudiantes pueden utilizar ChatGPT para hacer trampas, siendo necesario adaptar métodos de enseñanza interactivos que permitan evidenciar un correcto aprendizaje y adicionalmente herramientas de software de detección de plagio.

Hemos decidido utilizar ChatGPT debido a su nivel de accesibilidad, la capacidad de respuesta en distintos idiomas y la generación de respuestas relevantes en el ámbito de la programación. Además de su efectividad a la hora de generar código explicado y entendible para problemas con una lógica sencilla (Nascimento, Alencar & Cowan, 2023), pudiendo ser una herramienta de interés para estudiantes que se introducen a conceptos de programación y desarrollo de software. A través de este trabajo de investigación, presentaremos el efecto que puede llegar a tener la aplicación de ChatGPT con un enfoque de aula invertida en estudiantes de Ingeniería de Sistemas que cursan asignaturas de programación básica e identificar su efectividad a través de una encuesta y la comparación de la calificación obtenida durante ese periodo de tiempo por su rendimiento académico en comparación con la calificación obtenida por los estudiantes de años previos a partir del 2019.

2. Trabajos Relacionados

El impacto de las herramientas de IA se hace notar en la industria del desarrollo y la educación, un enfoque general de su impacto nos lo brindan (Adiguzel, Kaya & Cansu, 2023) quienes realizan una revisión de la literatura existente sobre las herramientas de IA capaces de generar respuestas similares a las de un ser humano basadas en modelos de lenguaje natural, concluyendo que el uso de estas herramientas en la educación podría significar la individualización de la misma, la retroalimentación y apoyo. Resalta también el gran avance que se da en la aplicación de estas herramientas en distintos entornos y la necesidad de conservar a los docentes como medio de apoyo y soporte pues existe la posibilidad de sesgo en estas herramientas. Por otro lado, (Karina, et al., 2020) resalta la necesidad de competencias digitales en el ámbito educativo, pues resultan fundamentales para adaptar a los estudiantes a la era digital, así como también el fortalecimiento del proceso educativo para adaptar estrategias pedagógicas responsables.

Por lo mismo se han dado investigaciones sobre el rendimiento de estas herramientas y eficiencia en distintos entornos, por ejemplo tenemos a los autores de (J. Bommarito, M. Bommarito, D. Katz, & J. Katz 2023), quienes realizan una evaluación de rendimiento de ChatGPT como un trabajador de conocimiento a través del examen CPA (Certified Public Accountants, por sus siglas en inglés), un examen estandarizado donde también evalúan a otros modelos de IA de OpenAI como: text-davinci-001, text-curie-001, text-babbage-001 y text-ada-001. Donde concluyen que estos modelos son capaces de resolver distintos tipos de ejercicios en distintas áreas de conocimiento, sin embargo, tienden a tener problemas cuando tratan problemas relacionados al cálculo.

Los autores de (Tian, Lu, On, Tang, Cheung, Klein & Bissyandé, 2023) nos presentan un enfoque alternativo para evaluar a ChatGPT, estos optan por realizar la evaluación de la herramienta con un enfoque en las actividades realizadas durante el desarrollo de software, seleccionan tres actividades del área: la generación de código, reparación del mismo y sumarización (resumen explicativo) del código, donde los autores concluyen que la herramienta puede realizar de estas actividades de forma efectiva, sin embargo identifican problemas durante la explicación de código, ya que no puede explicar problemas poco comunes y no logra abstraer la lógica general de los problemas que se le plantean. Además de indicar que utilizar entradas demasiado largas pueden resultar contraproducentes para obtener información de interés.

En lo que respecta a la educación, Halaweh (2023) nos brinda información relevante sobre la aplicación de ChatGPT en un entorno educativo de forma responsable, destacando entre sus principales ventajas: la calidad de escritura o el desarrollo de habilidades de investigación. Además, también resalta otras consideraciones a tener en cuenta, como la privacidad de los datos y el constante uso de la información que se le brinda para mejorar el modelo, la posible generación de dependencia de esta herramienta, consideraciones éticas y morales. Señalando además que pueden surgir nuevas consideraciones a tomar en cuenta en un futuro por lo tanto recomienda adoptar un enfoque proactivo y responsable.

Una de las principales herramientas de apoyo a los desarrolladores basadas en IA es GitHub Copilot. Y consideramos importante dar una revisión de la literatura al mismo. Se ha medido su efectividad desde distintos enfoques, (Asare, Nagappan & Asokan, 2023) evalúan la efectividad de esta herramienta en comparación con los seres humanos a la hora de codificar e introducir vulnerabilidades no previstas, dando como resultado que las sugerencias de esta herramienta no siempre resultan efectivas e incluso llegan a catalogarse como inseguras, a pesar de esto no resultan tan ineficientes como lo pueden llegar a ser los seres humanos. Otra evaluación realizada sobre GitHub Copilot es la de los autores de (Yetistiren, Ozsoy & Tuzun, 2022), quienes también evalúan la capacidad de generación de código en función de validez correctitud y eficiencia, donde determinan que puede brindar una solución correcta el 28.7% de las veces para el conjunto de datos que seleccionaron, además la eficiencia del código generado puede ser similar a uno desarrollado por el ser humano.

Varios estudios también se han enfocado en la aplicación de estas herramientas de IA en la educación como lo realiza el autor de (Qureshi, 2023) quien dividió a estudiantes en 2 grupos para la asignación de tareas de codificación, revelando que mientras más complejo era el problema, más tiempo les tomaba a los estudiantes que hacían uso de ChatGPT resolverlos, además de resolver dichos problemas con ciertas inconsistencias e imprecisiones que posteriormente serían difíciles o muy complejos de mantener o modificar. Algunas consideraciones a tomar en cuenta sobre el uso de ChatGPT en la educación nos la mencionan (Baidoo-anu & Owusu, 2023), y es que a pesar de señalar las distintas ventajas que pueden traer estos modelos de IA, una de sus principales desventajas es la falta de interacción humana, aquella que siempre se ha dado entre un docente o tutor y sus alumnos, además de señalar algunas consideraciones éticas a tomar en cuenta para utilizar estas herramientas.

Entre las metodologías disponibles, seleccionamos la de aula invertida, pues resulta flexible y sencilla de utilizar, siendo que (Contreras et al., 2017) la utiliza para evaluar y posteriormente mejorar materiales de aprendizaje, otro ejemplo nos lo brinda (Gren, 2020) quien utiliza encuestas y entrevistas para recopilar datos de los estudiantes, comparando los resultados de los estudiantes que no participaron en el enfoque de aula invertida concluyendo que resulta efectiva, pero no motivadora para los estudiantes, pues su actitud frente a los temas se conserva. También destaca la importancia de la retroalimentación y la necesidad de una correcta capacitación de los docentes con respecto al uso de esta metodología.

Otro estudio similar es dado por los autores de (Knutas, Herala, Vanhala, & Ikonen, 2016) donde aplica el enfoque de aula invertida en 2 cursos de programación distintos, uno enfocado en la enseñanza de POO, y el otro en aprendizaje sobre aplicaciones web. Durante el desarrollo de su investigación resaltan aspectos importantes a tomar en cuenta, como: una fuente adecuada de información, integración del material de aprendizaje con el del curso, soporte a los alumnos en clase, la importancia de la retroalimentación personalizada y la carga de trabajo que supone para el instructor diseñar las actividades, para finalmente concluir que este enfoque resulta más efectivo que el tradicional.

3. Metodología

El trabajo de investigación realizado sigue un diseño cuasiexperimental, donde estudiantes de primer año de universidad utilizaron ChatGPT como herramienta de apoyo para resolver prácticas elaboradas en base a tópicos específicos de programación relacionados con la asignatura de Fundamentos de Programación 1.

Las prácticas fueron aplicadas solamente a un grupo determinado de estudiantes, puesto que el dictado del curso se da entre varios docentes. Para los resultados de esta investigación, solamente se considera a los estudiantes que fueron asignados al docente que permitió el desarrollo de esta actividad. Para realizar la comparación de notas, se utilizaron las calificaciones académicas obtenidas por estos estudiantes con las de estudiantes de años previos que también trataron los mismos tópicos y con el mismo docente.

3.1. Población y muestra

Para realizar el estudio se tomó en cuenta a 26 estudiantes de primer año de Ingeniería de Sistemas a los que se les asignaron prácticas de refuerzo y se les indicó que debían de utilizar ChatGPT para solucionar los problemas de programación y además de completar un cuestionario sobre la actividad. Los tópicos seleccionados para la elaboración de las prácticas fueron los seleccionados en base al avance académico de las últimas etapas del curso, estos fueron: arreglos unidimensionales, arreglos bidimensionales e introducción a la Programación Orientada a Objetos (POO). Para el desarrollo de las actividades se optó por adoptar un enfoque de aula invertida (véase la figura 1). Esto implicó presentar a los estudiantes conceptos que no habían sido vistos en clases hasta el momento. Esto brindó a los estudiantes la oportunidad de interactuar con nuevos conceptos por adelantado, lo que podría mejorar su comprensión cuando se aborden durante la clase tradicional.

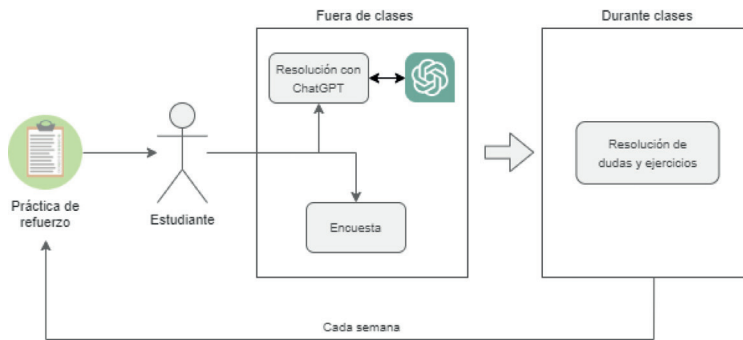


Figura 1 – Enfoque de aula invertida utilizando ChatGPT

También, debido a que las prácticas podrían ser resueltas simplemente utilizando la herramienta, la necesidad de garantizar la obtención de información relevante, la retención y entendimiento de los tópicos tratados, se solicitó a los estudiantes que utilicen los términos y conceptos aprendidos en el curso hasta el momento para explicar la solución obtenida. Asegurando que los estudiantes utilicen no solamente la herramienta para dar resolución a los problemas, sino que utilicen el conocimiento adquirido en los tópicos previos.

3.2. Procedimiento

A lo largo de tres semanas se prepararon prácticas que buscaban introducir a 26 estudiantes al siguiente tema según el cronograma académico planificado para el curso, el objetivo fue buscar un enfoque de aula invertida donde los estudiantes obtendrían información previa a las clases dictadas utilizando ChatGPT, para esto, se les asignaron prácticas durante las tres semanas de actividades académicas, se motivó a elaborar las consultas sobre los temas que desconocían y con el objetivo de resolver los ejercicios planteados, para al final solicitar completar una encuesta para conocer la opinión del estudiante sobre el método, si es que considera que logró la meta de aprendizaje e identificar la cantidad de consultas que el estudiante realizó en promedio por cada ejercicio. Las respuestas obtenidas para cada encuesta fueron reduciendo conforme se desarrollaba la investigación, se recibieron un total de 26 respuestas de los estudiantes en la primera sesión, en la segunda se recibieron 16 respuestas, para finalmente registrar un total de 11 respuestas en la última encuesta.

Para evaluar el nivel de mejora de los estudiantes, se realizó la comparación de las calificaciones obtenidas por los estudiantes en comparación con las obtenidas por estudiantes de años previos (2019-2022), tomando en cuenta los dos tipos de evaluación, además de utilizar la escala de 5 valores que se utilizan en los informes de desempeño en la institución educativa. Las 2 formas de evaluación que se dan en la institución son: evaluación continua y evaluación por examen, en un semestre académico normal se realizan un total de 6 evaluaciones y para esta investigación solo se consideran las 2 últimas de cada año, pues estas son las que representan las notas asignadas a los tópicos seleccionados para la investigación.

- Evaluación continua (EC): Evaluación realizada por cada etapa durante todo el semestre académico, tales como actividades, prácticas de laboratorio, evaluaciones en clase, etc.
- Evaluación parcial (EX): Examen de conocimientos realizado por cada etapa sobre los temas realizados.

3.3. Calificaciones

Como se mencionó anteriormente en esta sección se analizarán los resultados de las calificaciones de los 4 años previos en el curso de Fundamentos de Programación 1, considerando solamente las calificaciones relacionadas a los tópicos seleccionados para este trabajo, tomando calificaciones desde la enseñanza presencial en 2019 hasta el presente año. La institución toma un rango de valores para la calificación desde 0 (cero) para la nota mínima y 20 como máxima. Siendo las notas a entregar las antes mencionadas EC y EX.

En la tabla 1, se muestran las calificaciones de los estudiantes durante su tercera etapa entre los años 2019 y 2023, consideramos necesario utilizar un valor porcentual debido al cambio en la cantidad de estudiantes a lo largo de los años. Se representan ambos tipos de evaluación y el total de estudiantes por año. Los datos obtenidos para el año 2023 son los que representan los resultados de la aplicación de ChatGPT y un enfoque de aula invertida en los estudiantes, En la siguiente sección se explicarán los resultados obtenidos durante esta investigación, considerando principalmente los resultados de la evaluación EX, puesto que son relevantes para evaluar el correcto aprendizaje por parte de los estudiantes, sin embargo también se comentará sobre la calificación obtenida en la EC y cómo varía con respecto a años previos.

	2019		2020		2021		2022		2023		
Rango nota	EC	EX	EC	EX	EC	EX	EC	EX	EC	EX	Calificativo
AB	14.75%	15.57%	5.95%	13.10%	13.64%	13.64%	13.01%	16.26%	6.98%	6.98%	Abandono
1-7	2.46%	37.70%	9.52%	17.86%	6.36%	30.91%	7.32%	14.63%	18.60%	23.26%	Insatisfactorio
8-10	6.56%	18.85%	4.76%	16.67%	3.64%	17.27%	7.32%	15.45%	6.98%	23.26%	En proceso
11-14	27.87%	18.03%	35.71%	29.76%	26.36%	22.73%	25.20%	35.77%	30.23%	25.58%	Satisfactorio
15-20	48.36%	9.84%	44.05%	22.62%	50.00%	15.45%	47.15%	17.89%	37.21%	20.93%	Sobresaliente

Tabla 1 – Distribución porcentual de estudiantes en el 2023

Para brindar un contexto aproximado sobre la media de calificaciones obtenidas en la institución, se considera presentar las notas mínimas, máximas y el promedio de las notas de todos los estudiantes entre los años 2019 y 2023, observe la figura 2 y la figura 3 para observar las notas de EC y EX respectivamente entre los años mencionados.

El mayor promedio de calificación observado en este estudio para la EC se obtuvo en el 2019 siendo de 14.49, que posteriormente se fue reduciendo mínimamente hasta la fecha. Mientras que la menor media de EX se obtiene también en este mismo año e incrementa durante los siguientes años. Este cambio a través del tiempo puede deberse

a varios factores, desde el nivel de rigurosidad de la calificación, el docente asignado al curso, la transición a la enseñanza remota, etc. Para esta investigación reconocemos la existencia de estos factores, sin embargo, nos enfocamos en las calificaciones de los estudiantes disponibles y no en los factores que pudieron modificarlos.

