



ISSN: 1646-9895

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

S e t e m b r o 2 5 • S e p t e m b e r 2 5



©RISTI 2025 <http://www.risti.xyz>

Nº 59

Edição / Edición

Nº 59, 09/2025

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, ISI Web of Knowledge, Latindex, ProQuest, QUALIS, SciELO, SCImago, 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, University of Lisbon, PT

Coordenadores da Edição / Coordinadores de la Edición

Álvaro Rocha, ISEG, Universidade de Lisboa, PT

Conselho Editorial / Consejo Editorial

Abel Méndez, Technological Institute of Costa Rica, CR

Alejandro Peña, EAFIT University, CO

Alma Gomez-Rodríguez, University of Vigo, ES

Ana Rita Calvao, ESTGA, University of Aveiro, PT

Ania Cravero, Universidad de La Frontera, CL

António Abreu Silva, ISCAP, Polytechnic Institute of Porto, PT

Antonio Garcia, University of Santiago de Compostela, ES

António Godinho, ISEC, Polytechnic Institute of Coimbra, PT

Antonio Jiménez-Martín, Polytechnic University of Madrid, ES

Arturo J. Méndez, University of Vigo, ES

August Climent, Ramon LLull University, ES

Beatriz Rodríguez, Universidad de la Republica, UY

Borja Bordel, Polytechnic University of Madrid, ES

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

Bruno Miguel Ferreira Gonçalves, Polytechnic Institute of Braganca, PT

Carlos Alexandre Silva, Federal Institute of Minas Gerais, BR

Carlos Carreto, Polytechnic Institute of Guarda, PT

Carlos Morais, Polytechnic Institute of Braganza, PT

Carlos Rompante Cunha, UNIAG & CeDRI &

Polytechnic Institute Bragança, PT

Carlos Vaz de Carvalho, Polytechnic Institute of Porto, PT

Célio Marques, Polytechnic Institute of Tomar, PT

Círo Martins, University of Algarve, PT

Concepción Damián-Hernández, Instituto

Tecnológico Superior de Misantla, MX

Cristina Caridade, Polytechnic Institute of Coimbra, PT

Daniel Poland, University of Aveiro, PT
Dante Carrizo, University of Atacama, CL
Edwin Cedeño Herrera, Universidad de Panamá, PA
Fabio Marques, ESTGA, University of Aveiro, PT
Felipe Vasquez, Universidad de La Frontera, CL
Fernando Moreira, REMIT, IJP, Portucalense University & IEETA,
University of Aveiro, PT
Fernando Paulo Belfo, Coimbra Business School – ISCAC, Polytechnic
Institute of Coimbra, PT
Fernando Ribeiro, Polytechnic Institute of Castelo Branco, PT
Fernando Bandeira, PT
Filipe Cardoso, Polytechnic Institute of Viseu, PT
Flor Gomes de María Sánchez Aguirre, Universidad César Vallejo, PE
Francisco Javier Lena-Acebo, University of Cantabria, ES
Gabriel Guerrero-Contreras, University of Cádiz, ES
Gerardo González Filgueira, University of A Coruña, ES
Gloria Valencia, Universidad de las Fuerzas Armadas, EC
Helder Gomes, University of Aveiro, PT
Helder Zagalo, University of Aveiro, PT
Hélia Guerra, University of the Azores, PT
Henrique Gil, Polytechnic Institute of Castelo Branco, PT
Henrique S. Mamede, Open University, PT
Isaias Bianchi, Federal University of Santa Catarina, BR
Ismael Etxeberria-Agiriano, Universidad del País Vasco, ES
Ivan Garcia, Universidad Tecnológica de la Mixteca, MX
Jeimy Cano, Universidad de los Andes, CO
John Emilio Almeida, ISTECH - Porto, PT
João Paulo Ferreira, Polytechnic Institute of Coimbra, PT
Roberto de Toledo Quadros, CEFET/RJ, BR
Joao Tavares, University of Porto, PT
Joaquim Reis, ISCTE - University Institute of Lisbon, PT
Jorge Eduardo Ibarra Esque, Universidad
Autónoma de Baja California, MX

Jorge Hochstetter, Universidad de La Frontera, CL
José Álvarez-García, University of Extremadura, ES
Jose Felipe Cocon Juarez, Universidad Autónoma del Carmen, MX
José Lousado, Polytechnic Institute of Viseu, PT
José Luis Pastrana Brincones, University of Malaga, ES
Jose M. Molina, Carlos III University of Madrid, ES
José Ribeiro, Polytechnic of Leiria, PT
Jose Silvestre Silva, Military Academy, PT
Juan Angel Contreras Vas, University of Extremadura, ES
Juan M. Santos-Gago, University of Vigo, ES
Juan Pablo DAmato, National University of the Center of the
Province of Buenos Aires, AR
Leila Weitzel, Fluminense Federal University, BR
Leonardo Bermón Angarita, National University of Colombia, CO
Lilia Muñoz, Universidad Tecnológica de Panamá, PA
Lucila Romero, Universidad Nacional del Litoral, AR
Luis Alvarez Sabucedo, University of Vigo, ES
Luis Enrique Sánchez Crespo, University of 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, National Technological University, AR
Marcelo Zambrano Vizuite, Universidad Tecnica 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, University of Vigo, ES
Maria de los Milagros Gutierrez, National Technological University, AR
Maria do Rosario Bernardo, Open University, PT
Maria João Ferreira, Portucalense University, PT
Maria João Gomes, University of Minho, PT
Maria Sousa, ISCTE - University Institute of Lisbon, PT
Marisol B. Correia, ESGHT - University of Algarve & CiTUR, PT

Maristela Holanda, University of Brasilia, BR
Martin Llamas Nistal, University of Vigo, ES
Miguel Casquilho, University of Lisbon, PT
Miguel Ramón González Castro, Aimen Technological Center, ES
Mirna Ariadna Muñoz Mata, CIMAT, MX
Nelson Rocha, University of Aveiro, PT
Nuno Melão, Polytechnic Institute of Viseu, PT
Nuno Ribeiro, Fernando Pessoa University, Portugal
Patricia Dias, State University of Minas Gerais, BR
Paula Prata, University of Beira Interior, PT
Paulo Martins, University of Trás-os-Montes and Alto Douro, PT
Paulo Pinto, New University of Lisbon, PT
Paulo Rurato, Fernando Pessoa University, Portugal
Pedro Araújo, University of Beira Interior, PT
Pedro R. Palos-Sanchez, Universidad de Sevilla, ES
Pedro Sanz-Angulo, University of Valladolid, ES
Pedro Sobral, Fernando Pessoa University, Portugal
Pedro Sousa, University of Minho, Portugal
René Faruk Garzozi-Pincay, Santa Elena Peninsula State University, EC
Ruben Pereira, ISCTE – University Institute of Lisbon, PT
Rui Pedro Marques, University of Aveiro, PT
Rui Silva Moreira, Fernando Pessoa University, PT
Samuel Sepúlveda, Universidad de La Frontera, CL
Sandro Carvalho, Polytechnic of Cávado and Ave, PT
Santiago Raul Gonzales Sanches, Universidad Cesar Vallejo, PE
Sara Balderas-Díaz, University of Cádiz, ES
Sergio Araya Guzmán, Universidad del Bío-Bío, CL
Sergio F. Lopes, University of Minho, PT
Solange N Alves-Souza, University of Sao Paulo, BR
Telmo Silva, DigiMedia, University of Aveiro, PT
Thiago Dias, Federal Center for Technological Education of Minas Gerais, BR
Veronica Vasconcelos, Polytechnic Institute of Coimbra, PT
Vítor Carvalho, Polytechnic Institute of Cávado and Ave, PT

Vitor Santos, New University of Lisbon - NOVA IMS, PT

Wagner Tanaka Botelho, Federal University of ABC, BR

Yamid Hernández Julio, Universidad del Sinú, CO

Índice / Index

EDITORIAL

Avanços Recentes em Sistemas e Tecnologias de Informação	1
<i>Álvaro Rocha</i>	

ARTIGOS / ARTICULOS / ARTICLES

Modelos de Aprendizaje Automático para Clasificar el Riesgo de Corrupción en Contratación Pública: Caso Hospitales Públicos en Colombia.....	3
<i>Heriberto Felizzola, Carlos Arango, Yilber Erazo, Geraldine Camacho</i>	

Análise de metodologias para classificação de deepfakes	21
<i>Guilherme Veiga Santos Pinto, Rudimar Luiz Scaranto Dazzi, Anita Maria da Rocha Fernandes</i>	

Clasificación de Frases en Zapoteco mediante CNN con Espectrogramas	36
<i>Mariano Martínez Patiño, Sergio Juárez Vázquez, Efraín Dueñas Reyes, Francisco Javier Sol Sampedro, Nicolas Hernández Ruiz</i>	

Revisión sistemática de juegos virtuales en conocimiento numérico y científico para 1º a 5º de primaria.....	53
<i>Brehiner Alfredo Martinez Mora, Marlenny Guevara Guerrero</i>	

Educación Musical Temprana y Tecnología: Una revisión sistemática	68
<i>Vicenta Gisbert-Caudeli, Rosa María Díaz-Mayo</i>	

Caracterização espectral da cobertura do solo do Parque Estadual do Cocó a partir de imagens CBERS-4A	83
<i>Mauricio Alejandro Perea-Ardila</i>	

Avanços Recentes em Sistemas e Tecnologias de Informação

Recent Advances in Information Systems and Technologies

Álvaro Rocha¹

amr@iseg.ulisboa.pt

¹ ISEG, Universidade de Lisboa, Rua Miguel Lupi 20, 1200-109 Lisboa, Portugal.

DOI: 10.17013/risti.59.1-2

É com grande satisfação que apresentamos o número 59 da Revista Ibérica de Sistemas e Tecnologias de Informação (RISTI), publicação científica que continua a consolidar-se como um fórum de referência para investigadores, docentes e profissionais das áreas de sistemas, tecnologias e engenharia da informação. Neste número regular, mantemos o compromisso com a qualidade científica e a diversidade temática, refletindo a vitalidade da investigação ibero-americana no domínio das tecnologias de informação aplicadas à sociedade, à educação, à gestão e à sustentabilidade.

A taxa de aceitação deste número foi de 9,4%, o que evidencia o rigor do processo de revisão por pares e o elevado nível de exigência científica dos trabalhos aqui publicados.

O primeiro artigo, “Modelos de Aprendizaje Automático para Clasificar el Riesgo de Corrupción en Contratación Pública: Caso Hospitales Públicos en Colombia” propõe o uso de modelos de aprendizagem automática para analisar e prever riscos de corrupção na contratação pública de hospitais colombianos entre 2014 e 2019. Utilizando o percentual de contratação direta como variável de risco, foram aplicados os algoritmos Árvores de Decisão, Random Forest e Gradient Boosting. Este último apresentou o melhor desempenho, com níveis de precisão entre 46% e 59% e área sob a curva ROC entre 0,56 e 0,72. O estudo mostra que fatores contextuais, como indicadores socioeconómicos e de saúde, também influenciam o risco de corrupção. Os resultados demonstram o potencial da ciência de dados para promover maior transparência, integridade e eficiência na gestão dos recursos públicos.

O segundo artigo, “Análise de metodologias para classificação de deepfakes”, apresenta um estudo aprofundado sobre a sofisticação crescente dos vídeos manipulados e os

desafios na sua detecção. Com base em diferentes arquiteturas de Deep Learning, os autores comparam o desempenho de modelos como LSTM, CNN, GRU e CvT, bem como a aplicação da Lei de Benford como apoio estatístico. Os resultados destacam a relevância de métodos híbridos para garantir maior robustez na detecção de vídeos sintéticos e combater a desinformação.

O terceiro artigo, “Clasificación de Frases en Zapoteco mediante CNN con Espectrogramas”, propõe uma abordagem inovadora para a preservação linguística de línguas indígenas mexicanas. Através de redes neurais convolucionais aplicadas a espectrogramas de áudio, os autores alcançam elevadas taxas de precisão na classificação de frases em zapoteco do Istmo, evidenciando o potencial das tecnologias de inteligência artificial para apoiar a comunicação intercultural e a revitalização de línguas em risco de extinção.

O quarto artigo, “Revisión sistemática de juegos virtuales en conocimiento numérico y científico para 1º a 5º de primaria”, analisa como os jogos digitais contribuem para o desenvolvimento cognitivo nas áreas de matemática e ciências. Através de uma revisão sistemática de publicações entre 2017 e 2022, os autores concluem que as TIC são amplamente integradas nas práticas pedagógicas, mas que ainda é necessário aprofundar o estudo sobre o impacto cognitivo real destas ferramentas nos primeiros anos da educação básica.

O quinto artigo, “Educación Musical Temprana y Tecnología: Una revisión sistemática”, investiga o papel das tecnologias digitais na educação musical infantil. Os resultados apontam benefícios na criatividade, motivação e aprendizagem autónoma das crianças entre 2 e 6 anos, embora a escassez de estudos empíricos limite a generalização dos resultados. O trabalho reforça a necessidade de um modelo híbrido que conjugue práticas musicais tradicionais e recursos tecnológicos.

Por fim, o sexto artigo, “Caracterização espectral da cobertura do solo do Parque Estadual do Cocó a partir de imagens CBERS-4A”, centra-se no uso de imagens de satélite para o monitoramento ambiental de ecossistemas urbanos no Brasil. O estudo demonstra a eficácia do sensor WPM do satélite CBERS-4A na identificação de diferentes tipos de cobertura vegetal, evidenciando o papel da teledeteção como ferramenta essencial para a gestão sustentável e a conservação da biodiversidade em áreas metropolitanas.

Em conjunto, os artigos deste número reforçam a natureza multidisciplinar da RISTI, unindo ciência da computação, inteligência artificial, educação, sustentabilidade e ciências sociais numa perspetiva integradora e inovadora.

Agradecemos a todos os autores pela qualidade dos seus contributos, aos revisores científicos pela dedicação e rigor académico, e ao Conselho Editorial pelo apoio contínuo na consolidação desta revista como um espaço de excelência científica e de partilha de conhecimento.

Modelos de Aprendizaje Automático para Clasificar el Riesgo de Corrupción en Contratación Pública: Caso Hospitales Públicos en Colombia

Heriberto Felizzola¹, Carlos Arango¹, Yilber Erazo¹, Geraldine Camacho¹

healfelizzola@unisalle.edu.co; cararango@unisalle.edu.co; yerazo01@unisalle.edu.co;
gcamacho72@unisalle.edu.co

¹ Universidad de La Salle, Facultad de Ingeniería, Bogotá D.C., Colombia.

DOI: 10.17013/risti.59.3-20

Resumen: La contratación pública es particularmente vulnerable a la corrupción, lo que plantea desafíos significativos para la gestión pública. A nivel mundial, los gobiernos han implementado iniciativas de datos abiertos con el objetivo de promover la transparencia y fortalecer la integridad en la gestión de los recursos públicos. Sin embargo, el éxito de estas iniciativas depende en gran medida de la capacidad para analizar los datos disponibles y detectar patrones que puedan señalar riesgos de corrupción. Este estudio presenta el desarrollo y la evaluación de modelos de aprendizaje automático orientados a analizar y predecir el riesgo de corrupción en procesos de contratación pública. Para ello, se utilizaron datos abiertos de contratación en hospitales públicos colombianos entre 2014 y 2019. Como variable de riesgo, se empleó el porcentaje de contratación directa, ampliamente reconocido como un indicador de posibles irregularidades. Se implementaron y compararon tres algoritmos de aprendizaje automático: Árboles de Decisión, Random Forest y Gradient Boosting. Los resultados evidenciaron un desempeño aceptable, con niveles de exactitud entre el 46% y el 59% y un área bajo la curva ROC que oscila entre el 0.56 y el 0.72.

Palabras-clave: Corrupción en Contratación Pública; Modelos de Aprendizaje Automático; Riesgo de Corrupción; Hospitales Públicos; Datos Abiertos.

Machine Learning Models for Classifying Corruption Risk in Public Procurement: The Case of Public Hospitals in Colombia

Abstract: Public procurement is particularly vulnerable to corruption, posing significant challenges to public administration. Globally, governments have implemented open data initiatives to promote transparency and strengthen resource management's integrity. However, the success of these initiatives largely depends on the ability to analyze available data and identify patterns that may indicate corruption risks. This study presents the development and evaluation of machine learning models designed to analyze and predict the risk of corruption in public procurement processes. For this purpose, open procurement data from Colombian public hospitals between 2014 and 2019 were used. The percentage of

direct contracting, widely recognized as an indicator of potential irregularities, was employed as the risk variable. Three machine learning algorithms were implemented and compared: Decision Trees, Random Forest, and Gradient Boosting. The results demonstrated acceptable performance, with accuracy levels ranging from 46% to 59% and an area under the ROC curve between 0.56 and 0.72.

Keywords: Corruption in Public Procurement; Machine Learning Models; Corruption Risk; Public Hospitals; Open Data.

1. Introducción

Un problema persistente en la contratación pública es la corrupción, cuyo impacto económico es significativo, afectando tanto la eficiencia como la transparencia en la gestión de recursos (Neupane et al., 2014; Soreide, 2002). Según datos de la OCDE, las compras públicas representan aproximadamente el 12% del PIB en los países miembros, y cerca del 57% de los sobornos en el sector público son destinados para la obtención de contratos (OCDE, 2015; OECD, 2019). Dada la relevancia económica de la contratación pública y con el objetivo de combatir la corrupción, gobiernos, organizaciones, académicos y la sociedad civil han implementado diversas estrategias para mejorar la transparencia, integridad y rendición de cuentas en la gestión pública.

En este contexto, las iniciativas de datos abiertos han ganado relevancia en los últimos años. Estas iniciativas buscan proporcionar información estructurada sobre la gestión pública mediante plataformas accesibles y fáciles de usar (Adam et al., 2020; Camargo & Pinzon, 2022). Aunque los datos abiertos son esenciales para fomentar la transparencia, no son suficientes por sí solos; su verdadero impacto depende de su aprovechamiento en procesos analíticos que permitan extraer información clave para la toma de decisiones, la formulación de políticas públicas y el monitoreo de los procesos de contratación (Ansari et al., 2022; Duguay et al., 2019; Janssen et al., 2012).

La analítica aplicada a la contratación pública puede abordar problemas críticos como el análisis, la estimación y la detección de riesgos de corrupción (World Bank, 2022). Esto no solo permite identificar riesgos y los factores asociados, sino también desarrollar herramientas basadas en ciencia de datos capaces de procesar grandes volúmenes de información para automatizar tareas de control y vigilancia (Colonnelli et al., 2020). Estas herramientas ofrecen a los organismos de control, entidades públicas y la sociedad civil la capacidad de realizar una supervisión proactiva y efectiva, fortaleciendo así la integridad en los procesos de contratación (Gallego et al., 2020; Rabuzin & Modrušan, 2019).

El aprendizaje automático se presenta como una herramienta clave en la ciencia de datos, capaz de identificar patrones en los datos y realizar predicciones sobre características específicas (Domingos et al., 2016; Fazekas & Czibik, 2021; Fazekas & Dávid-Barrett, 2015). En el análisis de riesgos de corrupción en contratación pública, las técnicas de aprendizaje automático ofrecen un enfoque prometedor para detectar de manera automática posibles irregularidades. No obstante, el desarrollo de estos modelos enfrenta desafíos importantes. En primer lugar, la corrupción es un fenómeno complejo, difícil de medir y de detectar, lo que exige el diseño de mecanismos efectivos para identificar riesgos de manera confiable (Fazekas et al., 2013, 2016). En segundo lugar, la

contratación pública es un proceso multifacético que involucra múltiples etapas, actores y factores del entorno, lo que requiere la incorporación de un enfoque holístico en los modelos predictivos (Broms et al., 2019).

Este trabajo aborda las limitaciones existentes mediante el desarrollo de modelos de aprendizaje automático que predicen el riesgo de corrupción a partir de una de las señales de alerta más comunes: la contratación directa. Los modelos propuestos incorporan, además, predictores que trascienden las características intrínsecas de los procesos de contratación, integrando indicadores socioeconómicos y de contratación pública asociados a las regiones y localidades donde se llevan a cabo los procesos. Esto permite analizar factores exógenos que influyen en los riesgos de corrupción. Como resultado, se busca desarrollar un modelo predictivo capaz de caracterizar y clasificar a las entidades hospitalarias públicas en Colombia según su nivel de riesgo de corrupción, utilizando la contratación directa como indicador principal. Este enfoque propone una metodología para evaluar la falta de competencia en procesos de contratación pública mediante datos abiertos, apoyándose en estadísticas, indicadores de contratación y modelos de aprendizaje automático.

El artículo se organiza de la siguiente manera: la Sección 2 presenta una revisión del estado del arte sobre modelos de aprendizaje automático aplicados a la predicción de riesgos de corrupción en contratación pública. La Sección 3 describe los datos utilizados, su procesamiento y la definición de las variables. La Sección 4 detalla los modelos y técnicas de aprendizaje automático implementados, los escenarios de análisis y las métricas empleadas. La Sección 5 expone los resultados y el análisis de los modelos desarrollados. Finalmente, la Sección 6 presenta las conclusiones y posibles implicaciones de los hallazgos.

2. Revisión de Literatura

2.1. Medición del Riesgo de Corrupción en Contratación Pública

Los índices de corrupción existentes, como el Índice de Percepción de la Corrupción de Transparencia Internacional, ofrecen una evaluación general útil de los niveles de corrupción. Sin embargo, carecen del nivel de detalle necesario para abordar de manera efectiva las intervenciones anticorrupción en procesos específicos de contratación pública. Para llenar esta brecha, investigadores han explorado el uso de modelos de aprendizaje automático para predecir el riesgo de corrupción a un nivel más granular, considerando factores como el monopolio del poder, la asimetría de información, la transparencia y la rendición de cuentas en los procesos de contratación. Estos modelos pueden ayudar a responsables políticos y funcionarios de contratación a identificar contratos de alto riesgo y a implementar medidas específicas para mejorar la transparencia y la integridad en la gestión pública (Gallego et al., 2021; Modrušan et al., 2021).

El desafío de medir y detectar la corrupción en la contratación pública es de vital importancia, dado su impacto en la asignación eficiente de recursos, el crecimiento económico y la confianza pública. Como señalan (Fazekas & Kocsis, 2020), existe una notable carencia de indicadores confiables y accionables, especialmente para la corrupción de alto nivel.

Diversos autores han propuesto indicadores innovadores para evaluar el riesgo de corrupción en la contratación pública. Fazekas & Kocsis (2020) desarrollaron dos enfoques principales: la detección de licitaciones únicas en mercados que deberían ser competitivos y una puntuación compuesta de “señales de alerta” denominada Índice de Riesgo de Corrupción (CRI). Decarolis & Giorgiantonio (2022) validaron nuevos indicadores basados en el diseño de licitaciones, utilizando técnicas de aprendizaje automático para evaluar su capacidad predictiva. Por su parte, Fazekas et al. (2023) calcularon indicadores individuales de riesgo de corrupción y diseñaron un Índice de Riesgo de Corrupción (IRC) compuesto para Bulgaria.

Estos indicadores compuestos, como el CRI, combinan diferentes señales de alerta para medir de manera más precisa los riesgos de corrupción en la contratación pública. En conclusión, el desarrollo de indicadores objetivos y accionables para evaluar el riesgo de corrupción ha sido un campo en constante evolución. Estas herramientas no solo fortalecen la capacidad de los responsables de políticas públicas y los investigadores para combatir la corrupción, sino que también establecen un marco más sólido para la rendición de cuentas en el gasto público. La transición desde enfoques estadísticos tradicionales hacia técnicas avanzadas de aprendizaje automático, como las propuestas de Decarolis & Giorgiantonio (2022), señala un camino prometedor para futuras investigaciones en este ámbito.

2.2. Modelos de Aprendizaje Automático para la Predicción de Riesgos de Corrupción

Investigaciones previas han demostrado que los modelos de aprendizaje automático son herramientas efectivas para identificar patrones y anomalías en los datos de contratación pública que podrían señalar posibles casos de corrupción (Modrušan et al., 2021; Nai et al., 2022). Estos modelos analizan factores como la distribución de precios en las ofertas, la frecuencia con la que ciertos proveedores reciben adjudicaciones, y el cumplimiento de los objetos y plazos contractuales. Además, pueden detectar señales de alerta, como la falta de transparencia en los procesos de licitación, adjudicaciones a proveedores con vínculos políticos y contrataciones aceleradas o con escaso escrutinio público. La implementación de estos modelos permite identificar contratos de alto riesgo y tomar medidas preventivas y de detección temprana para combatir la corrupción en la contratación pública (Aldana et al., 2022).

Un estudio basado en más de 2 millones de contratos públicos en Colombia exploró el uso de modelos de aprendizaje automático para detectar corrupción, logrando precisiones superiores al 80% en algunos casos (Gallego et al., 2020). De manera similar, investigaciones en México demostraron que los modelos de conjunto pueden clasificar eficazmente contratos corruptos y no corruptos (Aldana et al., 2022). Estas investigaciones resaltan la importancia de la interpretabilidad de los modelos, un factor clave para su aplicación en el sector público. Sin embargo, su implementación práctica enfrenta desafíos significativos, como el desbalanceo de clases, donde los contratos conformes superan ampliamente a los problemáticos, lo que afecta la precisión de los modelos predictivos. Además, alinear estos modelos con las complejas dinámicas de la toma de decisiones públicas, que incluyen objetivos en conflicto y la necesidad de intervención humana, resulta complicado. Para mantener su efectividad, los modelos

deben ser adaptativos, capaces de actualizarse y evolucionar frente a nuevas estrategias de corrupción (Gallego et al., 2021). Este enfoque flexible es esencial para anticipar y mitigar las amenazas emergentes en un entorno dinámico y en constante cambio.

Diversos estudios han empleado la metodología CRISP-DM, ampliamente utilizada en proyectos de minería de datos y aprendizaje automático, que abarca las etapas de recopilación, preprocesamiento, modelado, evaluación e implementación (Torres-Berru et al., 2023; Torres-Berru & López Batista, 2021). Entre las actividades clave destacaron la ingeniería de características, el manejo del desbalanceo de clases y el refinamiento iterativo de modelos. Las investigaciones identificaron variables como el número de oferentes en las licitaciones, la diferencia entre la oferta ganadora y la segunda más alta, y la relación comprador-proveedor, como predictores relevantes de corrupción. Además, los modelos detectaron anomalías en los parámetros de calificación de adquisiciones públicas y revelaron un clúster denominado “Oferta Económica Nula”, caracterizado por un peso insignificante de la oferta económica, lo que podría sugerir manipulaciones (Torres-Berru & López Batista, 2021). Estos hallazgos resaltan el potencial de las técnicas de minería de datos para promover la transparencia y rendición de cuentas en la contratación pública. Asimismo, subrayan la necesidad de seguir investigando para optimizar los modelos y adaptarlos a contextos específicos, ampliando su aplicabilidad y efectividad en distintos escenarios.

Otros estudios han explorado el uso de clasificadores Naïve Bayes para evaluar riesgos en adquisiciones públicas, enfocándose en factores asociados con prácticas de colusión y corrupción. Este enfoque comienza definiendo el alcance del análisis, identificando tipos de contratos o agencias gubernamentales específicas, y seleccionando factores de riesgo como el tipo de licitación, renegociaciones posteriores y vínculos entre licitadores y funcionarios públicos (Balaniuk et al., 2013). Los factores se ponderan según su importancia para identificar casos de alto riesgo. Este método ha demostrado ser práctico y eficaz, logrando alinearse con evaluaciones de auditores expertos. Además, su capacidad para integrar datos de diversas fuentes con opiniones de expertos permite realizar evaluaciones más objetivas y completas, optimizando la asignación de recursos para auditorías y promoviendo la transparencia en los procesos de contratación pública.

En un trabajo desarrollado por Sales & Carvalho (2016) se utiliza Naïve Bayes para entrenar un modelo con indicadores de riesgo como variables de entrada y clasificaciones de contratos como “Bueno” o “Malo” como objetivo. El modelo mostró efectividad al calcular la probabilidad de que un contrato pertenezca al grupo “Malo”, priorizando así los contratos con mayor riesgo para auditorías. Este enfoque destacó por su precisión y capacidad para cuantificar riesgos en adquisiciones públicas.

García Rodríguez et al. (2022) concluyen que los algoritmos de aprendizaje automático, especialmente los métodos de conjunto como Extra Trees y Gradient Boosting, ofrecen resultados prometedores para detectar colusión en licitaciones de compras públicas. Estos modelos mostraron una alta precisión en la identificación de comportamientos colusorios, superando a otros algoritmos como las Máquinas de Vectores de Soporte y Random Forest. El estudio destaca la importancia de la ingeniería de variables y el preprocesamiento de datos como factores clave para optimizar el rendimiento de los modelos. Además, los autores señalan que las técnicas de aprendizaje automático son herramientas valiosas para las autoridades de competencia y las agencias de

vigilancia de compras públicas, ya que facilitan la detección y prevención de prácticas anticompetitivas, promueven la competencia justa y garantizan una gestión eficiente de los recursos públicos.

Por su parte, Domingos et al. (2016) destacan que los clasificadores bayesianos ofrecen un enfoque prometedor para medir el riesgo asociado a los contratos públicos. Su modelo, diseñado con base en diversos indicadores de riesgo, demuestra una precisión significativa al predecir la probabilidad de problemas con contratos específicos.

En conjunto, estos estudios subrayan el potencial de las técnicas de aprendizaje automático para fortalecer la transparencia y la rendición de cuentas en las adquisiciones públicas. Además, estas herramientas permiten detectar patrones asociados a la corrupción, contribuyendo a la mejora de los sistemas de compra y al uso responsable de los dineros públicos.

2.3. Factores Exógenos Asociados a los Riesgos de Corrupción en Contratación Pública

Las investigaciones han identificado una amplia gama de factores económicos y políticos que contribuyen a la prevalencia de la corrupción en la contratación pública. Por ejemplo, un estudio en Uganda destacó que los bajos salarios en el sector público, la falta de rendición de cuentas y la interferencia política en los procesos de contratación son predictores significativos de corrupción (Basheka, 2011). De manera similar, otros estudios han señalado que marcos institucionales débiles, la ausencia de transparencia y la influencia de grupos de interés poderosos agravan este problema (Soreide, 2002; Williams-Elegbe, 2018). La falta de mecanismos efectivos de supervisión y rendición de cuentas, combinada con una ejecución deficiente, crea un entorno propicio para prácticas corruptas.

Estos hallazgos revelan que la corrupción en la contratación pública no solo depende de cómo se diseñan, planifican, evalúan y ejecutan los contratos, sino también de factores externos que son determinantes para la eficiencia, integridad y transparencia de los procesos de contratación. Analizar y abordar estos elementos estructurales es esencial para mitigar los riesgos de corrupción y mejorar la gobernanza en este ámbito.

3. Metodología

3.1. Datos y Fuentes de Información

Para desarrollar los modelos predictivos, se utilizaron datos de contratación pública hospitalaria en Colombia disponibles en Sistema Electrónico de Contratación Pública (SECOPI), datos territoriales del Departamento Nacional de Planeación (Terridata) y estadísticas de salud proporcionadas por la Superintendencia de Salud (Supersalud).

Los datos extraídos de SECOPI corresponden al período 2014-2019. El conjunto inicial contiene 1.391.442 contratos y 44 variables, representando un valor total de contratos (incluyendo adiciones) de \$51.474 mil millones de pesos Colombianos (COP). Cada observación correspondía a un contrato entre una entidad hospitalaria gubernamental y un proveedor de bienes o servicios. Para el análisis, se seleccionaron 15 variables clave

relacionadas con la Identificación y ubicación de los contratistas y entidades; el tipo de contrato y tipo de proceso; los valores iniciales, adiciones y total de los contratos; las fechas de firma y registro; las adiciones en tiempo y enlaces URL de los contratos.

La preparación de los datos para los modelos incluyó una fase de depuración para garantizar su calidad y confiabilidad. En esta etapa se aplicaron filtros y se llevaron a cabo tareas de limpieza, entre las que destaca la corrección manual de registros que presentaban inconsistencias. Estas se encontraron principalmente en el valor del contrato, las adiciones y la identificación de los contratistas, y fueron subsanadas mediante la revisión y validación entre la base de datos abierta y los documentos contractuales y sus informes de cierre. En los filtros se eliminaron registros con contratos inferiores a COP \$1.000.000, valores faltantes en variables clave, convenios interadministrativos, transacciones bancarias, URLs inconsistentes y entidades con menos de tres contratos. Además, se revisaron manualmente contratos con cuantías y adiciones superiores a COP \$10 mil millones, detectándose errores en el 66,9% de los casos. Esto permitió corregir 90 contratos con errores en cuantías y 19 en adiciones, logrando una reducción del 32,42% en el valor total de los contratos, como se detalla en la Tabla 1, evidenciando el impacto positivo del procesamiento de datos en su calidad.

Del DNP, se emplearon datos extraídos del aplicativo Terridata. Este sistema gestiona y publica la información de indicadores agrupados en subcategorías y dimensiones para departamentos y municipios en Colombia. Para este estudio, se seleccionaron indicadores municipales y departamentales de cuatro dimensiones: Economía, Educación, Finanzas Públicas y Salud, correspondientes al período 2014-2019. Se priorizaron indicadores con menos del 40% de datos faltantes. Además, se incorporaron indicadores de la Medición del Desempeño Municipal (MDM), que evalúa el cumplimiento de metas de los planes de desarrollo, la provisión de servicios básicos, la ejecución presupuestal y la gestión administrativa y fiscal (Departamento Nacional de Planeación, 2020). De la Superintendencia de Salud (Supersalud), se utilizaron dos indicadores del Reporte de Circular Única, que proporciona información relevante sobre la gestión hospitalaria.

<i>Atributo</i>	<i>Valor Inicial (MM*)</i>	<i>Valor Final (MM*)</i>	<i>Reducción</i>
<i>Valor del Contrato</i>	\$ 42.580	\$ 28.719	32,55%
<i>Valor Adiciones</i>	\$ 4.644	\$ 3.194	31,22%
<i>Valor total con Adiciones</i>	\$ 47.225	\$ 31.913	32,42%

* Los datos se presentan en miles de millones de pesos colombianos.

Tabla 1 – Resumen del Ajuste en Valor de los Contratos

3.2. Variables e Indicadores de Contratación Pública

Tras la obtención de los datos de contratación, se consolidó una tabla maestra para el desarrollo de los modelos predictivos. En esta tabla, cada fila representa un hospital público, mientras que las columnas incluyen características específicas del hospital, estadísticas e indicadores de contratación, así como los indicadores proporcionados por el DNP (municipales y departamentales) y la Supersalud. Esta tabla contiene datos agregados por hospital y por año. Este nivel de granularidad de los datos se utilizó para

el cálculo de los indicadores de contratación y los indicadores tomados del DNP y la Supersalud.

Para representar las prácticas y características de la contratación en los hospitales, se emplearon indicadores basados en estudios sobre riesgos de corrupción en compras públicas realizados en Colombia (Zuleta et al., 2019), Chile (Jorquera, 2019), y México (Martínez & Torres, 2019). Cabe resaltar que todos los indicadores se normalizaron en una escala de 0-100. Los indicadores calculados son los siguientes:

Índice Herfindahl-Hirschman (HHI): Mide el nivel de concentración en la contratación pública de cada hospital, identificando posibles monopolios. Se calcula como:

$$HHI = \sum_{i=1}^C S_i^2 \quad (1)$$

Donde S_i es la participación porcentual (en cantidad y valor total) de cada contratista en la contratación del hospital.

Índice de diversidad de Simpson (ID): Originalmente diseñado para medir biodiversidad, en este contexto mide la diversidad de contratistas en un hospital. Su fórmula es:

$$ID = \frac{\sum_{i=1}^C q_i (q_i - 1)}{Q(Q - 1)} \quad (2)$$

Donde q_i es el volumen de contratación de cada contratista y Q es el volumen total de contratación de cada hospital.

Número de empresas ganadoras diferentes por cada 100 contratos: Mide el número de empresas diferentes adjudicatarias de los procesos de contratación. Dado que cada entidad tiene un volumen diferente de contratación, el índice se expresa en número de contratistas por cada 100 contratos.

Índice de concentración de las cuatro empresas con mayor número y valor de contratos - IC4k: Mide la proporción de procesos que el hospital adjudicó a los cuatro contratistas con más contratos. Un alto porcentaje, indica que la gran parte de la contratación del hospital está concentrada en pocos contratistas. Se calcula como:

$$IC4k = \sum_{c \in K} S_c \quad (3)$$

Donde S_c es la participación porcentual de cada contratista entre los cuatro principales.

Nivel de riesgo de corrupción: Para definir el nivel de riesgo de corrupción de un hospital público se utilizó el indicador de porcentaje de contratación directa de la entidad en función del número de contratos. La contratación directa es una modalidad de adquisición pública en la que una entidad estatal adjudica un contrato sin un proceso de

licitación abierta, generalmente por razones específicas como emergencias, exclusividad o conveniencia administrativa. Aunque este mecanismo está legalmente permitido en situaciones excepcionales, su uso excesivo o indebido representa un riesgo significativo de corrupción, ya que compromete principios fundamentales de la buena contratación, como la transparencia y la competencia justa. Este indicador se definió en tres niveles:

- Bajo o Moderado ($\leq 25\%$).
- Medio ($\leq 55\%$).
- Alto ($> 55\%$).

En la tabla maestra se eliminaron variables con un porcentaje de datos faltantes igual o superior al 40%. Para las variables numéricas restantes, los valores faltantes fueron imputados utilizando la mediana de cada atributo. El resultado final fue una tabla maestra con 892 observaciones (hospitales) y 345 atributos, de los cuales uno corresponde a la variable objetivo (Nivel de Riesgo de Corrupción) y 344 son predictores. Entre las variables predictoras, 342 son numéricas y 2 categóricas.

El nombre de todas las variables predictoras está formado de la misma manera: primero con la categoría a la que pertenece (indba = variable calculada a partir de SECOP, ine = educación, inec = economía, inf = finanzas públicas, inmdm = medición del desempeño municipal, ins = salud, y supers = Supersalud); segundo, la codificación del nombre del indicador; después el nivel de la variable (dto = departamento, mun = municipio y hos = hospital) y por último el año al que pertenece (si se establecía por años). En la Tabla 2 se presenta el listado de variables por categoría y una descripción general.

<i>Grupo de Variables en Indicadores</i>	<i>Indicadores</i>
2 variables categóricas de la base de datos de SECOP	Orden Entidad Departamento de la Entidad
79 estadísticas de total, porcentajes y de medidas de tendencia central, que indican patrones de comportamiento de los hospitales.	• Cantidad total de contratos (general y por años). • Valor inicial de los contratos • Valor (Cuantía) total de los contratos (general, mediana, media, desviación estándar y por años) • Valor (Cuantía) total con adiciones (general y por años) • Porcentaje de las adiciones sobre el valor inicial del contrato (general y por años) • Porcentaje de los contratos con adiciones en tiempo (general y por años) • Porcentaje de contratación directa en número. • Porcentaje de contratos por tipo de contrato. • Cantidad de contratos con alguna modificación. • Porcentaje de la cantidad y el valor de contratos por mínima cuantía por hospital y departamento. • Cantidad de contratos con adiciones • Cantidad de contratos a fin de año (por años) • Porcentaje de contratos del total de contratos del año (por años) • Cantidad de contratistas por entidad (por años) • Diferencia entre las fechas de cargue y firma (por años)

Grupo de Variables en Indicadores	Indicadores
18 indicadores de concentración de contratos e identificación de riesgos de corrupción.	• Índice Herfindahl-Hirschman para el número o valor de contratos de la entidad estatal (IHH) • El índice de diversidad de Simpson por número o valor de contratos (ID) • Índice de concentración de las cuatro empresas con mayor número y valor de adjudicaciones. (IC4K) • Cantidad de contratos por contratista por entidad (por años) • Valor de los contratos por contratista por entidad (por años) • Repitencia del mayor contratista en cantidad y proporción • Número de empresas ganadoras diferentes por cada 100 contratos • Índice de Riesgo de Corrupción
243 indicadores de desempeño municipal del aplicativo Terridata del DNP.	• Indicadores de la dimensión de Educación • Indicadores de la dimensión de Economía • Indicadores de la dimensión de Finanzas Públicas • Indicadores de la dimensión de Salud • Indicadores de Medición del Desempeño Municipal
2 Indicadores de SuperSalud	• Indicadores de Reporte de Circular Única

Tabla 2 – Listado de Variables e Indicadores de la Tabla Maestra

3.3. Modelos y Escenarios

3.3.1. Modelos de Machine Learning

Para predecir el nivel de riesgo de las entidades por contratación directa, se emplearon tres modelos de clasificación multiclase: Árboles de Decisión, Random Forest y Gradient Boosting (Extreme Gradient Boosting). Estos modelos se entrenaron utilizando la librería tidymodels e el lenguaje estadístico R. El desarrollo de todos los modelos incluyo las siguientes tareas de preprocesamiento: división de datos en conjuntos de entrenamiento (75%) y evaluación (25%), normalización de variables numéricas, creación de la categoría “otros” para los departamentos con menos del 5% del total de datos, generación de variables dummy para las variables categóricas y balanceo de clases.

Dado que la variable objetivo, nivel de riesgo de corrupción, presentaba una distribución desigual (540 observaciones en “bajo o moderado”, 83 en “medio” y 269 en “alto”), se aplicó balanceo de clases mediante el método SMOTE. Este procedimiento genera nuevas instancias sintéticas para las clases minoritarias utilizando los vecinos más cercanos de cada caso, mejorando así la representatividad de las clases en el conjunto de entrenamiento.

Se utilizó el método Boruta, basado en el algoritmo Random Forest, como estrategia para seleccionar el mejor conjunto de variables. Este enfoque genera en cada iteración variables “sombra” a partir de los predictores originales, calcula su importancia relativa y las compara con las variables sintéticas. Las variables con menor importancia que las sombras son rechazadas. Este proceso continúa hasta que las variables son clasificadas como importantes, no importantes o tentativas, con un límite de 500 iteraciones para este estudio debido a su alto costo computacional. Para el análisis se consideraron las variables importantes y tentativas.

3.3.2. Escenarios

Se definieron cuatro escenarios para evaluar el desempeño de los modelos en diferentes configuraciones de la tabla maestra:

- *Escenario 1. Solo variables de SECOP (indba):* Considera exclusivamente las variables internas relacionadas con contratación pública para determinar si son suficientes para predecir el nivel de riesgo.
- *Escenario 2. Variables de todas las fuentes:* Incluye indicadores de SECOP, Terridata, MDM y Supersalud, con el objetivo de evaluar la influencia de factores externos como los indicadores departamentales y municipales.
- *Escenario 3. Selección de variables en todas las fuentes:* Aplica técnicas de selección de variables a los datos combinados para eliminar aquellas que no aportan información relevante al modelo.
- *Escenario 4. Selección de variables sin MDM ni Supersalud:* Similar al escenario 3, pero excluyendo variables de estas fuentes, ya que no fueron seleccionadas como relevantes durante el proceso de reducción de dimensionalidad.

4. Resultados

4.1. Comparación de Modelos y Escenarios

Los resultados del desempeño de los modelos en los diferentes escenarios se presentan en la Tabla 3. Para la evaluación, se utilizaron las métricas de exactitud y el área bajo la curva ROC (ROC AUC). La exactitud se define como el porcentaje de predicciones que el modelo clasificó correctamente sobre el total de casos; se calcula dividiendo el número de predicciones correctas entre el número total de predicciones. Por su parte, el ROC AUC mide la capacidad del modelo para discriminar correctamente entre las clases; en este contexto, el nivel de riesgo de corrupción de la entidad (bajo, medio y alto). Esta métrica se obtiene de la curva ROC, que representa la relación entre la tasa de verdaderos positivos (sensibilidad) y la tasa de falsos positivos a través de diferentes umbrales de decisión. Un valor de AUC cercano a 1.0 indica una capacidad de discriminación perfecta, mientras que un valor de 0.5 sugiere un desempeño equivalente al azar.

Los resultados muestran que el modelo Random Forest y el Gradient Boosting presentan el mejor desempeño en términos de exactitud, mientras que el modelo Gradient Boosting sobresale en la métrica del área bajo la curva ROC (ROC AUC), obteniendo consistentemente los valores más altos en todos los escenarios.

En particular, los mejores puntajes se obtuvieron con Gradient Boosting, destacando que, independientemente de la métrica utilizada, este modelo sería el elegido, aunque en escenarios distintos: el escenario 2 para la exactitud y el escenario 3 para ROC AUC. Dentro de cada modelo, los escenarios 2 y 4 fueron los más destacados para la exactitud, mientras que los escenarios 3 y 4 sobresalieron en ROC AUC. Es importante resaltar que el escenario 1 no logró destacarse en ninguna métrica. En promedio, el escenario con mejor rendimiento en ROC AUC fue el escenario 3, mientras que para la exactitud fue el escenario 2. Además, los valores de la curva ROC superaron el 60% en la mayoría de los casos, evidenciando el poder predictivo de los algoritmos.

Modelo	Escenario	Exactitud	ROC AUC
Árbol de decisión	1	0,4642	0,5612
	2	0,5491	0,5991
	3	0,5089	0,6706
	4	0,4196	0,6030
Random Forest	1	0,5720	0,6044
	2	0,5855	0,6291
	3	0,5540	0,6336
	4	0,5945	0,6420
Gradient Boosting	1	0,5270	0,7059
	2	0,5990	0,6622
	3	0,5765	0,7202
	4	0,4909	0,7045

Tabla 3 – Desempeño de los modelos de acuerdo con cada escenario

4.2.Análisis del Modelo con Mejor Desempeño

El modelo Gradient Boosting aplicado al escenario 3 mostró el mejor desempeño global en términos de ROC AUC, como se ilustra en la Tabla 3. Esto resalta la importancia de utilizar todas las variables disponibles en la tabla maestra, aplicando algoritmos de selección como Boruta para reducir dimensionalidad y optimizar el modelo. Los resultados indican que no es suficiente trabajar únicamente con las variables de SECOP I; es fundamental incluir indicadores externos de fuentes como Terridata y Supersalud, además de aplicar técnicas de reducción de dimensionalidad para mejorar el rendimiento del modelo.

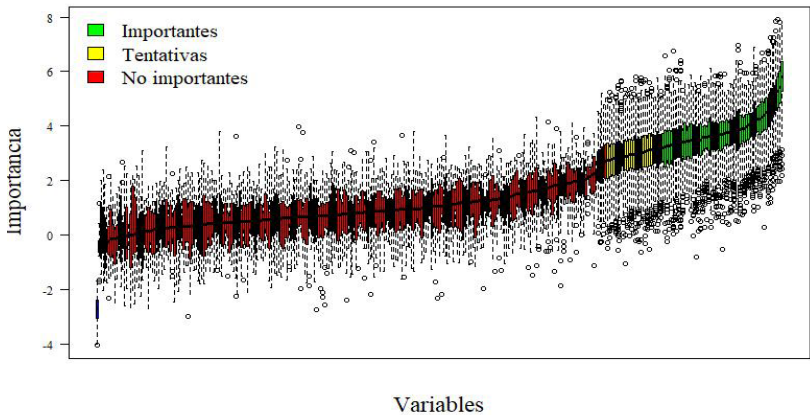


Figura 1 – Clasificación de variables según Boruta.

En la Figura 1 se muestra la clasificación de variables según el algoritmo Boruta, el cual identificó 81 de las 344 variables predictoras como importantes o tentativas. Este análisis reveló que muchas variables irrelevantes o “ruido” afectaban negativamente el rendimiento del modelo, lo que subraya la necesidad de una selección rigurosa. Es relevante señalar que ninguna variable de las fuentes MDM y Supersalud fue seleccionada como importante.

La Figura 2 presenta la curva ROC para cada categoría de la variable objetivo, Riesgo, mostrando que la categoría Medio fue la mejor clasificada. Todas las curvas ROC superan la diagonal, lo que confirma que el modelo clasifica mejor que la aleatoriedad, logrando una sensibilidad y especificidad superiores al 50% para cada clase.

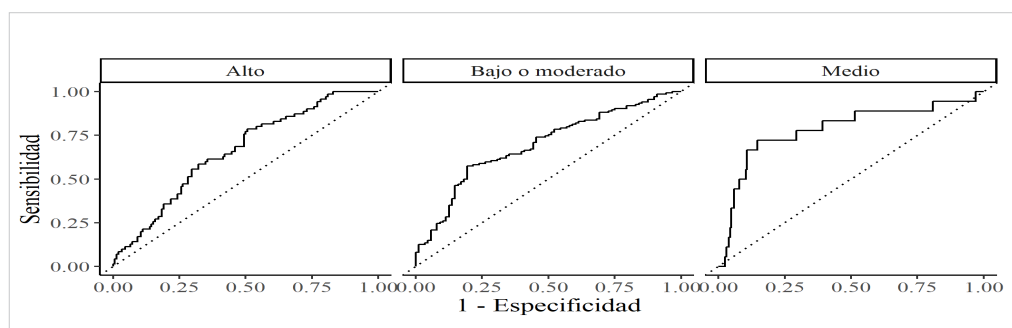


Figura 2 – Desempeño en curva ROC.

En la Figura 3 se destacan las 15 variables más importantes del modelo. Las primeras posiciones corresponden a variables internas de SECOP, seguidas por indicadores de educación, economía y salud de Terridata, excluyendo las dimensiones de finanzas públicas. Entre las variables más relevantes se incluyen los datos de pruebas de matemáticas (2015 y 2016), la cobertura de vacunación y la tasa de mortalidad neonatal. Además, el análisis muestra que las estadísticas relacionadas con la contratación al final del año y el tiempo entre la fecha de cargue y firma de los contratos son determinantes para predecir el nivel de riesgo de corrupción.

Es interesante destacar que variables aparentemente poco relacionadas, como la cobertura de vacunación o la mortalidad neonatal, resultaron altamente predictivas, mientras que algunos indicadores esperados de riesgo de corrupción fueron descartados. Este hallazgo sugiere que variables contextuales, como las relacionadas con la prestación de servicios, también influyen significativamente en los niveles de riesgo.

Para complementar el análisis de las variables importantes, se llevó a cabo un análisis de clúster utilizando las técnicas de PCA y UMAP, lo que permitió identificar agrupaciones de entidades según los indicadores de contratación. Como se observa en la Figura 4, las características de contratación varían significativamente entre departamentos. Departamentos como Antioquia, Valle del Cauca y Bogotá presentan comportamientos claramente diferenciados, separándose del resto del país en ambas técnicas. De manera similar, los departamentos de Santander y Boyacá tienden a formar agrupaciones

propias. Estos patrones indican que los modelos predictivos podrían beneficiarse de una segmentación por grupos de departamentos en lugar de una aplicación a nivel nacional, permitiendo incorporar las particularidades regionales y mejorar la precisión del análisis.

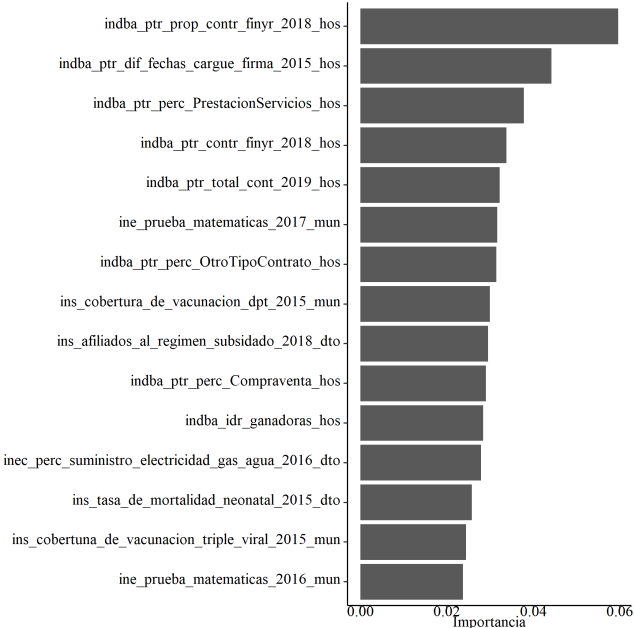


Figura 3 – Importancia de las Variables

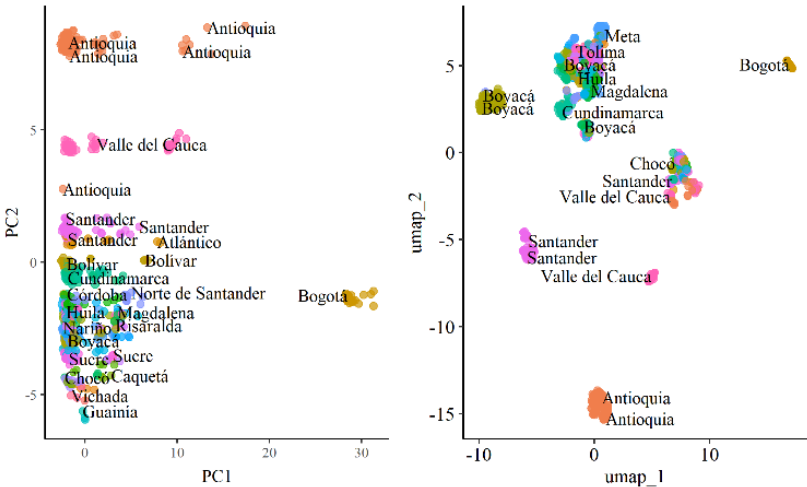


Figura 4 – Análisis de Clúster.

5. Conclusiones

Los modelos desarrollados permiten concluir que existe una relación significativa entre ciertos atributos y el nivel de riesgo de corrupción en la contratación pública. Variables como los contratos realizados al final del año, las diferencias en las fechas, el tipo de contrato y el número de contratistas, junto con características como el año y el nivel de la información, resultaron ser determinantes en la clasificación de las entidades hospitalarias. Además, se evidenciaron patrones de contratación específicos por departamento, lo que sugiere que los modelos predictivos podrían ser más efectivos al aplicarse de forma segmentada por regiones en lugar de a nivel nacional. Estos hallazgos destacan el potencial de los modelos como herramientas útiles para fomentar principios de buena contratación, apoyando la transparencia, la competencia justa y la rendición de cuentas en la contratación pública.

Las variables relacionadas con el entorno también contribuyen al desempeño de los modelos, siempre que se utilicen algoritmos de selección de variables como Boruta, ya que no solo mejoran la capacidad predictiva, sino que también reducen la complejidad del modelo mediante la disminución de su dimensionalidad. Este enfoque resulta esencial para optimizar los resultados y garantizar la eficiencia en el uso de datos heterogéneos.

En el análisis comparativo de los modelos de aprendizaje automático, el Gradient Boosting demostró ser el más efectivo, especialmente en la métrica ROC AUC, superando a los demás algoritmos evaluados. La implementación de estas técnicas de analítica representa un avance significativo en el desarrollo de herramientas destinadas a fortalecer los principios de la contratación pública eficiente y transparente.

Como trabajo futuro, sería valioso incluir información de otras fuentes tales como datos de entidades de vigilancia y control, como la Procuraduría y la Contraloría, para incorporar antecedentes sobre investigaciones relacionadas con entidades y contratistas. Asimismo, se recomienda desarrollar modelos predictivos segmentados por grupos de departamentos para evaluar su impacto en el desempeño, ya que los patrones regionales detectados en este estudio podrían mejorar la precisión y efectividad de los modelos.

Referencias

- Adam, I., Fazekas, M., & Tóth, B. (2020). Measuring the benefits of open contracting. Government Transparency Institute; http://redflags.govtransparency.eu/wp-content/uploads/2020/01/GTI-WP_OC-benefits-research_final-report_20200121.pdf
- Aldana, A., Falcón-Cortés, A., & Larralde, H. (2022). A machine learning model to identify corruption in México's public procurement contracts. arXiv. <https://doi.org/10.48550/arXiv.2211.01478>
- Ansari, B., Barati, M., & Martin, E. G. (2022). Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research. *Government Information Quarterly*, 39(1), 101657. <https://doi.org/10.1016/j.giq.2021.101657>

- Balaniuk, R., Bessiere, P., Mazer, E., & Cobbe, P. (2013). Collusion and Corruption Risk Analysis Using Naïve Bayes Classifiers. *Communications in Computer and Information Science*, 246, 89-100. https://doi.org/10.1007/978-3-642-42017-7_7
- Basheka, B. C. (2011). Economic and political determinants of public procurement corruption in developing countries: An empirical study from uganda. *Journal of Public Procurement*, 11(1), 33-60. <https://doi.org/10.1108/JOPP-11-01-2011-Boo2>
- Broms, R., Dahlström, C., & Fazekas, M. (2019). Political Competition and Public Procurement Outcomes. *Comparative Political Studies*, 52(9), 1259-1292. <https://doi.org/10.1177/0010414019830723>
- Camargo, E. A. R., & Pinzon, M. A. R. (2022). La importancia de la seguridad de la información en el sector público en Colombia. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, 46, 87-99. <https://doi.org/10.17013/risti.46.87-99>
- Colonnelli, E., Gallego, J. A., & Prem, M. (2020). What Predicts Corruption? SSRN Scholarly Paper ID 3330651. Social Science Research Network. <https://doi.org/10.2139/ssrn.3330651>
- Decarolis, F., & Giorgiantonio, C. (2022). Corruption red flags in public procurement: New evidence from Italian calls for tenders. *EPJ Data Science*, 11(1), Art. 1. <https://doi.org/10.1140/epjds/s13688-022-00325-x>
- Domingos, S. L., Carvalho, R. N., Carvalho, R. S., & Ramos, G. N. (2016). Identifying IT Purchases Anomalies in the Brazilian Government Procurement System Using Deep Learning. 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 722-727. <https://doi.org/10.1109/ICMLA.2016.0129>
- Duguay, R., Rauter, T., & Samuels, D. (2019). The Impact of Open Data on Public Procurement. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3483868>
- Fazekas, M., & Czibik, Á. (2021). Measuring regional quality of government: The public spending quality index based on government contracting data. *Regional Studies*, 55(8), 1459-1472. <https://doi.org/10.1080/00343404.2021.1902975>
- Fazekas, M., & Dávid-Barrett, E. (2015). Corruption Risks in UK Public Procurement and New Anti- Corruption Tools. November, 1-33.
- Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British Journal of Political Science*, 50(1), 155-164. <https://doi.org/10.1017/S0007123417000461>
- Fazekas, M., Poltoratskaia, V., & Tóth, B. (2023). Corruption Risks and State Capture in Bulgarian Public Procurement. World Bank. <https://doi.org/10.1596/1813-9450-10444>
- Fazekas, M., Tóth, I. J., & King, L. P. (2013). Corruption manual for beginners.
- Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research*, 22(3), 369-397. <https://doi.org/10.1007/s10610-016-9308-z>

- Gallego, J., Prem, M., & Vargas, J. F. (2021). Pandemic corruption: Insights from Latin America. En *Procurement in Focus: Rules, Discretion, and Emergencies* (p. 180). Centre for Economic Policy Research. <https://www.govtransparency.eu/wp-content/uploads/2022/01/Procurement-in-Focus.pdf#page=110>
- Gallego, J., Rivero, G., & Martínez, J. (2020). Preventing rather than punishing: An early warning model of malfeasance in public procurement. *International Journal of Forecasting*. <https://doi.org/10.1016/j.ijforecast.2020.06.006>
- García Rodríguez, M. J., Rodríguez-Montequín, V., Ballesteros-Pérez, P., Love, P. E. D., & Signor, R. (2022). Collusion detection in public procurement auctions with machine learning algorithms. *Automation in Construction*, 133, 104047. <https://doi.org/10.1016/j.autcon.2021.104047>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258-268. <https://doi.org/10.1080/10580530.2012.716740>
- Jorquera, M. (2019). *Compras públicas y Big Data: Investigación en Chile sobre índice de riesgo de corrupción*. <http://www.repository.fedesarrollo.org.co/handle/11445/3871>
- Martínez, A., & Torres, L. M. (2019). *Compras públicas y Big Data: El caso mexicano*. <http://www.repository.fedesarrollo.org.co/handle/11445/3874>
- Modrušan, N., Rabuzin, K., & Mršić, L. (2021). Review of Public Procurement Fraud Detection Techniques Powered by Emerging Technologies. *International Journal of Advanced Computer Science and Applications*, 12(2). <https://doi.org/10.14569/IJACSA.2021.0120272>
- Nai, R., Sulis, E., & Meo, R. (2022). Public Procurement Fraud Detection and Artificial Intelligence Techniques: A Literature Review. *The Knowledge Management for Law Workshop (KM4LAW)*, Bozen-Bolzano, Italy.
- Neupane, A., Soar, J., Vaidya, K., & Yong, J. (2014). Willingness to adopt e-procurement to reduce corruption: Results of the PLS Path modeling. *Transforming Government: People, Process and Policy*, 8(4), 500-520. <https://doi.org/10.1108/TG-03-2014-0007>
- OCDE. (2015). *Informe de la OCDE sobre el Soborno Internacional*. OECD Publishing. <https://doi.org/10.1787/9789264226654-es>
- OECD. (2019). *Government at a Glance 2019*. OECD. <https://doi.org/10.1787/8ccf5c38-en>
- Rabuzin, K., & Modrušan, N. (2019). Prediction of Public Procurement Corruption Indices using Machine Learning Methods: Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 333-340. <https://doi.org/10.5220/0008353603330340>
- Sales, L. J., & Carvalho, R. N. (2016). Measuring the Risk of Public Contracts Using Bayesian Classifiers. *Proceedings of the 13th UAI Bayesian Modeling Applications Workshop (BMAW 2016)*, 7-59.