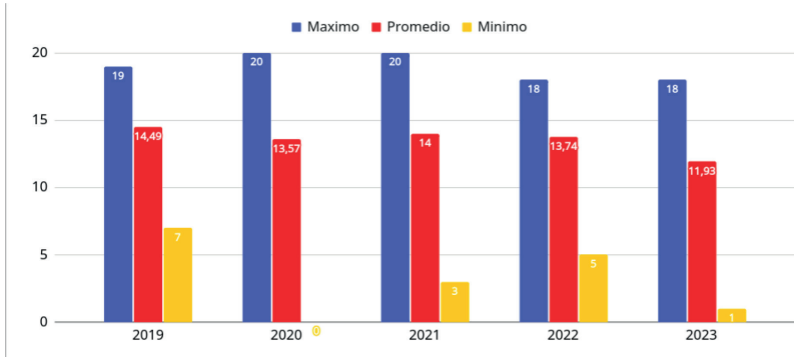


Figura 2 – Calificaciones de evaluación continua 2019 - 2023

En el caso de la media de evaluación continua, resaltamos una tendencia en el cambio de notas desde el año 2019, puesto que se da un decremento medianamente notorio en las calificaciones de los estudiantes (normalmente superando los 13.5 en años anteriores y obteniendo 11.93 en el 2023), esto puede deberse a la transición existente entre la enseñanza presencial y la remota, puesto que este cambio en la media también es notorio en la evaluación realizada mediante exámenes. Otro factor de importancia puede ser el cambio en el porcentaje de la nota final (calificación consolidada de las 6 notas obtenidas a lo largo del semestre), ya que para asegurar un correcto aprendizaje y desempeño, se realizó un cambio en la proporción de las notas, pudiendo suponer un mayor esfuerzo por obtener mejores notas en la EC que en la EX por parte de los estudiantes.

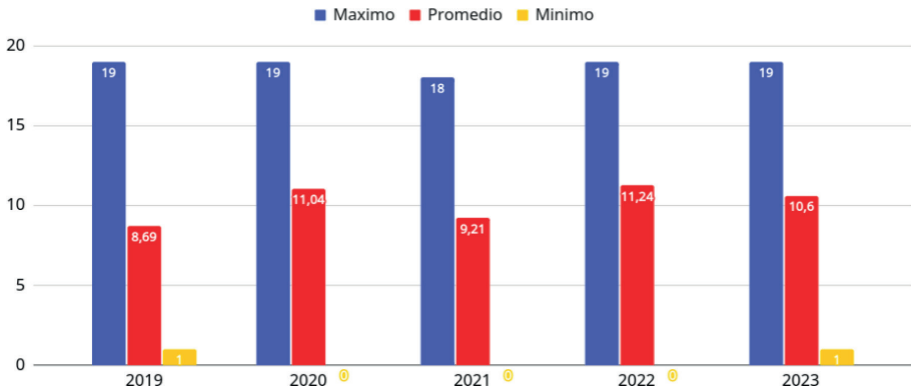


Figura 3 – Calificaciones obtenidas en los exámenes 2019 - 2023

Como se indicó anteriormente, en el caso de la media de la calificación mediante la evaluación por examen se distingue un cambio notorio, pues esta se reduce (siendo 14.49 en el 2019 y 11.04 en el 2020), y se conserva el cambio durante los años de pandemia tanto en el caso de la evaluación continua como por examen.

3.4. Resultados de las encuestas

Junto a cada práctica desarrollada, a lo largo de cada semana se han brindado encuestas a los estudiantes para obtener su opinión sobre su experiencia realizando esta actividad, a través de las siguientes preguntas:

- ¿Qué opinión tiene sobre el método de enseñanza aplicado en la actividad de avance? Calificado desde 1 (excelente) a 5 (deficiente)
- ¿Cuántas consultas por pregunta realizó a ChatGPT en promedio? Calificado con: 1, 2 o 3, 4 o 5 y más de 5.
- ¿Tiene alguna observación? Opinión no obligatoria del estudiante.

En la figura 4, se pueden observar los resultados de la encuesta, el interés de los estudiantes y su disposición para optar por un nuevo método de enseñanza se hace evidente, ya que ninguno consideró malo o pésimo el método de enseñanza, sin embargo, son varios los que consideran la forma de enseñanza algo regular, esto puede deberse a factores como la eficiencia de la herramienta, la sobreactividad que supone la adición de estos ejercicios a los estudiantes o simplemente la facilidad que brinda tener un docente en una clase presencial con quien atender las dudas que pudiera tener, explicaremos a mayor profundidad este punto al final de esta sección en base a las observaciones obtenidas por las mismas encuestas.

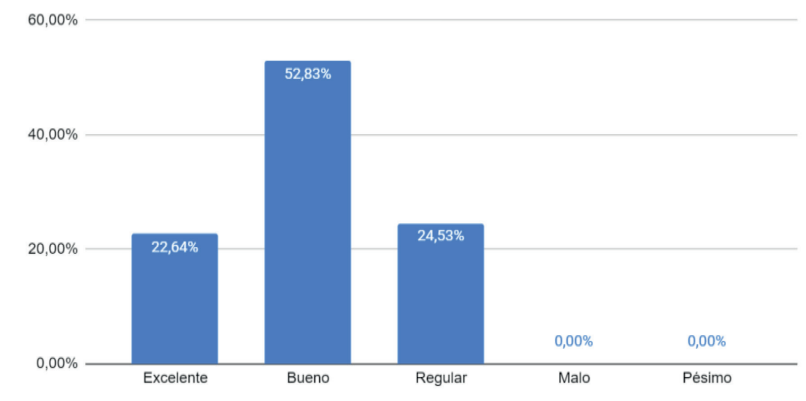


Figura 4 – Opinión de los estudiantes sobre la metodología aplicada.

El segundo punto de interés fue el saber si es que la herramienta resultó eficiente al punto de lograr enseñar sobre el tópico al estudiante, este se midió a través de 3 alternativas: si logro entender, más o menos, o no entendió. para lo cual se obtuvieron los resultados mostrados en la figura 5.

Más de la mitad de los estudiantes consideran que lograron comprender el tópico presentado, mientras que cierta cantidad no considera que lo haya entendido al completo y a su vez ninguno considera que no haya entendido nada, esto puede deberse a que para cada práctica se ha brindado sugerencias sobre qué debería de solicitar para obtener una respuesta entendible, y es que como mínimo se espera que se entienda un concepto básico del tema a tratar.

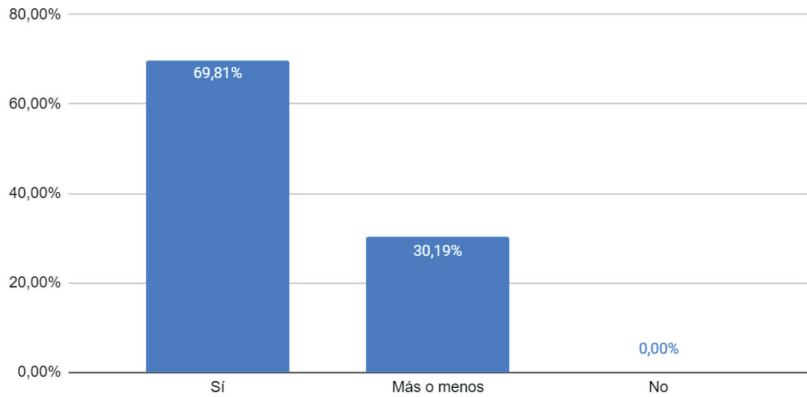


Figura 5 – Calificación sobre el entendimiento del tema.

Finalmente, la tercera pregunta estaba dirigida a identificar la cantidad de consultas que los estudiantes realizan por cada ejercicio, pues puede resultar relevante conocer si es que los estudiantes son capaces de elaborar instrucciones correctamente estructuradas o con la especificidad correcta para obtener respuestas satisfactorias a las consultas que realizaron, los resultados se pueden observar en la figura 6.

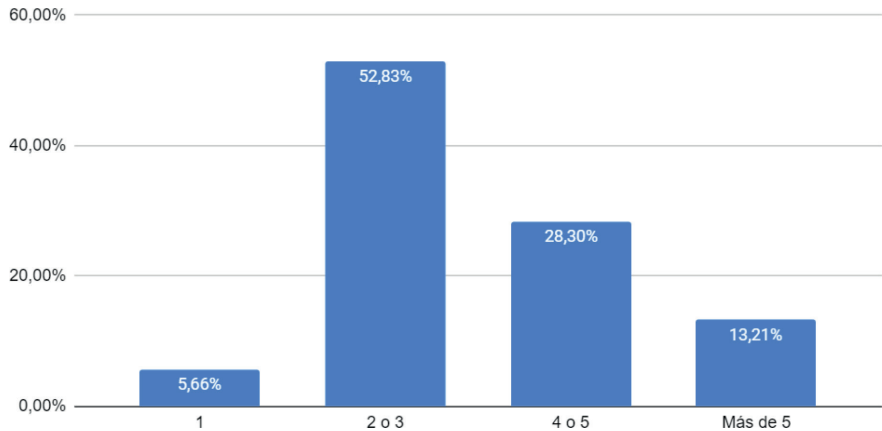


Figura 6 – Cantidad de consultas realizadas por cada ejercicio.

Esta estadística revela que la gran mayoría de estudiantes resuelve sus dudas con 2 o 3 consultas, seguido por aquellos que las resuelven en 4 o 5, para finalmente ser una minoría aquellos que las resuelven con 1 o más de 5 consultas. Estos resultados pueden deberse a la falta de experiencia con el uso de este tipo de herramientas, pues estas no funcionan de la misma manera que los buscadores en internet, y requieren un mayor contexto en la formulación de las entradas para obtener una salida útil para el usuario (White, Fu, Hays, Sandborn, Olea, Gilbert, Elnashar, Spencer-Smith & Schmidt, 2023).

3.5. Sobre los datos

El grado de mejora académica se dará en base a los datos de las calificaciones anteriores de estudiantes de años previos que cursaron la misma etapa de evaluación, considerando los 4 años previos a este estudio, además de esto, cabe recalcar que las notas obtenidas desde el año 2020 pudieron ser afectadas debido a que la institución adoptó un método de enseñanza remota, sin embargo, también se consideraron las calificaciones obtenidas durante el año 2019, año en el que se realizaron las actividades académicas con normalidad.

Además de la situación de enseñanza, también se encontraron desafíos en la obtención de la información previa al año 2019, puesto que el método de calificación planteada por la institución en años previos no se adecuaba a los datos obtenidos a partir del año mencionado hasta la actualidad, dificultando la obtención de notas de los tópicos seleccionados para este trabajo de investigación, razón por la cual no se contemplaron estas calificaciones en este estudio.

4. Resultados

En esta sección brindaremos los resultados de la investigación y la metodología aplicada a los estudiantes de Ingeniería de Sistemas, además de comparar su rendimiento académico con respecto a años previos, además de resumir los resultados de las encuestas.

La media de los resultados de los estudiantes a través de los años indica que se tenía cierta tendencia a mantenerse hasta la fecha de la presente investigación, el cambio más relevante se da durante los años trabajados de manera remota, aunque puede resultar relevante, en cierta medida se mantiene equilibrada (puesto que se dio un incremento en la media de la nota continua de los estudiantes y un decremento en la media de la calificación por examen). No se ha identificado una mejora considerable entre la media de las calificaciones obtenidas en el año 2023 en comparación con el resto de años, sin embargo, no descartamos la presencia de datos de interés.

La distribución de los estudiantes en base a su calificación por EX (véase la figura 7) a lo largo de los años ha sido variada, siendo generalmente las calificaciones “satisfactorias” e “insatisfactorias” las que ocupaban los picos más altos (mayores al 30%), sin embargo, en este último año cada categoría de calificación, a excepción de los estudiantes en situación de abandono, representan entre el 20.93% y 25.58% del total de estudiantes. En la tabla 2 representamos la diferencia en la proporción de estudiantes por cada

calificación, representamos el cambio de este último año en comparación con la media de los años previos.

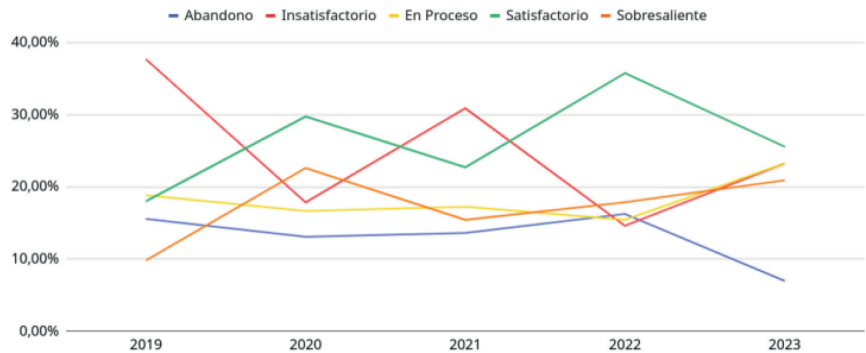


Figura 7 – Proporción de estudiantes según calificaciones de EX entre el 2019 y 2023

Calificación	Promedio (2019-2022)	2023	Diferencia
Abandono	14.64%	6.97%	-7.66%
Insatisfactorio	25.28%	23.26%	-2.02%
En proceso	17.06%	23.26%	+6.20%
Satisfactorio	26.57%	25.58%	-0.99%
Sobresaliente	16.45%	20.93%	+4.48%

Tabla 2 – Diferencia porcentual de calificaciones EX entre la media de 2019-2022 y 2023

Durante la investigación se aplicó una encuesta a los estudiantes sobre el uso de esta herramienta como fuente de información, la metodología y su experiencia durante su aplicación, a continuación, se enumeran los problemas más frecuentes que los estudiantes identificaron:

- Las consultas realizadas a ChatGPT pueden resultar complejas puesto que no siempre obtienen resultados que ayuden a entender los conceptos propuestos, sino que pueden resultar en respuestas complejas, incluir conceptos avanzados o simplemente brindar información que no resulta relevante.
- Las respuestas no resultan claras para el estudiante, ya que indicaron que fue necesario realizar varias consultas o elaborar una bastante específica para obtener información relevante.
- La metodología aplicada no resultó eficiente debido al tiempo de resolución que se les brindó para resolver la práctica y el desconocimiento del tema sobre el cual trata dicha práctica.

Las opinión de los estudiantes con respecto al método de enseñanza revela una efectividad media, puesto que si bien la mayoría indican que es “Bueno” puede que los problemas mencionados anteriormente sean los causantes de no obtener una calificación

“Excelente”, también, el total de estudiantes que respondieron la encuesta considera que ha logrado aprender al menos algo sobre el tema, finalmente la mayor frecuencia en la cantidad de consultas que realizaron a ChatGPT por cada ejercicio está entre 2 o 3.

5. Discusión

Las calificaciones obtenidas no revelaron cambios relevantes en la tendencia de las puntuaciones estudiantiles a lo largo de los años. Sin embargo, se identificó la posibilidad de mejorar el desempeño de aquellos estudiantes que muestran dificultades en su rendimiento académico. El uso de ChatGPT como fuente de conocimiento acompañado de un enfoque de aula invertida ha demostrado tener un impacto considerable entre los alumnos menos destacados, sin causar una perturbación sustancial en el rendimiento del resto de los estudiantes. Además, es esencial considerar una ampliación del alcance de este estudio. Esto implicaría analizar una extensa serie de datos históricos recopilados en un contexto similar para identificar tendencias más arraigadas.

Es importante mencionar que el uso de calificaciones obtenidas durante la pandemia puede tener un impacto considerable en los resultados de la investigación y, por tanto, es necesario abordar este aspecto con mayor profundidad en futuros análisis. Además, para trabajos futuros se recomienda realizar un estudio más duradero y dar un mayor seguimiento a los resultados de cada estudiante, puesto que debido a la planificación de la institución solamente se evaluó a los estudiantes durante parte del ciclo académico y se limitó el alcance de la evaluación a las calificaciones académicas.

Por otro lado, los resultados de las encuestas indican que los estudiantes consideran útil el método de enseñanza, e indican problemas como: la cantidad de actividades a realizar, la especificidad requerida para obtener resultados relevantes e incluso la percepción de dificultad en el aprendizaje o la claridad de las respuestas obtenidas de ChatGPT. Cabe recalcar que, si bien no se dio en esta encuesta, también es necesario considerar que las recomendaciones de los estudiantes pueden dejar de lado tópicos importantes que se tienen que tocar en el curso (Knutas, Herala, Vanhala & Ikonen, 2016).