- Soreide, T. (2002). Corruption in public procurement: Causes, consequences and cures. CMI.
- Torres-Berru, Y., & López Batista, V. F. (2021). Data Mining to Identify Anomalies in Public Procurement Rating Parameters. *Electronics*, 10(22), 2873. <https://doi.org/10.3390/electronics10222873>
- Williams-Elegbe, S. (2018). Systemic corruption and public procurement in developing countries: Are there any solutions? *Journal of Public Procurement*, 18(2), 131-147. <https://doi.org/10.1108/JOPP-06-2018-009>
- World Bank. (2022). Using Data Analytics in Public Procurement Operational Options and a Guiding Framework Equitable Growth, Finance & Institutions Insight. World Bank. <https://doi.org/10.1596/37467>
- Zuleta, M. M., Ospina, S., & Caro, C. A. (2019). Índice de riesgo de corrupción en el sistema de compra pública colombiano a partir de una metodología desarrollada por el Instituto Mexicano para la Competitividad. <http://www.repository.fedesarrollo.org.co/handle/11445/3872>

Análise de metodologias para classificação de deepfakes

Guilherme Veiga Santos Pinto¹, Rudimar Luiz Scaranto Dazzi²,
Anita Maria da Rocha Fernandes³

guilhermeveigarj@edu.univali.br; rudimar@univali.br; anita.fernandes@univali.br

^{1,2,3} Mestrado em Computação Aplicada, Universidade do Vale do Itajaí, R. Uruguai, 458 - Centro, Itajaí, CEP 88302-901, Itajaí, SC, Brasil.

DOI: 10.17013/risti.59.21-35

Resumo: Este trabalho investiga a crescente sofisticação dos deepfakes e a necessidade de aprimorar os métodos de detecção. O objetivo é avaliar o desempenho de diferentes modelos de Deep Learning na classificação de vídeos manipulados, buscando uma acurácia mínima de 80%. Foram analisadas as arquiteturas LSTM, CNN, GRU e CvT, além da aplicação da Lei de Benford como método estatístico auxiliar. A base Celeb-DF V2 foi utilizada por sua fidelidade e complexidade. O modelo LSTM apresentou a melhor acurácia no treinamento, enquanto o GRU obteve melhor desempenho na avaliação prática. Já o CvT mostrou limitações tanto em desempenho quanto em custo computacional. A Lei de Benford mostrou-se promissora, mas inconclusiva sem um parâmetro de correlação. Os resultados reforçam a importância de atualizar continuamente os mecanismos de detecção para acompanhar a evolução dos vídeos sintéticos e mitigar riscos associados à desinformação.

Palavras-chave: Detecção de DeepFake; Manipulação de Vídeo; Aprendizado Profundo; Redes Neurais; Lei de Benford.

Analysis of Deepfake Classification Methodologies

Abstract: This study investigates the increasing sophistication of deepfakes and the need to improve detection techniques. The objective is to evaluate the performance of different Deep Learning models in classifying manipulated videos, aiming for a minimum accuracy of 80%. The architectures LSTM, CNN, GRU, and CvT were analyzed, along with the application of Benford's Law as a statistical aid. The Celeb-DF V2 dataset was selected for its realism and complexity. The LSTM model achieved the highest training accuracy, while GRU showed superior performance during evaluation. The CvT model underperformed in both accuracy and computational cost. Although promising, Benford's Law yielded inconclusive results without a correlation metric. The findings highlight the importance of continually updating detection methods to keep pace with synthetic video advancements and mitigate risks linked to misinformation.

Keywords: DeepFake Detection; Video Manipulation; Deep Learning; Neural Networks; Benford's Law.

1. Introdução

O avanço das redes neurais artificiais tem impulsionado significativamente a manipulação de conteúdos multimídia, como imagens e vídeos. Ferramentas de inteligência artificial amplamente acessíveis, como FaceApp, Stable Diffusion WebUI e Roop-Unleashed, permitem a criação de deepfakes cada vez mais realistas, facilitando a produção de vídeos sintéticos de alta fidelidade com poucos recursos técnicos (Rana et al., 2022). Embora frequentemente utilizadas de forma recreativa, essas tecnologias vêm sendo empregadas em golpes e fraudes sofisticadas, como revelado por reportagens da BBC e por autoridades de Hong Kong, que relataram prejuízos causados por chamadas de vídeo falsas utilizando técnicas de FaceSwap em tempo real (Machado, 2024; Chen & Magramo, 2025). Tais avanços têm como base técnica o uso de Redes Adversárias Generativas (GANs), que combinam redes generativas e discriminativas para criar conteúdos altamente realistas (Goodfellow et al., 2014; Oberoi, 2018; Hui, 2018).

Nesse contexto, torna-se essencial questionar a eficácia dos métodos de detecção e classificação presentes na literatura científica. Apesar da rápida evolução das pesquisas em resposta às ameaças à privacidade e à segurança da informação, os modelos atuais precisam ser constantemente reavaliados diante do aumento do realismo e da acessibilidade das ferramentas de criação de deepfakes. Técnicas como Machine Learning, Deep Learning e Blockchain vêm sendo amplamente estudadas (Rana et al., 2022), mas mesmo com bons resultados estatísticos, há dúvidas quanto à sua robustez frente aos vídeos sintéticos mais recentes e sofisticados.

Este estudo propõe avaliar diferentes abordagens de classificação de deepfakes, utilizando o Deep Learning (DL) como base principal. Foram testados modelos como Long Short-Term Memory (LSTM) com arquitetura ResNext50, Convolutional Neural Network (CNN) baseada na arquitetura Xception, arquitetura Gated Recurrent Unit (GRU), Convolutional Vision Transformer (CvT), além da aplicação da Regra de Benford para análise de padrões numéricos. A base de dados utilizada foi a Celeb-DF V2, composta por vídeos reais e manipulados de celebridades (Li et al., 2019). O objetivo é investigar se essas técnicas ainda oferecem precisão na distinção entre vídeos autênticos e manipulados, considerando os desafios impostos pelas falsificações modernas e hiper-realistas.

O restante do artigo encontra-se estruturado nas seguintes seções: a seção 2 apresenta a fundamentação teórica; a seção 3 trata dos trabalhos relacionados; a seção 4 descreve o desenvolvimento do trabalho; a seção 5 apresenta os resultados do trabalho; e por fim a seção 6 traz as conclusões da pesquisa e os trabalhos futuros.

2. Fundamentação Teórica

Esta seção apresenta uma síntese dos principais temas que fundamentaram a elaboração deste estudo.

2.1. Deepfakes

O termo “deepfake” refere-se a mídias sintéticas geradas por técnicas avançadas de aprendizado profundo, especialmente GANs. Desde sua introdução, a tecnologia

evoluiu significativamente, tornando-se cada vez mais difícil distinguir conteúdos reais de sintéticos. De acordo com Pei et al. (2024), essas redes consistem em dois modelos neurais: o gerador, que cria conteúdo falso, e o discriminador, que avalia a autenticidade deste conteúdo. O treinamento contínuo entre esses modelos resulta em mídias altamente realistas, dificultando a distinção entre o real e o falso.

Recentemente, modelos de difusão têm emergido como uma alternativa promissora na geração de deepfakes. Esses modelos aprendem a transformar dados de uma distribuição simples para uma complexa, permitindo a criação de conteúdos ainda mais convincentes. Ferramentas como o Stable Diffusion exemplificam essa abordagem, oferecendo resultados impressionantes na síntese de imagens e vídeos. Além das GANs e dos modelos de difusão, técnicas de transferência de estilo e mapeamento de características faciais são empregadas para aprimorar a qualidade dos deepfakes. Essas técnicas permitem a substituição de rostos em vídeos, mantendo expressões faciais e movimentos labiais coerentes com o conteúdo original. O Roop-Unleashed é um exemplo de ferramenta que utiliza essas metodologias para criar vídeos sintéticos de alta fidelidade (Pei et al., 2024).

Com o avanço das técnicas e o aumento da capacidade computacional, a geração de deepfakes tornou-se mais acessível, permitindo que indivíduos sem expertise técnica produzam conteúdos falsificados com relativa facilidade. A evolução contínua dessas tecnologias levanta preocupações sobre a capacidade dos modelos atuais de detecção em identificar deepfakes. À medida que os métodos de geração se tornam mais sofisticados, é de extrema importância desenvolver técnicas de detecção igualmente avançadas para mitigar os riscos associados à disseminação de mídias sintéticas maliciosas (Lee et al., 2024).

2.2. Fundamentos e Arquiteturas de Deep Learning

O DL é um subcampo do Machine Learning baseado em redes neurais artificiais, com estrutura inspirada no cérebro humano, e tornou-se um pilar da inteligência artificial moderna graças ao avanço computacional e à abundância de dados (Goodfellow, Bengio, & Courville, 2016). Essas redes aprendem representações hierárquicas dos dados por meio de múltiplas camadas, como destacam LeCun, Bengio e Hinton (2015), utilizando métodos como o Stochastic Gradient Descent para otimizar funções de perda. As principais arquiteturas incluem as CNNs, as Redes Recorrentes (RNNs) e os Transformers, todas amplamente aplicadas em tarefas como classificação de imagens, reconhecimento de fala e processamento de linguagem natural.

Dentre as RNNs, destaca-se a arquitetura LSTM, proposta por Hochreiter e Schmidhuber (1997), que foi desenhada para resolver o problema do vanishing gradient e modelar dependências sequenciais de longo prazo. A LSTM realiza este controle através de uma célula de memória e três mecanismos de portas (entrada, esquecimento e saída), permitindo seu uso em tarefas que exigem a retenção seletiva de informações em sequências extensas (Gers, Schmidhuber, & Cummins, 2000; Graves, 2012). Para simplificar essa estrutura, surgiram variantes como as GRUs, que alcançam desempenho comparável com um número reduzido de parâmetros (Cho et al., 2014). As GRUs utilizam apenas duas portas (atualização e reinicialização) para controlar o fluxo de

informações, oferecendo uma alternativa mais leve e eficiente para tarefas sequenciais (Chung, Gulcehre, Cho, & Bengio, 2014; Bahdanau, Cho, & Bengio, 2014).

As CNNs, por sua vez, são fundamentais no processamento de imagens, utilizando camadas convolucionais que extraem características como bordas e texturas (LeCun, Bottou, Bengio, & Haffner, 2002). Sua estrutura é composta por camadas convolucionais, pooling e totalmente conectadas, com funções de ativação como ReLU (Simonyan & Zisserman, 2014). Para tarefas de detecção de deepfakes, que exigem alta precisão e eficiência, destacam-se arquiteturas avançadas como a Xception (Extreme Inception). Este modelo substitui as convoluções padrão por convoluções separáveis em profundidade, que desacoplam a filtragem espacial e a combinação de canais, reduzindo a complexidade computacional e otimizando o fluxo de informação por meio de conexões residuais (Chollet, 2017).

Combinando o melhor das CNNs e dos Transformers, a arquitetura CvT propõe uma abordagem híbrida para processar imagens, integrando convoluções no processo de tokenização e nos mecanismos de atenção (Wu et al., 2021). A inovação central do CvT reside na utilização do Convolutional Token Embedding e em blocos hierárquicos. Esta estrutura permite a extração eficiente de padrões locais antes da autoatenção e a modelagem global de dependências, proporcionando um modelo mais robusto e eficiente em comparação com modelos puramente baseados em Vision Transformer (Dosovitskiy et al., 2020; Tolstikhin et al., 2021).

2.3. Regra de Benford

A Regra de Benford, também conhecida como Regra do Primeiro Dígito, descreve a distribuição não uniforme dos primeiros dígitos em conjuntos de dados numéricos naturais. Segundo Benford (1938), essa distribuição indica que o dígito 1 tende a aparecer como primeiro dígito com maior frequência do que o dígito 9. Essa característica é expressa por (1) que determina a probabilidade de cada dígito entre 1 e 9 aparecer na primeira posição. Hill (1995) complementa afirmando que cerca de 30,1% dos números legítimos começam com o dígito 1, enquanto apenas 4,6% iniciam com o dígito 9. Essa distribuição é comum em dados que abrangem múltiplas ordens de magnitude, com crescimento exponencial, sendo aplicável a áreas como contabilidade, finanças e ciências naturais.

$$P(d) = \log_{10} \left(1 + \frac{1}{d} \right) \quad (1)$$

A aplicação da Regra de Benford na detecção de fraudes se baseia na ideia de que conjuntos de dados manipulados tendem a se desviar dessa distribuição esperada. Nigrini (1996) destaca que padrões suspeitos podem indicar manipulações, sendo essa abordagem utilizada em auditorias de relatórios financeiros, registros contábeis e dados eleitorais (Durtschi et al., 2004). Nigrini e Mittermaier (1997) apontam que empresas fraudulentas frequentemente apresentam relatórios com distribuições anômalas, tornando a Regra de Benford uma ferramenta valiosa para auditores. No entanto, Hill (1998) adverte que a regra não se aplica a todos os conjuntos de dados, especialmente os

compostos por valores fixos, limitados ou artificialmente definidos, exigindo assim o uso complementar de outras técnicas estatísticas.

Na literatura recente, a Regra de Benford também vem sendo aplicada à detecção de manipulações em imagens digitais. Vishnu (2021) descreve uma metodologia que envolve a conversão dos pixels da imagem para o domínio da frequência por meio da Transformada Discreta de Cosseno (DCT). Em seguida, extrai-se o primeiro dígito dos coeficientes da DCT e verifica-se, por meio de uma visualização gráfica, se a distribuição desses dígitos segue o padrão esperado da Regra de Benford.

3. Trabalhos Relacionados

Esta seção apresenta a síntese dos sete trabalhos analisados que exploram metodologias de DL e técnicas estatísticas aplicadas à detecção de deepfakes. A seleção destes estudos foi definida por um processo sistemático de busca em bases de dados bibliográficas. As strings de pesquisa foram construídas para identificar publicações recentes (2018-2025) focadas nas arquiteturas de redes neurais (RNNs, CNNs, CvT) e no método da Regra de Benford para classificação de deepfakes. A aplicação de critérios de inclusão e exclusão resultou no conjunto de sete artigos, que servem de base para a comparação de desempenho e custo computacional.

Os estudos de Kularkar et al. (2023) e Mogulkoc e Yuce (2024) enfatizam modelos que combinam aprendizado profundo. Kularkar propõe uma abordagem híbrida com ResNext e LSTM, alcançando 94,09% de precisão ao analisar quadros de vídeo para capturar inconsistências sutis. Já Mogulkoc e Yuce comparam CNN, LSTM e GRU, observando melhor desempenho com GRU (85% de acurácia), o que destaca a importância do uso de modelos capazes de explorar as dependências temporais em vídeos manipulados.

Complementando essa linha de pesquisa, os trabalhos de Wodajo e Atnafu (2021, 2023) investigam o uso de arquiteturas baseadas em Transformers. Em 2021, os autores introduzem o CvT, que combina CNN e Vision Transformer (ViT), atingindo 91,5% de acurácia. Em 2023, evoluem para o GenConViT, uma arquitetura mais complexa que incorpora Autoencoders e Swin Transformers, alcançando impressionantes 95,8% de acurácia e excelente generalização em diversos datasets. Ambos os trabalhos reforçam a tendência atual de utilização de arquiteturas híbridas com mecanismos de atenção para melhor desempenho.

Por outro lado, três estudos analisam a aplicação da Lei de Benford como ferramenta estatística para detecção de manipulações em imagens digitais e deepfakes. Fernandes e Antunes (2023) utilizam a Transformada Rápida de Fourier (FFT) e a métrica Cramer-Von Mises (CVM), para extrair padrões numéricos de imagens e obtêm um F1-score médio de 90,47%, com destaque para a correlação de Pearson como melhor métrica. Já Vishnu (2021) emprega DCT e DFT para transformar as imagens e compara várias métricas de desvio estatístico, apontando a DCT como mais eficaz. Bonettini et al. (2020) ampliam essa abordagem ao utilizar a DCT em imagens divididas em blocos e extrair vetores de características alimentados em classificadores Random Forest, alcançando acurácia superior a 99% em grandes conjuntos de imagens de GANs.

Em resumo, os trabalhos revisados se agrupam em duas grandes vertentes: modelos de aprendizado profundo com redes neurais (CNN, LSTM, GRU, Transformers) e abordagens estatísticas com a Lei de Benford. Os primeiros apresentam alta acurácia, mas com maior custo computacional e necessidade de treinamento robusto. Já os métodos baseados em Benford destacam-se pela eficiência com recursos limitados. A tabela 1 a seguir detalha essas diferenças em termos de abordagem, datasets, desempenho e limitações.

Trabalho	Abordagem	Desempenho	Dataset
Kularkar et al. (2023)	LSTM	94,09%	FaceForensics++, CelebDF e DFDC
Mogulkoc e Yuce (2024)	CNN, LSTM e GRU	85% (GRU) 83% (LSTM) 60% (CNN)	DFDC
Wodajo e Atnafu (2021)	CvT	91,5%	DFDC
Wodajo e Atnafu (2023)	GenConViT	98,5% (DFDC) 98,28% (TIMIT) 97% (FF++) 90,94% (Celeb-DF V2)	DFDC, DeepfakeTIMIT, FaceForensics++ e Celeb-DF v2
Fernandes e Antunes (2023)	Regra de Benford, FFT e Cramer-Von Mises (CVM)	90,47%	Dataset do próprio autor.
Vishnu (2021)	Regra de Benford, DCT e FFT	Imagens manipuladas apresentaram 0,0173 de desvio (distância euclidiana), contra 0,0175 para imagens originais.	Dataset do próprio autor.
Bonettini et al. (2020)	Regra de Benford e DCT	87.96%	Datasets gerados via Cycle-Gan e ProGAN
Este Trabalho	CNN, LSTM, GRU, CvT e Regra de Benford	97,32%	Celeb-DF v2

Tabela 1 – Análise comparativa dos trabalhos relacionados.

Os estudos revisados agrupam-se em duas vertentes: modelos de aprendizado profundo, que demonstram alta acurácia com maior custo computacional, e a Lei de Benford, que oferece eficiência, mas cuja aplicação requer métricas de correlação para a validação dos resultados. Observa-se, no entanto, a ausência de uma avaliação comparativa direta destas arquiteturas e do método estatístico empregados no mesmo dataset de alta fidelidade (Celeb-DF V2) e sob idênticas condições de teste. O objetivo deste trabalho é preencher esta lacuna, estabelecendo um *benchmark* de desempenho entre estas metodologias na tarefa de classificação de deepfakes.

4. Análise Do Celeb-DF v2 e De Outros Datasets Disponíveis

A eficácia na detecção de deepfakes depende fortemente de datasets com alta qualidade e diversidade. Entre as bases mais relevantes, o Celeb-DF v2 se destaca pelo realismo

nas trocas faciais de celebridades, preservando expressões, iluminação e movimentos naturais. Como ilustrado na Figura 1, essas características tornam o dataset especialmente desafiador para os modelos de detecção, sendo ideal para simular cenários do mundo real.

Além de sua qualidade visual, o Celeb-DF v2 apresenta variedade de rostos, gêneros e contextos, o que contribui para o treinamento de modelos mais robustos e generalistas. Ao todo, a base contém 5.639 vídeos manipulados e 890 vídeos reais, extraídos de vídeos públicos do YouTube, totalizando mais de 2 milhões de frames com cerca de 10 segundos cada, no formato MP4.

Comparativamente, o DFDC, criado para o desafio promovido pelo Facebook, oferece grande volume e diversidade de manipulações, mas com qualidade facial inferior ao Celeb-DF v2, como também demonstrado na Figura 1. Já o FaceForensics++ inclui manipulações por autoencoders e GANs em diversos níveis de compressão, sendo útil para testar robustez, embora com realismo inferior, conforme observado na Figura 2.



Figura 1 –Comparativo da base de dados Celeb-DF v2 a esquerda, com a base de dados DFDC a direita. Adaptado de Li et al. (2020) e Dolhansky et al. (2020).

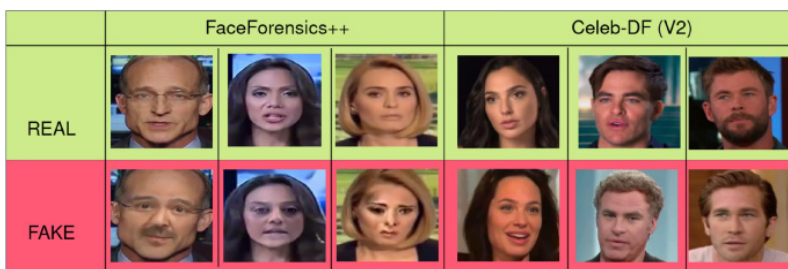


Figura 2 – Comparativo visual entre o FaceForensics++ e o Celeb-DF v2. Reimpresso de *Visual attention-based deepfake video forgery detection* por Ganguly et al., 2022.

Diante dessas considerações, o Celeb-DF v2 foi escolhido como principal base deste estudo por oferecer o melhor equilíbrio entre fidelidade visual e diversidade, sendo mais adequado para treinar e avaliar modelos de detecção em contextos realistas e complexos.

4.1. Pré-processamento e Treinamento

Após a escolha da base de dados, iniciou-se o pré-processamento. Os vídeos foram convertidos em trechos de quatro segundos, com resolução padronizada de 224x224 pixels, por meio de recorte facial (*face crop*) utilizando a biblioteca *face recognition*. Essa etapa garantiu foco nos rostos, facilitando o treinamento dos modelos.

Vídeos corrompidos e com áudio foram removidos, por não contribuírem para os objetivos do estudo. Com os dados limpos, os vídeos foram carregados na memória do ambiente de execução para início do treinamento.

O dataset Celeb-DF V2 apresentava desequilíbrio entre classes, com predominância de vídeos manipulados. Para evitar viés de classificação, foi aplicado balanceamento, igualando-se a quantidade de vídeos reais e sintéticos, totalizando 1.869 vídeos (50% reais e 50% falsos), conforme recomendações de Krishna et al. (2022).

O treinamento com os modelos ResNext50 (LSTM), Xception (CNN) e GRU foi realizado no Google Colab, com GPU NVIDIA Tesla T4. Os vídeos foram extraídos do Google Drive e processados com OpenCV, resultando em média de 148 frames por vídeo. As classes foram definidas com base em um arquivo CSV contendo os nomes e rótulos dos vídeos (REAL ou FAKE).

A divisão do conjunto de dados foi feita com a biblioteca *Scikit-learn*: 80% para treinamento, 10% para teste e 10% para validação, totalizando 1.495 vídeos para treino e 374 para teste e validação. Os modelos foram treinados por 20 épocas, com taxa de aprendizado de 0.00001. Em geral, os treinamentos duraram entre duas e quatro horas, considerando o tempo de carregamento dos vídeos (cerca de 40 minutos).

Para o modelo Convolutional Vision Transformer (CvT), houve uma diferença metodológica: a extração dos frames e a separação dos dados foram feitas manualmente, mantendo as proporções de 80/10/10. Isso resultou em um total de 276.460 imagens. O treinamento foi realizado no Google Colab com GPU NVIDIA Tesla A100, com duração aproximada de oito horas, sendo o processo mais demorado entre os modelos avaliados.

5. Resultados

Conforme especificado na seção 4.1, quatro diferentes algoritmos de DL foram treinados com uma distribuição de dados de 80% para treinamento, 10% para validação e 10% para teste. Os modelos treinados consistiram no modelo LSTM com a arquitetura ResNext50, o modelo CNN com a arquitetura Xception2, que incorpora uma variante da arquitetura Xception, o modelo GRU com a arquitetura ResNet-50, e o modelo CvT.

A principal métrica avaliada foi a acurácia, que consiste na divisão das previsões corretas pelo número total de previsões conforme observado em (2). Em outras palavras, é a divisão da soma dos verdadeiros positivos (VP) e verdadeiros negativos (VN) pela soma dos verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN).

$$\text{Acurácia AC} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2)$$

Os modelos foram avaliados com base em sua matriz de confusão e acurácia, além de serem submetidos a um teste utilizando 10 vídeos reais e 10 vídeos sintéticos da base *Celeb-DF V2*, os quais não fizeram parte da etapa de treinamento. Nesta etapa de teste, é feita uma extração de um dos frames desses vídeos, retornando-o com um mapa de calor, e um valor de confiança de predição com o rótulo categorizado pelo modelo.

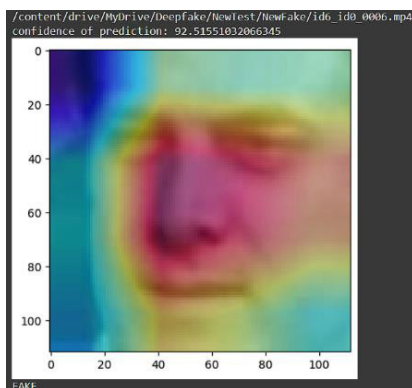


Figura 3 – Etapa de avaliação do modelo com vídeos separados da base.

No treinamento do modelo LSTM com a arquitetura ResNext50 por 20 épocas, foi alcançada uma acurácia máxima de 97,32% na fase de validação, com uma queda consistente nos valores de perda de treinamento e validação conforme a Figura 4. A matriz de confusão na Tabela 2 revelou que, entre 200 vídeos sintéticos e 174 vídeos reais, houve uma maior ocorrência de falsos negativos do que de falsos positivos. Na etapa de avaliação, o modelo demonstrou bom desempenho na classificação de frames, acertando todos os 10 vídeos sintéticos testados. No entanto, ao classificar os 10 vídeos reais, dois foram incorretamente rotulados como falsos positivos, conforme a Tabela 3.

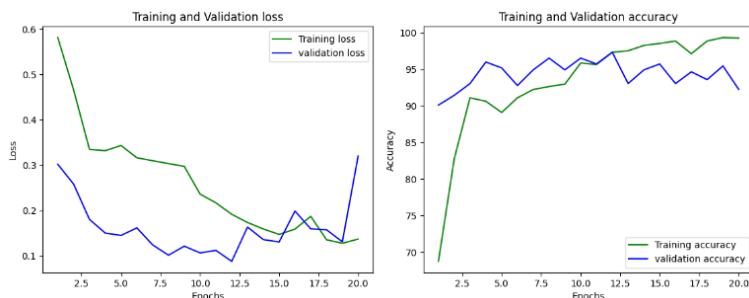


Figura 4 – Gráficos de Perda e de Acurácia para Treinamento e Validação da arquitetura LSTM.

Já no treinamento do modelo CNN com a arquitetura Xception2 por 20 épocas, foi registrada na Figura 5 uma acurácia máxima de 78,34% na validação, com picos de perda

variando entre 0,47 e 0,59. Na Tabela 2, a matriz de confusão indicou uma quantidade maior de falsos negativos em comparação aos falsos positivos. Na fase de avaliação, o desempenho do modelo CNN foi menos consistente em relação ao modelo anterior, apresentando menor confiança nas predições, além da ocorrência de um verdadeiro negativo e dois falsos positivos, conforme a Tabela 3.

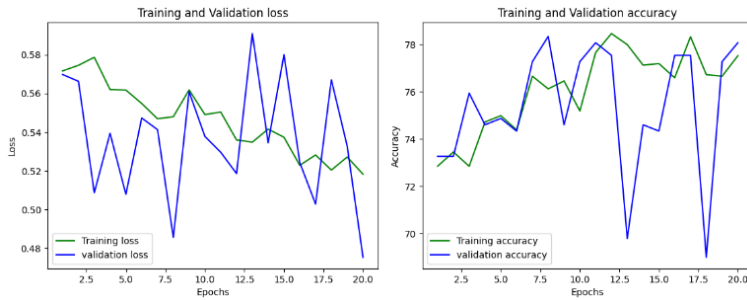


Figura 5 – Gráficos de Perda e de Acurácia para Treinamento e Validação da arquitetura CNN.

O modelo GRU, treinado por 20 épocas, alcançou uma acurácia máxima de 96,52% na validação e apresentou na Figura 6, uma redução consistente nos valores de perda de treinamento e validação, comportamento semelhante ao modelo LSTM. Sua matriz de confusão na Tabela 2 indicou uma leve predominância de falsos negativos sobre falsos positivos. Na etapa de avaliação, o GRU demonstrou desempenho destacado tanto nos níveis de confiança das predições quanto na taxa de acertos, registrando apenas um falso positivo na classificação de vídeos reais, Conforme a Tabela 3.

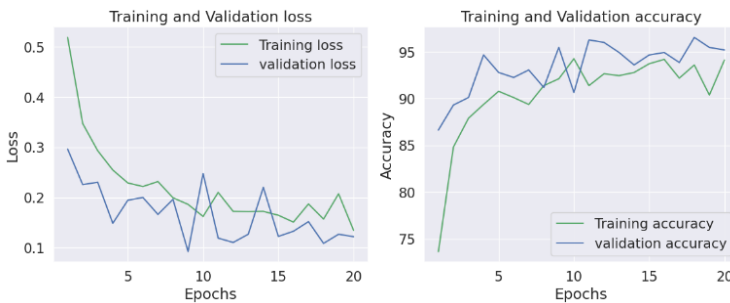


Figura 6 – Gráficos de Perda e de Acurácia para Treinamento e Validação da arquitetura GRU.

O modelo CvT, apesar de exigir mais tempo de treinamento em comparação aos demais, apresentou o pior desempenho entre os modelos avaliados, com uma acurácia de apenas 50,53% na classificação de deepfakes e perdas estagnadas durante o treinamento e validação. Conforme a Tabela 3 na etapa de avaliação com 10 vídeos reais e 10 sintéticos, observou-se que o modelo classificou todos os vídeos como deepfakes, utilizando uma abordagem baseada em valores de predição — abaixo de 0,5 para REAL e igual ou

acima de 0,5 para FAKE. Todos os valores de predição foram idênticos, indicando um comportamento inconsistente.

Classe	LSTM	CNN	GRU
Verdadeiro Positivo (VP)	186	189	198
Falso Positivo (FP)	3	11	2
Falso Negativo (FN)	26	71	16
Verdadeiro Negativo (VN)	159	103	158

Tabela 2 – Matriz de Confusão das arquiteturas LSTM, CNN e GRU.

Classe	Classe	Acertos	Erros	Tipo de Erro	Média da Confiança de Predição (%)
LSTM	FAKE	10	0	-	97,57%
LSTM	REAL	8	2	Falso Positivo	95,74%
CNN	FAKE	9	1	Falso Negativo	75,42%
CNN	REAL	8	2	Falso Positivo	71,10%
GRU	FAKE	10	0	-	98,49%
GRU	REAL	9	1	Falso Positivo	93,56%
CvT	FAKE	10	0	-	51,04%
CvT	REAL	0	10	Falso Positivo	51,04%

Tabela 3 – Desempenho dos modelos para classificação de vídeos.

Adicionalmente, foi realizada uma análise em frames de vídeos do dataset Celeb-DF v2 utilizando a Regra do Primeiro Dígito, ou Regra de Benford, com o objetivo de identificar possíveis manipulações. O processo envolveu a extração dos valores RGB das imagens, normalização e aplicação da Transformada Discreta de Cosseno (DCT), que converte os dados para o domínio da frequência, separando componentes de baixa e alta frequência. A partir dos coeficientes da DCT, foram extraídos os primeiros dígitos e analisada sua distribuição, comparando-a com o padrão teórico da Regra de Benford.

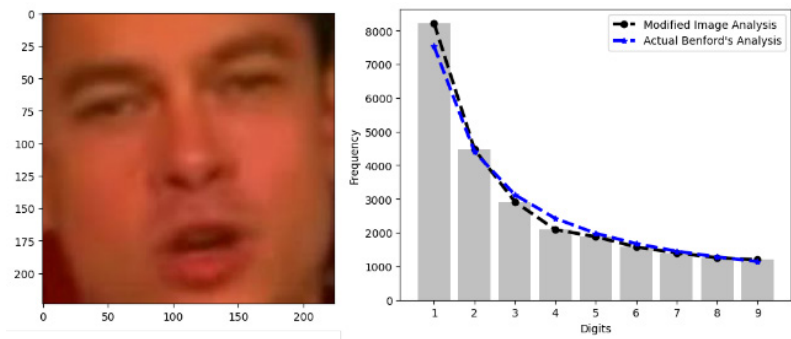


Figura 7 – Análise de um frame pertencente à classe FAKE, com variação à Regra de Benford.

A conformidade com a regra foi visualizada por meio de gráficos que comparam a frequência dos dígitos observados com a distribuição esperada. Como observado na Figura 7, imagens de vídeos manipulados apresentaram, em alguns casos, menor aderência à lei do que as imagens reais. No entanto, também ocorreram casos contrários, em que deepfakes se aproximaram mais da lei do que imagens reais, o que indica que os resultados ainda não são conclusivos sem o cálculo das distâncias entre as linhas nos gráficos. Além disso, foram realizadas análises com frames completos dos vídeos, sem focar apenas nos rostos.

6. Conclusões

Este trabalho abordou a crescente relevância da detecção de deepfakes avaliando a eficácia de diferentes modelos de DL e a aplicação da Regra de Benford no desafiador dataset Celeb-DF V2. Os resultados demonstram a alta eficácia das arquiteturas recorrentes no contexto de alta fidelidade do Celeb-DF V2, com o modelo LSTM (97,32% de acurácia) e o GRU (96,52%) superando o objetivo mínimo de 80%. Este desempenho é notavelmente superior ao encontrado em trabalhos relacionados que utilizaram arquiteturas similares em outros datasets (por exemplo, Kularkar et al., 2023, e Mogulkoc e Yuce, 2024) e, em comparação direta, superou a acurácia de modelos mais complexos como o GenConViT de Wodajo e Atnafu (2023) no mesmo dataset. Esta superioridade sugere que modelos DL especializados em sequências temporais, quando otimizados com backbones CNN, fornecem o desempenho mais robusto para a detecção de manipulações sutis em vídeos de alta qualidade.

Em contraste, o modelo CvT apresentou baixa acurácia (50,53%) e alto custo computacional, falhando em generalizar para a classe REAL. Essa queda abrupta de desempenho, contrastando com o resultado de 91,5% obtido por Wodajo e Atnafu (2021) no dataset DFDC, sugere que, embora os Vision Transformers sejam promissores, eles são particularmente sensíveis às características de alta complexidade do Celeb-DF V2, indicando que podem exigir maior volume de dados ou ajustes mais finos no treinamento para estabilização. Observou-se também uma tendência dos modelos recorrentes a encontrar mais falsos positivos (classificar real como falso) do que verdadeiros negativos.

Adicionalmente, a análise com a Regra de Benford mostrou-se promissora, mas ainda inconclusiva sem métricas estatísticas formais. A inconsistência observada, com deepfakes se aproximando da Lei em alguns casos, ressalta a advertência de Hill (1998) sobre a aplicabilidade limitada da regra. Embora outros trabalhos (Fernandes e Antunes, 2023; Bonettini et al., 2020) tenham obtido alta performance com o método, a ausência de um parâmetro de correlação clara aqui reforça a necessidade de integrar a análise com testes estatísticos robustos para validar o desvio.

Como trabalhos futuros, propõe-se testar os modelos com vídeos produzidos pelos próprios autores — respeitando aspectos éticos e legais — gerados via *Roop-Unleashed*, e aplicar métodos como a correlação de Pearson para aprimorar a análise estatística dessas manipulações.

Referências

- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American philosophical society*, 551-572.
- Bonettini, N., Bestagini, P., Milani, S., & Tubaro, S. (2021). On the use of Benford's law to detect GAN-generated images. In *2020 25th international conference on pattern recognition (ICPR)* (pp. 5495-5502). IEEE.
- Chen, H., & Magramo, K. (2025). *Funcionário de multinacional paga US\$ 25 mi a golpista que usou deepfake para simular reunião*. CNN Brasil. <https://www.cnnbrasil.com.br/internacional/funcionario-de-multinacional-paga-us-25-mi-a-golpista-que-usou-deepfake-para-simular-reuniao/>
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., & Ferrer, C. C. (2020). The deepfake detection challenge (dfdc) dataset. *arXiv preprint arXiv:2006.07397*.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting*, 5(1), 17-34.
- Fernandes, P., & Antunes, M. (2023). Benford's law applied to digital forensic analysis. *Forensic Science International: Digital Investigation*, 45, 301515.
- Ganguly, S., Ganguly, A., Mohiuddin, S., Malakar, S., & Sarkar, R. (2022). ViXNet: Vision Transformer with Xception Network for deepfakes based video and image forgery detection. *Expert Systems with Applications*, 210, 118423.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
- Graves, A. (2012). *Supervised sequence labelling with recurrent neural networks* (Vol. 385). Studies in Computational Intelligence.

- Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 6645-6649). Ieee.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical science*, 354-363.
- Hill, T. P. (1998). The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist*, 86(4), 358-363.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Hui, J. (2018). How deep learning fakes videos (Deepfake) and how to detect it?. *Medium Corporation*, 28.
- Krishna, P., Arukala, S., Battula, V. R., & Sri, M. B. (2022). Deepfake Detection using LSTM and ResNext. *Journal of Engineering Sciences*, 13(07).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Kularkar, T., Jikar, T., Rewaskar, V., Dhawale, K., Thomas, A., & Madankar, M. (2023). Deepfake detection using LSTM and ResNext.
- Lai, G., Chang, W. C., Yang, Y., & Liu, H. (2018). Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st international ACM SIGIR conference on research & development in information retrieval* (pp. 95-104).
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2002). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Lee, H., Lee, C., Farhat, K., Qiu, L., Geluso, S., Kim, A., & Etzioni, O. (2024). The Tug-of-War Between Deepfake Generation and Detection. *arXiv preprint arXiv:2407.06174*.
- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2019). Celeb-df (v2): a new dataset for deepfake forensics [j]. *arXiv preprint arXiv*.
- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2020). Celeb-df: A large-scale challenging dataset for deepfake forensics. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 3207-3216).
- Machado, S. (2024). 'Eram meu rosto e minha voz, mas era golpe': Como criminosos 'clonam pessoas' com inteligência artificial. BBC News Brasil. <https://www.bbc.com/portuguese/articles/cd1jv45dq3go>

- Mogulkoc, Z., & Yuce, B. (2024, 25 de março). *A comparative analysis of state-of-the-art algorithms for robust deep fake detection*. Abdullah Gul University. Acessado em 15 de fevereiro de 2025.
- Nigrini, M. J. (1996). A taxpayer compliance application of Benford's law. *The Journal of the American Taxation Association*, 18(1), 72.
- Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's Law as an Aid in Analytical Procedures. *Auditing: A journal of practice & theory*, 16(2).
- Oberoi, G. (2018). *Exploring DeepFakes*. <https://goberoi.com/exploring-deepfakes-20c9947c22d9>
- Pei, G., Zhang, J., Hu, M., Zhang, Z., Wang, C., Wu, Y., & Tao, D. (2024). Deepfake generation and detection: A benchmark and survey. *arXiv preprint arXiv:2403.17881*.
- Rana, M. S., Nobil, M. N., Murali, B., & Sung, A. H. (2022). Deepfake detection: A systematic literature review. *IEEE access*, 10, 25494-25513.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sundermeyer, M., Schlüter, R., & Ney, H. (2012). Lstm neural networks for language modeling. In *Interspeech* (Vol. 2012, pp. 194-197).
- Tolstikhin, I. O., Houlsby, N., Kolesnikov, A., Beyer, L., Zhai, X., Unterthiner, T., & Dosovitskiy, A. (2021). Mlp-mixer: An all-mlp architecture for vision. *Advances in neural information processing systems*, 34, 24261-24272.
- Vishnu, U. (2021). *Deepfake detection using Benford's Law and Distribution Variance Statistic*. *International Research Journal of Engineering and Technology (IRJET)*, 8(10), 712-719. Acessado em 13 de fevereiro de 2025.
- Wodajo, D., & Atnafu, S. (2021). Deepfake video detection using convolutional vision transformer. *arXiv preprint arXiv:2102.11126*.
- Wodajo, D., Atnafu, S., & Akhtar, Z. (2023). Deepfake video detection using generative convolutional vision transformer. *arXiv preprint arXiv:2307.07036*.
- Wu, H., Xiao, B., Codella, N., Liu, M., Dai, X., Yuan, L., & Zhang, L. (2021). Cvt: Introducing convolutions to vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 22-31).
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9, 611-629.

Clasificación de Frases en Zapoteco mediante CNN con Espectrogramas

Mariano Martínez Patiño¹, Sergio Juárez Vázquez¹, Efraín Dueñas Reyes¹,
Francisco Javier Sol Sampedro¹, Nicolas Hernández Ruiz¹

mariano971149@gmail.com; sjuarez@sandunga.unistmo.edu.mx;
eduenas@sandunga.unistmo.edu.mx; franciscosol@sandunga.unistmo.edu.mx;
nicolas@sandunga.unistmo.edu.mx

¹ Universidad del Istmo, Ciudad Universitaria S/N, Santa Cruz, 70760 Santo Domingo Tehuantepec, Oax., México.

DOI: 10.17013/risti.59.36–52

Resumen: Este estudio se enfoca en la clasificación de espectrogramas, representaciones visuales del audio para aplicar aprendizaje automático. Los métodos tradicionales, como los MFCCs con clasificadores clásicos, presentan limitaciones en lenguas con pocos recursos, como el zapoteco del Istmo. Modelos avanzados como RNNs y transformers requieren grandes volúmenes de datos, difíciles de obtener en contextos indígenas. Como alternativa, se propone una red neuronal convolucional profunda de 28 capas, entrenada con 10 frases comunes convertidas en espectrogramas y aumentadas artificialmente. El modelo logró un 100% de precisión en entrenamiento y 96.2% en validación. Aunque prometedor, se destaca la necesidad de ampliar el conjunto de datos. El trabajo evidencia el potencial del aprendizaje profundo para mejorar la comunicación intercultural y preservar lenguas indígenas en peligro.

Palabras-clave: Comunicación intercultural; lenguas indígenas zapoteca; imágenes espectrales; red neuronal profunda.

Application of CNN in Audio Recognition of Isthmus Zapotec

Abstract: This study focuses on the classification of spectrograms, visual representations of audio, for the application of machine learning. Traditional methods, such as MFCCs with classical classifiers, have limitations in resource-poor languages such as Isthmus Zapotec. Advanced models, such as RNNs and transformers, require large volumes of data, which are often difficult to obtain in indigenous contexts. As an alternative, a 28-layer deep convolutional neural network is proposed, trained with 10 common phrases converted into spectrograms and artificially augmented. The model achieved 100% training accuracy and 96.2% validation accuracy. Although promising, the need to expand the dataset is highlighted. This work demonstrates the potential of deep learning to improve intercultural communication and preserve endangered indigenous languages.

Keywords: Intercultural communication; zapotec indigenous languages; spectral images; deep neural network.

1. Introducción

La lengua indígena zapoteco del Istmo, conocida localmente como Diidxazá, es una de las variantes más representativas de la familia lingüística zapoteca, hablada principalmente en la región del Istmo de Tehuantepec, Oaxaca, México. Esta lengua no solo cumple una función comunicativa, sino que se erige como pilar fundamental de la identidad, las tradiciones y la historia de los pueblos zapotecos.

Sin embargo, actualmente se encuentra en una situación crítica. De acuerdo con el Censo de Población y Vivienda 2020 del Instituto Nacional de Estadística y Geografía (INEGI, 2020), el número de hablantes de zapoteco del Istmo ha disminuido drásticamente en la última década, con reducciones que van del 1% y el 30% en municipios como Unión Hidalgo, Tehuantepec, Matías Romero, entre otros. Esta pérdida representa una amenaza para el patrimonio cultural y lingüístico de México, especialmente porque muchas personas mayores de estas comunidades, que no hablan español, se enfrentan a barreras de comunicación en contextos esenciales como los servicios médicos o administrativos (Delgadillo et al., 2020).

Esta dramática pérdida de hablantes representa una amenaza directa para el patrimonio cultural y lingüístico de México. La situación es particularmente crítica porque muchas personas mayores de estas comunidades, que no hablan español, se enfrentan a barreras de comunicación en contextos esenciales como los servicios médicos, administrativos y de atención social (Delgadillo et al., 2020). En particular, los adultos mayores monolingües en zapoteco enfrentan importantes barreras cuando intentan acceder a servicios básicos que requieren comunicación en español, lo que compromete significativamente su calidad de vida y acceso a atención adecuada. Por lo tanto, es crucial desarrollar herramientas de comunicación efectivas que puedan atender adecuadamente las necesidades de estos hablantes.

En respuesta a esta problemática, las tecnologías digitales se presentan como soluciones prometedoras para la preservación de las lenguas indígenas. Por ejemplo, plataformas de juegos serios y aplicaciones móviles con realidad aumentada han facilitado la enseñanza de lenguas ancestrales (Panamá-Mazhenda & Robles Bykbaev, 2024). De manera similar, el proyecto Etnoenglish Cultural Exchange promueve la revitalización de lenguas indígenas colombianas mediante contenidos digitales creados colaborativamente, integrando las TIC en la educación intercultural (Villa et al., 2024).

Entre 2010 y 2020, la brecha digital en México se ha reducido considerablemente, principalmente gracias a la adopción generalizada de teléfonos móviles, aunque las comunidades marginadas aún permanecen rezagadas. Paralelamente, investigaciones recientes han demostrado que las plataformas web pueden contribuir de manera efectiva a la revitalización de las lenguas indígenas (Medina et al., 2023).

Dentro de este panorama tecnológico, las redes neuronales convolucionales (CNN) son una solución eficaz para reconocer patrones complejos en señales acústicas mediante el análisis de espectrogramas, que son representaciones visuales en 2D de señales de audio (Franzoni, 2023). Un espectrograma muestra cómo varía la intensidad de las diferentes frecuencias de un sonido a lo largo del tiempo, convirtiendo una señal de audio en una imagen bidimensional que puede ser procesada por las CNN.

Las CNN aprovechan esta transformación aplicando filtros en capas convolucionales para identificar y extraer características relevantes, como patrones frecuenciales, texturas espectrales y formas específicas dentro de estos espectrogramas. Gracias a su estructura jerárquica y capacidad para aprender representaciones en distintos niveles de abstracción, las CNN son especialmente adecuadas para clasificar patrones vocales y lingüísticos en grabaciones de audio.

Por lo tanto, el presente trabajo explora el potencial de las CNN para la clasificación de frases en zapoteco del Istmo mediante la transformación de audios en espectrogramas, contribuyendo así a la preservación tecnológica de esta lengua indígena amenazada. Este enfoque permite convertir señales de audio en imágenes espectrales que pueden ser procesadas por la red neuronal, facilitando el reconocimiento automático de expresiones comunicativas. El modelo presentado tiene como objetivo principal realizar una clasificación automática de espectrogramas de audio correspondientes a frases comunes en zapoteco del Istmo, facilitando así la comunicación entre hablantes de esta lengua indígena y personas que no la dominan.

El documento está organizado de la siguiente forma: la Sección 2 expone los antecedentes y trabajos relacionados, proporcionando el contexto para el enfoque planteado. La Sección 3 describe el conjunto de datos utilizado y las muestras que contiene, incluyendo el proceso de transformación a espectrogramas. En la Sección 4 se explica la metodología, detallando la arquitectura de la red neuronal convolucional propuesta, las técnicas de preprocesamiento de datos y los parámetros de entrenamiento utilizados. En la Sección 5 se presentan los resultados experimentales, incluyendo métricas de precisión, matrices de confusión y análisis del rendimiento del modelo, seguidos de una interpretación de los hallazgos. Finalmente, la Sección 6 concluye el artículo, resumiendo las principales contribuciones y delineando perspectivas para futuras investigaciones.

2. Trabajos relacionados

Diversos estudios han demostrado que las redes neuronales convolucionales (CNN) son altamente eficaces para la clasificación de señales de audio, especialmente cuando estas se transforman en representaciones visuales como espectrogramas (Franzoni, 2023). El uso de espectrogramas permite que las CNN aprovechen su capacidad para detectar patrones espaciales, facilitando el reconocimiento de características acústicas relevantes en tareas como la clasificación de sonidos ambientales, música y habla.

Modelos basados en CNN han alcanzado precisiones superiores al 90% en conjuntos de datos estándar, y su rendimiento se potencia al combinarse con técnicas como Mel-Frequency Cepstral Coefficients (MFCC) o espectrogramas Mel (Mushtaq et al., 2021). Además, la transferencia de aprendizaje con arquitecturas preentrenadas en imágenes, como ResNet o DenseNet, ha demostrado ser efectiva para tareas de audio, incluso cuando los datos disponibles son limitados (Zaman et al., 2023).

Por otra parte, se han desarrollado modelos híbridos que combinan CNN con unidades recurrentes controladas (GRU), mejorando la captura de características espaciales y temporales en señales de audio. Estos modelos han sido aplicados con éxito en la identificación de hablantes (Ye & Yang, 2021). Asimismo, los sistemas de identificación

de idiomas han utilizado estas técnicas, alcanzando alta precisión incluso con segmentos cortos de audio (Rammo & Al-Hamdani, 2022).

En el ámbito de la clasificación de dialectos y lenguas regionales, recientes investigaciones han reportado resultados prometedores mediante enfoques de aprendizaje profundo. Por ejemplo, Al-Anzi & Thankaleela (2025) utilizaron un modelo preentrenado basado en redes neuronales convolucionales 1D y 2D para identificar dialectos árabes, clasificando clips de audio en dialectos como árabe egipcio, árabe del Golfo, árabe levantino, árabe moderno estándar y árabe saudí, entre otros, logrando una precisión de prueba del 94.28% y de validación del 95.55%.

Por su parte, Lai et al. (2024) emplearon coeficientes cepstrales en frecuencia Mel (MFCC) y una red neuronal convolucional 1D para extraer y clasificar características de género y dialecto regional en señales de habla, alcanzando precisiones superiores al 90% y altos valores en métricas de recall y F1 score, demostrando eficacia y robustez en reconocimiento automático.

En cuanto al reconocimiento de emociones en dialecto árabe saudí, Aljuhani et al. (2021) utilizaron MFCC y espectrogramas Mel para extraer características, aplicando clasificadores SVM, MLP y KNN, con SVM alcanzando la mejor precisión del 77.14%, mejorando el reconocimiento automático de emociones en árabe.

Sin embargo, una revisión sistemática destacó sesgos en la investigación, como la preferencia por variedades regionales más estudiadas y la escasez de recursos para dialectos urbanos y lenguas vernáculos (Elnagar et al., 2021). Estos vacíos evidencian la necesidad de ampliar los estudios hacia lenguas y dialectos menos explorados, como las lenguas indígenas.

En el contexto específico de lenguas indígenas y preservación lingüística, Salau et al. (2020) implementaron un modelo CNN-LSTM para la clasificación de acentos en tres lenguas indígenas nigerianas, utilizando características espectrales extraídas de grabaciones de audio. Su trabajo demuestra la viabilidad de aplicar técnicas de aprendizaje profundo a lenguas con recursos limitados, logrando una precisión del 78% en la clasificación de variantes dialectales. Esta investigación establece un precedente importante para el uso de CNN en el análisis de lenguas indígenas.

Asimismo, Noda et al. (2015) desarrollaron un sistema audio-visual basado en CNN aplicado a espectrogramas, logrando una precisión del 89% en tareas de reconocimiento de fonemas. Este enfoque subraya la riqueza informativa de las representaciones espectrales para la clasificación de patrones vocales, lo que es especialmente relevante para lenguas con características fonológicas específicas.

Finalmente, en el ámbito de la preservación de lenguas amenazadas mediante tecnología, Dueck (2024) presenta un marco comprehensivo para el uso de inteligencia artificial en la preservación de historias orales indígenas, enfatizando la importancia de sistemas automáticos de reconocimiento de habla adaptados a las particularidades fonológicas y estructurales de estas lenguas.

Los estudios revisados en esta sección demuestran la eficacia de las redes neuronales convolucionales aplicadas a espectrogramas en diversas aplicaciones de análisis de señales de audio, particularmente en tareas de clasificación y reconocimiento de habla.

Sin embargo, se destaca la limitada investigación específica sobre la aplicación de CNN a espectrogramas para la clasificación de frases completas en lenguas indígenas mesoamericanas como el zapoteco del Istmo, lo que sugiere un área de investigación valiosa y poco explorada que puede contribuir significativamente a la preservación de estas lenguas en peligro de extinción.

Por tanto, el presente trabajo busca abordar esta brecha, proponiendo un modelo basado en CNN para la clasificación de frases en zapoteco mediante el análisis de espectrogramas, aprovechando las ventajas del aprendizaje profundo en contextos con recursos limitados.

3. Descripción del conjunto de datos

En este estudio, la calidad y la cantidad de los datos son fundamentales para el éxito de la clasificación de frases en zapoteco mediante redes neuronales convolucionales (CNN). Por lo tanto, es crucial describir detalladamente cómo se recolectaron, preprocesaron y aumentaron los datos utilizados en este trabajo. A continuación, se presenta una explicación paso a paso de este proceso.

El conjunto de datos está formado por 5,055 grabaciones de 10 frases comunes en zapoteco del Istmo, seleccionadas para representar situaciones cotidianas de solicitud de ayuda y necesidades básicas de comunicación. Cada frase cuenta con entre 465 y 570 muestras. Esta base es fundamental para entrenar y evaluar la red neuronal convolucional (CNN). La distribución detallada se presenta en la Tabla 1.

Frase en español	Frase en zapoteco	Número de muestras
Tengo hambre	Candaana'	480
Tengo frio	Cayaca' nanda'	525
Tengo sed	Cayate' nisa	570
Tengo sueño	Cadi candaana dia'	495
No tengo Hambre	Cadi Cayaca nanda dia'	465
No tengo frio	Cadi Cayate dia' nisa	525
No tengo sed	Hracaladxe' gahua'	525
Quiero comer	Hracaladxe' nisa	495
Quiero agua	Hracaladxe' gaaze'	495
Quiero bañarme	Dxa' bacaanda lua'	480

Tabla 1 – Lista de frases utilizadas en el conjunto de datos.

Para organizar los datos, se creó una carpeta por cada frase (o clase), asignando un nombre que identifica la frase correspondiente, como se muestra en la Figura 1. Este paso aseguró una estructura clara para el manejo de los datos durante el entrenamiento del modelo.

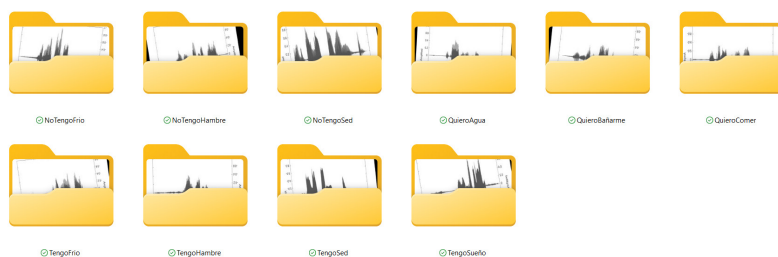


Figura 1 – Carpetas por clase del conjunto de datos.

Para que las CNN procesen las grabaciones de audio, estas se transformaron en espectrogramas utilizando MATLAB. Un espectrograma es una representación visual que muestra cómo varían las frecuencias de un audio a lo largo del tiempo, con las amplitudes indicadas mediante colores o intensidades. Estas imágenes capturan características acústicas clave, como tonos y patrones fonéticos, esenciales para distinguir frases en una lengua tonal como el zapoteco.

Posteriormente, los espectrogramas se convirtieron a escala de grises y se estandarizaron a un tamaño uniforme de 500×500 píxeles para simplificar el procesamiento y garantizar la consistencia en las entradas del modelo, lo cual es esencial para el entrenamiento efectivo de la CNN. Un ejemplo de espectrograma generado se presenta en la Figura 2.

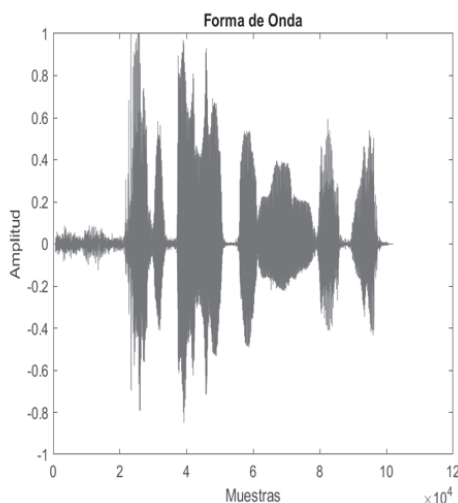


Figura 2 – Espectrograma de frase en zapoteco del Istmo

Dado que el número inicial de grabaciones por frase era limitado, se aplicó un aumento de datos para enriquecer el conjunto y mejorar la robustez del modelo. Utilizando un

software especializado para análisis espectral, se realizaron transformaciones en los espectrogramas, como rotaciones y traslaciones, para generar nuevas versiones de las imágenes sin alterar las características acústicas esenciales. Por ejemplo, una rotación de la imagen no cambia el contenido fonético del audio, pero introduce variabilidad que ayuda al modelo a generalizar mejor. La Tabla 2 detalla la cantidad de imágenes generadas por clase tras este proceso, mostrando un conjunto de datos más amplio y diverso.

Frases	Clase	Número de muestras
<i>Tengo hambre</i>	TengoHambre	960
<i>Tengo frío</i>	TengoFrio	1050
<i>Tengo sed</i>	TengoSed	1140
<i>Tengo sueño</i>	TengoSueño	990
<i>No tengo Hambre</i>	NoTengoHambre	930
<i>No tengo frío</i>	NoTengoFrio	1050
<i>No tengo sed</i>	NoTengoSed	1050
<i>Quiero comer</i>	QuieroComer	990
<i>Quiero agua</i>	QuieroAgua	990
<i>Quiero bañarme</i>	QuieroBañarme	960

Tabla 2 – Información del conjunto de datos.

Este proceso permitió crear un conjunto de datos robusto y estructurado, partiendo de un total inicial de 5,550 grabaciones. Tras aplicar aumento de datos mediante transformaciones en los espectrogramas, como rotaciones y traslaciones, se duplicó aproximadamente el número de muestras por clase, alcanzando un total de 10,110 grabaciones. Esta ampliación enriqueció la diversidad del conjunto, optimizando el rendimiento de la red neuronal y fortaleciendo la preservación y comunicación en esta lengua indígena.

4. Metodología

En este estudio se exploraron diversas arquitecturas de redes neuronales convolucionales (CNN) para optimizar la precisión en la clasificación de audios de frases en zapoteco del Istmo, evaluando variaciones en el número de capas, estructuras simétricas y asimétricas, y estrategias de reducción dimensional. Tras un análisis comparativo, se seleccionó una arquitectura eficiente de 28 capas, que inicia con una capa de entrada de imágenes de 500 por 500 píxeles con un canal.