Fuera del foco de la investigación, también se identifica cierta variación en las notas entre los años 2019 y 2020 ya que la evaluación a los estudiantes durante los años afectados por la pandemia y la modalidad de enseñanza remota afecta mínimamente su rendimiento, siendo reducida la media cuando se trata de EC e incrementando cuando se trata de EX. Con respecto a la evaluación continua, consideramos que se denota un crecimiento debido a la facilidad de investigación y trabajo desde el hogar, sin embargo, esta comodidad también afecta al rendimiento y la adquisición de conocimientos siendo evidencia de esto la media de notas del examen durante estos años.

6. Conclusiones

En este estudio se motivó a estudiantes de Ingeniería de Sistemas a que hicieran uso de ChatGPT para la resolución de ejercicios de programación con una base previa sobre conceptos básicos de programación tales como: variables, librería Math de Java, bucles y condicionales, etc. Las actividades se realizaron con un enfoque de aula invertida, brindándoles prácticas de programación en base a los tópicos académicos asignados a

su etapa, además de una encuesta para obtener información sobre su experiencia con el desarrollo de la actividad y su opinión sobre el método utilizado. Para finalmente analizar los resultados obtenidos en base a su calificación académica en comparación con los obtenidos por estudiantes de años previos.

Las calificaciones obtenidas para el periodo seleccionado (2023) no indican una mejora relevante en todos los estudiantes, sin embargo, si se sale a relucir una tendencia en las notas obtenidas de los años previos con respecto a la evaluación de conocimiento (EX), se puede identificar una mejora en la distribución de los estudiantes con respecto a sus calificaciones. Viéndose reducida considerablemente la cantidad de estudiantes que tienden a obtener notas bajas con respecto a la media de los años anteriores.

La opinión de los estudiantes sobre la actividad fue positiva en su mayoría, sin embargo, considerando las observaciones obtenidas, se recomienda adoptar algunas medidas, como una capacitación previa sobre el enfoque de aula invertida, la reducción de la cantidad de actividades sin comprometer el objetivo de aprendizaje e incluso también considerar una capacitación previa a los estudiantes sobre la estructuración de consultas a los grandes modelos de lenguaje como ChatGPT. Además, los estudiantes consideran que han logrado los objetivos de aprendizaje de las prácticas ya que, en las respuestas registradas, la mayoría indica que comprendió el tópico asignado.

En relación al uso de ChatGPT como fuente de información para conceptos de programación, la consideramos una herramienta idónea por su nivel de accesibilidad y su capacidad de respuesta, pues es capaz entender cuestiones de distintas complejidades y ofrecer la información pertinente respecto al tema. No obstante, es importante señalar que la estructura de las consultas requiere un conocimiento previo de la herramienta. Como ya mencionamos, esta limitación puede ser superada mediante una capacitación preliminar de los usuarios o a través de la exploración de otros modelos de lenguaje o implementaciones específicas centradas en la resolución de consultas relacionadas exclusivamente con la programación.

7. Trabajos futuros

Para finalizar, en base a estos resultados obtenidos, como trabajo futuro se plantea incluir una capacitación sobre la adopción de técnicas de “Prompt Engineering” para mejorar la búsqueda de resultados en modelos de lenguaje considerando aspectos como las limitaciones, contextos, foco de la consulta, etc. Esto debido a que la opinión de los estudiantes revela que consideran útil esta forma de enseñanza, pero requieren mayor experiencia con el uso de esta herramienta ya que muchas veces las consultas realizadas por los estudiantes no resultan útiles debido a entradas confusas, respuestas demasiado avanzadas o el hecho de realizar varias consultas con especificaciones cada vez más exactas para obtener una respuesta útil. Además, contemplamos estudiar la aplicación de otros modelos de lenguaje liberados hasta la fecha, con distintos enfoques, permitiendo así evaluar en base a qué combinación de estos se pueden dar una solución óptima a las necesidades de los estudiantes en lo que respecta a su aprendizaje enfocado en conceptos de programación.

Referencias

- Adiguzel, T., Kaya, H., & Cansu, F. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15, ep429. <https://doi.org/10.30935/cedtech/13152>
- Asare, O., Nagappan, M., & Asokan, N. (2023). Is GitHub's Copilot as bad as humans at introducing vulnerabilities in code?. *Empirical Software Engineering*, 28. <https://doi.org/10.48550/arXiv.2204.04741>
- Baidoo-anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, 7(1), 52-62. <https://dx.doi.org/10.2139/ssrn.4337484>
- Basit Q. (2023). Exploring the Use of ChatGPT as a Tool for Learning and Assessment in Undergraduate Computer Science Curriculum: Opportunities and Challenges. <https://doi.org/10.48550/arXiv.2304.11214>
- Bommarito, J., Bommarito, M., Katz, D. M., & Katz, J. (2023). GPT as Knowledge Worker: A Zero-Shot Evaluation of (AI)CPA Capabilities. Disponible en: <https://doi.org/10.48550/arXiv.2301.04408>
- Contreras, J. Á., Masa, J. A., Andrade, M. G. M., & Espada, R. M.. (2017). Uso del modelo de aprendizaje inverso para mejorar materiales educativos universitarios. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*. 17-32. 10.17013/risti.23.17-32
- Daysi, F., & Bedon, A. (2020). Competencias Digitales y Educación: Aproximaciones a los Consumos de Estudiantes en Formación Docente. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, E31, 318-327.
- Gren, L. (2020). A Flipped Classroom Approach to Teaching Empirical Software Engineering. *IEEE Transactions on Education*, 63(3), 155-163. <https://doi.org/10.1109/TE.2019.2960264>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15. <https://doi.org/10.30935/cedtech/13036>
- King, M. (2023). A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. *Cellular and Molecular Bioengineering*, 16, 1-2. <https://doi.org/10.1007/s12195-022-00754-8>
- Knutas, A., Herala, A., Vanhala, E., & Ikonen, J. (2016). The Flipped Classroom Method: Lessons Learned from Flipping Two Programming Courses. In *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016* (pp. 423-430). Association for Computing Machinery. <https://doi.org/10.1145/2983468.2983524>

- Mahmoud Abu Zohair, L. (2018). The Future of Software Engineering by 2050s: Will AI Replace Software Engineers?. *International Journal of Information Technology and Language Studies*, 2, 1-13. Disponible en: <https://journals.sfu.ca/ijitls/index.php/ijitls/article/view/23>.
- Morgan, R., Frank, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences*, 116(14), 6531-6539. <https://doi.org/10.1073/pnas.1900949116>
- Nascimento, N., Alencar, P., & Cowan, D. (2023). Comparing Software Developers with ChatGPT: An Empirical Investigation. <https://doi.org/10.48550/arXiv.2305.11837>
- Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The Impact of AI on Developer Productivity: Evidence from GitHub Copilot. <https://doi.org/10.48550/arXiv.2302.06590>
- Reini, N. (2022). The Impact of AI powered code completion in the software engineering field. (Tesis de bachillerato en Administración de Empresas, Business Information Technology, Haaga-Helia University of Applied Sciences). Theseus. Disponible en: <https://www.theseus.fi/handle/10024/781700>.
- StackOverflow. (2023). 2023 Developer Survey. Disponible en: <https://survey.stackoverflow.co/2023/>
- Tian, H., Lu, W., Li, T. O., Tang, X., Cheung, S. C., Klein, J., & Bissyandé, T. F. (2023). Is ChatGPT the Ultimate Programming Assistant – How far is it?. <https://doi.org/10.48550/arXiv.2304.11938>
- Vorobeva, D., Fassi, Y., Costa Pinto, D., Hildebrand, D., Herter, M., & Mattila, A. (2022). Thinking Skills Don't Protect Service Workers from Replacement by Artificial Intelligence. *Journal of Service Research*, 25, 109467052211043. <https://doi.org/10.1177/10946705221104312>
- Yetistiren, B., Ozsoy, I., & Tuzun, E. (2022). Assessing the Quality of GitHub Copilot's Code Generation. In *Proceedings of the 18th International Conference on Predictive Models and Data Analytics in Software Engineering* (pp. 62–71). Association for Computing Machinery. <https://doi.org/10.1145/3558489.3559072>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. <https://doi.org/10.48550/arXiv.2302.11382>

As literacias mediática e digital e a participação cívica como tecnologias para dinamizar os territórios do interior de Portugal

Paulo Bruno Alves¹, Lúdia Oliveira²

paulobruno@esev.ipv.pt; lidia@ua.pt

¹ Escola Superior de Educação de Viseu, Rua Dr. Maximiano Aragão, 41, 3500-155 Viseu, Portugal

² Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

DOI: 10.17013/risti.52.113-125

Resumo: As ações de promoção nos territórios do interior de Portugal – ao nível das literacias mediática e digital e da participação cívica dos seus habitantes – são oportunidades para captar diversas capacidades e dinamizar esses territórios, tornando-os desejáveis, em termos de captação de investimentos, de estruturas tecnológicas, de turismo e até de novos residentes. Esses elementos dinamizadores constituem – ou deverão constituir – fontes motivacionais para os cidadãos, de forma a desenvolver e a expandir o seu sentimento de pertença aos territórios e à valorização dos mesmos. O objetivo deste texto passa por apresentar duas dinâmicas (literacias mediática e digital e participação cívica), no quadro interpretativo dos territórios do interior. Elas têm como cenário as 21 Comunidades Intermunicipais, com muitos municípios e freguesias de baixa densidade. São ferramentas importantes, que poderão funcionar como *inputs* dos próprios cidadãos, de forma a estarem mais capacitados tecnologicamente para almejar melhores condições para eles e para esses territórios.

Palavras-chave: Territórios; Interior; Literacia mediática e digital; Participação cívica

Media and digital literacy and civic participation as technologies to boost the Portugal interior territories

Abstract: Promotion actions in Portugal interior territories – in terms of media and digital literacy and the civic participation of their inhabitants – are opportunities to capture different capabilities and boost these territories, making them desirable, in terms of attracting investment, technological structures, tourism and even new residents. These driving elements constitute – or should constitute – motivational sources for citizens, in order to develop and expand their sense of belonging to the territories and to their appreciation. The objective of this text is to present two dynamics (media and digital literacy and civic participation), within the interpretative framework of interior territories. They are set against the backdrop of 21 Intermunicipal Communities, with many low-density municipalities

and parishes. These are important tools, which can act as input from the citizens themselves, so that they are more technologically capable to aim for better conditions for themselves and these territories.

Keywords: Territories; Interior; Media and digital literacy; Civic participation

1. Enquadramento geral dos territórios do interior de Portugal

Os “territórios de baixa densidade”, designação que foi recentemente transformada em “territórios do interior”, pelo XXIII Governo nacional, ocupam a maioria da sua superfície total (92.212 Km²)¹, de acordo com a Cooperação e Desenvolvimento Económico (OCDE). Este organismo internacional identificou, sobre Portugal, três áreas definidoras: as áreas metropolitanas de Lisboa e do Porto, que são identificadas como sendo uma zona “predominantemente urbanizada”; parte da faixa litoral do país, desde Viana do Castelo (norte) até Lisboa (centro), como uma “região intermédia”; a denominação “Predominantemente Rural”, que concentra quase a globalidade do território nacional.

É nesse último e extenso território que um conjunto de variáveis marca o quotidiano das regiões e de quem vive nelas. A realidade é clara: a população escasseia, sobretudo a mais jovem, que abandona os territórios em busca de melhores condições de vida, num quadro nacional de grande emigração. De acordo com os dados do Relatório Estatístico de 2022, do Observatório de Emigração e Rede Migra, houve um aumento de portugueses a emigrar, em 2021, na ordem dos 60 mil, cerca de 15 mil a mais do que em 2020 (Pires, R. P., et. al., 2022). Rui Pena Pires, coordenador do Observatório de Emigração e Rede Migra², admite que muitos desses portugueses são jovens licenciados que emigram. As razões são várias, mas sustentam-se, sobretudo, nos baixos salários, nos altos preços da habitação e na elevada taxa de desemprego jovem, que se fixou nos 17.2%, no segundo semestre de 2023³.

No território nacional e, concretamente, nos territórios do interior, permanece uma população com baixa qualificação escolar, iliterata tecnologicamente e cada vez mais envelhecida. Em pouco mais de seis décadas, o índice de envelhecimento passou de 27,3%, em 1960, para 181,2%, em 2021⁴. Nesses territórios faltam infraestruturas e serviços. As empresas que aí existem, ainda que façam parte de um importante tecido económico, são, em termos de dimensão, maioritariamente micro, pequenas e médias. É

¹ Cf. <https://ue.missaoportugal.mne.gov.pt/pt/portugal/sobre-portugal/dados-gerais> (consultado em 1 de setembro de 2023).

² As declarações de Rui Pena Pires foram prestadas à Rádio Renascença, em março de 2022. Cf. <https://rr.sapo.pt/noticia/pais/2023/03/22/baixos-salarios-e-precos-altos-da-habitacao-favorecem-saida-de-jovens-licenciados-do-pais/324831/> (consultado em 1 de setembro de 2023).

³ Cf. Dados do BPstat, o portal de estatísticas do Banco de Portugal: <https://bpstat.bportugal.pt/serie/5739365>. (consultado em 1 de setembro de 2023).

⁴ Cf. <https://www.pordata.pt/portugal/indice+de+envelhecimento+e+outros+indicadores+de+envelhecimento+segundo+os+censos-525> (consultado em 1 de setembro de 2023).

uma realidade que cruza o país. Com base nos dados da Pordata, existiam em Portugal, em 2021, mais de um milhão de empresas registadas (1.357.657), sendo 1.305.100 micro, 45.036 pequenas e 7.521 médias⁵. A agricultura (muita de subsistência) é ainda a principal atividade que domina a paisagem, mas onde permanece um nível baixo de produtividade e a quase inexistência de uma real competitividade com os mercados internacionais (Almeida, 2021).

O “círculo vicioso” que essa realidade transparece não ajuda a promover políticas de atração e de fixação das pessoas nos territórios. A falta de infraestruturas tecnológicas e outras condiciona o desenvolvimento dos territórios do interior (figura n.º 1), o que promove a saída da população local (Mota, 2019).

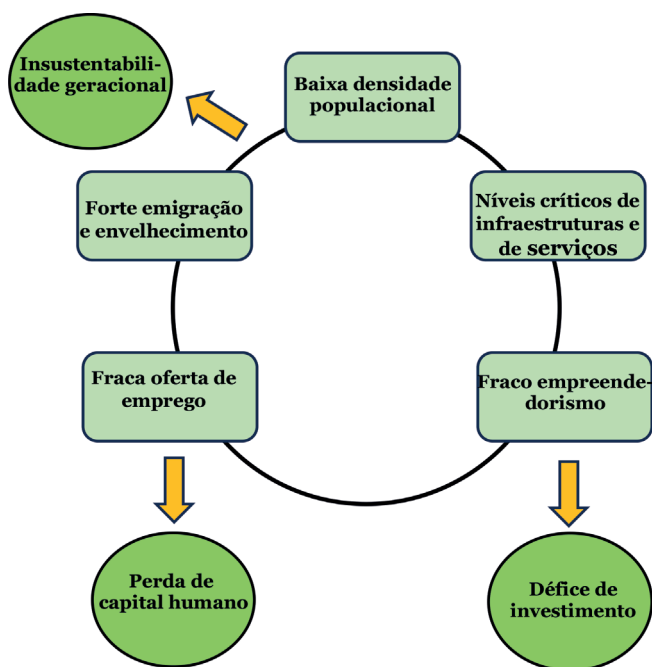


Figura n.º 1 – “Ciclo Vicioso” dos territórios do interior. Fonte: Criação própria a partir de <https://revista.aps.pt/pt/territorios-do-interior-coesao-territorial-e-modelos-de-governanca-a-proposito-do-programa-nacional-para-a-coesao-territorial/>

Em paralelo, é notório o afastamento desses territórios dos grandes centros de decisão. Urge renovar políticas que desenvolvam os territórios do interior, no quadro de um novo modelo de “governança”. O maior exemplo é o *Programa Nacional de Ordenamento do Território* (PNPOT). Como salienta Alcides Monteiro (2019), este é um instrumento que tem como objetivo principal inverter o ciclo vicioso que cruza os territórios do

⁵ Cf. <https://www.pordata.pt/portugal/pequenas+e+medias+empresas+total+e+por+dimensao-2927>

interior, com a inclusão de medidas que sejam capazes de alavancar esses territórios, como sejam: Conhecer as especificidades dos vários territórios; Apostar na inversão de tendências demográficas; Fixar populações e formar capital humano; Produzir e captar conhecimento; Potenciar a inovação e o desenvolvimento territorial; Atrair investimentos sustentáveis; Valorizar os recursos endógenos; Territorizar e ordenar as políticas; Articular os poderes de decisão nacionais e locais; Abranger a participação ativa de agentes externos.