La red incluye seis bloques convolucionales con sus respectivas capas de normalización por lotes, funciones de activación ReLU y capas de max pooling que reducen progresivamente la dimensión espacial y aumentan la profundidad de los mapas de características. Estas capas convolucionales tienen tamaños que van desde 500 x 500 x 8 hasta 10 x 10 x 256.

Finalmente, la arquitectura incluye dos capas totalmente conectadas y una capa softmax para la clasificación, con salida de tamaño 1 por 1 por 2, correspondiente a las clases definidas. Esta red profunda está diseñada específicamente para maximizar el rendimiento mediante un proceso efectivo de extracción de características y clasificación de frases en zapoteco. La Figura 3 ilustra el diagrama general de esta arquitectura.

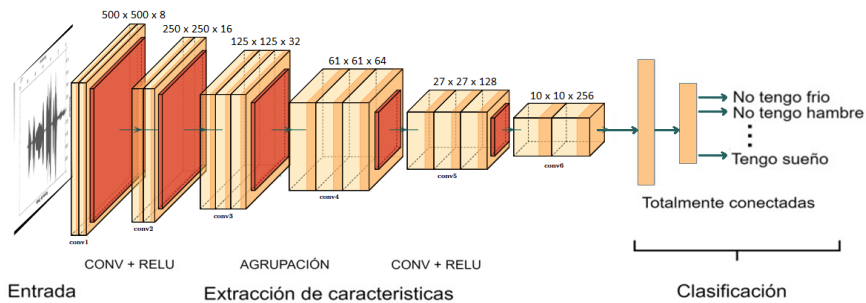


Figura 3 – Arquitectura de CNN para clasificación de espectrogramas

Las capas convolucionales son fundamentales para extraer características relevantes de los espectrogramas de audio. La arquitectura propuesta incluye seis capas convolucionales con un número creciente de filtros y tamaños de núcleo, lo que permite detectar patrones cada vez más complejos. La primera capa procesa espectrogramas en escala de grises de 500×500 píxeles utilizando ocho filtros de tamaño 3×3 . Las siguientes capas aplican filtros de 16, 32, 64, 128 y 256, incrementando el tamaño de los núcleos hasta 6×6 en las capas más profundas. Esta progresión mejora la capacidad de la red para captar características acústicas abstractas. La figura 4 muestra ocho kernels representativos de cada una de estas capas convolucionales.

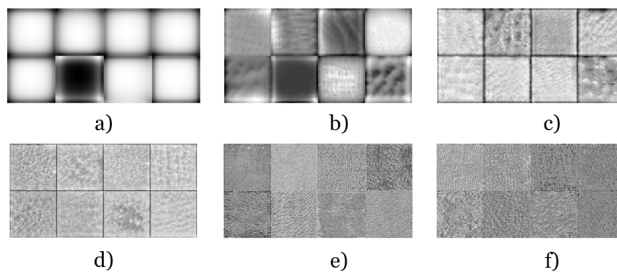


Figura 4 – Extracción de características en las capas: a) convolución 1, b) convolución 2, c) convolución 3, d) convolución 4, e) convolución 5 y f) convolución 6.

Las capas de activación ReLU introducen no linealidad al eliminar valores negativos y conservar los positivos, ayudando a la red a aprender las variaciones tonales propias del

zapoteco. El modelo incluye seis capas ReLU. La figura 5 ilustra algunos de los kernels asociados a estas capas ReLU.

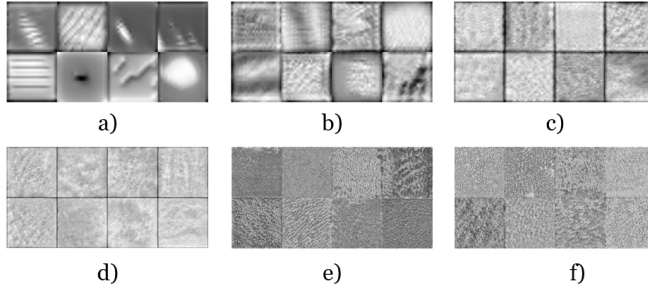


Figura 5 – Extracción de características en las capas: a) ReLU 1, b) ReLU 2, c) ReLU 3, d) ReLU 4, e) ReLU 5 y f) ReLU 6.

Por otro lado, las operaciones de max pooling, aplicadas en cinco etapas, reducen la dimensionalidad al seleccionar el valor máximo en regiones de 2×2 , preservando las características esenciales y evitando el sobreajuste. La figura 6 presenta ocho kernels de algunas de estas capas de max pooling.

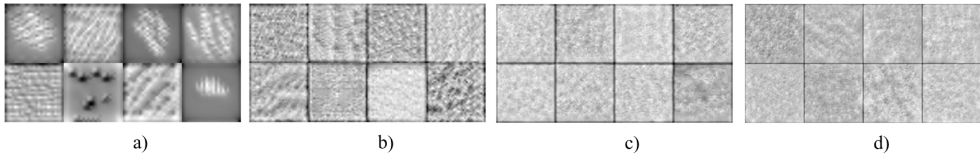


Figura 6 – Extracción de características en las capas: a) MaxPooling 1, b) MaxPooling 2, c) MaxPooling 3, d) MaxPooling 4.

5. Resultados

Esta sección presenta los resultados de la evaluación del modelo de red neuronal convolucional (CNN) entrenado para clasificar frases en zapoteco utilizando espectrogramas. El análisis incluye la distribución de los datos, los parámetros de entrenamiento, la precisión del modelo y las métricas de rendimiento basadas en matrices de confusión. El objetivo es evaluar la capacidad del modelo para generalizar y clasificar con precisión espectrogramas no vistos correspondientes a frases habladas en zapoteco.

Para el entrenamiento del modelo, el conjunto de datos se dividió en dos subconjuntos: un conjunto de entrenamiento que contiene el 80% de las imágenes y un conjunto de validación que contiene el 20% restante.

Clase	Entrenamiento	Validación
NoTengoFrio	840	210
NoTengoHambre	744	186
NoTengoSed	840	210
QuieroAgua	792	198
QuieroBañarme	768	192
QuieroComer	792	198
TengoFrio	840	210
TengoHambre	768	192
TengoSed	912	228
TengoSueño	792	198
Total	8088	2022

Tabla 4 – Número de muestras en conjunto de validación y entrenamiento.

Los parámetros utilizados para el proceso de entrenamiento fueron los siguientes: se empleó el algoritmo de gradiente descendente estocástico, un tamaño de lote de 40 muestras, un máximo de 20 épocas, una tasa de aprendizaje inicial de 0.0001 y una frecuencia de validación de 6 iteraciones. El entrenamiento se llevó a cabo durante 20 épocas, con 202 iteraciones en cada una, sumando un total de 4040 iteraciones en aproximadamente 627 minutos, utilizando un conjunto de 8088 imágenes. Estas pruebas se realizaron en una computadora equipada con un procesador Intel i5, GPU NVIDIA GeForce RTX 3050 y 32 GB de memoria RAM. El resultado obtenido fue un porcentaje de precisión del 96.24%, como se muestra en la figura 7.

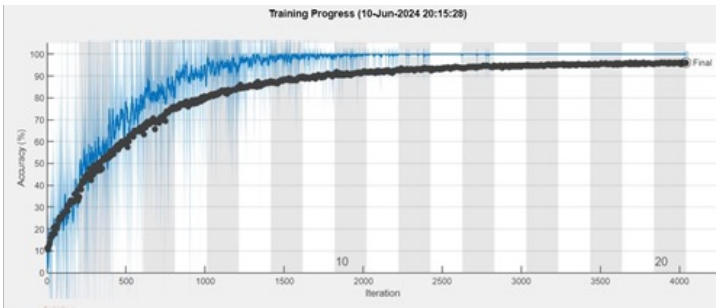


Figura 7 – Curva de aprendizaje resultante del proceso de entrenamiento.

Para evaluar el desempeño del modelo se usaron matrices de confusión en los conjuntos de entrenamiento y validación. Se analizaron tres métricas clave: Verdaderos Positivos (VP), que indican frases correctamente clasificadas; Falsos Positivos (FP), que son frases de otras clases erróneamente asignadas a una clase; y Falsos Negativos (FN),

que corresponden a frases mal clasificadas fuera de su clase original. Estas métricas permiten medir con detalle la precisión del modelo en la clasificación de las 10 clases de frases en zapoteco del Istmo.

En el conjunto de entrenamiento, el modelo alcanzó una precisión del 100%, clasificando correctamente todas las 8088 imágenes de espectrogramas correspondientes a las 10 frases en zapoteco del Istmo. Como se muestra en la Figura 8, la matriz de confusión presenta valores altos en la diagonal principal, indicando que no hubo errores de clasificación. Este rendimiento perfecto sugiere que la red neuronal convolucional (CNN) aprendió eficazmente los patrones acústicos de los espectrogramas. Sin embargo, una precisión del 100% en el conjunto de entrenamiento puede indicar sobreajuste, lo que resalta la importancia de evaluar el modelo con datos no vistos en el conjunto de validación para garantizar su capacidad de generalización.

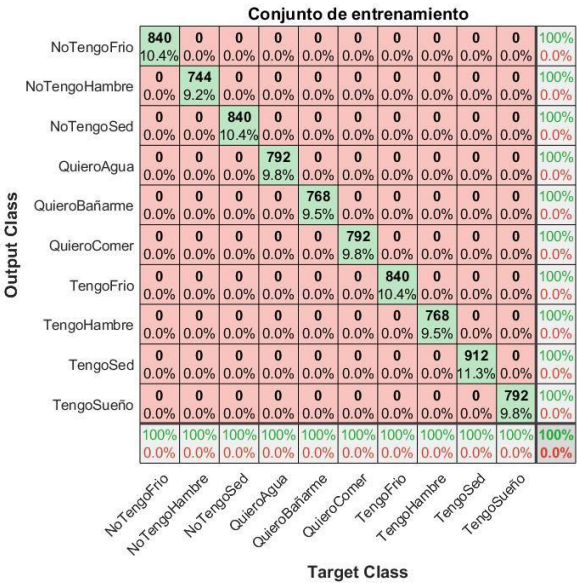


Figura 8 – Matriz de confusión del conjunto de entrenamiento.

Por otra parte, en el conjunto de validación, un total de 199 imágenes de la clase NoTengoFrio fueron clasificadas correctamente dentro de la misma clase, representando casos verdaderos positivos. Mientras tanto, 11 imágenes de otras clases fueron clasificadas incorrectamente como NoTengoFrio, correspondiendo a falsos positivos, y 5 imágenes que pertenecían a esta clase fueron clasificadas erróneamente como otras, representando falsos negativos. Del mismo modo, para la clase NoTengoHambre, el modelo identificó 180 verdaderos positivos, junto con 6 falsos positivos y 5 falsos negativos. La matriz de

confusión que resume estos resultados se muestra en la Figura 9, donde se observa una exactitud del 96,2% en el conjunto de validación.

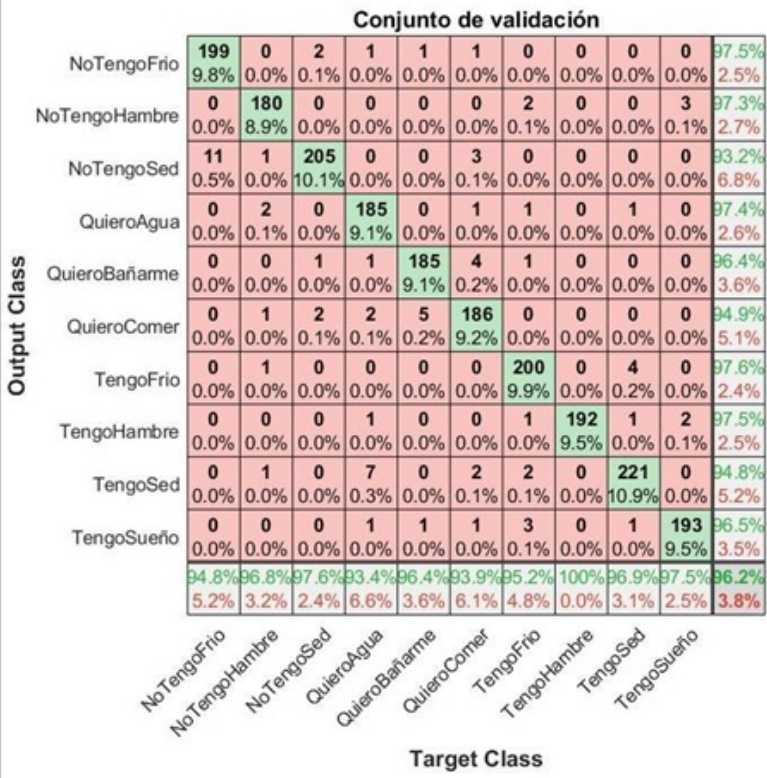


Figura 9 – Matriz de confusión del conjunto de validación.

En total, el conjunto de datos produjo un total de 1946 verdaderos positivos, 76 falsos positivos y 76 falsos negativos, como se detalla en la Tabla 5.

Clase	VP	FP	FN
NoTengoFrio	199	11	5
NoTengoHambre	180	6	5
NoTengoSed	205	5	15
QuieroAgua	185	13	5
QuieroBañarme	185	7	7
QuieroComer	186	12	10
TengoFrio	200	10	5

Clase	VP	FP	FN
TengoHambre	192	0	5
TengoSed	221	7	12
TengoSueño	193	5	7
Total	1946	76	76

Tabla 5 – Resultados de la matriz de confusión del conjunto de validación.

Además, se calcularon la precisión, la sensibilidad y la especificidad del modelo para proporcionar una evaluación más completa de su rendimiento. La precisión mide la proporción de predicciones correctas, calculada como $(TP + TN) / (TP + TN + FP + FN)$. La sensibilidad (o recuerdo) refleja la capacidad del modelo para identificar correctamente los casos positivos y se calcula como $TP / (TP + FN)$. La especificidad indica la capacidad del modelo para identificar correctamente los casos negativos y se calcula como $TN / (TN + FP)$. Estas métricas se resumen en la Tabla 6.

Conjunto	Exactitud	Sensibilidad	Especificidad
Entrenamiento			
99.32%	100%	100%	100%
99.29%			
99.35%			
Validación	96.21%	96.23%	96.23%

Tabla 6 – Resultados Exactitud, Sensibilidad y Especificidad

6. Conclusiones

La preservación y facilitación de la comunicación en lenguas indígenas, como el zapoteco del Istmo, es fundamental para mejorar la integración y calidad de vida de las comunidades indígenas. En este estudio, se propuso un modelo de red neuronal convolucional (CNN) de 28 capas para la clasificación de espectrogramas de 10 frases en zapoteco del Istmo. El modelo fue entrenado y validado con un conjunto de datos que contiene 10,110 imágenes (8088 de entrenamiento y 2022 de validación), generadas a partir de un número limitado de audios originales, normalizadas y ajustadas a un tamaño de 500 x 500 píxeles.

Los resultados demuestran que el modelo CNN propuesto clasifica con alta precisión frases en zapoteco a partir de espectrogramas, alcanzando un 100% de exactitud en entrenamiento y un 96.2% en validación. Aunque el conjunto de datos es pequeño y limitado a 10 frases, estos resultados son comparables e incluso superiores a los obtenidos en otros trabajos de clasificación de audio con aprendizaje profundo en contextos de recursos limitados.

De manera similar, Telmem et al. (2025) lograron un 85% de precisión en el reconocimiento del habla amazigh utilizando CNNs basadas en espectrogramas. Wang et al. (2020) alcanzaron un 92.1% en la clasificación de escenas acústicas combinando

múltiples representaciones espectrales. De forma similar, Binjaku et al. (2022) desarrollaron un sistema para identificar el idioma albanés en grabaciones de voz mediante espectrogramas y CNN, obteniendo una precisión del 94%.

Otros estudios relevantes incluyen a Dwivedi et al. (2023), quienes lograron más del 95% de precisión en la identificación de tartamudez en habla hindi usando modelos secuenciales de CNN entrenados con espectrogramas. Lesnichaia et al. (2022) alcanzaron entre 96.4% y 98.7% de precisión en la clasificación de acentos en inglés, empleando CNNs con mel-espectrogramas de amplitud en escala lineal, superando modelos previos con el mismo conjunto de datos. Por su parte, Dayal et al. (2022) lograron un 95.2% de precisión en la clasificación de 11 sonidos de fondo incrustados en señales de habla humana, utilizando un modelo CNN ligero basado en espectrogramas, que además superó a modelos de referencia en tiempo de inferencia.

Por su parte, Demir et al. (2020) desarrollaron dos modelos, uno basado en redes neuronales artificiales (ANN) y otro en CNN, para la identificación del idioma albanés a partir de espectrogramas, alcanzando precisiones del 85% y 94%, respectivamente.

Una limitación importante de este estudio es el reducido tamaño del conjunto de audios originales por clase, lo cual puede afectar la capacidad de generalización del modelo CNN en la clasificación de frases en zapoteco. La literatura especializada destaca que tanto el tamaño como la diversidad del conjunto de datos son factores clave para el rendimiento y la robustez de los modelos de aprendizaje profundo aplicados a la clasificación de audio.

Diversos estudios recientes han confirmado que aumentar el tamaño del dataset mejora la precisión y la generalización de los modelos CNN, aunque este beneficio tiende a estabilizarse tras cierto umbral (Paul et al., 2025; Shoumy et al., 2021). Además, no solo la cantidad sino también la diversidad de los datos —incluyendo diferentes hablantes, entonaciones y contextos acústicos— resulta fundamental para que el modelo sea efectivo en aplicaciones prácticas, como sistemas de traducción o asistencia en comunidades monolingües (Paul et al., 2025; Pandian et al., 2024).

Esta limitación puede reducir la capacidad del modelo para distinguir entre las variantes del zapoteco, afectando su precisión y utilidad en aplicaciones prácticas. Por ello, es necesario contar con un conjunto de datos amplio y representativo que incluya la diversidad lingüística y acústica de estas variantes para mejorar la generalización del modelo.

En futuras investigaciones, se recomienda explorar técnicas modernas como redes neuronales recurrentes o transformadores para mejorar la precisión y capacidad de generalización de nuestro modelo multiclase. Además, contar con un conjunto de datos más amplio y diverso que refleje la variedad lingüística y acústica del zapoteco fortalecerá su aplicabilidad en contextos reales.

Referencias

- Al-Anzi, F. S., & Thankaleela, B. S. S. (2025). Region-Wise Recognition and Classification of Arabic Dialects and Vocabulary: A Deep Learning Approach. *Applied Sciences*, 15(12), 6516. <https://doi.org/10.3390/app15126516>

- Aljuhani, R. H., Alshutayri, A., & Alahdal, S. (2021). Arabic speech emotion recognition from Saudi dialect corpus. *IEEE Access*, 9, 127081-127085. <https://doi.org/10.1109/ACCESS.2021.3110992>
- Binjaku, K., Janku, J., & Meçe, E. K. (2022). Identifying Low-Resource Languages in Speech Recordings through Deep Learning. In *2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)* (pp. 1-6). IEEE. <https://doi.org/10.23919/SoftCOM55329.2022.9911376>
- Dayal, A., Yeduri, S. R., Koduru, B. H., Jaiswal, R. K., Soumya, J., Srinivas, M. B., & Cenkeramaddi, L. R. (2022). Lightweight deep convolutional neural network for background sound classification in speech signals. *The Journal of the Acoustical Society of America*, 151(4), 2773-2786. <https://doi.org/10.1121/10.0010257>
- Delgadillo, L. G., Arce, J. & Pastrana, S. A. (2020). Vulnerabilidad de la lengua en hablantes indígenas, el caso de México. *Circula*, (12), 19-40. <https://doi.org/10.17118/11143/18441>
- Demir, F., Abdullah, D. A., & Sengur, A. (2020). A new deep CNN model for environmental sound classification. *IEEE Access*, 8, 66529-66537. <https://doi.org/10.1109/ACCESS.2020.2984903>
- Dueck, G. W. (2024). Using AI to help preserve indigenous oral histories. *Proceedings of the 2024 IEEE International Humanitarian Technologies Conference*, 1-6. <https://doi.org/10.1109/IHTC61819.2024.10855026>
- Dwivedi, S., Ghosh, S., & Dwivedi, S. (2023). Binary classifier for identification of stammering instances in Hindi speech data. *International Journal of Speech Technology*, 26(3), 765-774. <https://doi.org/10.1007/s10772-023-10046-9>
- Elnagar, A., Yagi, S. M., Nassif, A. B., Shahin, I., & Salloum, S. A. (2021). Systematic literature review of dialectal Arabic: identification and detection. *IEEE Access*, 9, 31010-31042. <https://doi.org/10.1109/ACCESS.2021.3059504>
- Franzoni, V. (2023). Cross-domain synergy: Leveraging image processing techniques for enhanced sound classification through spectrogram analysis using CNNs. *Journal of Autonomous Intelligence*, 6(3), 1-14. <https://doi.org/10.32629/jai.v6i3.678>
- INEGI (2020). Censo de Población y Vivienda 2020. <https://www.inegi.org.mx/programas/ccpv/2020/#documentacion>
- Lai, H. Y., Hu, C. C., Wen, C. H., Wu, J. X., Pai, N. S., Yeh, C. Y., & Lin, C. H. (2024). Mel-Scale Frequency Extraction and Classification of Dialect-Speech Signals with 1D CNN based Classifier for Gender and Region Recognition. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3430296>
- Lesnichaia, M., Mikhailava, V., Bogach, N., Lezhenin, Y., Blake, J., & Pyshkin, E. (2022). Classification of Accented English Using CNN Model Trained on Amplitude Mel-Spectrograms. In *Interspeech* (pp. 3669-3673). <https://doi.org/10.21437/Interspeech.2022-462>

- Medina, M. A. G., Jiménez, J. L. M., Meza, D. D. J. A., Vergara, J. T., & Cantero, C. L. (2023). Resignificación de la lengua materna Zenú mediante la plataforma web Tozì. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (E59), 24-38. <https://dialnet.unirioja.es/servlet/articulo?codigo=10079679>
- Mushtaq, Z., Su, S. F., & Tran, Q. V. (2021). Spectral images based environmental sound classification using CNN with meaningful data augmentation. *Applied Acoustics*, 172, 107581. <https://doi.org/10.1016/j.apacoust.2020.107581>
- Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H. G., & Ogata, T. (2015). Audio-visual speech recognition using deep learning. *Applied Intelligence*, 42(4), 722-737. <https://doi.org/10.1007/s10489-014-0629-7>
- Panamá-Mazhenda, K., & Robles-Bykbaev, V. (2024). Revisión sistemática de literatura de metodologías para el diseño y desarrollo de juegos serios: análisis MLR. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E66), 515-527. <https://dialnet.unirioja.es/servlet/articulo?codigo=10003357>
- Pandian, J. A., Thirunavukarasu, R., & Kotei, E. (2024). A novel convolutional neural network model for automatic speaker identification from speech signals. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3385858>
- Paul, B., Phadikar, S., Bera, S., Dey, T., & Nandi, U. (2025). Isolated word recognition based on a hyper-tuned cross-validated cnn-bilstm from mel frequency cepstral coefficients. *Multimedia Tools and Applications*, 84(17), 17309-17328. <https://doi.org/10.1007/s11042-024-19750-3>
- Rammo, F. M., & Al-Hamdani, M. N. (2022). Detecting the speaker language using CNN deep learning algorithm. *Iraqi Journal for Computer Science and Mathematics*, 3(1), 43-52. <https://doi.org/10.52866/ijcsm.2022.01.01.005>
- Salau, A. O., Olowoyo, T. D., & Akinola, S. O. (2020). Accent classification of the three major Nigerian indigenous languages using 1D CNN LSTM network model. In *Advances in Computational Intelligence and Robotics* (pp. 1-15). Springer. https://doi.org/10.1007/978-981-15-2620-6_1
- Shoumy, N. J., Ang, L. M., Rahaman, D. M. M., Zia, T., Seng, K. P., & Khatun, S. (2021). Augmented Audio Data in Improving Speech Emotion Classification Tasks. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, 12799 LNAI, 360-365. https://doi.org/10.1007/978-3-030-79463-7_30
- Telmeme, M., Laaidi, N., & Satori, H. (2025). The impact of MFCC, spectrogram, and Mel-Spectrogram on deep learning models for Amazigh speech recognition system. *International Journal of Speech Technology*, 1-14. <https://doi.org/10.1007/s10772-025-10183-3>
- Villa, M. G. R., Zapata, J. A. S., Ospina-Giraldo, M. N., Holguin, M. M. O., Cataño, D. F. G., & Buitrago, J. D. R. (2024). Proyecto Etnoenglish Cultural Exchange: Una Experiencia Pedagógica de Digiculturalidad y Educación Inclusiva en la escuela. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (E72), 370-381. <https://dialnet.unirioja.es/servlet/articulo?codigo=9852073>

- Wang, H., Zou, Y., & Chong, D. (2020). Acoustic scene classification with spectrogram processing strategies. *arXiv preprint arXiv:2007.03781*. <https://doi.org/10.48550/arXiv.2007.03781>
- Ye, F., & Yang, J. (2021). A deep neural network model for speaker identification. *Applied Sciences*, 11(8), 3603. <https://doi.org/10.3390/APP11083603>
- Zaman, K., Sah, M., Direkoglu, C., & Unoki, M. (2023). A survey of audio classification using deep learning. *IEEE Access*, 11, 106620-106649. <https://doi.org/10.1109/ACCESS.2023.3318015>

Revisión sistemática de juegos virtuales en conocimiento numérico y científico para 1° a 5° de primaria

Brehiner Alfredo Martinez Mora¹, Marlenny Guevara Guerrero²

brehmert.martinez@campusucc.edu.co; marlenny.guevara@correounivalle.edu.co

¹ Universidad Cooperativa de Colombia, Carrera 13 # 37-43 piso 8, Edificio Cavipetrol. Bogotá, Colombia.

DOI: 10.17013/risti.59.53-67

Resumen: El presente trabajo indaga como los juegos en formato virtual contribuyen al desarrollo cognitivo en las áreas de conocimiento matemático y científico dirigidos a grados de 1° a 5° de primaria. Se realizó una revisión de literatura entre el periodo (2017 – 2022) en tres revistas referentes en publicaciones de tecnología y educación. Como resultado, se identificaron que la gran mayoría de estudios usaron diseños de investigación cuasi experimental y frente a las situaciones mediadas por TIC hubo una tendencia hacia los video juegos generalistas. El análisis de los estudios (diseños de investigación - videojuegos) evidenció que el uso de las TIC se relaciona a los aspectos educativos, dejando de lado un poco, la discusión acerca del desarrollo cognitivo.

Palabras-clave: Tecnología educacional; video juegos; diseños; desarrollo de habilidades; procesos cognitivos.

Systematic review of virtual games in numerical and scientific knowledge for grades 1 to 5

Abstract: The present work investigates how games in virtual format contribute to cognitive development in the areas of mathematical and scientific knowledge aimed at grades 1 to 5 of primary school. A literature review was carried out during the period (2017 - 2022) in three leading technology and education publications journals. As a result, it was identified that the vast majority of studies used quasi-experimental research designs, and in the face of ICT-mediated situations, there was a tendency towards generalist video games. The analysis of the studies (research designs - video games) showed that the use of ICT is related to educational aspects, leaving aside the discussion about cognitive development.

Keywords: Educational technology; video games; designs; skill development; cognitive processes.

1. Introducción

El uso de las tecnologías de la información y la comunicación en los procesos de enseñanza y aprendizaje, se han incrementado en los dos últimos años, debido a la emergencia sanitaria COVID 19 (CEPAL -UNESCO, 2020). Cabe aclarar, que la implementación de las TIC en el contexto educativo no solo tuvo lugar a raíz de la pandemia, sino que se ha venido implementando durante los últimos diez años, las instituciones educativas requieren una inversión económica alta, en infraestructura tecnológica, dispositivos móviles y software, por lo tanto, esto limita ofrecer a los docentes procesos de formación orientado a las TIC (Manches y Plowman, 2015). El manejo de las TIC, por parte de los profesores se ha convertido en un desafío, porque no siempre los docentes cuentan con procesos de formación orientados al diseño de actividades pedagógicas mediadas por herramientas tecnológicas o plataformas digitales. A esto se le suma, que las creencias y actitudes de los profesores frente a los recursos digitales, inciden en el diseño y planeación de prácticas pedagógicas basadas en la tecnología (Clements, 2002; Lara et al., 2021; Mertala, 2017).

Ante el auge de la tecnología y las restricciones de los docentes para acceder a ellas y usarlas con propósitos educativos (Hatzigianni y Kalaitzidis, 2018). Esto ha conllevado al surgimiento de lineamientos o directrices para aplicar las TIC a favor de la calidad de las prácticas educativas (UNESCO, 2020). Sin embargo, es necesario mencionar, que realizar actividades académicas que implican lo virtual, no es sinónimo de adquisición de conocimiento. Es por ello, que se requiere integrar lo tecnológico y los objetivos de aprendizaje, aludiendo, a la elaboración de un diseño previo, que tenga como finalidad, mejorar habilidades o desempeños en la práctica pedagógica de los docentes y en el aprendizaje de los estudiantes (Beltrán et al., 2021; Laranjeiro, 2021).

Integrar las TIC a los procesos educativos implica tener en cuenta una serie de elementos que consideren la elección de plataformas digitales, los aspectos curriculares a abordar y la práctica pedagógica que promueva aprendizaje autónomo, resolución de problemas y seguimiento permanente al desempeño del estudiante (Dubé y Dubé, 2020; Flynn et al., 2021; Mayer, 2016). Con respeto, a la elección de plataformas digitales, Dubé y Dubé (2020) proponen características en la preparación de los profesores para implementar situaciones mediadas por la virtualidad en el salón de clase: la alfabetización mediática, la cual, consiste en la formación dirigida a los profesores para diferenciar entre los juegos diseñados para enseñar y los juegos recreativos. Y, por otra parte, la transferencia de aprendizaje, la cual, consiste en brindar directrices a los profesores para adoptar juegos virtuales que sean efectivos al promover habilidades en un dominio de conocimiento. Frente al tema de la práctica pedagógica, Flynn et al. (2021) proponen tres características para realizar una actividad virtual en el aula de clase: 1) la claridad en el nuevo conocimiento que se va a implementar en la situación virtual, 2) indicar el objetivo de la actividad virtual y 3) la identificación de cuáles y que tipo de ayudas se le van a brindar al estudiante durante la situación virtual. En términos generales, el uso de las TIC implica una planeación por parte del docente para llevar a cabo la práctica pedagógica y generar procesos de aprendizaje en sus estudiantes. De modo similar, en el tema de los aspectos curriculares, Mayer (2016) sugiere cuatro aspectos para tener en cuenta para implementar las TIC en los contextos educativos: 1) hacer un proceso de sensibilización a los actores académicos para transformar prácticas pedagógicas tradicionales a prácticas

educativas centradas en lo virtual. 2) Los planes de estudio deben brindar claridad sobre los juegos virtuales mencionando los objetivos de aprendizaje y procesos cognitivos implicados en la situación virtual. 3) Las actividades de aprendizaje mediadas por TIC deben ser usadas para fortalecer metas específicas de adquisición de conocimiento en el aula de clase y 4) seleccionar situaciones virtuales orientadas a impactar los resultados de aprendizaje de los estudiantes en lugar de acceder a su gusto por cierto tipo de juegos. En esta idea, el uso de las TIC implica una reflexión conjunta por parte de los directivos docentes y profesores acerca de cómo van a implementar las tecnologías educativas en el salón de clases, esto implica una evaluación de la institución educativa desde infraestructura tecnológica hasta el conocimiento que tienen los docentes al manejar las TICS, para desarrollar una propuesta que responda a las necesidades que requieren los estudiantes en cuanto a su aprendizaje.

En el tema de las TIC hay unas perspectivas teóricas que reflejan unos aspectos pedagógicos para que el estudiante adquiera el conocimiento, promoviendo el aprendizaje autónomo, el aprendizaje por descubrimiento, la resolución de problemas, entre otros. Para Abdsamad et al. (2020) la perspectiva Mobile – Learning, se soporta en el aprendizaje autónomo porque son actividades que generan conocimiento al interactuar con aplicaciones móviles, juegos virtuales o redes sociales. Frente al aprendizaje por descubrimiento Hwang et al. (2012) proponen Actividades de resolución de problemas basadas en la web, se refiere a un conjunto de pasos para resolver preguntas de un tema específico que exige identificar el origen del problema, definir las acciones para resolver el problema y hacer seguimiento a la resolución del problema. Con respecto a la perspectiva Aprendizaje basado en juegos, Hwang, et al. (2022) proponen situaciones de juego virtual que implican la resolución de un problema, a partir, de una instrucción que exige unos conocimientos previos y habilidades para desempeñarse en la situación. De modo similar, la Gamificación¹, se considera una estrategia de aprendizaje activo que exige desempeños específicos de acuerdo con las metas de aprendizaje propuestas (Zainuddina et al., 2020).

Con respecto a la investigación en TIC, se han llevado a cabo, estudios sistemáticos de literatura (Flynn et al., 2020; Zainddin et al., 2020) orientados a conocer los diseños metodológicos, más usados en las investigaciones para promover aprendizaje y habilidades cognitivas, en niños de grado escolar por medio del uso de TIC.

En ese sentido, Flynn et al. (2020) al revisar la literatura en las bases de datos Scopus, PsychInfo y Ovid Medline, en el periodo entre 2010-2019, encuentran que los diseños experimentales son los más usados para explorar el que y como aprenden los niños en las situaciones de juego virtual. Por su parte, Zainddin et al. (2020) al revisar la literatura en las bases de datos Science Direct, EbscoHost, Taylor & Francis Online, y Springer Link, en el periodo entre 2016 -2019, evidencian que las intervenciones con grupo control y experimental son las más usadas para evaluar la incidencia de las TIC en

¹ La diferencia entre Gamificación y Aprendizaje basado en juegos consiste en la elaboración de la actividad de aprendizaje. En el primer caso, usa el diseño del juego en actividades cotidianas motivando a los participantes a partir de puntos, medallas o desbloqueo de nuevo contenido. Mientras, en el segundo caso, al juego se le agrega problemas del mundo real que facilita al estudiante la comprensión y adquisición de conocimiento de situaciones aplicables a la vida real (Sjoberg y Arroyos, 2022).

el aprendizaje. En términos generales, se puede plantear que la evidencia mencionada anteriormente, indica que las TIC se convierten en un recurso pedagógico interesante para promover procesos de aprendizaje en las intervenciones educativas. Sin embargo, se dice poco acerca de la relación entre los diseños de investigación implementados en los estudios de TICS y la concepción o idea que muestran sobre el desarrollo cognitivo.

El propósito de la presente revisión de literatura está orientado a identificar y analizar cómo a partir del 2017 al 2022 se han implementado metodológicamente los juegos en formato virtual en el dominio de conocimiento numérico y científico en estudiantes de 1° a 5° de primaria. Para esto se toma como referencia las publicaciones de tres revistas de tecnología y educación (Computer & Education, British Journal of Education Technology, y Education Technology & Society), que se consideran referentes en la investigación en este campo de estudio, como resultado de la revisión bibliométrica de Hwang y Chen (2022) quienes reportaron alto índice de citación, la participación de autores más reconocidos en esta área de conocimiento y el abordaje de temas de aprendizaje basado en juegos, tales como: ambientes interactivos de aprendizaje, innovación educativa, aprendizaje asistido por computadora y tecnologías y desarrollo.

A partir del análisis metodológico de los estudios publicados en las tres revistas seleccionadas, la presente revisión indaga ¿Cómo los juegos en formato virtual contribuyen al desarrollo cognitivo en estudiantes de 1° a 5° grado de primaria, en el dominio de conocimiento numérico y científico?

De manera específica, para responder la pregunta que guía este estudio, la presente revisión busca identificar: (1) los diseños de investigación implementados en los juegos virtuales en el dominio de conocimiento numérico y científico; (2) los tipos de juegos en formato virtual empleados para el dominio de conocimiento numérico y científico; (3) las formas de presentar los juegos en formato virtual.

2. Método

La presente revisión de literatura toma como referencia el proceso de búsqueda sistemática de bibliografía propuesta por (Medina-López et al., 2010), a partir de la cual se llevaron a cabo las siguientes fases:

- I. Identificación del campo de estudio y periodo analizar: se definió como campo de estudio las “situaciones de juego en formato virtual” en los dominios de conocimiento numérico y científico, dentro del periodo 2017 – 2022.
- II. Selección de las fuentes de información: incluyeron tres revistas (Computer & Education, British Journal of Education Technology, y Education Technology & Society). La elección de estas revistas surge a partir del estudio realizado por Hwang y Chen (2022) quienes a través de un análisis bibliométrico consultando la base de datos Web of Science dentro del periodo 1990 a 2019, indicaron que de (34) revistas analizadas, (7) presentaron el mayor número de publicaciones e índice de citación
 - Computer & Education (citaciones: 15,210 y publicaciones: 305)
 - British Journal of Educational Technology (citaciones: 3,212 y publicaciones: 110)
 - Educational Technology & Society (citaciones: 2,106 y publicaciones: 110)

- Interactive Learning Environments (citaciones: 632 y publicaciones: 81)
 - Journal of Computer Assisted Learning (citaciones: 2,350 y publicaciones: 66)
 - Educational Technology Research and Development (citaciones: 2, 011 y publicaciones: 66)
 - Innovations in Education and Teaching International (citaciones:182 y publicaciones:11)
- III. Realización de la búsqueda: se realizaron búsqueda manuales y automáticas consultando directamente la página electrónica de cada revista, usando ocho descriptores de búsqueda: “digital games”, “video games”, “games interactive”, “digital games-based learning”, “educational computer games”, “digital games numerical cognition”, “digital games scientific reasoning”, y “digital games mathematical.
- IV. Gestión y depuración de los resultados de la búsqueda: para las tres revistas en las que se basa el presente estudio se llevó a cabo la lectura del resumen y metodología de cada uno de los artículos (2017-2022). Luego, se depuró la selección de artículos a partir de los siguientes criterios de inclusión: población de 1º a 5º grado de primaria, artículos de investigación que contienen situaciones de juego en formato virtual relacionados con los dominios de conocimiento numérico y científico, y sin restricción de país ni idioma). En consecuencia, se descartó investigaciones con población de estudiantes de bachillerato, educación superior, docentes o padres, artículos de revisiones temáticas, comunicaciones cortas, reportes de caso, informes breves de investigación o artículos teóricos y de reflexión.
- V. Análisis de los resultados: En esta etapa a partir de la lectura completa de los artículos seleccionados, se realiza una reseña de cada artículo que permite la elaboración de una matriz de datos en Excel, donde se consignó la información que hace alusión a los aspectos metodológicos a identificar (diseños de investigación, tipos de juego en formato virtual y formas de presentar los juegos). Luego, se procede a describir cada uno de estos aspectos y, por último, se realiza un análisis orientado a comprender cómo las situaciones de juegos en formato virtual aportan al desarrollo cognitivo de estudiantes de 1º a 5º de primaria en el conocimiento numérico y científico.

3. Resultados

De las tres revistas consultadas de tecnología y educación (Computer & Education, British Journal of Education Technology, y Education Technology & Society) entre el periodo 2017 – 2022 se identificaron un total de 2.257 artículos. En la búsqueda fueron seleccionados 28 artículos (originales completos) que cumplieran con los criterios de inclusión (ver Figura 1). De estos 12 fueron de Computer & Education; 8 de British Journal of Education Technology; y 8 de Education Technology & Society. Los artículos abordaron temas relacionados a investigaciones que implican juegos virtuales, modelado computacional, realidad aumentada y herramientas digitales, que son usadas para trabajar en los dominios de conocimiento numérico y científico. En ese sentido, la figura 1 indica el número de artículos seleccionados en cada dominio de conocimiento para las revistas.

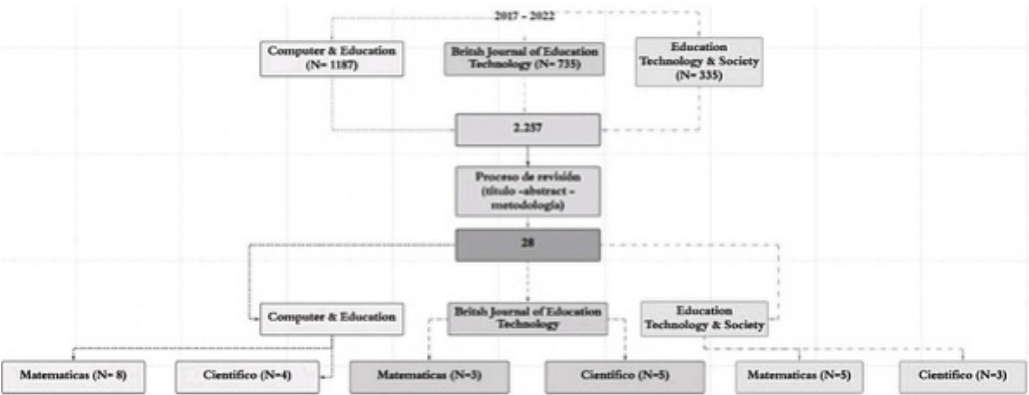


Figura 1 – Síntesis de artículos seleccionados por dominios de conocimiento para las revistas

A continuación, se presenta un balance de los estudios seleccionados respondiendo a los objetivos planteados en la presente revisión de literatura.

3.1. Diseños de investigación implementados en los juegos virtuales en el dominio de conocimiento numérico y científico

La descripción de los diseños de investigación empleados en los artículos seleccionados sobre TIC y educación se realizó a partir de León y Montero (2002) quienes consideraron el tipo de diseño, según el grado de manipulación de las variables, la posibilidad de asignación de los sujetos a los grupos y las veces que se mide una variable en el tiempo. De modo específico, de los 28 artículos seleccionados acerca de las situaciones de juego virtual para el dominio de conocimiento matemático y científico, se encontró la siguiente tendencia frente los tipos de diseños: el diseño cuasi experimental (n =15 artículos), el diseño experimental (n =7 artículos), el diseño intra-sujetos (n =4 artículos) y el diseño evolutivo (n= 2 artículos) (Tabla 1)

En el diseño cuasi experimental, se evidencio varios aspectos en términos de las características de los diseños: (1) el número de participantes en los estudios de tecnología y educación osciló entre 20 a 259 estudiantes; (2) todos los estudios realizaron intervenciones pretest y posttest; (3) todos los estudios usaron medidas repetidas en sus intervenciones; y (4) gran parte de los estudios trabajaron con grupo control y grupo experimental, excepto, los estudios de Dickes et al., (2019); Herodotou, (2018); Sun et al.,(2020) quienes trabajaron con un solo grupo de participantes y tomaron un subconjunto de la muestra, como grupo de referencia para observar los efectos de la intervención.

En el marco del diseño experimental, se evidencio varios aspectos en términos de las características de los diseños (1) el número de participantes en los estudios de tecnología y educación osciló entre 46 a 1808 estudiantes; (2) gran parte de los estudios realizaron intervenciones pretest y posttest, excepto, el estudio de Faber et al., (2017) quienes para medir la eficacia de un programa formativo digital tuvieron una muestra representativa de 1808 estudiantes de 3º de primaria, este alto número de participantes valido el diseño

de la intervención; (3) todos los estudios usaron medidas repetidas en sus intervenciones; y (4) todos los estudios trabajaron con grupo control y grupo experimental.

Por su parte, los diseños intra-sujeto, evidencio varios aspectos en términos de las características de los diseños: (1) el número de participantes en los estudios osciló entre 2 a 61 estudiantes; (2) se encontró diversidad en el tipo de intervención, algunas investigaciones realizaron intervenciones pretest y posttest (Luo et al., 2020; Sýkora et al., 2021), mientras, Mercier et al. (2017); Outhwaite et al. (2017) diseñaron intervenciones sin pretest y posttest; (3) todos los estudios usaron medidas repetidas en sus intervenciones; y (4), todos los estudios sustentaron la comparación de los grupos, a partir de variables, como, colegio, grado escolar y género.

Por último, los diseños evolutivos, evidenciaron varios aspectos en términos de las características de los diseños: (1) el número de participantes en los estudios osciló entre 38 a 147 estudiantes; (2) frente al tipo de intervención, Fallon (2019) realizó una intervención sin pretest y posttest, mientras I-Hua Chen et al. (20220) llevaron a cabo, la intervención con solo pretest; (3) los dos estudios usaron medidas repetidas; y (4) los estudios hicieron énfasis en comparar los cambios en el desempeño que presentaron los sujetos a lo largo de las sesiones.

Tipo de Diseño	Artículo	Tipos de Intervención	Conformación y Tamaño de la Muestra
Diseño experimental	Lai, A. F., Chen, C. H., & Lee, G. Y. (2019).	pre-post	Aleatorio 46
	Cheng, S. C., Hwang, G. J., & Chen, C. H. (2019).	pre-post	Aleatorio 50
	Faber JM, Luyten H. & Visscher AJ (2017).	Intervención	Aleatorio 1808
	Shao-Chen Chang, Ting-Chia Hsu, & Morris Siu-Yung Jong. (2020).	pre-post	Aleatorio 51
	Wu, H. M., Kuo, B. C., & Wang, S. C. (2017).	pre-post	Aleatorio 118
	Yang, K.-H., Chu, H.-C., & Chiang, L.-Y. (2018).	pre-post	Aleatorio 58
Diseño cuasiexperimental	Herodotou, C. (2018).	pre-post	Asignado 32
	Dickes, A. C., Kamarainen, A., Metcalf, S. J., Gün-Yildiz, S., Brennan, K., Grotzer, T., & Dede, C. (2019).	pre-post	Asignado 43
	Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2020).	pre-post	Asignado 62
	Liang, H. Y., Hsu, T. Y., & Hwang, G. J. (2021).	pre-post	Asignado 43
	Chester S.J. Huang, Addison Y.S. Sub, Stephen J.H. Yang, Hsin-Hun Liou, (2017)	pre-post	Asignado 64
	Stoyle K.L. & Morris B.J. (2017)	pre-post	Asignado 134
	Volk, M., Coti, M., Zajc, M. & Starcicun, A. I. (2017).	pre-post	Asignado 259
	Killia, K, Moellerb, k & Ninausb M. (2018).	pre-post	Asignado 99
	Jon-Chao Hong, Hsien-Sheng Hsiao, Po-Hsi Chen, Chow-Chin Lu, Kai-Hsin Tai, & Chi-Ruei Tsai c. (2021).	pre-post	Asignado 180
	Reeves, J. L., Gunter, G. A., & Lacey, C. (2017).	pre-post	Asignado 28
	Hwang, W.-Y., Chen, H.-R., Chen, N.-S., Lin, L.-K. & Chen, J.-W. (2018).	pre-post	Asignado 52
	Hwang, G.-J., Tu, N.-T. & Wang, X.-M. (2018).	pre-post	Asignado 72
	Siew, P. H. (2018).	pre-post	Asignado 20
	Sun, D., Looi, C. K., Yang, Y., & Sun, J. (2020).	pre-post	Asignado 37
	Chu, H. C., Chen, J. M., Kuo, F. R., & Yang, S. M. (2021).	pre-post	Asignado 116
Diseño intra-sujetos	Mercier, E, Vourloumi, G. & Higgins, S. (2017).	Intervención	Asignado 32
	Sýkora, T., Stárková, T., & Brom, C. (2021).	pre-post	Asignado 95
	Luo, F, Antonenko, P, & Davis, Ch. (2020).	pre-post	Asignado 2
	Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2017).	Intervención	Asignado 61
	Fallon, G. (2019).	Intervención	Asignado 38
Diseño evolutivo (transferencia de habilidades)	I-Hua Chen, Jeffrey Hugh Gamble, Zeng-Han Lee, & Qian-Lan Fu. (2020).	Solo Pretest	Asignado 178

Tabla 1 – Tipos de diseños

3.2. Tipo de juegos virtuales y sus características al implementarlos en el dominio de conocimiento numérico y científico

La clasificación de la tipología de juegos virtuales empleados en los artículos seleccionados sobre TIC y educación se realizó a partir de Rodríguez et al. (2022) quienes consideraron el tipo de juego según la acción del jugador al interactuar con la situación de juego virtual. En ese sentido, se analiza modalidad de juego, procedimiento para obtener la meta en el juego, número de sesiones y duración por sesión.

De modo específico, de los 28 artículos seleccionados acerca de las situaciones de juego virtual, para el dominio de conocimiento matemático y científico, se encontraron los siguientes tres tipos de juego: video juego de entretenimiento generalista (n =14 artículos), video juego de estrategia (n =9 artículos), y video juego fotorrealista (n =5 artículos) (tabla 2,3,4).

En el tipo de video juegos de entrenamiento generalista: se encontró para la modalidad de juego, que la gran mayoría de las situaciones de juego virtual son de tipo individual, excepto los estudios de I-Hua Chen et al. (2020); Chester et al. (2017) quienes proponen actividades orientadas a ser resueltas de manera individual y grupal. Frente, al procedimiento meta del juego, gran parte de los estudios requieren que los niños den respuestas acertadas en las actividades de juego virtual, por ejemplo, la aplicación “Maths 3-5” y “Maths 4-6” los niños deben aprobar el 100% en cada una de las pruebas (Outhwaite et al., 2020) o en la aplicación “Snappet” deben completar diez lecciones y el programa indica el acierto y el error (Faber et al., 2017). Frente al número de sesiones, osciló entre 3 a 10 sesiones con un tiempo de duración entre 20 a 40 minutos.

En el marco del tipo de video juego de estrategia: en la modalidad de juego, gran parte de las situaciones de juego virtual son de corte individual, excepto el estudio de Liang y Hwang, (2021) quienes a partir del sistema móvil “CoboFun” propusieron que unas misiones se resuelvan de manera individual y otras grupal. Frente, al procedimiento meta del juego, se hace énfasis en promover la reflexión del niño al desempeñarse en la actividad de juego, además, de llevarlo a niveles más avanzados de conocimiento, por ejemplo, el juego “Matemág” o el juego “Reino de la suma y la resta” (Sýkora et al., 2021; Yan y Chiang, 2018). Frente, al número de sesiones, osciló entre 2 a 10 sesiones con un tiempo de duración entre 8 minutos a 4 horas.

Por último, el tipo de video juego fotorrealista: en esta modalidad, la mayoría de las situaciones de juego virtual son individual, excepto, el estudio de Cheng et al. (2019) quienes propusieron que las actividades requeridas usando el sistema de aprendizaje móvil de acuaponía sean grupales. Frente, al procedimiento meta del juego, la mayoría de estos juegos fueron orientados a que los estudiantes usaran la simulación virtual para comprender el funcionamiento real de los conceptos que están trabajando y a partir de allí puedan surgir posibles soluciones o elaborar sus proyectos de aula, por ejemplo, se trabajó con realidad virtual a partir del “Parque Geológico Yehliu” o el programa “EcoMOD” (Dickes et al., 2019; Shao-Chen et al., 2020). Frente, al número de sesiones, osciló entre 2 a 15 sesiones con un tiempo de duración entre 40 a 80 minutos. Vale la pena mencionar que todas las situaciones de juego fotorrealista fueron dirigidas a la intervención en el dominio de conocimiento científico.

Tipo de Juego	Artículo	procedimientos para obtener la meta en el juego	No. Sesiones / Duración
Generalista	Outhwaite, L. A., Guilford, A., & Pitchford, N. J. (2020).	Los niños en la aplicación "Maths 3-5" y "Maths 4-6", deben de aprobar el 100% en cada una de las pruebas que contiene 10 preguntas (número, forma, espacio y medida)	10 NR
	Faber J.M., Luyten H. & Visscher A.J. (2017).	Los niños en la aplicación Snapper deben completar diez lecciones relacionadas con suma, resta y comparación de cantidades, el programa indica el acierto con bloques verdes y el error con bloques rojos retroalimentando sus respuestas.	10 NR
	Chester S.J., Hwang, Addison Y.S. Sub, Stephen J.H. Yang, Hsin-Hun Liao, (2017)	Los niños usan un software DFLS bolígrafo digital el cual funciona a través de una plataforma que hace seguimiento a las respuestas de los niños frente a las operaciones aritméticas rojas, respuesta reprobadas y verdes respuestas aprobadas.	4 / 30 min
	Outhwaite, L. A., Guilford, A., & Pitchford, N. J. (2017).	Los niños usan un software en el cual un maestro virtual indica las instrucciones que se deben realizar para desempeñarse en las tareas de (patrones y formas, conteo, y suma y resta) para avanzar de un tema a otro debe aprobar el 100% en cada tarea.	6 / 30 min
	Stoyle K.L. & Morris B.J. (2017)	Los niños trabajan en la plataforma de blogs que se conectó a un dispositivo móvil protegido con contraseña kidblog en la cual se les pide resolver problemas de suma, resta, división y multiplicación de fraccionarios, una vez culminan las tareas, deben publicar sus respuestas para explicarlas y justificarlas	3 / 30 min
	Vols.M., Cotti, M., Zajc, M. & Staricman, A. I. (2017).	Los niños trabajan con tablas de interacción fácil las cuales contienen aplicaciones (app clock, grid, y map) que evalúan componentes matemáticos como el tiempo, formas geométricas, y espacio con cuadrículas de coordenadas. Para desempeñarse en las tareas, se debe seguir las instrucciones de la aplicación y esta informa si está realizando la tarea de manera correcta e incorrecta.	3 / NR
	Jon-Chao Hong, Hsien-Cheng Hsiao, Po-Hsi Chen, Chow-Chin Lu, Kai-Hsin Tai, & Chi-Ruei Tsai c. (2021).	Los niños trabajan con el software Enriched Thinking Cloud para estudiar la perspectiva de la luna, la ubicación cambiante de la luna y la forma de la luna. Para trabajar con el programa se dan preguntas guías que implican al estudiante hacer predicciones acerca de la ubicación de la luna, realizar observaciones mediante Google Earth, que brinden recursos para explicar sus respuestas acerca de la ubicación y funcionamiento de la luna. Finalmente, si el estudiante no responde las preguntas el investigador da las respuestas y explicaciones.	3 / 40 min
	Reeves, J. L., Gunter, G. A., & Lacey, C. (2017).	Los niños jugaron con los iPads en diversas aplicaciones educativas de conteo, asociaciones numéricas y razonamiento matemático, se les dio la libertad a los estudiantes que escogieran la aplicación de acuerdo a sus habilidades y necesidades, igual las aplicaciones tuvieron un control de sonido, cuando daban una respuesta incorrecta para alertar al estudiante.	6 / NR
	I-Hua Chen, Jeffrey Hugh Gamble, Zeng-Han Lee, & Qian-Lan Pu.(2020).	Los niños trabajaron con el software EasiCare, el cual, consiste en proponer juegos matemáticos de manera individual y grupal, para esto, los estudiantes debían responder de manera correcta a las pruebas de operaciones aritméticas, además, de participar activamente en las discusiones de la intervención, el programa permite dar puntuaciones en tiempo real donde los estudiantes ven sus puntajes individuales como grupales.	4 / NR
	Wu, H. M., Kuo, B. C., & Wang, S. C. (2017).	Los niños trabajaron con un sistema de pruebas adaptativas dinámicas informatizadas con PHP y MySQL en la web APACHE servidores. Este programa consta de 20 preguntas acerca de suma y resta de fraccionarios, si el estudiante responde de manera errónea un ítem el programa brinda indicaciones de instrucción con diversos niveles de acuerdo a número de veces que haya respondido ese ítem incorrecto, es decir el programa retroalimenta al estudiante de acuerdo a sus necesidades.	5 / 40 min
	Hwang, W.-Y., Chen, H.-R., Chen, N.-S., Lin, L.-K. & Chen, J.-W. (2018).	Los niños trabajaron con un E-book que consta de un GPS y cámara fotográfica, con este sistema abordaron temas relacionados con los incendios y los metales (hierro, aluminio y cobre) una vez los estudiantes interactuaron con el libro electrónico se les pide a los estudiantes que relacionen lo aprendido con su experiencia diaria para responder a la causa de los incendios y hacer algunos registros fotográficos. A demás se pide que clasifiquen los óxidos que se encuentran en los diversos tipos de metales.	4 / NR
	Hwang, G.-J., Tu, N.-T. & Wang, X.-M. (2018).	Los niños trabajaron con el software de diseño de libros electrónicos, ShineCue, donde abordaron temas relacionados con 16 tipos de plantas y la ecología, por medio de este programa se les pide a los estudiantes que den cuenta de las características ambientales que inciden en el crecimiento de cierto tipo de plantas. Finalmente, los estudiantes reciben retroalimentación al presentar sus respuestas por parte del investigador y sus pares.	8 / 40 min
	Siew, P. H. (2018).	Los niños trabajaron con el software DigGEMS (Digital Games para la Educación en Matemáticas) el cual contiene tareas de comparar cantidades, forma y espacio, suma, resta y conteo. Este programa se adapta a la capacidad que tienen cada estudiante para construir su conocimiento matemático de manera gradual.	5 / 20 min
	Sun, D., Looi, C. K., Yang Y., & Sun, J. (2020).	Los niños trabajaron con El sistema de aprendizaje nQuire-it, a través de la aplicación Sense-it. Este programa es usado para abordar temas relacionados con la energía y para esto los estudiantes deben responder preguntas orientadas a la comprensión de la intensidad de la luz solar, la Fuentes de luz solar en la vida diaria y tiempo y lugares en los que la energía solar es más fuerte. Todas las respuestas son monitoreadas por los investigadores a través del programa y se promueve la interacción entre investigador y grupos de trabajo.	4 / NR