Alcides Monteiro (2019) apresenta, no seu estudo *“Territórios do Interior, Coesão Territorial e Modelos de Governança – a propósito do Programa Nacional para a Coesão Territorial”*, diversas considerações para se olhar para o interior do país com mais atenção. Uma delas é a constatação sobre o caminho que o Governo deverá seguir na afirmação do seu plano para a “governança”, que envolve a revitalização dos territórios do interior. Isto é, segundo o autor, o programa governamental não terá um grande impacto no desenvolvimento dos territórios e da coesão territorial e, naturalmente, nos de baixa densidade, por ficar abaixo das expectativas e *quicá* das necessidades. Daí que Alcides Monteiro (2019) admite que o Governo irá seguir um modelo de ação multinível, delegando competências às entidades locais.

De qualquer forma, nos últimos anos, o discurso político europeu e nacional tem enfatizado a necessidade de se olhar, com outra atenção, para os territórios de baixa densidade. Em Portugal, no início de 2023, o Ministério da Coesão Territorial (criado em 2019) apontou a necessidade de se descentralizar o país, tornando-o mais autónomo e em que o interior assumisse uma importância cada vez maior. A *Estratégia de Valorização do Interior*, uma das medidas apresentadas pelo Governo, demonstrou a necessidade de serem aplicados vários recursos, assentes no conhecimento e na tecnologia, para aumentar a competitividade através de ações como a criação de emprego mais qualificado, a diversificação da base económica, a atração e fixação de população nos territórios do interior⁶.

A estratégia do Governo para a valorização dos territórios do interior português sustenta-se no desenvolvimento de soluções para inverter o “ciclo vicioso” (figura n.º 3). Para isso, o Governo criou uma “nova área de governação” designada “Coesão Territorial” e desenvolveu quatro eixos de ação: 1) Valorizar os Recursos Endógenos e a Capacidade Empresarial do Interior; 2) Promover a Cooperação Transfronteiriça para a Internacionalização de Bens e Serviços; 3) Captar Investimento e Fixar Pessoas no Interior; 4) Tornar os Territórios do Interior Mais Competitivos⁷.

Com as medidas desenvolvidas a partir desses quatro eixos, o Governo pretende responder aos desafios demográficos, reduzir desigualdades e construir uma sociedade digital mais eficiente e inovadora. Para isso, foram delineados dois pilares de intervenção: 1) “Transversalidade do plano”, capaz de identificar várias medidas para desenvolver

⁶ Cf. XXII Governo Constitucional de Portugal (2019-2022). *Estratégia para a Coesão Territorial*, p. 11. <https://www.portugal.gov.pt/gc22/programa-de-valorizacao-do-interior/estrategia-para-a-coesao-territorial-pdf.aspx> (consultado em 2 de setembro de 2023).

⁷ Cf. <https://www.portugal.gov.pt/gc22/programa-de-valorizacao-do-interior/estrategia-para-a-coesao-territorial-pdf.aspx>, p. 11. (consultado em 2 de setembro de 2023).

os territórios, num princípio de cooperação interministerial; 2) “Articulação com os agentes presentes no território”, ou seja, com os decisores e as “forças vivas” locais e regionais, como Autarquias, Comunidades Intermunicipais, instituições de Ensino Superior, empresas e associações⁸.

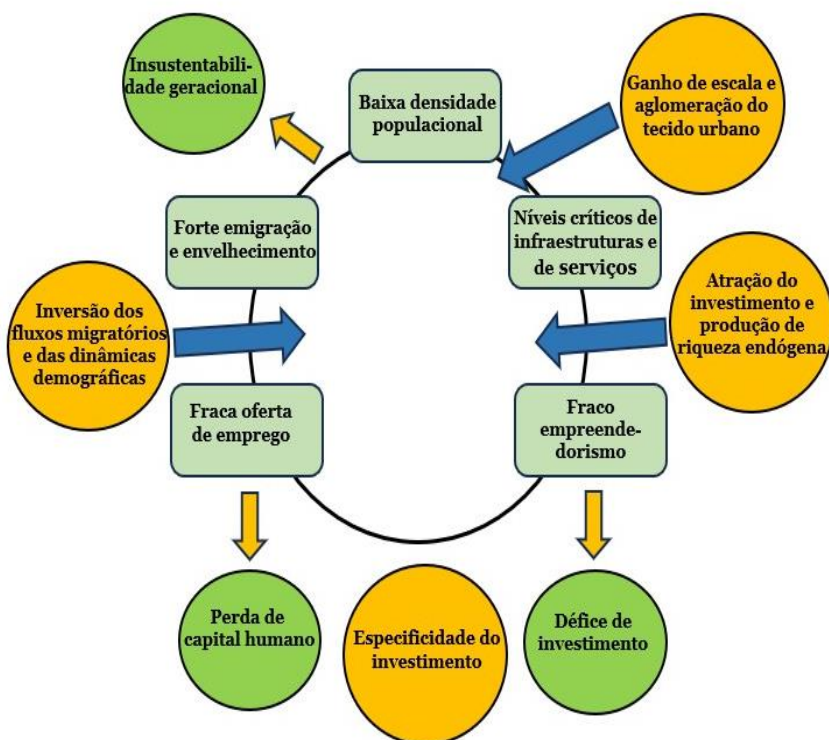


Figura n.º 2 – “Modelo para revitalizar” os territórios do interior. Fonte: criação própria a partir de <https://revista.aps.pt/pt/territorios-do-interior-coesao-territorial-e-modelos-de-governanca-a-proposito-do-programa-nacional-para-a-coesao-territorial/>

O modelo para “revitalizar os territórios”, em oposição ao “círculo vicioso” é sugestivo. De forma a inverter a forte emigração, o envelhecimento e a fraca oferta de emprego, o Governo aposta na “inversão dos fluxos migratórios e das dinâmicas demográficas”. Face à baixa densidade populacional do país onde faltam infraestruturas, tecnologias, serviços e onde o empreendedorismo é fraco, sugere-se a “atração do investimento e produção de riqueza endógena”. Por seu turno, perante a fraca oferta de emprego e o fraco empreendedorismo, de que resultam uma “perda de capital humano” e um “défice de investimento” é sugerido uma “especificidade do investimento”. Os problemas dos territórios do interior estão identificados e as soluções também. O investimento já está previsto e em implementação. Falta a inversão real das tendências demográficas.

⁸ Cf. Idem, *ibidem*.

2. Comunidades Intermunicipais, muitos territórios do interior

De acordo com a lei n.º 75/2013, de 12 de setembro, sobre o regime jurídico da transferência de competências do Estado para as autarquias locais e para as entidades intermunicipais, existem em Portugal 21 Comunidades Intermunicipais (CIM), mais as referidas Áreas Metropolitanas de Lisboa e do Porto, que agrupam 278 municípios⁹.

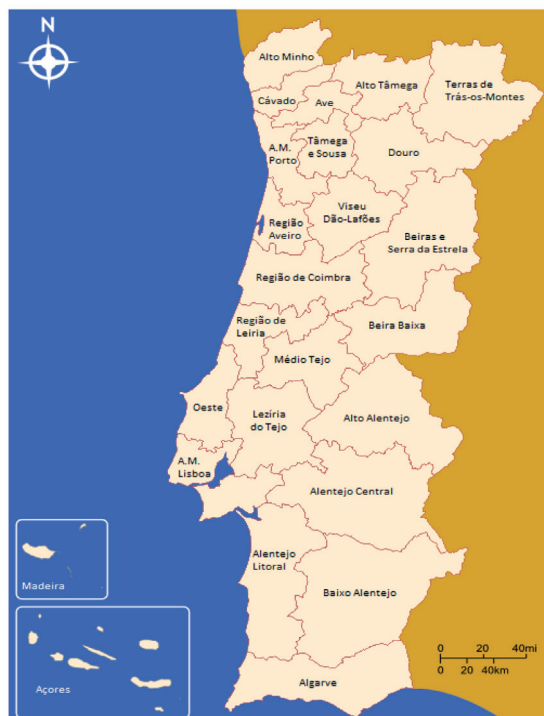


Figura n.º 3 – Mapa das Comunidades Intermunicipais e Áreas Metropolitanas.
Fonte: Wikipédia

Essas Comunidades Intermunicipais são constituídas pelos seguintes municípios (sentido norte-sul, como publicado no decreto-lei): *Comunidade Intermunicipal do Alto Minho* (10 municípios); *Comunidade Intermunicipal do Cávado* (6 municípios); *Comunidade Intermunicipal do Ave* (8 municípios); *Área Metropolitana do Porto* (17 municípios); *Comunidade Intermunicipal do Alto Tâmega* (6 municípios); *Comunidade Intermunicipal do Tâmega e Sousa* (11 municípios); *Comunidade Intermunicipal do Douro* (19 municípios); *Comunidade Intermunicipal das Terras de Trás-os-Montes* (9 municípios); *Comunidade Intermunicipal da Região de Aveiro* (11 municípios); *Comunidade Intermunicipal da Região de Coimbra* (19 municípios); *Comunidade*

⁹ Cf. <https://files.dre.pt/1s/2013/09/17600/0568805724.pdf> (consultado em 4 de setembro de 2023).

Intermunicipal da Região de Leiria (10 municípios); *Comunidade Intermunicipal Viseu Dão Lafões* (14 municípios); *Comunidade Intermunicipal das Beiras e Serra da Estrela* (15 municípios); *Comunidade Intermunicipal da Beira Baixa* (6 municípios); *Comunidade Intermunicipal do Oeste* (12 municípios); *Comunidade Intermunicipal do Médio Tejo* (13 municípios); *Área Metropolitana de Lisboa* (18 municípios); *Comunidade Intermunicipal do Alentejo Litoral* (5 municípios); *Comunidade Intermunicipal do Alto Alentejo* (15 municípios); *Comunidade Intermunicipal do Alentejo Central* (14 municípios); *Comunidade Intermunicipal do Baixo Alentejo* (13 municípios); *Comunidade Intermunicipal da Lezíria do Tejo* (11 municípios); *Comunidade Intermunicipal do Algarve* (16 municípios).

Em paralelo, estão identificados 327 concelhos como sendo territórios de baixa densidade (do interior), pelo Programa Nacional para a Coesão Territorial. Estes estão agrupados na NUTS III, a Nomenclatura das Unidades Territoriais para Fins Estatísticos, um sistema hierárquico de tripla divisão do território em regiões. Há, ainda, a juntar a essas mais de três centenas de municípios, 76 freguesias e união de freguesias que são consideradas territórios de baixa densidade, mesmo que incluídas em municípios que não o são. Isso sucede em 12 NUTS III: *Algarve* (7 freguesias); *Alto Minho* (28 freguesias); *Área Metropolitana do Porto* (2 freguesias); *Ave* (1 freguesia); *Cávado* (4 freguesias); *Lezíria do Tejo* (3 freguesias); *Médio Tejo* (9 freguesias); *Região de Aveiro* (2 freguesias); *Região de Coimbra* (1 freguesia); *Região de Leiria* (1 freguesia); *Região de Viseu Dão Lafões* (6 freguesias); *Tâmega e Sousa* (12 freguesias)¹⁰.

A cooperação intermunicipal tem sido apresentada, mormente no espaço europeu, como uma solução para reduzir as dificuldades que os municípios têm em gerir, de forma solitária. Por seu turno, a cooperação intermunicipal, perante a realidade portuguesa, afigura-se como uma opção muito válida (Tavares, 2015). Segundo Hulst e Monfort (2007), essa realidade assenta em dois pilares importantes que se combinam: 1) a “governança” e as redes entre os municípios funcionam melhor; 2) Há uma maior qualidade do “governo local”, em especial em termos dos serviços públicos. Assim se afigura o panorama nacional atual: duas Áreas Metropolitanas, 21 Comunidades Intermunicipais e 53 Associações Intermunicipais (Teles, 2016).

3. Duas dinâmicas para alavancar os territórios do interior

De acordo com o programa nacional para a coesão territorial, o interior de Portugal apresenta “dinâmicas demográficas regressivas”, nomeadamente: perda acentuada da população e envelhecimento desta; baixa taxa de fertilidade sustentada na “fuga de jovens”; baixos níveis de instrução (e tecnológica) da população residente face à média nacional; taxas migratórias que não conseguem equilibrar positivamente as saídas e as entradas¹¹.

¹⁰ Cf. <https://business.turismodeportugal.pt/SiteCollectionDocuments/financiamento/lista-territorios-de-baixa-densidade-mai-2018.pdf> (consultado em 4 de setembro de 2023).

¹¹ Cf. <https://www.portugal.gov.pt/pt/gc21/governo/programa/programa-nacional-para-a-coesao-territorial-/ficheiros-coesao-territorial/programa-nacional-para-a-coesao-territorial-o-interior-em-numeros-os-territorios-do-interior-pdf.aspx> (consultado em 5 de setembro de 2023).

A coesão territorial tem merecido, por parte dos governantes, no quadro da União Europeia, uma atenção crescente nos últimos anos. No caso específico de Portugal, o Orçamento de Estado, para o ano de 2023, alocou uma fatia financeira importante para os territórios do interior. Uma das medidas propostas pelo ministério da Coesão Territorial passa pela valorização dos territórios do interior, através de um investimento de 950 milhões de euros¹². Para isso, estão a ser (ou serão) criados benefícios fiscais para as empresas e cidadãos que criem postos de trabalho nesses territórios. Um desses programas é o *Emprego Interior MAIS – Mobilidade Apoiada para Um Interior Sustentável*, que o Governo lançou em 2020 e que se prolongará até ao final de 2023. O objetivo do Governo é apoiar financeiramente os trabalhadores, com contratos de trabalho (por conta de outrem ou criação do próprio emprego) ou empresas, cujo local de trabalho seja nos territórios interiores. No fundo, fica implícito o desejo de fazer crescer o número de empresas nessas regiões, atrair mais pessoas e fixar as que aí residem¹³. Investimentos e benefícios (diferenciação positiva) como esses são considerados necessários, de forma a estimular os territórios do interior, em vários parâmetros, e de promover neles a “boa governança”. A governança territorial é considerada essencial para a revitalização territorial e estabelece-se num princípio alargado de responsabilidade, de coordenação, de cooperação, de coerência e de informação (Tsukamoto, 2017).

As dinâmicas territoriais também funcionam como alavancas para o desenvolvimento dos territórios do interior. Entre os diversos elementos que ajudam a promover o seu crescimento estão dinâmicas como as literacias mediática e digital e a participação cívica. Esses dois elementos poderão, como outros, agir como *inputs* dos cidadãos e das comunidades, num processo integrado de mais-valias para ambos. É nesse ambiente tripartido que atores como os nómadas digitais podem ser verdadeiros agentes para ajudar a incrementar o desenvolvimento local¹⁴. Os nómadas digitais estão cada vez mais presentes no tecido laboral, à escala planetária, e constituem uma parcela importante da população em idade ativa. Eles possuem competências tecnológicas, recursos financeiros e muitos nómadas digitais procuram o bem-estar nos territórios do interior, longe dos centros urbanos¹⁵.

Mas, então, como é que se ligam as literacias mediática e digital e a participação cívica com os territórios do interior? E como é que esses dois elementos podem ou devem ajudar a dinamizar esses territórios? Os dados fornecidos pelo Instituto Nacional de Estatística e Pordata não permitem, por exemplo, o cruzamento entre o nível de ensino e o local de residência em Portugal. O mesmo se passa com o ato de exercer cidadania. Qualquer indicação será meramente especulativa e esse vazio invalida uma correta análise científica. Porém, há dados que poderão indiciar uma certa tendência futura.

¹² Cf. <https://oe2023.gov.pt/areas-governativas/coesao-territorial/> (consultado em 5 de setembro de 2023).

¹³ Cf. <https://www.dgert.gov.pt/emprego-interior-mais> (consultado em 5 de setembro de 2023).

¹⁴ Cf. <https://eco.sapo.pt/2023/04/21/nomadas-digitais-sao-fundamentais-diz-ministra-do-trabalho/> (consultado em 5 de setembro de 2023).

¹⁵ Cf. <https://www.ruraldigitalnomads.com/pt/2023/08/18/impulsionando-o-desenvolvimento-do-interior-atraindo-nomadas-digitais/> (consultado em 5 de setembro de 2023).

3.1. Literacias mediática e digital

As literacias mediática e digital são essencialmente um conjunto de competências tecnológicas e de conhecimentos adquiridos pelas pessoas, de forma a acederem, criarem, avaliarem e compreenderem melhor as mensagens dos diversos meios de comunicação (Lopes, 2018). Saber lidar com os dispositivos tecnológicos é tão importante como a formação de um pensamento crítico. Isso ajuda a fazer escolhas mais informadas num mundo marcado pela desinformação e pós-verdade (Barriga, 2023).

Nesse âmbito, em 2022, foi promovido um estudo pelo *Open Society Institute – Sofia e Media Literacy Index da European Policies Initiative (EuPI)*¹⁶. Um dos objetivos do estudo foi avaliar a vulnerabilidade das populações face à desinformação, nomeadamente as notícias falsas (*fake news*), mas aspetos como a liberdade dos Média e a confiança na Sociedade também foram considerados. Portugal ocupou o 14.º lugar no índice da literacia mediática, num universo de 41 países europeus. A consciência e a educação dos cidadãos foram apresentadas como duas das soluções a longo prazo, para aumentar o rácio das literacias mediática e digital, como aludido no estudo.

A educação é um dos indicadores que ajuda a compreender o papel das literacias mediática e digital. Nesse sentido, a Comissão Europeia apresentou, em 2020, no âmbito do Plano de Ação para a Educação Digital (2021-2027), as “Orientações para professores e educadores sobre o combate à desinformação e a promoção da literacia digital através da educação e da formação”¹⁷.

Já no caso português foi criada, em 2017, a *Iniciativa Nacional Competências Digitais e.2030, Portugal INCoDe.2030*. O objetivo é promover as competências digitais nos cidadãos, de forma a tornarem-se mais ativos no mercado de trabalho, que está cada vez mais digitalizado. Mas também a serem, por via das competências digitais, melhores cidadãos, mais conhecedores e mais críticos na sociedade¹⁸.

Todavia, como referido anteriormente, Portugal ainda apresenta valores baixos ao nível das literacias mediática e digital (14.º lugar). Considera-se que essa realidade tem relação com os dados da educação. De acordo com os Censos em Portugal de 2021, do Instituto Nacional de Estatística, expostos pela Pordata, ainda existem 528.088 pessoas sem qualquer nível de escolaridade. Um outro indicador aponta a existência de 292.809 pessoas que não sabem ler/escrever, uma taxa de 3.1%¹⁹. Ora, esses dados são uma antítese com os princípios das literacias mediáticas e digital: capacitação e desenvolvimento de ferramentas que promovam o pensamento crítico, face ao que é lido nos Média e, nomeadamente, nas plataformas digitais²⁰. O ensino básico, com 49.7% é o

¹⁶ Cf. https://osis.bg/wp-content/uploads/2022/10/HowItStarted_MediaLiteracyIndex2022_ENG_.pdf (consultado em 5 de setembro de 2023).

¹⁷ Cf. https://learning-corner.learning.europa.eu/learning-materials/tackling-disinformation-and-promoting-digital-literacy_pt (consultado em 6 de setembro de 2023).

¹⁸ Cf. <https://www.incode2030.gov.pt/> (consultado em 6 de setembro de 2023).

¹⁹ Cf. <https://www.pordata.pt/portugal/populacao+analfabeta+residente+com+10+e+ mais+ anos+segundo+os+censos+total+e+por+sexo-2516> (consultado em 6 de setembro de 2023).

²⁰ Cf. https://repositorium.sdum.uminho.pt/bitstream/1822/44959/1/Pinto_Manuel_coord_2011_epm-Portugal.pdf (consultado em 6 de setembro de 2023).

mais representativo, seguindo-se o ensino secundário, com 23.5%. O número de pessoas com diploma ou frequência no ensino superior, em 2021, fixou-se em 1.782.888. São menos de dois milhões de residentes em Portugal, num universo de 10.407.7²¹.

Noutro prisma, é notório o crescimento progressivo de estrangeiros, anos antes da pandemia de covid-19. Referindo apenas o caso dos nómadas digitais, de acordo com a *Nomad List*, um dos *sites* mais conhecidos no meio, Lisboa era, em 12 de setembro de 2023, o destino mais popular na Europa e no Mundo. A *Nomad List* tem flutuações diárias (com base na opinião dos nómadas digitais), mas nesse dia havia seis lugares nacionais entre os 50 primeiros lugares de uma lista de 1359 locais: Açores (4.º lugar); Madeira (6.º); Portimão (18.º); Lagos (19.º); Ericeira (34.º); Porto (36.º)²². Entre os diversos parâmetros de análise estão os cinco essenciais: Aspetos totais; Custo; Internet; Divertimento; Segurança. Há um elemento que salta à vista: esses seis locais não são territórios do interior, pelo que não ajudam ao aumento da taxa nacional do nível das literacias mediática e digital em Portugal. De resto, eles poderão contribuir para a tendência futura da desertificação daqueles territórios, que o Governo português tenta travar com várias medidas²³.

Anos	Sexo		
	Total	Masculino	Feminino
1980	50.750	30.018	20.732
.....			
2017	421.711	205.874	215.837
2018	480.300	237.466	242.834
2019	590.348	296.417	293.931
2020	662.095	336.123	325.972
2021	698.887	359.862	339.025

Tabela 1 – População estrangeira com permanência regular em Portugal. Fonte: <https://www.pordata.pt/portugal/populacao+estrangeira+com+permanencia+regular+total+e+por+sexo-623>

Há um outro indicador que merece atenção. O relatório “*Education at a Glance 2023*” regista que Portugal gastou, por estudante/ano, menos 14% por cento do que a média dos países da Organização para a Cooperação e Desenvolvimento Económico (OCDE). Face à média anual da OCDE, de 11.766 euros anuais por aluno, Portugal gastou menos 1703 euros por estudante/ano, isto é, 10.063 euros²⁴. Este dado poderá indiciar uma

²¹ Cf. <https://www.pordata.pt/portugal> (consultado em 6 de setembro de 2023).

²² Cf. <https://nomadlist.com/europe> (consultado em 7 de setembro de 2023).

²³ Cf. <https://www.portugal.gov.pt/pt/gc21/governo/programa/programa-nacional-para-a-coesao-territorial-/ficheiros-coesao-territorial/programa-nacional-para-a-coesao-territorial-o-interior-em-numeros-os-territorios-do-interior-pdf.aspx> (consultado em 7 de setembro de 2023).

²⁴ Cf. <https://www.oecd-ilibrary.org/docserver/e13bef63-en.pdf?expires=1694550201&id=id&accname=guest&checksum=6B58B776919653B6C40C3080B3D3BDAC> (consultado em 7 de setembro de 2023).

tendência no futuro, em particular ao nível das literacias mediática e digital em Portugal e, naturalmente, nos territórios do interior.

3.2. Participação cívica

A participação dos cidadãos na sociedade alterou-se muito com o desenvolvimento da Internet e, sobretudo, das redes sociais. A utilização massiva das tecnologias abriu espaço para que a sociedade civil passasse a dispor de meios para participar na vida pública de uma forma mais ativa e visível. Isso possibilitou aos cidadãos aumentar a capacidade de intervir em áreas como a participação cívica, num novo estreitar de relações entre o Estado e a Sociedade (Araújo, 2015).

Segundo o relatório “*Estado Global das Democracias*”, do Instituto Internacional para a Democracia e a Assistência Eleitoral (IDEA), relativo ao ano de 2021, refere-se que Portugal é considerado um país democraticamente estável, mas onde há um défice na participação dos portugueses na sociedade civil. Ou seja, concretamente ao nível dos referendos e outras consultas populares, o relatório regista que o grau de abertura do Governo português é menor, face a outros países europeus, em proporcionar aos cidadãos nacionais essa tomada de decisões políticas²⁵.

Os cidadãos têm geralmente dois modos de exercer a sua participação cívica: 1) por via da *democracia representativa*, em que participam nas eleições, votando nos partidos políticos, em referendos, etc. Neste modo, a participação dos cidadãos reduz-se geralmente apenas a esse ato cívico, pelo que outras formas de participação não existem ou são reduzidas; 2) através da *democracia participativa*, em que têm uma presença mais frequente nos momentos de decisão. Tal verifica-se, por exemplo, nos *orçamentos participativos*, em que os cidadãos se sentem atores e decisores democráticos, pois agem em benefício das próprias comunidades em que vivem (Magalhães, 2017).

Nesse segundo princípio de “democracia direta” e tendo como referência o relatório “*Estado Global das Democracias*”, em que moldes se constrói a participação cívica em Portugal? E pode ela, especificamente, refletir uma dinâmica nos territórios do interior? A participação dos portugueses na sociedade civil é, como referido anteriormente, deficitária. A sua explicação tem sido apontada por vários fatores, entre os quais: democracia tardia em Portugal, cultura cívica incipiente, práticas de participação e importância cívica baixas (Pedro-Rêgo e Simplicio, 2004).

Há, em Portugal, um histórico de abstenção em atos eleitorais²⁶. Cancela e Vicente (2019) reforçam, no estudo *Abstenção e Participação Eleitoral em Portugal: Diagnóstico e hipóteses de Reforma* que a abstenção eleitoral dos cidadãos portugueses é maior

²⁵ Cf. <https://observador.pt/2022/11/30/portugal-apresenta-defice-democratico-na-participacao-civica/> (consultado em 8 de setembro de 2023). Ver também: https://www.idea.int/sites/default/files/news/news-pdfs/IDEA_GSoD_2022_GLOBAL_PRESS_RELEASE-EMBARGOED_Portuguese.pdf O relatório em inglês: <https://www.idea.int/> (consultado em 8 de setembro de 2023).

²⁶ Cf. <https://www.pordata.pt/tema/portugal/participacao+eleitoral-44> (consultado em 8 de setembro de 2023).

nas zonas menos populosas (interior) do que nas regiões mais povoadas (litoral)²⁷. Há outros dois aspetos que merecem atenção: 1) Os círculos eleitorais mais pequenos concentram-se, *grosso modo*, no interior do país. Sociologicamente, os seus eleitores tendem a considerar que o seu voto não tem o mesmo peso do voto dos eleitores do litoral, onde há círculos eleitores maiores, que concentram 70% dos deputados eleitos para a Assembleia da República; 2) As desigualdades socioeconómicas entre o litoral (mais rico) e o interior (mais pobre) têm aumentado nos últimos anos e isso interfere no momento de votar, como salientou João Cancela (2022), em declarações ao jornal *Observador*²⁸. O politólogo e autor do estudo (referido em cima) assevera que há um “fosso de participação entre as pessoas que têm mais recursos financeiros e as pessoas que têm menos recursos”²⁹, e que essas desigualdades são perceptíveis no momento de votar. De tal forma que, para João Cancela, os cidadãos que estão na base da pirâmide socioeconómica participam menos nos atos eleitorais do que os que ocupam o topo.

Ora, estas são duas variáveis que podem ajudar a perceber os níveis de abstenção nos territórios do interior; E essa realidade pode, assim, ser um travão à participação cívica dos cidadãos no interior do país e, mormente, nos territórios do interior. Daí que seja notório que há ainda muito a fazer nesse nível.

4. Conclusões

A valorização dos territórios do interior é, desde os últimos anos, um tema emergente no seio da União Europeia e, também, em Portugal, com a aplicação de investimentos avultados, como é o caso do Programa de Valorização do Interior, em Portugal, com a dotação financeira total de 6.6 mil milhões de euros. Uma das áreas desse investimento são as pessoas. A aplicação de benefícios fiscais para as empresas, que se fixem e criem postos de trabalho nos territórios de baixa densidade, tem as pessoas como aspeto fundamental.

De forma a contrariar a desertificação do interior, não bastará ‘apenas’ injetar dinheiro nas empresas ou noutras estruturas nos territórios. Dinâmicas tecnológicas como as literacias mediática e digital e a participação cívica são peças-chave para também alavancar os territórios do interior que constituem quase a globalidade das comunidades intermunicipais estabelecidas em Portugal. Ora, uma população mais informada e digitalmente conhecedora, mais atenta ao panorama geral e mais participativa constitui, sem dúvida, um benefício essencial para o desenvolvimento global. Há que envidar esforços políticos relevantes para subtrair as diferenças entre litoral e interior, de forma a promover, efetivamente, uma desigualdade positiva em favor dos territórios de baixa densidade, isto é, do interior. Também se deverá atender ao nível da abstenção que se verifica em Portugal, em cada ato eleitoral. E, como se referiu, com reais incidências em diferenças socioeconómicas da população e dos votantes, que importa nivelar. De facto,

²⁷ Cf. https://research.unl.pt/ws/portalfiles/portal/16817196/Estudo_Portugal_Talks_Absten_o_e_Participa_o_Eleitoral_em_Portugal_2019_1.pdf (consultado em 6 de setembro de 2023).

²⁸ <https://observador.pt/especiais/da-participacao-massiva-a-abstencao-eleitoral-como-chegamos-ate-aqui/> (consultado em 6 de setembro de 2023).

²⁹ Idem, *ibidem*.

as pessoas são, verdadeiramente, o ‘motor’ dos territórios do interior e os reais agente de mudança. Sem elas, o país não tem futuro e o interior ainda menos.

Referências

- Almeida, M. (2021). A nova paisagem rural portuguesa e o trabalho na agricultura superintensiva. *Da Escravidão ao Trabalho Digno: Nos 150 anos da abolição da escravidão em Portugal e nos 100 anos da criação da OIT*. Seixas, M., Rodrigues, C. (orgs). Cadernos Sociedade e Trabalho, XXI, (289-302). Edições GEP.
- Barriga, A. C. (2023). Lugares (ausentes) de literacia mediática e digital. *Configurações*, 31, consultado em 9 de setembro de 2023. <http://journals.openedition.org/configuracoes/17005>
- Cancela, J., Vicente, M. (2019). *Abstenção e Participação Eleitoral em Portugal: Diagnóstico e Hipóteses de Reforma*. Câmara Municipal de Cascais.
- Lopes, P. (2018). Avaliação de competências de literacia mediática: instrumentos de recolha de informação e opções teórico metodológicas. Imprensa da Universidade de Coimbra.
- Lopes, P. (2011). *Literacia mediática – conceito e orientações: a abordagem europeia. Literacia(s) e literacia mediática*. CIES-IUL. Disponível em [http://repositorio.ual.pt/bitstream/11144/195/1/literacia\(s\)%20e%20literacia%20mediatica.pdf](http://repositorio.ual.pt/bitstream/11144/195/1/literacia(s)%20e%20literacia%20mediatica.pdf)
- Monteiro, A. (2019). *Territórios do Interior, Coesão Territorial e Modelos de Governança: a propósito do Programa Nacional para a Coesão Territorial*. ISCTE-IUL.
- Mota, B. (2019). *A Problemática dos Territórios de Baixa Densidade: Quatro Estudos de Caso*. ISCTE-IUL.
- OECD (2023). *Education at a Glance 2023 Sources, Methodologies and Technical Notes*. OECD Publishing. Disponível em <https://doi.org/10.1787/d7f76adc-en>.
- Pires, R. P., Vidigal, I., Pereira, C., Azevedo, J., & Veiga, C. M. (2022). *Emigração Portuguesa 2022: Relatório Estatístico*. Observatório da Emigração e Rede Migra, CIES-ISCTE.
- Magalhães, R. (2017). *Mecanismos de Participação Cívica em Portugal: O Caso dos Referendos Locais*. Instituto Politécnico de Bragança. Disponível em https://bibliotecadigital.ipb.pt/bitstream/10198/15934/1/Magalh%C3%A3es_Rui.pdf
- Santos, D. (2005). *Dinâmicas de inovação em territórios de baixa densidade*. IPCB.
- Tsukamoto, M. (2017). *Governança Multiníveis em Territórios de Baixa Densidade: as Comunidades Intermunicipais do Alto Alentejo e da Beira Baixa*. FCSH-UNL.
- Pedro-Rêgo P., Simplício, D. (2004). Redes Localizadas: a Participação Cívica, a partir do Meio Local. In *V Congresso da Geografia Portuguesa: Territórios e Protagonistas*. Disponível em https://apgeo.pt/files/docs/CD_V_Congresso_APG/web/index.html

Análisis bibliométrico de la educación virtual y rendimiento académico universitario periodo 1996-2021, utilizando datos Scopus

Ignacio Romero Cruz¹

iromero77cruz@gmail.com

¹ Universidad Autónoma Gabriel Rene Moreno, Santa Cruz-Bolivia.