Tabla 2 – Tipología y características del juego generalista

Estrategia	Haradotos, C. (2018).	A los niños se les entrega una Tablet en modo fuera de línea con el juego Angry Bird, este contiene un software grabador de pantalla que guarda el nivel alcanzado y la puntuación obtenida por cada niño. El juego Angry Bird contiene 3 mundos de 21 niveles cada uno y los niños deben liberar a los cerdos de sus jaulas, lanzando objetos con una cañonera, para esto deben tener en cuenta el ángulo que les permite determinar qué tan lejos va llegar el objeto. El puntaje se obtiene cuando se consiguen tres estrellas en cada nivel, a medida que va ganando un nivel se van habilitando otros.	7 / 8 min
	Mercier, E, Voukroum, G. & Higgins, S. (2017).	Los niños debían completar tres misterios matemáticos con problemas verbales complejos con técnica solución correcta. Estos misterios evalúan diversas habilidades numéricas: 1. Sneaky Sydney (conocimiento sobre el número), 2. Walter (cálculo), y 3. Dinner Disaster (problema de lógica). El juego "Matemag" brinda una historia que contextualiza las actividades que los niños deben realizar en el juego. Matemag consta de 23 niveles conformado de 7 tareas por nivel. Cuando se completa un nivel se habilita el siguiente. Los personajes del juego hacen un viaje a través de una tierra mágica y las tareas matemáticas están incrustadas en el paisaje como obstáculos para los personajes: 1. Number Web (habilidades aritméticas básicas usando grifos simples), 2. la Hungry Plant (dirigir el alimento de cada planta de forma equitativa), y 3. Weighing Machine (equilibrar el peso en balanzas resolviendo ecuaciones).	6 / NR
	Liang, H. Y., Hsu, T. Y., & Hwang, G. J. (2021).	El sistema móvil CoboFun consta de 12 misiones. En primera instancia los niños deben resolver 5 misiones de manera individual, a partir de indagar sobre temas ambientales. En un segundo momento, los niños de manera grupal deben resolver 5 misiones, acerca de la biodiversidad para resolver problemas ambientales haciendo uso de los conocimientos adquiridos en las primeras misiones.	2 / 4 hrs
	Kilka, K, Moeller, K & Nimmo M. (2018).	La versión beta de Semidun School consta de seis mundos de juego que incluían 62 niveles. En el estudio se usó el primer mundo con su respectivo nivel. Cada nivel representa un camino a la cima de la montaña que consta de 10 o 12 tareas. Después de completar una tarea, el jugador pasa a la siguiente tarea, esto implica desempeñarse lo suficientemente bien para llegar a la cima de la montaña y pasar al siguiente nivel. Las tareas aumentan su dificultad según el progreso de los niños en el juego. 1. tareas de estimación en rectas numéricas que iban del 0 al 1 y del 0 al 5. El jugador tiene que localizar el número decimal 0.50 en la recta numérica para desenterrar las monedas escondidas, y 2. tareas de comparación / ordenación de magnitudes, los jugadores tenían que colocar las piedras en orden ascendente de acuerdo con las magnitudes numéricas representadas en ellas, además, la ubicación exacta en la recta numérica (que va desde los números más pequeños a la izquierda hasta números más grandes a la derecha).	5 / NR
	Brezovcizka, B., McMillen, J., Veermans, K, Hammille-Sormunen, M, Rodríguez-Velázquez, G, Pongskidit, N, Laakkonen, E & Lehtinen, E. (2019).	El juego de navegación matemática consta de 64 mapas que son la unidad básica del juego. El mapa tiene 4 lugares de destino que deben recorrerse al interior de cada mapa. A medida que los jugadores avanzan en el juego, los mapas se vuelven más difíciles, es decir, que el progreso en el juego desbloquea nuevos niveles, siempre y cuando los niños tengan cierto rendimiento al establecer relaciones numéricas y dar cuenta de las operaciones aritméticas requeridas para completar el mapa.	10 NR
	Lao, F, Antonello, F. & Davis, Ch. (2020).	Se usó la aplicación móvil Blobsby, la cual, opera con un robot denominado DASH. A partir de esta aplicación se presentan a los niños el ciclo de vida de los bichos y de los musgos en la secuencia correcta, con esta información los niños deben armar rompecabezas, descubrir dibujos que hagan referencia a la vida de los bichos y musgos. Por último, se les pide que reflexionen sobre las diferencias entre los bichos y los musgos.	8 / 2 hr
	Yang, K.-H., Chu, H.-C., & Chiang, L.-Y. (2018).	El juego matemático se llama "Reino de la suma y la resta". El juego consiste en que el usuario debe organizar un plan para salvar el reino, para esto cuenta con cuatro mapas (aldeas, bosques, laboratorios y habitación) para usar los mapas el guerrero debe aprender el concepto de suma y resta. Para completar la tarea asignada según los diferentes niveles, el guerrero debe responder y obtener el audífono, así como traer suficiente dinero antes de que puedan ser entregados a la princesa. Si la tarea se completa con éxito, entonces el reino se salvará por completo.	3 / 2 hr
	Chu, H. C., Chen, J. M., Kuo, F. R., & Yang, S. M. (2021).	Se usó el juego matemático basado en la relación concepto-efecto adaptativo (CER) que consta de 9 unidades y 48 tareas. En este juego el niño hace el papel de jugador principal, el juego requiere que se encuentren todos los tesoros y completen las tareas para avanzar al siguiente desafío. La trama del juego exige que los niños sumen, resta, comparen cantidades, sumen, resta y multipliquen fraccionarios. Por otra parte, si los estudiantes responden incorrectamente, el sistema dará la respuesta y proporcionará los pasos de resolución de problemas para que el estudiante lo pueda desarrollar.	3 / 30 min

Tabla 3 – Tipología y características del juego de estrategia

Fotorrealista	Dicks, A. C., Kamarainen, A., Mettall, S. J., Qiu-Yidie, S., Brennan, K., Grotzer, T., & Dedo, C. (2019).	Se usó el programa EcoMOD. En primer momento se presentó un entorno virtual inmersivo en 3D, donde los estudiantes a través de esta herramienta simulaban una de las especies dentro del ecosistema y realizaban acciones correspondientes a esa especie. En segunda instancia, momento los estudiantes debían realizar un mapa casual de las relaciones dentro del ecosistema, analizando y estableciendo un diálogo frente a los cambios presentados en ese ecosistema y que los ocasionaba.	15 / 50 min
	Lai, A. F., Chen, C. H., & Lee, G. Y. (2019).	Se usó el mecanismo de aprendizaje basado en realidad aumentada. Cada estudiante recibe un dispositivo móvil el cual permitirá interactuar con el sistema de aprendizaje basado en realidad aumentada y un libro de texto sobre topografía. Los estudiantes deben leer todo el libro, en el momento que el estudiante se encuentre en el libro una frase realzada en rojo, debe hacer uso del dispositivo móvil y el sistema de aprendizaje basado en realidad aumentada escanea la frase para obtener el material multimedia. Al finalizar con la lectura del libro cada estudiante deberá realizar unas tareas de aprendizaje, durante ese desarrollo recibirán comentarios rápidos y obtendrán puntos del módulo de respuesta y del módulo de calificación simultáneamente.	4 / 50 min
	Fallon, G. (2019).	La situación de juego se llamó Descubriendo la electricidad. Para el juego se usan cuatro simulaciones (<i>Electronics for Kids, circuits builder, experiments simple circuits y panel bulb</i>) utilizando la AppStore calificó en 4+ para intervenir con los estudiantes. Posterior, al juego se les pide a los estudiantes diseñar sus propios circuitos usando componentes virtuales y protoboards	7 / 40 min
	Chang, S. C., Hwang, G. J., & Chen, C. H. (2019).	Se usó el sistema de aprendizaje acupuntura mediado por tecnología móvil. El experimentador presenta varios entornos de acupuntura en realidad aumentada y le pide a los estudiantes que realicen observación, recopilación y análisis de los datos. Después de observar y estudiar el entorno de acupuntura, debían registrar lo que sucedió en una semana y realizar una reflexión para compartir con los demás miembros del equipo. Cada equipo de acuerdo con la observación realizada podía hacer cambios en la frecuencia de alimentación y el fotoperíodo.	4 / NR
	Shao-Chan Chang, Ting-Chia Hsu, & Morris Szu-Yung Jong (2020).	Se usó el sistema de aprendizaje de realidad virtual a partir del paisaje rocoso del Parque Geológico Yehliu que es una de las maravillas más famosas del mundo. El contenido temático fue el concepto de erosión y deposición. Una vez que los estudiantes se familiarizaron con el funcionamiento del sistema de realidad virtual en sus tablets, los estudiantes debían diseñar preguntas sobre el Parque Geológico Yehliu y encontrar las respuestas correctas. Finalmente, podían hacer sus proyectos de realidad virtual con los materiales existentes o recursos en línea diseñando sus propios objetos.	2 / 50 min

Tabla 4 – Tipología y características del juego fotorrealista

4. Discusión

La presente revisión de literatura destaca varios aspectos claves a nivel metodológico para educadores y psicólogos cognitivos interesados en examinar situaciones de juegos virtuales como un elemento importante para mejorar habilidades cognitivas y adquisición de conocimiento numérico y científico en niños entre 1º y 5º grado de primaria. Muchas de las situaciones de juego virtual reportadas en nuestro estudio fueron juegos orientados a un contenido temático específico, donde algunos investigadores, Hong et al. (2021); Siew (2018) hicieron énfasis en el acierto y el error validando conocimientos en los estudiantes frente a lo numérico y lo científico. Por ejemplo, juegos como “DigiGEMS” o “Enriched Thinking Cloud” brindan instrucciones a los niños para ir desarrollando una secuencia de acciones que al final el acierto o el error lo califican, a partir, de un porcentaje. Consideramos que el interés desde la cognición es hacer énfasis en el estudiante usando las TIC en aula de clase o en investigaciones, para generar demandas cognitivas que impliquen resolver un problema, razonamiento científico o matemático, procesos de interacción y motivación entre otros.

Desde la cognición una desventaja de usar juegos desarrollados con un contenido temático específico, tiene que ver con que limita la generalización de cómo los niños aprenden o mejoran sus habilidades cognitivas durante la actividad de juego. Sin embargo, usar juegos de estrategias, los cuales, su énfasis está en el estudiante, tiene una ventaja porque permite adaptar las situaciones de juego virtual para corroborar hipótesis en términos de los procesos de desarrollo cognitivo, hacer seguimiento a las estrategias empleadas para resolver un problema de tipo científico o matemático y que aspectos del juego generan aprendizaje o habilidades cognitivas para la adquisición de conocimiento.

En este contexto, nuestra revisión evidencio una amplia variedad de diseños de investigación usados en los estudios para las situaciones de juego virtual en el dominio de conocimiento numérico y científico. Es claro, que la decisión para elegir el tipo de diseño consiste en la pregunta de investigación formulada por el investigador, los recursos económicos, humanos y tecnológicos con los que cuenten para ser consistentes con los procedimientos requeridos para ejecutar el estudio. De los estudios seleccionados muchos usaron diseño experimental y cuasi experimental orientados a probar la efectividad que tiene una situación de juego virtual en un contenido temático específico, esto implica aludir a intervenciones pre – post con diferentes condiciones de control para

dar confiabilidad y validez al recurso digital. Por último, se encontró muy pocos estudios que usaron un diseño evolutivo, los cuales, se hacen preguntas orientadas a indagar por el aporte que hicieron las situaciones de juego virtual al desempeño, a la comprensión o la evolución que presenta un sujeto al interactuar en actividades digitales.

En términos generales, tanto los tipos de juegos en situaciones virtuales como los diseños de investigación, analizados en los estudios seleccionados para la revisión de literatura, permiten mencionar, que realmente el uso de las TIC está muy vinculado a los aspectos educativos, es decir, validar juegos virtuales que puedan ser aplicados a la enseñanza de contenidos matemáticos y científicos, dejando de lado un poco, la discusión acerca de los aspectos cognitivos, el aprendizaje y los procesos de desarrollo.

Es de considerar que la revisión de literatura realizada en este estudio tuvo algunas limitaciones, particularmente porque se hizo énfasis en solo tres revistas, la estrategia pudo limitar encontrar otros estudios, que dieran una apreciación más cercana asuntos relacionados al desarrollo cognitivo que promueve el uso de las TIC en los procesos de enseñanza y aprendizaje. Teniendo en cuenta, la limitación presentada, no se pueden hacer afirmaciones finales sobre el propósito de los tipos de juego y los diseños de investigación usados en estos estudios. Sin embargo, es necesario que en futuras investigaciones acerca del uso de situaciones de juego virtual en los dominios de conocimiento numérico y científico sean abordadas desde un enfoque más centrado en el sujeto aludiendo a una perspectiva micro genética, considerando la variabilidad, respetando las representaciones del sujeto entre otras.

Referencias

- Amankwah-Amoah, J., Khan, Z., Wood, G., & Knight, G. (2021). COVID-19 and digitalization: the Great Acceleration. *Journal of Business Research*, 136, 602-611. <https://doi.org/10.1016/j.jbusres.2021.08.011> .
- Aslan, S. A., & Duruhan, K. (2020). The effect of virtual learning environments designed according to problem-based learning approach to students' success, problem-solving skills, and motivations. *Education and Information Technologies*, 26(2), 2253-2283. <https://doi.org/10.1007/s10639-020-10354-6>.
- Behnamnia, N., Kamsin, A., & Ismail, M. A. (2020). The Landscape of research on the use of digital game-based learning apps to Nurture creativity among young Children: a review. *Thinking Skills and Creativity*, 37 (3), 1-12. <https://doi.org/10.1016/j.tsc.2020.100666> .
- Behnamnia, N., Kamsin, A., Ismail, M. A., & Hayati, A. (2020). The Effective Components of creativity in digital game-based learning among young children: a case study. *Children and Youth Services Review*, 116, 1-13. <https://doi.org/10.1016/j.childyouth.2020.105227> .
- Beltrán, J., Sanchez, H., & Rico, M. (2021). Aprendizaje divertido de programación con gamificación. *RISTI - Revista Ibérica de Sistemas y Tecnologias de informação*, (41),17-43. <https://doi.org/10.17013/risti.41.17-33>.

- Beltrán-Pellicer, P., Begué, N., & De Hierro, A. F. R. L. (2020). Experiencias y recursos TIC en la enseñanza y aprendizaje de la probabilidad. *Investigación en Entornos Tecnológicos en Educación Matemática*, 1, 15-22. <https://doi.org/10.7203/ietem.1.16287>.
- Cepal & Unesco. (2020). La educación en tiempos de la pandemia de Covid-19. *Cepal.org*. <https://hdl.handle.net/11362/45904>
- Chen, S. A., Tsai, J., Liu, S. Y., & Chang, C. (2021). The effect of a scientific board game on improving creative problem solving skills. *Thinking Skills and Creativity*, 41, 13-22. <https://doi.org/10.1016/j.tsc.2021.100921>.
- Cieza, E. C., & Lujan, D. (2018). Educational mobile application of augmented reality based on markers to improve the learning of vowel usage and numbers for children of a kindergarten in Trujillo. *Procedia Computer Science*, 130, 352-358. <https://doi.org/10.1016/j.procs.2018.04.051>.
- Colliver, Y., & Veraksa, N. (2019). The aim of the game: a pedagogical tool to support young children's learning through play. *Learning, Culture and Social Interaction*, 21, 1-35. <https://doi.org/10.1016/j.lcsi.2019.03.001>.
- Connolly, T., Boyle, E., MacArthur, E., Hainey, T., & Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661-686. <https://doi.org/10.1016/j.compedu.2012.03.004>.
- Dubé, A. K., & Dubé, N. J. (2020). Policies to guide the adoption of educational games into classrooms. *Educational Technology Research and Development*, 69(1), 167-171. <https://doi.org/10.1007/s11423-020-09835-9>.
- Flynn, R., Kleinknecht, E., Ricker, A., & Blumberg, F. C. (2021). A narrative review of methods used to examine digital gaming impacts on learning and cognition during middle childhood. *International Journal of Child-Computer Interaction*, 30, 1-13. <https://doi.org/10.1016/j.ijcci.2021.100325>
- Gunawan, G., Suranti, N. M. Y., Nisrina, N., & Herayanti, L. (2018). Students' problem-solving skill in physics teaching with virtual labs. *International journal of pedagogy and teacher education*, 2, 87-96. <https://doi.org/10.20961/ijpte.v2i0.24952>.
- Herrewijn, L., De Jans, S., Hudders, L., & Cauberghe, V. (2021). Leveling Up Advertising Literacy! Investigating the cognitive and motivational effectiveness of a digital game for learning aimed at improving children's advertising literacy. *Electronic Commerce Research and Applications*, 46, 1-16. <https://doi.org/10.1016/j.elerap.2021.101036>.
- Hoareau, L., Thomas, A., Tazouti, Y., Dinet, J., Luxembourger, C., & Jarlégan, A. (2021). Beliefs about digital technologies and teachers' acceptance of an educational app for preschoolers. *Computers & education*, 172, 1-32. <https://doi.org/10.1016/j.compedu.2021.104264>.
- Hussein, M. A. H., Ow, S. H., Cheong, L. S., & Thong, M. (2019). A digital game-based learning method to improve students' critical thinking skills in elementary science. *IEEE Access*, 7, 1-10. <https://doi.org/10.1109/access.2019.2929089>.

- Hussein, M. A. H., Ow, S. H., Cheong, L. S., Thong, M., & Ebrahim, N. A. (2019). Effects of digital Game-based learning on Elementary science Learning: A Systematic review. *IEEE Access*, 7, 1-17. <https://doi.org/10.1109/access.2019.2916324>.
- Hwang, G., & Chen, P. (2022). *Interweaving gaming and Educational technologies: Clustering and forecasting the trends of game-based learning research by bibliometric and visual analysis. Entertainment Computing*, 40, 1-12. <https://doi.org/10.1016/j.entcom.2021.100459>.
- Hwang, G., Wu, P., & Chen, C. (2012). An online game approach for improving students' learning performance in web-based problem-solving activities. *Computers & education*, 59(4), 1246-1256. <https://doi.org/10.1016/j.compedu.2012.05.009>.
- Jossan, K. S., Gauthier, A., & Jenkinson, J. (2021). Cultural implications in the acceptability of game-based learning. *Computers & Education*, 174, 1-16. <https://doi.org/10.1016/j.compedu.2021.104305>.
- Kinnula, M., & Iivari, N. (2021). Manifesto for Children's Genuine Participation in Digital Technology design and making. *International Journal of Child-Computer Interaction*, 28, 1-14. <https://doi.org/10.1016/j.ijcci.2020.100244>.
- Laranjeiro, D. (2021). Development of game-based m-learning apps for preschoolers. *Education Sciences*, 11, 1-25. <https://doi.org/10.3390/educsci11050229>.
- León, O. G., & Montero, I. (2002). *Métodos de investigación en psicología y educación* (3ª edición). Mc Graw Hill.
- Li, M., & Tsai, C. C. (2013). Game-based Learning in Science Education: A review of relevant research. *Journal of Science Education and Technology*, 22(6), 877-898. <https://doi.org/10.1007/s10956-013-9436-x>.
- Liao, C., Chen, C., & Shih, S. J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. *Computers & education*, 133, 43-55. <https://doi.org/10.1016/j.compedu.2019.01.013>.
- Lin, L., Shadiev, R., Hwang, W. Y., & Shen, S. (2020). From Knowledge and Skills to Digital Works: An application of design Thinking in the Information Technology course. *Thinking Skills and Creativity*, 36, 1-28. <https://doi.org/10.1016/j.tsc.2020.100646>.
- Lin, S., Chien, S., Hsiao, C. T., Hsia, C., & Chao, K. (2020). Enhancing computational thinking capability of preschool children by game-based smart toys. *Electronic Commerce Research and Applications*, 44, 1-33. <https://doi.org/10.1016/j.elerap.2020.101011>.
- Linderoth, J., Lantz-Andersson, A., & Lindström, B. (2002). Electronic Exaggerations and Virtual Worries: Mapping research of computer games relevant to the understanding of children's game play. *Contemporary Issues in Early Childhood*, 3(2), 226-250. <https://doi.org/10.2304/ciec.2002.3.2.6>.

- Martínez-Domínguez, M., & Fierros-González, I. (2022). Determinants of internet use by school-age children: The Challenges for Mexico during the COVID-19 Pandemic. *Telecommunications Policy*, 46, 1-18. <https://doi.org/10.1016/j.telpol.2021.102241>.
- Mayer, R. E. (2016). What should be the role of computer games in education? *Policy insights from the behavioral and brain sciences*, 3(1), 20-26. <https://doi.org/10.1177/2372732215621311>.
- Medina-López, C., Marin-Garcia, J. A., & Alfalla-Luque, R. (2010). Una propuesta metodológica para la realización de búsquedas sistemáticas de bibliografía. *Working papers on operations management*, 1(2), 13-30. <https://doi.org/10.4995/wpom.vii2.786>.
- Mertala, P. (2017). Digital Technologies in Early Childhood Education – A frame Analysis of preservice teachers’ perceptions. *Early Child Development and Care*. <https://doi.org/10.1080/03004430.2017.1372756>
- Mertala, P. (2017a). Wonder Children and Victimizing Parents – Preservice early childhood teachers’ beliefs about children and technology at home. *Early Child Development and Care*, 189(3), 392-404. <https://doi.org/10.1080/03004430.2017.1324434>.
- Mertala, P., & Meriläinen, M. (2019). The Best Game in the World: exploring young children’s digital game-related meaning-making via design activity. *Global Studies of Childhood*, 9(4), 275-289. <https://doi.org/10.1177/2043610619867701>.
- Nizam, D. N. M., & Law, E. L. (2021). Derivation of young children’s interaction strategies with digital educational games from Gaze sequences analysis. *International journal of human-computer studies*, 146, 1-32. <https://doi.org/10.1016/j.ijhcs.2020.102558>.
- Pinedo, R., García-Martín, N., Rascón, D., Caballero, C., & Cañas, M. (2021). Reasoning and Learning with Board game-based Learning: a case study. *Current Psychology*, 41(3), 1603-1617. <https://doi.org/10.1007/s12144-021-01744-1>.
- Rajendran, D., & Santhanam, P. (2021). Towards digital game-based learning content with multi-objective reinforcement learning. *Materials Today Proceedings*, 1-2. <https://doi.org/10.1016/j.matpr.2021.03.156>.
- Samad, M. R. A., Haji, Z., & Khalid, F. (2020). A Systematic Review: The Integration of Mobile Applications in Science Education. Are you ready? *International Journal of Advanced Science and Technology*, 29, 846-862. <https://Downloads/mysecondpublishedarticle.pdf>.
- Sevin, R., & DeCamp, W. (2016). From playing to programming: the effect of video game play on confidence with computers and an interest in computer science. *Sociological Research Online*, 21(3), 14-23. <https://doi.org/10.5153/sro.4082>
- Sharma, K., Torrado, J. J., Gómez, J., & Jaccheri, L. (2021). Improving girls’ perception of computer science as a viable career option through game playing and design: lessons from a systematic literature review. *Entertainment Computing*, 36, 1-13. <https://doi.org/10.1016/j.entcom.2020.100387>.

- Sjöberg, J., & Brooks, E. (2022). Collaborative interactions in problem-solving activities: school children's orientations while developing digital game designs using smart mobile technology. *International Journal of Child-Computer Interaction*, 33, 1-13. <https://doi.org/10.1016/j.ijcci.2022.100456>.
- Tobias, S., Fletcher, J. D., Bediou, B., Wind, A. P., & Chen, F. (2014). Multimedia learning with computer games. En *Cambridge University Press eBooks* (pp. 762-784). <https://doi.org/10.1017/cbo9781139547369.037>.
- Torres, P., Ulrich, P., Cucuiat, V., Cukurova, M., De La Presa, M. C. F., Luckin, R., Carr, A., Dylan, T., Durrant, A., Vines, J., & Lawson, S. (2021). A systematic review of physical-digital play technology and developmentally relevant child behaviour. *International Journal of Child-Computer Interaction*, 30, 1-13. <https://doi.org/10.1016/j.ijcci.2021.100323>.
- Tu, C., Nurymov, Y., Umirzakova, Z., & Berestova, A. (2021). Retracted: Building an online educational platform to promote creative and affective thinking in special education. *Thinking Skills and Creativity*, 40, 1-10. <https://doi.org/10.1016/j.tsc.2021.100841>.
- Valenzuela, S., Batanero, C., & Begué, N. (2021). Recursos en internet para el estudio de los estadísticos de orden. *Revista Sergipana de Matemática e Educação Matemática*, 6(1), 1-22. <https://doi.org/10.34179/revisem.v6i1.14419>.
- Vásquez, C., & Pincheira, N. (2021). ¿Qué procesos matemáticos se movilizan cuando se enseña probabilidad? un estudio de caso en el aula de Educación Infantil. *Revista Sergipana de Matemática e Educação Matemática*, 6(2), 62-84. <https://doi.org/10.34179/revisem.v6i2.16007>.
- Vega, V., & Robb, M. B. (2019). *The Common sense Census: Inside the 21st-century classroom*. *Common Sense Media*. Recuperado 5 de junio de 2023, de <https://www.common sense media.org/research/the-common-sense-census-inside-the-21st-century-classroom>.

Educación Musical Temprana y Tecnología: Una revisión sistemática

Vicenta Gisbert-Caudeli¹, Rosa María Díaz-Mayo¹

vicenta.gisbert@uam.es; rosamaria.diaz@uam.es

¹ Universidad Autónoma de Madrid, Ciudad Universitaria Cantoblanco, Carretera de Colmenar, km. 15, 28049 Madrid, Spain.

DOI: 10.17013/risti.59.68–82

Resumen: La educación musical en la primera infancia y la integración de recursos tecnológicos han despertado un interés creciente por sus potenciales beneficios en el desarrollo cognitivo, la motivación y la interacción social. Con el objetivo de analizar los efectos de la tecnología aplicada a la práctica musical en niños de 2 a 6 años, se realiza una revisión sistemática en las bases de datos *Web of Science y Scopus*. Se identificaron 151 registros, se seleccionaron ocho artículos que cumplieran con los criterios de inclusión, con una muestra total de 697 participantes. Los resultados sugieren que la incorporación de herramientas tecnológicas en la práctica musical colectiva favorece la creatividad, la motivación y el aprendizaje autónomo. La escasez de investigaciones empíricas limita la generalización de los hallazgos, lo que evidencia la necesidad de ampliar la producción científica y de avanzar hacia un modelo educativo híbrido que integre equilibradamente tecnología y métodos tradicionales.

Palabras-clave: educación musical temprana; educación infantil; aplicaciones tecnológicas; revisión sistemática; PRISMA.

Early Music Education and Technology: A Systematic Review

Abstract: Early childhood music education and the integration of technological resources have sparked growing interest due to their potential benefits for cognitive development, motivation, and social interaction. To analyze the effects of technology applied to musical practice in children aged 2 to 6, a systematic review was conducted in the Web of Science and Scopus databases. One hundred and fifty-one records were identified, and eight articles that met the inclusion criteria were selected, with a total sample of 697 participants. The results suggest that the incorporation of technological tools into collective musical practice fosters creativity, motivation, and autonomous learning. The scarcity of empirical research limits the generalization of the findings, highlighting the need to expand scientific production and move toward a hybrid educational model that smoothly integrates technology and traditional methods.

Keywords: Early music education; early childhood education; technological applications; systematic review; PRISMA.

1. Introducción

El uso tecnológico y de la música durante la etapa de Educación Infantil ha sido objeto estudio (Bodrova & Leong, 2024; De Vries, 2021; Fernández-Company et al., 2021; Hallam & Himonides, 2022; Kim, 2020), hasta la fecha no se han identificado revisiones sistemáticas que integren los hallazgos de investigaciones que apliquen de manera conjunta estos recursos. En este sentido, se considera que la ausencia de un marco de referencia consolidado limita la comprensión de cómo la combinación de herramientas tecnológicas y prácticas musicales puede enriquecer el proceso educativo durante la primera infancia.

Las últimas décadas han mostrado un escenario educativo en el que, de forma progresiva, se ha integrado la utilización de tecnología digital transformando las dinámicas de enseñanza-aprendizaje (Vasil et al., 2018), fomentando un uso creativo de la tecnología (Fernández-Scagliusi y Llorente-Cejudo, 2024). La educación infantil incorpora la tecnología como elemento enriquecedor y motivador conducente al desarrollo individual y potenciador del aprendizaje autónomo (Bers et al., 2019). Desde las aplicaciones móviles hasta las pizarras interactivas, se aprecia un proceso de re-diseño de metodologías pedagógicas, tanto en el aprendizaje musical como en lo concerniente al proceso cognitivo y la integración social. Algunos implicados en el proceso educativo pretenden despejar dudas sobre la eficacia del recurso tecnológico mediante un análisis comparativo respecto de la eficacia de los métodos tradicionales de aprendizaje (Lee & Huai, 2015).

Entre las virtudes del uso de tecnologías en la educación, encontramos conexiones con el desarrollo creativo y la visión interdisciplinar, la innovación educativa encuentra un interesante potenciador motivacional en los recursos tecnológicos (Bucura & Weissberg, 2017). Estas innovaciones han mejorado también los procesos de evaluación, incorporando retroalimentación inmediata en la práctica de refuerzo de contenidos o destrezas (Galindo-Dominguez et al., 2024). Las pizarras interactivas y las aplicaciones móviles facilitan la práctica lúdica auditiva mejorando la discriminación sonora y reconocimiento de patrones musicales en la etapa infantil (Cano & Sánchez-Iborra, 2015), se ha observado también la mejora en la capacidad de abstracción y del pensamiento computacional (Duarte & Sigal, 2019). Entre las mejoras observadas destacan también las vinculadas al proceso de reconocimiento y regulación emocional, así como la comunicación e interacción social, la capacidad cooperativa y empática (Hanrahan et al., 2019).

Aunque resultan evidentes los beneficios de la implicación tecnológica en el aula, se observan algunos aspectos que precisan especial atención para obtener mejores experiencias educativas: capacitación del profesorado (Arvelo-Rosales et al., 2021), adaptación de contenidos, compatibilidad entre métodos tradicionales y recursos tecnológicos, y adaptabilidad o personalización de los recursos con relación a las cualidades del alumnado (Gisbert-Caudeli et al., 2023). La efectividad de los métodos tradicionales y los enfoques tecnológicos plantea el interrogante sobre su complementariedad o carácter indispensable, este es un aspecto mencionado en la literatura existente hasta el momento.

Diversos estudios exploran el impacto que la tecnología tiene en su aplicación en la educación infantil, tanto en las habilidades cognitivas de los menores como específicamente en el área de música (Mosquera et al., 2022). El interés en conocer los distintos análisis y estudios desarrollados al respecto de la temática que nos ocupa consiste en mostrar las diversas aplicaciones y utilidades que la tecnología proporciona en el aula e inspirar en los lectores el espíritu crítico para sopesar las limitaciones y virtudes que los elementos digitales pueden aportar al proceso educativo. Para ello, se considera que el estudio sistemático puede proporcionar una amplia y sintetizada visión del estado de la cuestión y la evolución temática (Quispe et al., 2021; Valenzuela-Romero et al., 2025), arrojando luz sobre aspectos no contemplados en las investigaciones realizadas y abriendo nuevas líneas de investigación para futuros trabajos relacionados.

La educación musical temprana y la integración de la tecnología son cruciales para el desarrollo infantil, ya que fomentan la creatividad, el aprendizaje cognitivo y la interacción social. Esta revisión investiga los efectos de las herramientas tecnológicas en la práctica musical en niños de 2 a 6 años, analizando cómo estas tecnologías pueden aumentar la motivación y enriquecer la experiencia musical colectiva. La investigación no solo pretende ser una aportación al campo académico, sino que también ofrece estrategias para optimizar la educación. La metodología que se consideró más conveniente fue la revisión sistemática, pues permite conocer la literatura más relevante del objeto de estudio y también analizar los enfoques, limitaciones y conclusiones obtenidas. Por tanto, el objetivo central de este estudio es ofrecer una visión global del impacto que los recursos tecnológicos presentan en la Educación Infantil.

2. Metodología de la revisión

La presente investigación se desarrolló a partir de una revisión sistemática de la literatura, considerada la estrategia metodológica más adecuada para identificar, evaluar y sintetizar los hallazgos de investigaciones previas en torno a la relación entre educación musical y tecnología en la etapa de Educación Infantil. Este enfoque resulta pertinente al permitir no solo la recopilación exhaustiva de la producción científica, sino también la comparación crítica de los diferentes enfoques y resultados obtenidos, aportando una visión integradora y fundamentada del estado del conocimiento.