DOI: 10.17013/risti.52.126–141

Resumen: En la presente investigación se realizó un análisis descriptivo de la producción científica de la educación virtual y el rendimiento académico universitario, la metodología utilizada es de bibliometría, se utilizó ocho indicadores para explicar el objeto de estudio. Los resultados, indican un crecimiento exponencial de artículos, los países que lideran con coautoría y producción son: España, Estados Unidos y Arabia Saudí, las instituciones de mayor producción están: Estados Unidos, España y Arabia Saudí, los autores más relevantes son: Aggarwal Rajesh, Abdel Wahab, Hussein M, y Al Abdulmonem, Waleed. Los temas de mayor interés están relacionados al aprendizaje, motivación, satisfacción del estudiante, rendimiento académico y universidad. Se concluye, que la educación virtual tiende a tener mayor interés por parte de investigadores, y que el mismo está asociado a los avances tecnológicos, lo que implica un espectro más amplio a realizar investigaciones en el área de educación mediados por la tecnología.

Palabras-clave: Bibliometría, producción científica, educación virtual.

Bibliometric analysis of virtual education and university academic performance from 1996 to 2021, using Scopus data

Abstract: In the present study, a descriptive analysis of the scientific production in virtual education and university academic performance was conducted. The methodology employed is bibliometrics, utilizing eight indicators to elucidate the object of study. The results indicate an exponential growth in articles, with leading co-authorship and production from Spain, the United States, and Saudi Arabia. The institutions with the highest production are in the United States, Spain, and Saudi Arabia, and the most relevant authors include Aggarwal Rajesh, Abdel Wahab Hussein M, and Al Abdulmonem Waleed. The topics of greatest interest are linked to learning, motivation, student satisfaction, academic performance, and university. In conclusion, virtual education tends to garner increased interest from researchers, and this interest is associated with technological advancements, implying a broader spectrum for conducting research in technology-mediated education.

Keywords: Bibliometrics, scientific production, virtual education.

1. Introducción

La evolución constante de las tecnologías de la información y comunicación (TIC) han cambiado todos los sectores de la economía, entre ellos la educación virtual, modificando varios aspectos de la vida moderna (Ahmad et al., 2023); no obstante, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura UNESCO (2022), afirma que la tecnología mejora el acceso a la educación y promueve un aprendizaje de calidad para todos.

La educación virtual sobre el rendimiento académico superior ha sido estudiada por diferentes investigadores. Entre ellos Algahtani et al. (2021), describe las autopercepciones académicas de los estudiantes y la satisfacción del curso durante los estudios virtuales, destaca que estos factores son importantes para retener la motivación de los estudiantes en el aprendizaje y el rendimiento. Se debe considerar que el aprendizaje virtual es una buena práctica, si las TIC en la educación el proceso y sus principios se adhieren de manera más eficiente (Abdullah et al., 2021).

Otras variables surgen a partir de la relación entre la educación virtual y el rendimiento académico como manifiesta Cervero et al. (2020), quien estudia la influencia en la calidad de los procesos de enseñanza-aprendizaje en entornos virtuales, los resultados encontrados sugieren cuatro variables: respuestas satisfactorias de los docentes a las preguntas y observaciones de los estudiantes, la actitud positiva de los docentes hacia el uso de las tecnologías de la información y la comunicación, que los estudiantes tengan habilidades digitales adecuadas y actividades que fomenten las ideas y el debate.

Sin embargo, otros estudios se enfocan en la mejora a la infraestructura tecnológica, determinando que es un factor que contribuye a mejorar la comunicación entre docentes y estudiantes (El Said, 2021). De forma similar, Majid (2019) propone que, una mejora de las instalaciones y la infraestructura del equipo tecnológico (TIC) incrementa la eficacia en la implementación de los entornos de aprendizaje virtual (VLE). Es evidente que contar con infraestructura tecnológica en la educación virtual hace del proceso enseñanza-aprendizaje más significativo.

Dentro de este orden de ideas, otros estudios realizados en entornos de la educación virtual como de Torres Martín (2021), Makransky (2019), Mishra (2020) establecen una insatisfacción generalizada de los estudiantes en las metodologías aplicadas en el diseño de los contenidos de aprendizaje. Sin embargo, los trabajos de Hernández (2023), Huang (2016), Radu (2020) coinciden que los cursos mediados por la realidad virtual tienen un impacto positivo en la utilidad percibida cuando existe un entorno donde se incentive la creatividad, se brinde claridad con respecto a las tareas creativas y se de asesoramiento por parte de los profesores en el proceso creativo. En contraste o contraposición (Zambrano-Vacacela, 2020) menciona, que existe pocas habilidades en el uso de las tecnologías de información y comunicación (TIC) y la inteligencia emocional IE de parte de los docentes, llegando establecer una baja correlación entre el manejo de las TIC y la IE. Entonces que el profesor cuente con habilidades didácticas, metodológicas y tecnológicas si pudiera tener un impacto positivo en el proceso de enseñanza-aprendizaje de los estudiantes (Aguayo et al., 2021).

En este contexto, el campo de la educación ha tratado de desarrollar un marco teórico referente a la educación virtual o a distancia. Algunos autores como Simonson et al. (2019), se refiere a la educación virtual “como la educación formal basada en una institución donde los grupos de estudio están separados y los estudiantes están conectados entre sí mediante un sistema de telecomunicaciones interactivo” (p.31). Sin embargo, García-Aretio (2003), enfatiza que la educación se sustenta en un dialogo pedagógico en el que intervienen un profesor (institución) y estudiante, quienes se sitúan en un espacio diferente al anterior aprendiendo de forma independiente y/o colaborativa.

Los trabajos realizados en el área de la educación a distancia que más contribuyeron son de: Otto Peters, Michael Moore, Borje Holmberg, Desmond Keegan, D.R. Garrison, Myra Baynton, Doug Shale y John Verduin y Thomas Clark. (Díaz, 2021). Un aporte importante que hizo en este campo fue Keegan (2013), en su libro “Fundamentals of Distance Education”, propuso tres enfoques teóricos: independencia y autonomía, industrialización de la enseñanza e interacción y comunicación.

De acuerdo con Peters (2020), plantea que la teoría de la educación a distancia es explicada en función a conceptos y principios asociados a la teoría de la producción industrial y que estos procesos de enseñanza y aprendizaje guardan relación con estos esquemas de la producción industrial. Sin embargo, Wedemeyer (1975) afirma que la investigación sobre el aprendizaje abierto se basa en estudios independientes que enfatizan la enseñanza y el aprendizaje de los estudiantes. Ambas teorías abordan los roles de profesores y estudiantes en la educación a distancia mediados a través de interacciones comunicativas.

Por lo expuesto, conocer el campo investigativo de la educación virtual contribuirá a tener un conocimiento más amplio de la misma y así plantear alternativas que permitan mejorar el rendimiento académico. En ese sentido se planteó en el presente trabajo de investigación un estudio bibliométrico para realizar un análisis cuantitativo de la producción científica del objeto de estudio (Tomás-Górriz & Caster, 2018), con el objetivo de realizar un análisis descriptivo de la producción científica sobre la educación virtual en el rendimiento académico universitario, en base a datos de Scopus en el periodo 1996-2021, utilizando indicadores bibliométricos.

Para ello se utiliza el método descriptivo, comparativo y retrospectivo con la finalidad de responder a la pregunta de investigación, ¿Cuál es la tendencia de la producción científica de la educación virtual en el rendimiento académico universitario en el periodo 1996-2021?

2. Metodología y diseño de la investigación

La presente investigación utiliza el análisis bibliométrico para dar evidencia y presentar el conocimiento existente sobre la educación virtual. De acuerdo con Paul (1934), la terminología de bibliometría “es el campo de estudio enfocado en la obtención de datos estadísticos utilizando métricas de libros y documentos” (p.13-14). Posteriormente se reafirma la definición de la bibliometría centrado en la búsqueda de información sobre un tema específico aplicando métodos matemáticos y estadísticos (Pritchard, 1969).

A través del uso de la metodología de la bibliometría se puede evaluar el rendimiento de las publicaciones científicas a nivel internacional de campos específicos de estudio (Li, 2019). Por otro lado, para tener una comprensión sobre un tema específico e identificar las variables existentes en los estudios de investigación, el análisis bibliométrico es un instrumento que permite visualizar las colecciones electrónicas existentes en ese campo (Ali-Shoaib et al., 2021). Del mismo modo, con la enorme cantidad de información documental existente un análisis bibliométrico permite identificar los cambios que se generan en el conocimiento a partir del rastrear y ver las etapas de evolución (Mokhnacheva & Tsvetkova, 2020). De forma similar el análisis bibliométrico tiene como meta mapear el estado del arte sobre un tema de investigación (El Mohadab et al., 2020).

2.2. Recopilación de datos

En la identificación de las publicaciones científicas para el análisis bibliométrico se utilizó la base de datos Scopus por las siguientes cinco razones:

a) Contiene datos de todos los autores en las referencias citadas, lo que hace que el análisis de citas y citas conjuntas basado en el autor sea más preciso; b) tiene una cobertura más amplia y es útil para mapear áreas de investigación más pequeñas que no están cubiertas por Web of Science; c) Está respaldado por paquetes de software bibliométricos más utilizados (Zupic & Cater, 2015); d) Las publicaciones de investigaciones son cuidadosamente seleccionadas por un equipo de asesores y tiene mayor amplitud en cobertura que WOS (Aksnes & Slvertsén, 2019); e) Es la base de datos más grande que alberga todas las disciplinas de la ciencia y conocimiento (Hernández-González et al., 2016).

La recopilación de datos se efectuó el mes de mayo del año 2023 utilizando las siguientes palabras de búsqueda clave: “*Virtual Education*” “*University Academic Performance*”, conectados por operadores booleanos AND.

Se identificaron inicialmente 278 artículos científicos en los intervalos de 1996-2021, y se utilizó el siguiente procedimiento de exclusión: se descartaron publicaciones del 2022 (n=35) para trabajar con año completo y no generar sesgo, también se utilizó otros criterios de exclusión como: *conference paper* (n=84), *conference review* (n=6), *Book Chapter* (n=4), *Editorial* (n=4), *Review* (n=4); *Editorial* (n=1), *Article in Press* (n=2), *Lenguaje Turkish* (n=1), *Undifened* (n=4). Quedando 137 artículos identificados en la base de datos de Scopus para el estudio.

Quedando la ecuación final de la siguiente forma: *TITLE-ABS-KEY((virtual AND education*) AND (university AND academic AND performance)) AND (EXCLUDE(PUBYEAR, 2022)) AND (EXCLUDE(DOCTYPE, “cp”) OR EXCLUDE(DOCTYPE, “cr”) OR EXCLUDE(DOCTYPE, “ch”) OR EXCLUDE(DOCTYPE, “re”) OR EXCLUDE(DOCTYPE, “ed”)) AND (EXCLUDE(PUBSTAGE, “aip”), AND (EXCLUDE(LANGUAGE, “TURKISH”)) AND (EXCLUDE(AFFILCOUNTRY, “Undefined”))*.

La búsqueda de las unidades documentales se basó en las directrices del enfoque Preferred Reporting Items for Systematic Reviews and Meta- Analyzes. (Véase Figura I).

El método PRISMA está compuesto por 27 elementos de verificación que son esenciales, incluye un diagrama de proceso de cuatro etapas para que los informes sean transparentes en las revisiones sistemáticas (Nawijn et al., 2019). En otros términos, el método prisma legitima la transparencia en los procesos de recopilación e inclusión en el análisis de artículos proporcionando puntos de referencias claros para otras diciplinas (Booth et al., 2020), la técnica PRISMA son procesos que otorgan confiabilidad a los lectores de cómo se realizó la revisión de informes conteniendo detalles para su evaluación y confiabilidad de resultados y hallazgos (Page et al., 2021). El procedimiento PRISMA describe paso a paso de cómo se realizó la búsqueda de información en los procesos de selección e identificación de los estudios con el fin de que los trabajos de investigación sean transparentes y valiosos para los usuarios (Yepes-Nuñez et al., 2021).

Como se muestra en la figura 1, se siguió el método PRISMA, que incluye cuatro etapas: I) Identificación como registro identificado a través de búsqueda en la base de datos; II) Selección de los documentos en la base de datos; III) Documentos de elegibilidad; IV) Selección de estudios.

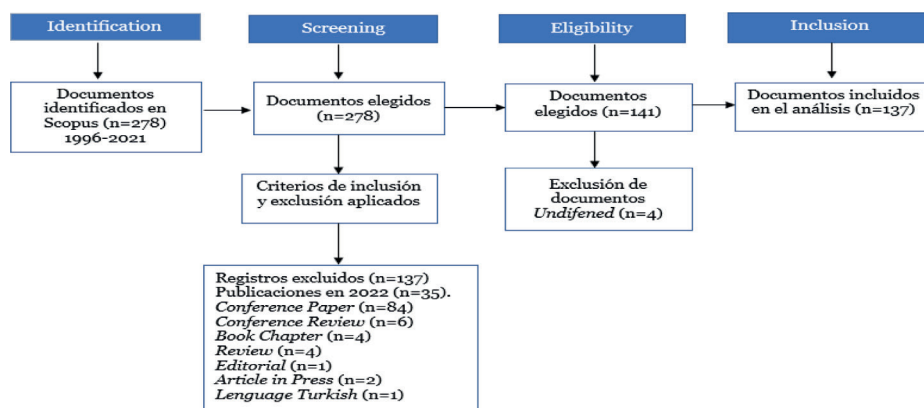


Figura 1 – Procedimiento PRISMA para la identificación de documentos.

2.3. análisis de datos

Esta investigación se realiza basada en la teoría de grafos, para Barrero (2010) los grafos “son estructuras no lineales que consisten en un conjunto de vértices o nodos y arcos, cada uno de los cuales se unen un vértice con otro” (p.1).

Para el análisis se utilizó los softwares Stata versión 17, Vosviewer versión 1.16.8, QGIS versión .32.3: el primero permitió elaborar la tendencia de la producción científica y las tasas de crecimiento anual, el segundo se utilizó para mapear y ver las relaciones de los trabajos de investigación, el tercer software se utilizó para visualizar la contribución por países en el conocimiento de la educación virtual. Los indicadores analizados son: productividad y tasas de crecimiento anual, coocurrencia, red de colaboración entre países, red de colaboración entre investigadores, producción científica por países, países más citados, las 10 revistas más productivas en la investigación de educación virtual y el

artículo más citado, y los 15 autores más relevantes en la investigación de la educación virtual.

3. Resultados

3.1. Productividad y tasas de crecimiento científico de la educación virtual por año

Se produjeron 137 artículos revisados entre los años 1996-2021. La gráfica evidencia cuatro momentos: 2010, 2013, 2015 y 2021 en la cual la producción científica es mayor en relación con los periodos anteriores, sobre todo el último año presenta un incremento exponencial en la elaboración de artículos asociados a la educación virtual. (Véase Figura 2).

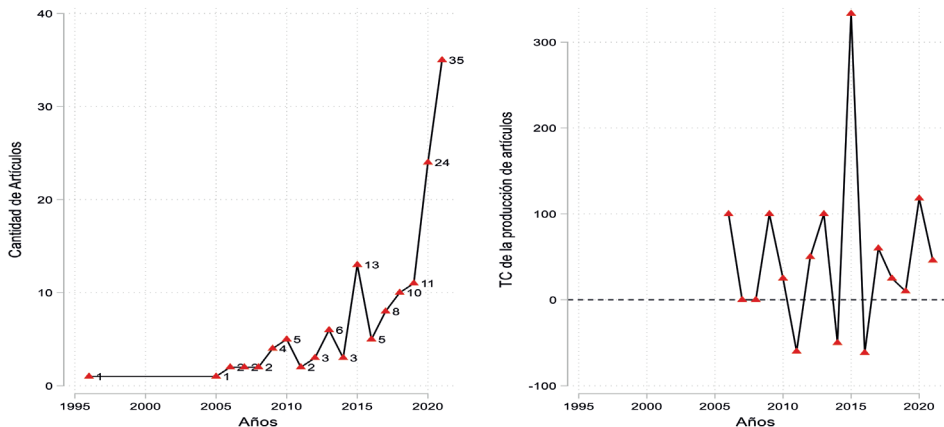


Figura 2 – Evolución y tasas de crecimiento de la producción científica.

Las tasas de crecimiento de la producción científica de la educación virtual en los periodos 1996 - 2021 presenta dos comportamientos: uno es positivo por encima de cero y dos son negativos por debajo de cero, ello implica que los que están por encima de cero, presentan un incremento significativo en la producción de artículos, sin embargo, en los lugares donde se observa tasas negativas se puede entender una disminución en elaboración de artículos, siendo que los años iniciales de 1996 -2005 y 2007 - 2008 la producción se mantuvo constante sin variaciones y el resto de los años presenta variaciones, la media de crecimiento es de 49.73 %. (Véase Figura 2).

La figura 3, está formada por círculos y líneas: los círculos indican un nodo, el tamaño refleja la frecuencia y las líneas representan las interconexiones. La intensidad de color indica los años que aparece las palabras clave.

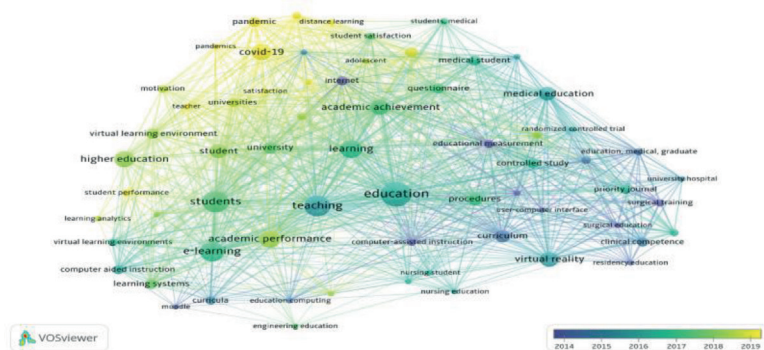


Figura 3 – Coocurrencia de palabras clave del autor.

Los resultados muestran la formación de tres clústeres con 66 ítem: el primer clúster está formado por 24 ítem comprendido entre el 2014-2015, se identifica las siguientes palabras más frecuentes utilizadas por los autores: *Education* (37 apariciones y 415 enlaces), *Virtual reality* (19 apariciones, 217 enlaces), *Medical education* (17 apariciones, 239 enlaces), *Procedures* (15 apariciones, 244 enlaces), y *Curriculum* (15 apariciones, 202 enlaces). El segundo clúster tiene 20 ítem, representada por los de color de menor intensidad comprendida entre 2016-2017, las palabras más usadas en este periodo por los autores son: *Students* (30 apariciones, 305 enlaces), *e-learning* (27 apariciones, 217 enlaces), *academic performance* (22 apariciones, 190 enlaces), *higher education* (21 apariciones, 67 enlaces), *virtual learning environment* (12 apariciones, 109 enlaces), *curricula* (9 apariciones, 71 enlaces). El tercer clúster compuesta por 22 ítems y marcada con una tonalidad más clara, vinculados a los periodos 2018-2019, las palabras con mayor frecuencia son: *Learning* (24 apariciones, 253 enlaces), *Covid-19* (21 apariciones, 137 enlaces), *Academic achievement* (20 apariciones, 301 enlaces), *Student* (18 apariciones, 191 enlaces), *University* (14 apariciones, 193 enlaces), *Motivation* (8 apariciones, 63 enlaces), *Student satisfaction* (7 apariciones, 92 enlaces).

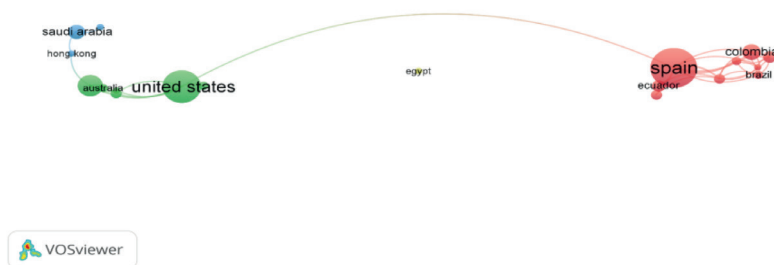


Figura 4 – Red de colaboración entre países.

La figura de red de colaboración entre países, cuanto mayor es la cercanía entre dos países es más fuerte su relación, si las líneas son más gruesa mayor es el vínculo.

Los datos indican que España es el país que presenta mayores coautorías: 28 coautorías, y colabora con 6 países, Colombia (9 coautorías, 8 enlaces), México (6 coautorías, 7 enlaces), Ecuador (6 coautorías, 4 enlaces), Venezuela (5 coautorías, 9 enlaces), Brasil (4 coautorías, 6 enlaces). En este mismo grupo esta Perú con 5 coautorías y llega a colaborar a Ecuador y Venezuela. En segundo lugar, esta Estados Unidos, presenta 22 coautorías llegando a cooperar con 5 países, Reino Unido con (13 coautorías, 7 enlaces), Australia (4 coautorías, 3 enlaces), Canadá (5 coautorías, 3 enlaces), China (4 coautorías, 1 enlaces), France (3 coautorías, 3 enlaces). Un tercer país es Arabia Saudí, con 8 coautorías y se asocia con, Hong Kong (3 coautorías, 2 enlaces), y Malaysia (3 coautorías, 1 enlaces), y por último esta Egipto que no está afiliada a otro país. (Véase Figura 4).

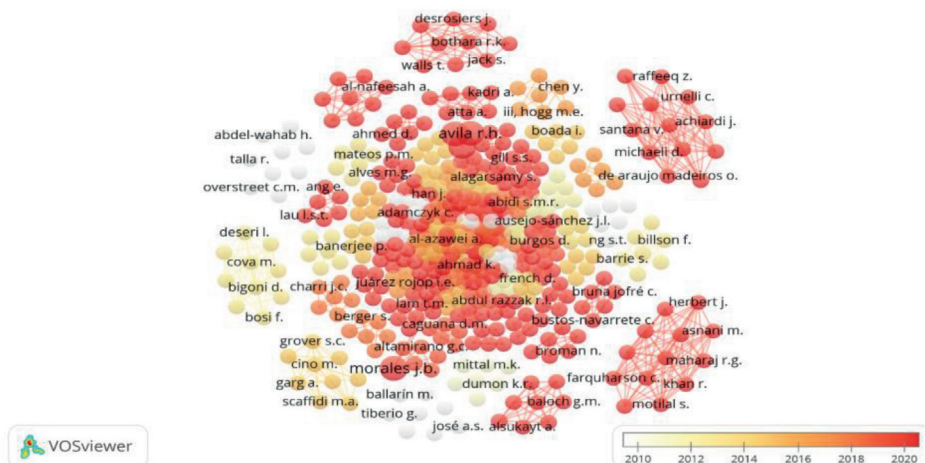


Figura 5 – Red de colaboración entre investigadores.

La Figura 5, Sobre la red de colaboración entre investigadores se identificó 18 clúster distribuidos de la siguiente forma: clúster 1 (13 ítem), clúster 2 (13 ítem), clúster 3 (10 ítem), clúster 4 (10 ítem), clúster 5 (9 ítem), clúster 6 (9 ítem), clúster 7 (8 ítem), clúster 8 (8 ítem), clúster 9 (8 ítem), clúster 10 (8 ítem), clúster 11 (7 ítem), clúster 12 (7 ítem), clúster 13 (7 ítem), clúster 14 (7 ítem), clúster 15 (6 ítem), clúster 16 (6 ítem), clúster 17 (6 ítem) y clúster 18 (6 ítem). Los resultados evidencian que en los años 2018-2020 presenta un incremento significativo en la producción científica en esta temática y a la vez se da una mayor colaboración entre autores.

La figura 6: Sobre producción científica por países permite visualizar que país es el que lidera en la elaboración de artículos científicos en el área de educación virtual y ver su influencia en este campo de estudio.

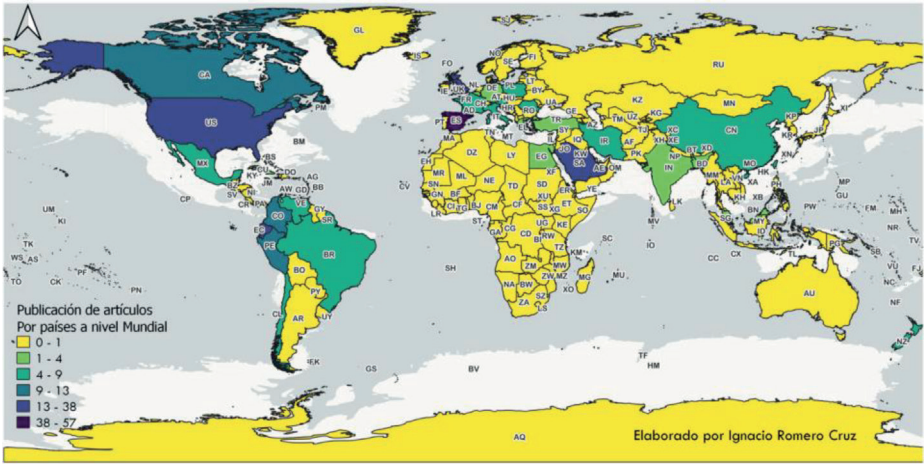


Figura 6 – Producción científica por países.

Los resultados indican que los 15 países más productivos en la investigación sobre educación virtual son: España y Estados Unidos son los que más contribuyeron, el primero con el 23 % y el segundo con 16 %, de un total de 137 artículos, lo que implica que estos dos países son los referentes y los que más estudios realizaron en este campo de estudio. Posterior a ellos esta Arabia Saudí, Ecuador y Reino Unido (7%), Canadá, Colombia y Perú (5%), Australia (4%), Francia, México, Brasil, China, Italia y Venezuela (3 %) y Chile con (2 %).

Nº	Países	CT	CPA	Nº	Países	CT	CPA
1	Spain	320	17.78	11	Jordan	54	54
2	Saudi Arabia	291	72.75	12	Ecuador	51	51
3	United Kingdom	221	27.62	13	Germany	39	39
4	United States	183	12.2	14	Malta	34	34
5	China	155	38.75	15	Pakistan	25	25
6	Canada	98	24.5	16	France	23	7.67
7	Egypt	92	46	17	Sweden	19	19
8	Hong Kong	81	40.5	18	Malaysia	17	8.5
9	Romania	68	34	19	Belgium	14	14
10	Australia	58	14.5	20	Chile	14	7

Tabla 1 – Países más citados.

La tabla 1, de un total de 20 países y 1857 citas, el país más citado es España con 17 %, Arabia Saudita 16 %, Reino unido 12 %, Estados Unidos 10 %, China 8%, Canadá y Egipto

(5 %), Hong Kong y Rumania (4%), Australia, Jordania y Ecuador (3 %), Alemania y Malta (2 %), el resto de los países: Pakistán, Francia, Suecia, Malaysia, Bélgica y Chile (1%).

Revista	PT	TC	PC 2021	El artículo más citado	VC	Editorial
BMC Medical Education	627	7103	3.7	Effects of the COVID-19 pandemic on medical students: a multicenter quantitative study	51	BioMed Central Ltd
Nurse Education Today	435	6443	5.3	Clinical placement experience of nursing students during the COVID-19 pandemic: A cross-sectional study	53	Churchill Livingstone
International Journal of Emerging Technologies In Learning	427	5682	3.8	Remote Academic Platforms in Times of a Pandemic	28	International Association of Online Engineering
Risti Revista Ibérica de Sistemas e Tecnologias de Informação	404	866	0.6	Factors that influence the adoption of the internet of things in the hotel sector	15	Associacao Iberica de Sistemas e Tecnologias de Informação
Computers In Human Behavior	360	23983	14.9	Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study	125	Elsevier Ltd

Tabla 2 – Las 5 revistas más productivas en investigación de la Educación Virtual.

La tabla 2, sobre las revistas más productivas en educación virtual, en primer lugar, está la revista BMC Medical Education es la que tiene la mayor producción con 627 artículos y un total de citaciones de 7103 con puntuación de la cita 3.7 y el artículo Effects of the COVID-19 pandemic on medical students: a multicenter quantitative study, esta 51 veces citado. En 2do lugar la revista Nurse Education Today, con una producción total de 435 artículos y un total de citaciones de 6443 con una puntuación de la cita de 5.3 y el artículo Clinical placement experience of nursing students during the COVID-19 pandemic: A cross-sectional study está 53 veces citado. En tercer lugar, la revista International Journal of Emerging Technologies In Learning tiene una producción de 427 artículos y un total de citaciones de 5682 y su puntuación de la cita es 3.8 y el artículo Clinical placement experience of nursing students during the COVID-19 pandemic: A cross-sectional study, está citado 23 veces. En cuarto lugar, está la revista Risti Revista Ibérica de Sistemas E Tecnologias de Informacao con 404 artículos y 866 total de citaciones, con una puntuación de 0.6 y el artículo Factors that influence the adoption of the internet of things in the hotel sector, está 15 veces citado. El 5to lugar, está la revista Computers In Humann Behavior con una producción total de 360 artículos y tiene un

total de citaciones de 23983 con puntuación de la cita de 14.9 y el articulo Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study, está 125 veces citado.

Autor	Scopus Autor ID	APP	TP	Index-H	TC	Actual afiliación	País
Aggarwal, Rajesh	8616911800	2004	275	58	10651	Panda Health Inc	EE. UU
Abdel-Wahab, Hussein M.	7006481768	1993	87	14	865	Old Dominion University	EE. UU
Al Abdulmonem, Waleed	57213606186	2019	49	7	266	Al Qassim University	Saudí Arabia
Wilches, Fernando Jove	57215843578	2020	37	2	6	Universidad de Sucre	Colombia
Ahmad, Khurshid	57198191812	2010	21	7	149	The Islamia University of Bahawalpur	Pakistan
Al-Azawei, Ahmed	56523409000	2015	18	9	349	University of Babylon	Iraq
Abdul Razzak, Rima L.	6508355887	2011	16	6	91	Arabian Gulf University	Bahrain
Feria-Díaz, Jhon Jairo	56913276000	2015	15	5	75	Tecnológico Nacional de México	México
Adamczyk, Christopher	24166196100	2006	14	10	409	Klinikum der Universität München	Germany
Ahmed, Danish	57217208973	2014	11	2	13	Prince Mohammad Bin Fahd University	Saudí Arabia

Tabla 3 – Los 10 autores más relevantes en la investigación de Educación Virtual.

La tabla 3, sobre los 10 autores más relevantes en investigación en educación virtual en el periodo 1996-2021, en primer lugar, está el investigador Aggarwal Rajesh que inicio en el 2004 a investigar en este campo y es el que más contribuyó con un total de producciones de 275 trabajos investigativos y el index H es de 58 y total de citaciones es de 10651 afiliado a Panda Health Inc de EE. UU. En segundo lugar, está el autor Abdel Wahab, Hussein M, que trabaja desde 1993 en este campo con un total de producciones de 87 trabajos, con index H de 14, total citaciones de 865 con filiación de Old Dominion University de EE. UU. En tercer lugar, de autores más relevantes está Al Abdulmonem Waleed que incursiono el 2019, tiene un total de producciones de 49 trabajos, index H de 7, total de citaciones de 266, afiliado a Al Qassim University de Arabia Saudita. El cuarto lugar, los autores Wilches, Fernando Jove, su primera publicación en este campo investigativo data el 2020 con total de producciones de 37, index H de 2, total citaciones

6, afiliado a la Universidad de Sucre de Colombia. Quinto lugar de autores relevantes están Ahmad, Khurshid que vienen trabajando desde el 2010, con total de producciones de 21, index H de 7, total de citaciones de 149, afiliado a The Islamia University of Bahawalpur de Pakistan y posteriormente se ubican los demás autores.

4. Discusión

El objetivo de este estudio ha sido analizar cuál es la tendencia de la producción científica de la educación virtual y el rendimiento académico universitario en el periodo 1996-2021. Para ello se utilizó la base de datos Scopus. De los resultados obtenidos, se muestra un panorama general sobre la productividad y tasas de crecimiento anual, la coocurrencia, la red de colaboración entre países, la red de colaboración entre investigadores, la producción científica por países, las instituciones con mayor producción científica en educación virtual, países más citados, las 5 revistas más productivas en la investigación de educación virtual y artículo más citado y los 10 autores más relevantes en la investigación de la educación virtual.

La producción de artículos en los periodos analizados tuvo una tendencia cíclica en los periodos 2005-2016, y entre el 2016-2021 la tendencia fue exponencial lo que implica que hubo mayor interés por realizar trabajos vinculados a la educación superior mediados por la tecnología. Lo cual es corroborado por las tasas de crecimiento de la producción de artículos alcanzando una media de 49.73 % de crecimiento.

Debe señalarse por otra parte que la coocurrencia de palabras claves del autor encuentra tres clúster, e identifica las palabras más frecuentemente usadas: El primer clúster, el cual muestra el periodo 2014-2015, realiza más énfasis a palabras como: virtual reality, medical education, procedures y curriculum, el segundo clúster comprende desde 2016-2017, y los términos más frecuentes esta: student, academic performance, higher education, virtual learning environment y curricula, el tercer clúster alcanza los años 2018-2019, las palabras claves son: learning, academic achievement, student, university, motivation y student satisfaction. En estos tres periodos temporales, las palabras claves van evolucionando: desde la comprensión y esquemas, a los procedimientos, y la satisfacción del estudiante en entornos de la educación virtual.

De este modo, la red de colaboración entre países, los datos evidencian que España es el país que presenta mayores coautorías, el segundo país es Estados Unidos y el tercero es Arabia Saudí. Sobre la red de colaboración entre investigadores se identificó 18 clúster, en los años 2018-2020, presenta un incremento significativo y como efecto de ello la colaboración de autores crece.

De este modo, la red de colaboración entre países, los datos evidencian que España es el país que presenta mayores coautorías, el segundo país es Estados Unidos y el tercero es Arabia Saudí. Sobre la red de colaboración entre investigadores se identificó 18 clúster, en los años 2018-2020, presenta un incremento significativo y como efecto de ello la colaboración de autores crece.

Ahora bien, la producción científica por países: el primer lugar ocupa España, segundo lugar Estado Unidos y tercero Arabia Saudí. Los datos guardan correlación con las coautorías de la red de colaboración entre países. De las instituciones con mayor

producción científica en educación virtual, de 30 instituciones que poseen filiación con los autores, los datos nos indican que en EE. UU. del 100 % concentra el 20 % de las instituciones, España 13 %, Arabia Saudita 10 %. De igual forma sobre países más citados, de un total de 20 países y 1857 citas, el país más citado es España con 17 %, Arabia Saudita 16 %, Reino unido 12 %, Estados Unidos 10 %. Estos dos últimos resultados indican, que universidad es la que más trabajos realiza en educación virtual y los países a los que más se cita.

En relación con las 5 revistas más productivas en investigación de la educación virtual esta: BMC Medical Education con 627 artículos, Nurse Education Today 435, International Journal of Emerging Technologies In Learning 427, Risti Revista Iberica de Sistemas E Tecnologías de Informacao 404, Computer In Human Behavior 360. Finalmente se determinó los 10 autores más relevantes en investigación de la educación virtual, la tabla 3 contempla datos del año de la primera publicación, la afiliación y a que país pertenece. El autor más relevante es Aggarwal Rajesh con 275 trabajos en coautoría, entre otros autores esta Abdel Wahab, Hussein M, con 87 producciones, Abdulmonem, Waleed 49 producciones, Wilches, Fernando Jove 37 producciones, Ahmad, Khurshid 21 producciones, Al-Azawei, Ahmed 18 producciones. Estos resultados afirman la supremacía de EE. UU. en el estudio de la educación virtual.

5. Conclusiones

El estudio bibliométrico es un análisis en intervalos de espacios temporales de datos cuantitativos y de variables vinculados a la producción científica, nos proporcionan una visión general acerca de un tema específico. A partir de la aplicación de esta técnica, el presente estudio identifica sobre cuál es la tendencia de la producción científica de la educación virtual en los periodos de 1996-2021.

Por consiguiente, claramente existe evidencia de que hubo un incremento gradual en la producción de estudios científicos en la educación virtual alcanzando una media de producción de artículos de 7.61 con tasas de crecimiento de 49.73 %. Los temas de mayor interés en los últimos años están vinculados al aprendizaje, motivación, satisfacción del estudiante, rendimiento académico y universidad.

Se observa que las coautorías y la producción científica por países la mayor parte se concentraron en: España, Estados Unidos y Arabia Saudí. Consecuentemente las instituciones con mayor producción científica en educación virtual son: Estados Unidos, España y Arabia Saudí. Sin embargo, entre las revistas más productivas esta: BMC Medical Education artículos, Nurse Education Today, International Journal of Emerging Technologies In Learning, Risti Revista Iberica de Sistemas E Tecnologías de Informacao, Computer y In Human Behavior. De ahí que se identificó a los autores más relevantes entre los que podemos citar a: Aggarwal Rajesh, Abdel Wahab, Hussein M, y Al Abdulmonem, Waleed

Dentro de este marco las limitaciones del estudio se encuentran en que solo se tomó la base de datos de Scopus, para otros estudios se podría complementar con otras bases de datos como de WOS. Por lo tanto, se puede concluir que las áreas de investigación en educación virtual tienden a tener mayor interés por parte de investigadores y este crecimiento está asociado a los avances tecnológicos que se vienen dando en los últimos

años, lo que implica un espectro más amplio de realizar investigaciones en el área de educación mediados por la tecnología.

Referencias

- Abdullah, A. S. (2021). Assessment of Academic Performance with the e-mental health interventions in virtual learning environment using Machine learning Techniques: A Hybrid approach. *Journal of Engineering Education Transformations*, 34, 79-85. <https://doi.org/10.16920/jeet/2021/v34i0/157109>
- Aguayo, R., Lizarraga, C., & Quiñonez, Y. (2021). Evaluación del desempeño académico en entornos virtuales utilizando el modelo PNL. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*(41), 34-49. <https://doi.org/10.17013/risti.41.34-49>
- Ahmad, S. M. (2023). eLearning acceptance and adoption challenges in Higher Education. *Sustainability*, 15(7), 6190. <https://doi.org/10.3390/su15076190>
- Aksnes, D. W. (2019). A criteria-based assessment of the coverage of Scopus and Web of Science. *Journal of Data And Information Science*, 4(1), 1-21. Obtenido de <https://sciendo.com/downloadpdf/journals/jdis/4/1/article-p1.pdf>
- Algahtani, F. D. (2021). Academic Self-Perception and Course Satisfaction among University Students Taking Virtual Classes during the COVID-19 Pandemic in the Kingdom of Saudi-Arabia (KSA). *Education Sciences*, 11(3), 134. <https://doi.org/10.3390/educsci11030134>
- Ali, N. S. (2021). Trends of Research Visualization of Digital Collections and Resources in Academic Libraries from 2001 to 2020: A Bibliometric Analysis. *Library Philosophy & Practice*.
- Barrero, A. C. (2010). *Introducción a la Teoría de Grafos*. Armenia, Quindio: Elizcom sas.
- Booth, P. C. (2020). Entrepreneurship in island contexts: A systematic review of the tourism and hospitality literature. *International Journal of Hospitality Management*, (85), 102438. <https://doi.org/10.1016/j.ijhm.2019.102438>
- Cervero, A. C.-L.-B. (2020). Evaluation of educational quality performance on virtual campuses using fuzzy inference systems. *Plos one*, e0232802. <https://doi.org/10.1371/journal.pone.0232802>
- Diaz, D. M. (2021). *AIC*. Obtenido de Marco Teórico de la Educación a Distancia. <http://contadores-aic.org/marco-teorico-de-la-educacion-a-distancia/>
- El Mohadab, M. B. (2020). Bibliometric method for mapping the state of the art of scientific production in Covid-19. *Chaos, Solitons & Fractals*, 139, 110052. <https://doi.org/10.1016/j.chaos.2020.110052>
- El Said, G. R. (2021). How Did the COVID-19 Pandemic Affect Higher Education Learning Experience? An Empirical Investigation of Learners' Academic Performance at a University in a Developing Country. *Advances in Human-Computer Interaction*, 1-10. <https://doi.org/10.1155/2021/6649524>

- García Aretio, L. (2003). La educación a distancia. Una visión global. *Boletín Ilustre Colegio de Doctores y Licenciados de España*, (146), 13-27. Obtenido de http://espacio.uned.es/fez/eserv/bibliuned:20182/educdistanc_visionglobal.pdf
- Hernández, J. P. (2023). Diseño de un entorno virtual de aprendizaje para promover la creatividad colaborativa en universitarios. *RIED Revista Iberoamericana de Educación a Distancia*, 26(2), 177-197. <https://doi.org/10.5944/ried.26.2.36209>
- Hernández-González, V. S.-R.-D.-M. (2016). Comparación entre Web of Science y Scopus, estudio bibliométrico de las revistas de anatomía y morfología. *International Journal of Morphology*, 34(4), 1369-1377. <https://doi.org/10.4067/S0717-95022016000400032>
- Huang, H. M. (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. *Interactive Learning Environments*, 24(1), 3-19. <https://doi.org/10.1080/10494820.2013.817436>
- Keegan, D. (2013). *Foundations of Distance Education*. Routledge. <https://doi.org/10.4324/9781315004822>
- Li, J. A. (2019). Trends and issues in multimedia learning research in 1996–2016: A bibliometric analysis. *Educational Research Review*, 28, 100282. <https://doi.org/10.1016/j.edurev.2019.100282>
- Liberati, A. A. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology*, 10, 62. <https://doi.org/10.7326/0003-4819-151-4-200908180-00136>
- Majid, R. A. (2019). The effectiveness of Frog VLE implementation: students' perspective. *Indonesian Journal of Electrical Engineering and Computer Science*, 14(1), 381-387. <https://doi.org/10.11591/ijeecs.v14.i1.pp381-387>
- Makransky, G. T. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction*, (60), 225-236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>
- Mishra, L. G. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *International journal of educational research open*, (1), 100012. <https://doi.org/10.1016/j.ijedro.2020.100012>
- Mokhnacheva, Y. V. (2020). Development of bibliometrics as a scientific field. *Scientific and Technical Information*, 47, 158-163. <https://doi.org/10.3103/S014768822003003X>
- Nawijn, F. H. (2019). Quality of reporting of systematic reviews and meta-analyses in emergency medicine based on the PRISMA statement. *BMC emergency medicine*, 1-8. <https://doi.org/10.1186/s12873-019-0233-6>
- Otlet, P. (1934). *Traité de documentation. Le livre sur le livre, théorie et pratique*. Bruxelles: Editions Mundaneum.

- Page, M. J. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/bmj.n160>
- Peters, O. (2020). Distance teaching and industrial production A comparative interpretation in outline. In: Distance education, (pp. 95-113). *Routledge*.
- Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of documentation*(25), 348.
- Radu, M. C. (2020). The impact of the COVID-19 pandemic on the quality of educational process: A student survey. *International journal of environmental research and public health*, 17(21), 7770. <https://doi.org/10.3390/ijerph17217770>
- Simonson, M. Z. (2019). *Teaching and learning at a distance: Foundations of distance education*. Unites States of America: IAP. Obtenido de https://www.academia.edu/39818858/Teaching_and_Learning_at_a_Distance_Foundations_of_Distance_Education_SIXTH_EDITION
- Tomás-Górriz, V. &-C. (2018). La Bibliometría en la evaluación de la actividad científica. *Hospital a domicilio*, 2(4), 145-163. <https://doi.org/10.22585/hospdomic.v2i4.51>
- Torres Martín, C. A. (2021). Impact on the virtual learning environment due to COVID-19. *Sustainability*, 13(2), 582. <https://doi.org/10.3390/su13020582>
- UNESCO. (2022). *Transformar la educación para el futuro*. París: UNESCO. Obtenido de https://unesdoc.unesco.org/ark:/48223/pf0000382765_spa
- Wedemeyer, C. A. (1975). Implications of Open Learning for Independent Study.. Obtenido de <https://eric.ed.gov/?id=ED112766>
- Yepes Nuñez, J. J.-F. (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-799. Obtenido de <https://diposit.ub.edu/dspace/bitstream/2445/180116/1/713477.pdf>
- Zambrano Vacacela , L. (2020). Uso de la Tecnología de la Información y Comunicación en educación virtual y su correlación con la Inteligencia Emocional de docentes en el Ecuador en contexto COVID-19. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (40), 31-44. <https://doi.org/10.17013/risti.40.31-44>
- Zupic, I. (2015). Bibliometric methods in management and organization. *Organizational research methods*, 18(3), 429-472. Obtenido de <https://research.gold.ac.uk/id/eprint/26859/1/Zupic%20Cater%202015%20-%20Bibliometric%20methods%20in%20management%20and%20organization.pdf>

Critérios Editoriais

A RISTI (Revista Ibérica de Sistemas e Tecnologias de Informação) é um periódico científico, que foca a investigação e a aplicação prática inovadora no domínio dos sistemas e tecnologias de informação.

O Conselho Editorial da RISTI incentiva potenciais autores a submeterem artigos originais e inovadores para avaliação pelo Conselho Científico.

A submissão de artigos para publicação na RISTI deve realizar-se de acordo com as chamadas de artigos e as instruções e normas disponibilizadas no sítio Web da revista (<http://www.risti.xyz/>).

Todos os artigos submetidos são avaliados por um conjunto de membros do Conselho Científico, não inferior a três elementos.

Em cada número da revista são publicados entre cinco a oito dos melhores artigos submetidos.

Criterios Editoriales

La RISTI (Revista Ibérica de Sistemas y Tecnologías de la Información) es un periódico científico, centrado en la investigación y en la aplicación práctica innovadora en el dominio de los sistemas y tecnologías de la información.

El Consejo Editorial de la RISTI incentiva autores potenciales a enviar sus artículos originales e innovadores para evaluación por el Consejo Científico.

El envío de artículos para publicación en la RISTI debe hacerse de conformidad con las llamadas de los artículos y las instrucciones y normas establecidas en el sitio Web de la revista (<http://www.risti.xyz/>).

Todos los trabajos enviados son evaluados por un número de miembros del Consejo Científico de no menos de tres elementos.

En cada número de la revista se publican cinco a ocho de los mejores artículos enviados.



Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

©RISTI 2023 <http://www.risti.xyz>