El procedimiento se estructuró conforme a las directrices metodológicas de la Declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher, 2009), lo cual garantiza transparencia, reproducibilidad y rigor científico en el desarrollo del estudio.

2.1. Criterios de inclusión y de exclusión

Los artículos elegidos para su inclusión en esta revisión sistemática debían cumplir con los siguientes criterios: a) publicados en la última década, desde 2014 hasta el año 2024; b) publicados en inglés o español; c) publicados en revistas de impacto revisadas por pares; d) investigaciones con trabajo de campo; e) artículos de acceso abierto. Los criterios de exclusión fueron: los artículos en los que se abordan estudios realizados en otros contextos o etapas vitales distintas a la educación temprana o aquellos en los que no se combinaban la educación musical y los recursos tecnológicos no se han incluido.

Se contemplaron tres fases en la estrategia de búsqueda: la primera consistió en una búsqueda preliminar en las bases de datos Web of Science y Scopus. A continuación, se realizó un análisis de acceso a documento completo, áreas de aplicación del estudio y términos clave. Finalmente se aplicaron los criterios de inclusión y exclusión. Esta investigación se desarrolló siguiendo las directrices establecidas en la Declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) (Moher, 2009). En la Figura 1, se muestran las diferentes estrategias y mecanismos de búsqueda de los artículos seleccionados.

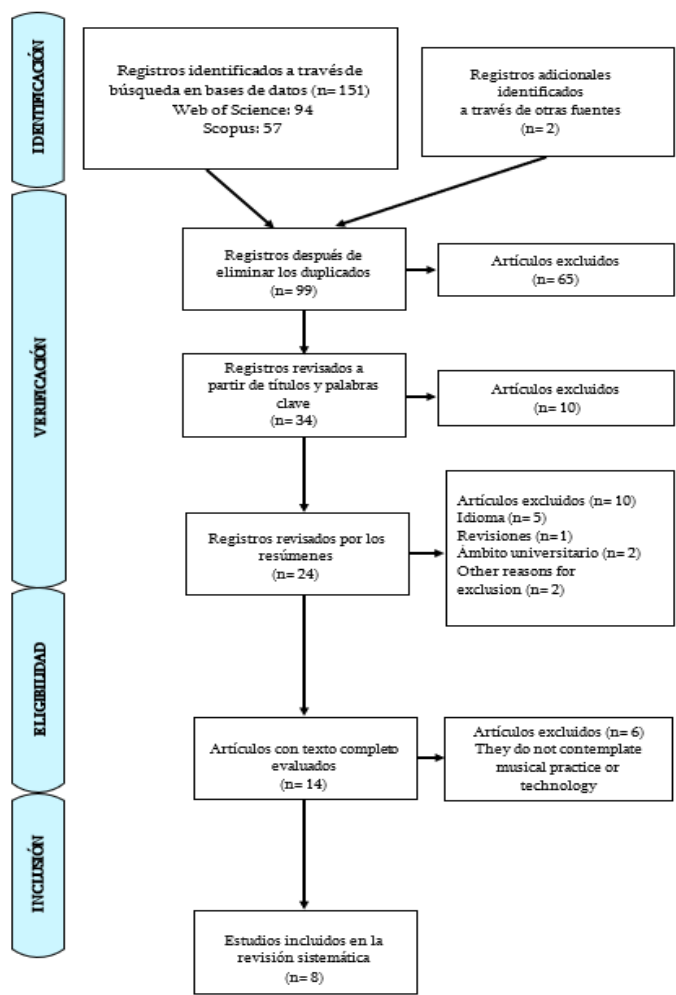


Figura 1 – Procedimiento de selección de artículos según PRISMA. Basado en Moher et al., (2009)

2.2. Extracción y gestión de datos

Los datos más relevantes de los artículos se extrajeron utilizando diversos formularios estandarizados en formato tabla. Se incluyeron los siguientes campos: autoría, año, país, muestra participante, contexto del estudio, técnicas empleadas, número de sesiones o duración, diseño del estudio, herramientas de medida y resultados. Tras la sistematización de datos de la revisión se incorporaron los datos bibliográficos atendiendo al sistema de referenciado APA 7ª edición.

3. Análisis y resultados

La incorporación de artículos en esta revisión implica un total de ocho estudios dedicados a la educación musical temprana mediante recursos tecnológicos. A lo largo del proceso de selección de artículos 10 de ellos fueron descartados: cuatro de ellos por estar escritos en coreano (Ajmahs, 2017; Kwon, 2019; Lee et al., 2016; Young & Suk, 2022), otro por estar escrito en húngaro (Kata, 2016). Otros dos artículos quedaron fuera de la selección final por estar enmarcados en el ámbito universitario (Berrón-Ruiz et al., 2023; Palazón-Herrera & Soria-Vilchez, 2021). Asimismo, se descartó un artículo porque no se implicaba recurso tecnológico en el estudio (Pérez-Moreno & Reverté, 2019). Se encontró un artículo no empírico que tampoco figura entre los seleccionados (Sverre, 2019) y otro más por ser un artículo de revisión (Cuervo et al., 2022).

	País	Participantes	Duración	Contexto
Lee & Huai (2015)	Taiwán	14 niños preescolar	12 semanas con sesiones semanales	Programas musicales en entorno educativo
Paule-Ruiz, et al. (2016)	España	43 menores	5 sesiones semanales durante el segundo trimestre del curso	Escuela pública
Sullivan & Bers (2018)	Singapur	98 menores entre 3 y 6 años	7 semanas	Contexto escolar multicultural
Chao-Fernández et al. (2018)	España	23 escolares	1 curso escolar	Contexto escolar inclusivo con alumnado con limitación auditiva
Preka & Rangoussi, (2019)	Grecia	20 alumnos	2 semanas	Escuela pública
Rahiem, (2021)	Indonesia	72 menores	No se especifica	Club de narración de relatos
Lee & Ho (2023)	Taiwán	389 menores con edades entre 36 y 72 meses	16 semanas	Enfoque educativo interdisciplinario
Campollo-Urkiza (2023)	España	38 alumnos de 3 años	25 semanas	Centro privado-concertado con diversidad cultural

Tabla 1 – Estudios que cumplen los requisitos de inclusión

Todos los artículos incorporados en este estudio han sido publicados entre 2014 y 2024, se puede acceder a ellos en español o inglés y se encuentran publicados en revistas científicas sometidas al proceso de revisión por doble par ciego. Se presentan los ocho estudios que cumplen los criterios de selección en la Tabla 1, mostrados en orden cronológico.

3.1. Participantes e intervenciones

Partiendo de los datos informados en los artículos, el número total de participantes en los estudios incluidos fue de 697 personas con edades comprendidas entre los dos y los seis años de edad. Se encuentran en el entorno escolar, multicultural e inclusivo. Se trata de experiencias desarrolladas en ámbito público y privado, con duración variable y diversidad geográfica (España, Grecia, Taiwán, Singapur...). Los estudios seleccionados contemplan intervenciones musicales combinadas con utilización de diferentes recursos tecnológicos. En todos los artículos se incorpora la medición de resultados combinando el enfoque cuantitativo y el cualitativo. Se encuentra en todos los trabajos apartado de discusión o conclusiones donde se muestra impacto de los métodos implementados en el plano emocional, cognitivo o en la actitud de los participantes. Las intervenciones han contemplado actividades interactivas que contaban con recursos tecnológicos como Scratch, pizarra digital, software musical, aplicaciones móviles, recursos procedentes de la robótica o del pensamiento computacional o realidad aumentada.

3.2. Aspectos metodológicos y efectividad de las intervenciones

El primer estudio, Lee & Huai (2015), desarrollado en Taiwán, se dedica al análisis comparativo de los efectos del uso de la pizarra interactiva (IWB) y los métodos tradicionales de enseñanza en el aula de música, con el objetivo de evaluar sus beneficios y desventajas en la actitud de aprendizaje y los logros académicos de los niños. Los estímulos incluyeron lecciones interactivas con software musical, como “SFSKIDS FUN WITH MUSIC” (San Francisco Symphony, s.f.) para el grupo que utilizó la IWB, y actividades tradicionales como tarjetas de ritmo y fichas musicales para el grupo que siguió el método tradicional. Para medir el impacto de ambos enfoques, se utilizaron dos escalas: la Escala de Actitud de Aprendizaje y la Escala de Resultados de Aprendizaje, las cuales evaluaron tanto las actitudes como los logros antes y después de la implementación de los métodos. Los resultados mostraron mejoras significativas en la actitud de los niños hacia el aprendizaje y en sus resultados académicos. Sin embargo, no se encontraron diferencias significativas entre los dos métodos en cuanto a la actitud y los resultados de aprendizaje, lo que sugiere similitud entre la efectividad del uso de la IWB y de los métodos tradicionales para fomentar el aprendizaje musical en los niños.

Dos años después en España, Paule-Ruiz et al. (2016), se desarrolló un estudio, el segundo seleccionado, con el objetivo de evaluar la utilización de la aplicación móvil SAMI en la enseñanza de la discriminación auditiva y el aprendizaje de notas musicales en niños de educación preescolar. Encontramos en este caso grupo control y grupo experimental. Los estímulos utilizados incluyen juegos interactivos en tabletas para el grupo experimental, y campanas de colores para el grupo control. La medición de resultados comprende pruebas estadísticas como el test de Friedman (Friedman, 1937), análisis de varianza y entrevistas con alumnado y equipo docente para evaluar

el rendimiento, interés y motivación. Los resultados muestran mejores resultados en el grupo experimental, que utilizó SAMI, con óptimo reconocimiento de notas musicales y mayor motivación y creatividad, comparado con el grupo control que utilizó el método tradicional. El estudio concluye que el uso de tecnologías móviles potencia el aprendizaje autónomo e incrementa el interés en la educación musical, sugiriendo la necesidad de adaptar el diseño de las aplicaciones educativas a las características específicas de los niños.

En tercer lugar, el estudio desarrollado en Singapur por Sullivan & Bers (2018) tuvo como objetivo evaluar el impacto del uso de la robótica (KIBO) en el aprendizaje de conceptos de programación y habilidades de ingeniería en niños de preescolar. El estímulo principal fue la implementación del currículo “Danzas del Mundo”, el cual integró arte, música y robótica para enseñar programación. La medición de resultados se realizó mediante varias herramientas, incluyendo la evaluación “Solve-Its” (Strawhacker & Bers, 2015) para determinar el dominio de conceptos de programación básicos y avanzados, y la lista de verificación PTD (Bers, 2012), para evaluar el desarrollo de habilidades como comunicación, colaboración y creatividad. Los resultados mostraron un alto nivel de dominio en la mayoría de los conceptos de programación, con mejoras significativas entre el pre-test y el post-test. Los estudiantes lograron integrar con éxito aspectos de ingeniería y diseño en sus proyectos finales, demostrando la efectividad del uso de la robótica para fomentar el aprendizaje interdisciplinario desde la primera infancia.

El cuarto estudio contempla la combinación del enfoque cuantitativo y cualitativo, evaluando habilidades musicales y de integración social de alumnado con discapacidad auditiva a través de cuestionarios y observaciones de los docentes, antes y después de implementar una intervención educativa. Chao-Fernández et al. (2018) pretenden determinar si la realización de tareas específicas que combinan el uso de Tecnologías de la Información y la Comunicación (TIC) y la educación musical pueden mejorar el aprendizaje y la integración de alumnado con discapacidad auditiva en el aula. Los estímulos utilizados fueron un cuento tradicional, presentado de manera interactiva y mediante lenguaje de señas y extractos musicales de obras con cierta conexión temática con el relato. Se diseñaron actividades interactivas que involucraban la visualización activa, juegos y acompañamiento musical con instrumentos de percusión no afinados. Las herramientas de medida incluyeron cuestionarios administrados a los docentes antes y después de la intervención, evaluando la integración (IC), conocimiento de cuentos tradicionales (TT), habilidades musicales adquiridas (MS) y conocimiento de lenguaje de señas (SL), así como un análisis estadístico de los resultados registrados en tablas de datos. Se obtienen mejoras significativas en el alumno con discapacidad auditiva en su integración en el aula, autoestima, habilidades musicales, conocimiento de cuentos tradicionales, empatía y sensibilidad del alumnado hacia la discapacidad, facilitando la integración. Se observa incremento motivacional, calidad educativa y mejora del resultado posterior a la implementación de la propuesta.

Enmarcado en Grecia, Preka & Rangoussi (2019) encontramos este quinto estudio, con diseño experimental mixto, con la intención de evaluar la viabilidad y efectividad de la tecnología de realidad aumentada (códigos QR) y dispositivos móviles como herramientas de aprendizaje en alumnos de preescolar y jardín de infantes. La intervención se diseñó como un juego de “búsqueda del tesoro”, utilizando estímulos como códigos

QR y actividades interactivas que conectan música y elementos lúdicos (rompecabezas, premios y pruebas musicales). Para la intervención se utilizó contenido musical atractivo para los menores. La medición se realizó mediante entrevistas estructuradas (pre y post-test) para evaluar sus conocimientos sobre elementos musicales y su experiencia con la tecnología. Además, se grabaron en video todas las sesiones de intervención para analizar la colaboración e interacción entre los miembros del equipo. Entre los resultados obtenidos encontramos mejora en el dominio cognitivo (discriminación de tempos -lentos y rápidos- y -agudos y graves-; mejora de la interacción, participación y colaboración en el juego. En el plano emocional se observó mayor motivación y disfrute con los códigos QR, evidenciando el potencial de los dispositivos móviles para mejorar los resultados de aprendizaje en la educación infantil.

El sexto artículo seleccionado, Rahiem (2021) presenta un trabajo de valor teórico y empírico sobre el uso de la narración digital y la interacción musical y artística en la educación infantil. Se basa en un enfoque cualitativo que se centra en un caso único, específicamente en las experiencias de los maestros en un club de narración, arte y ciencia. Se indagó sobre la integración de la tecnología digital en la narración, así como las percepciones y prácticas relacionadas en un entorno de educación infantil, buscando comprender la dinámica entre las estructuras pedagógicas y prácticas en la narración digital y su impacto en el aprendizaje de los niños. La metodología del estudio se sustentó en observaciones directas, permitiendo a la investigadora ver cómo los maestros y los niños interactuaban durante las actividades de narración digital involucrando sonoridades y estímulos artísticos. Los resultados indican que la narración digital hace que la experiencia de contar historias sea más atractiva y rica para los menores. No obstante, la investigación también subraya la necesidad de ampliar la muestra. Se comenta el interés de ampliar la visión a la familia para obtener una visión más holística del impacto de la narración digital en la educación infantil.

Seguidamente incorporamos el estudio de enfoque cuantitativo y diseño experimental, el séptimo en la selección, realizado por Lee & Ho (2023), que incluye un grupo experimental (EG) y un grupo de control (CG). El objetivo principal fue analizar el efecto de la tecnología musical en las habilidades de comprensión del lenguaje (LC) y control personal (SC) en niños con retrasos en el desarrollo. Los estímulos consistieron en la implementación de tecnología musical en actividades educativas con recursos interactivos (luces musicales y cubos parlantes). La toma de datos se realizó mediante pre y post-test utilizando el coeficiente de correlación de Pearson, así como análisis descriptivos y de covarianza para observar las relaciones entre las variables, junto con índices de ajuste del modelo en el análisis de ecuaciones estructurales. Los resultados mostraron el efecto positivo de la tecnología musical en las habilidades de LC y SC de los participantes, encontrándose correlaciones significativas entre las puntuaciones de pre y post-test en ambos grupos. El análisis de covarianza reveló que la tecnología musical predijo positivamente las mediciones de LC y SC, concluyendo que su uso en el aula puede ser una intervención efectiva para mejorar estas habilidades en niños con retrasos en el desarrollo.

Concluye esta descripción de métodos y recursos implicados en los estudios seleccionados con este octavo estudio, desarrollado en España por Campollo-Urkiza (2023), de tipo

cualitativo que presenta la elaboración e implementación de un programa de actividades didáctico-musicales dirigido al desarrollo del pensamiento computacional en aulas de educación infantil para niños de 3 años. Su objetivo principal es evaluar la pertinencia y efectividad de estas actividades en la adquisición de habilidades y destrezas precursoras del pensamiento computacional. Los estímulos utilizados incluyen actividades lúdicas que promueven la orientación espacial, la psicomotricidad y el control postural, así como ejercicios que involucran comandos y programación desenchufada. Las herramientas de medida empleadas son los registros de observaciones de los docentes participantes, así como las aportaciones de los informantes sobre el desempeño de los niños durante las actividades. Los resultados indican que las actividades propuestas favorecen el desarrollo de habilidades como la organización lógica, la representación simbólica y las destrezas óculo-manuales. Además, se observa una mejora en la socialización y colaboración entre los niños, así como la transferencia de conocimientos hacia el uso de herramientas tecnológicas como Scratch, evidenciando así la efectividad del enfoque utilizado en el programa.

4. Discusión

La integración de la tecnología en la educación infantil ha sido objeto de creciente análisis en las últimas décadas, poniendo de relieve tanto sus aportaciones como sus limitaciones. Vasil et al. (2018) señalan que la transformación digital ha permitido adaptar las metodologías educativas a los avances tecnológicos, promoviendo una enseñanza más inclusiva, autónoma y motivadora. En la misma línea, Bers et al. (2019) destacan que herramientas como las aplicaciones móviles y las pizarras interactivas contribuyen al desarrollo individual, al aprendizaje autónomo y al fomento de la creatividad, aspectos vinculados con la innovación educativa y la visión interdisciplinar (Bucura & Weissberg, 2017).

Diversos estudios han identificado beneficios concretos en el desarrollo de habilidades específicas. Cano y Sánchez-Iborra (2015) evidencian que pizarras y aplicaciones interactivas favorecen la discriminación auditiva y el reconocimiento de patrones musicales, mientras que Duarte y Sigal (2019) vinculan estos recursos con mejoras en el pensamiento computacional y la abstracción. En paralelo, Hanrahan et al. (2019) señalan que el uso de tecnologías puede reforzar la comunicación, la cooperación y la empatía en la primera infancia, favoreciendo tanto el desarrollo emocional como la interacción social.

No obstante, persisten retos que condicionan la efectividad de estas herramientas. Gisbert-Caudeli et al. (2023) advierten de la necesidad de capacitación docente, de la adaptación de contenidos y de la búsqueda de un equilibrio entre métodos tradicionales y tecnológicos. El debate sobre si los recursos digitales deben considerarse como un complemento o como un posible sustituto de los enfoques pedagógicos tradicionales continúa abierto (Lee & Huai, 2015). En este sentido, Mosquera et al. (2022) enfatizan la relevancia de evaluar los efectos de la tecnología no solo en el ámbito cognitivo, sino también en el musical y expresivo de los niños, mientras que Quispe et al. (2021) destacan el valor de los estudios sistemáticos para ofrecer una visión amplia y crítica sobre el estado de la cuestión.

5. Conclusiones

En conclusión, la incorporación de la tecnología en la educación infantil presenta un notable potencial para enriquecer el aprendizaje musical y cognitivo, así como para fortalecer las competencias sociales y emocionales. Sin embargo, su implementación debe ser equilibrada, entendiéndose como un complemento y no como un sustituto de los métodos tradicionales de enseñanza. La adecuada formación del profesorado resulta esencial, ya que son los docentes quienes median entre la herramienta digital y el aprendizaje, garantizando un uso pedagógicamente significativo.

Asimismo, es necesario que los recursos tecnológicos sean inclusivos y accesibles para todo el alumnado, reduciendo la brecha digital y ofreciendo alternativas adaptadas a las distintas necesidades educativas. A pesar de los beneficios evidenciados, como la mejora de la creatividad y la autonomía, la personalización limitada de los recursos y las desigualdades de acceso siguen constituyendo desafíos relevantes. Por ello, futuras investigaciones deben ampliar su alcance empírico y metodológico, al tiempo que se promueve la inversión en tecnologías inclusivas y la capacitación docente continua. Solo desde esta perspectiva podrá garantizarse que la tecnología en la educación infantil cumpla su promesa de favorecer el aprendizaje y contribuir al desarrollo integral de los niños.

En suma, la integración de la tecnología en la educación musical infantil constituye un recurso de gran potencial para estimular la creatividad, la motivación y el desarrollo integral de los niños; no obstante, su efectividad depende de una implementación equilibrada y crítica que complemente, y no sustituya, a los métodos pedagógicos tradicionales. La capacitación docente, la accesibilidad tecnológica y la personalización de los recursos se perfilan como condiciones indispensables para maximizar los beneficios identificados y reducir las limitaciones observadas. A futuro, resulta imprescindible ampliar la investigación empírica en este campo con enfoques diversos y muestras más representativas, a fin de consolidar un marco teórico y práctico que oriente la aplicación pedagógica de la tecnología en la educación musical temprana.

5.1. Limitaciones y prospectiva

Si bien los hallazgos obtenidos permiten vislumbrar tendencias relevantes en torno a la integración de la tecnología en la educación musical durante la primera infancia, es necesario señalar que el número de artículos seleccionados resulta insuficiente para extraer conclusiones generalizables o con un grado elevado de robustez científica. La escasez de investigaciones empíricas que aborden de manera conjunta la tecnología y la educación musical en este nivel educativo restringe la posibilidad de establecer comparaciones sólidas entre contextos y metodologías, lo que puede derivar en interpretaciones parciales o sesgadas del fenómeno. Esta limitación pone de relieve la necesidad de ampliar la producción científica en este campo, tanto en diversidad de muestras como en enfoques metodológicos, de modo que los futuros estudios permitan confirmar, matizar o incluso cuestionar los resultados preliminares identificados en la presente revisión.

Aunque el uso de tecnologías en la educación infantil presenta múltiples ventajas, encontramos ciertas condiciones que deben ser consideradas para maximizar su impacto.

Un primer reto es la dependencia tecnológica, que podría disminuir el uso de métodos pedagógicos tradicionales, afectando habilidades básicas como la memorización y el cálculo mental. Esto puede llevar a que tanto estudiantes como educadores dependan demasiado de las herramientas digitales, desplazando prácticas que han sido efectivas durante mucho tiempo (Gisbert-Caudeli et al., 2023).

Otro desafío importante es el acceso desigual a la tecnología, lo que puede aumentar la brecha educativa entre distintos grupos sociales o regiones geográficas. Mientras algunos estudiantes tienen acceso constante a dispositivos avanzados y conexión a internet, otros carecen de estas herramientas, lo que compromete su aprendizaje y participación en las actividades tecnológicas propuestas (Quispe et al., 2021).

Además, la formación insuficiente de los docentes es una limitación crucial. Sin la capacitación adecuada, las herramientas tecnológicas se usan de manera superficial o rutinaria, sin aprovechar su verdadero potencial para fomentar la creatividad y el aprendizaje significativo. Particularmente, el profesorado ha de contar con habilidades técnicas y pedagógicas actualizadas para maximizar el uso de la tecnología en el aula (Gisbert-Caudeli et al., 2023). Otra dificultad es la personalización limitada de los recursos digitales, ya que muchos no están diseñados para adaptarse a los ritmos y capacidades de aprendizaje individuales. Esto afecta especialmente a estudiantes con necesidades específicas o estilos de aprendizaje únicos, disminuyendo la eficacia de las herramientas tecnológicas (Quispe et al., 2021).

Existe también el riesgo de que las tecnologías sean vistas como un reemplazo de los métodos tradicionales, en lugar de complementarias. Este enfoque puede llevar a la pérdida de prácticas pedagógicas valiosas, como el juego y las interacciones cara a cara, esenciales para el desarrollo social y emocional de los niños en sus primeras etapas (Lee & Huai, 2015)

La proyección del uso de la tecnología en la educación infantil debe enfocarse en varios aspectos clave para maximizar su impacto positivo. En primer lugar, es esencial promover una integración equilibrada de la tecnología en el aula. Las herramientas digitales son valiosas, pero no deben desplazar las interacciones humanas ni los métodos tradicionales que fomentan el aprendizaje social y emocional de los niños. Un modelo educativo híbrido que combine lo mejor de ambos ámbitos permitirá que la tecnología complemente las experiencias de aprendizaje directas en lugar de reemplazarlas.

Otro aspecto crucial es la mejora de la capacitación docente. Para que la integración tecnológica sea efectiva, es necesario que los sistemas educativos inviertan en la formación continua de los educadores, asegurándose de que no solo dominen el uso de las herramientas, sino que también sepan integrarlas pedagógicamente. Esta capacitación debe ser adaptativa, permitiendo personalizar las experiencias de aprendizaje según las necesidades de cada estudiante y sus contextos socioculturales, ya que los docentes son el vínculo fundamental entre la tecnología y el aprendizaje.

Finalmente, se considera fundamental desarrollar tecnologías inclusivas y realizar evaluaciones rigurosas del impacto de la tecnología en la educación infantil. Concretamente, las herramientas digitales deben ser accesibles y adaptables para todo tipo de alumnado, incluyendo aquellos con necesidades específicas. Además, las investigaciones deben evaluar el impacto a largo plazo de estas herramientas en áreas

cognitivas y sociales. Un enfoque interdisciplinario en el uso de la tecnología permitirá conectar distintos campos del saber, potenciando las habilidades creativas de los estudiantes y preparándolos mejor para los desafíos futuros.

En síntesis, aunque la revisión realizada permite reconocer el potencial transformador de la tecnología en la educación musical infantil, la limitada cantidad de estudios disponibles impide extraer conclusiones de carácter generalizable y con suficiente solidez empírica. Los beneficios señalados —como el fomento de la creatividad, la motivación y el aprendizaje autónomo— se ven contrarrestados por desafíos significativos relacionados con la dependencia tecnológica, la desigualdad en el acceso, la insuficiente capacitación docente y la escasa personalización de los recursos digitales. Esta situación evidencia que, más allá de una aceptación acrítica de la tecnología como sustituto de las prácticas tradicionales, resulta indispensable avanzar hacia un modelo híbrido que combine de manera equilibrada lo digital y lo presencial, potenciando la dimensión social, emocional y cooperativa del aprendizaje en la primera infancia. Asimismo, la producción científica en este ámbito demanda una expansión tanto en número como en diversidad metodológica y contextual, de modo que futuras investigaciones puedan ofrecer una comprensión más completa, crítica y fundamentada del verdadero impacto de la integración tecnológica en la educación musical temprana.

Referencias

- Ajmahs (2017). Development of an Integrated Model for Curriculum for Early Childhood Music Practice Subject. *Asia-Pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, 7(8), 597-607. <http://doi.org/10.35873/ajmahs.2017.7.8.056>
- Arvelo-Rosales, C.N., Alegre de la Rosa, O.M., Guzmán-Rosquete, R. (2021). Initial Training of Primary School Teachers: Development of Competencies for Inclusion and Attention to Diversity. *Education Sciences*, 11, 413. <https://doi.org/10.3390/educsci11080413>
- Berrón-Ruiz, E., Arriaga-Sanz, C., & Campayo-Muñoz, E.A. (2023). Recursos tecnológicos para la formación inicial del profesorado de Música: una intervención en el contexto universitario español. *Revista Electrónica de LEEME*, 51, 16-35. <https://doi.org/10.7203/LEEME.51.25680>
- Bers, M. U. (2012). PTD Engagement Checklist. DevTech Research Group, Universidad de Tufts.
- Bers, M., González, C., & Armas, B. (2019). Coding as a playground: promoting positive learning experiences in childhood classrooms. *Computers & Education*, 138(3), 130-145. <https://doi.org/10.1016/j.compedu.2019.04.013>
- Bodrova, E., & Leong, D. (2024). Tools of the mind: The Vygotskian approach to early childhood education. Taylor & Francis. <https://doi.org/10.4324/9781003164920>
- Bucura, E., & Weissberg, J. (2017). Children's musical empowerment in two composition task designs. *Research & Issues in Music Education*, 13(1), 4. <https://commons.lib.jmu.edu/rime/vol13/iss1/4>

- Campollo-Urkiza, A. (2023). Development of a Programme of Musical Activities for the Contribution of Unplugged Computational Thinking in Early Childhood Education. *Revista Electrónica Educare*, 27(3), 1-17. <https://doi.org/10.15359/ree.27-3.17180>
- Cano, M., & Sánchez-Iborra, R. (2015). On the use of a multimedia platform for music education with handicapped children: a case study. *Computers and Education*, 87, 254-276. <https://doi.org/10.1016/j.compedu.2015.07.010>
- Chao-Fernandez, R., Román-García, S., & Chao-Fernandez, A. (2017) Online Interactive Storytelling as a Strategy for Learning Music and for Integrating Pupils with Hearing Disorders into Early Childhood Education (ECE), *Procedia - Social and Behavioral Sciences*, 237, 17-22. <https://doi.org/10.1016/j.sbspro.2017.02.005>
- Cuervo, L., Bonastre, C., & García, D. (2022). Tecnología digital en la educación musical infantil. *Praxis & Saber*, 13(32), e13201. <https://doi.org/10.19053/22160159.v13.n32.2022.13201>
- De Vries, P. (2021). Music at home with the under-fives: What is happening? En *Music in the Lives of Young Children* (pp. 59-72). Routledge. <https://doi.org/10.1080/03004430802691914>
- Duarte, M., & Sigal, J. (2019). Working with electroacoustic music in rural communities: the use of an interactive music system in the creative process in primary and secondary school education. *Organised Sound*, 24(3), 228-239. <https://doi.org/10.1017/S135577181900030X>
- Fernández-Company, J.F., García-Rodríguez, M., Moreno, A., & Gamella, D. (2021). Identificación emocional a través de la música en la infancia. En Elke Castro León (Coord.), *Oportunidades y retos para la enseñanza de las artes, la educación mediática y la ética en la era postdigital*, (pp. 975-1006). Dykinson.
- Fernández-Scagliusi, M. V., & Llorente-Cejudo, C. (2024). Evaluación de Competencias Digitales en Estudiantes de Educación: un estudio en la Universidad de Bolonia. *EduTec, Revista Electrónica de Tecnología Educativa*, 90, 92-110. <https://doi.org/10.21556/edutec.2024.90.3257>
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675-701. <https://doi.org/10.2307/2279372>
- Galindo-Domínguez, H., Delgado, N., Sainz de la Maza, M., & Expósito, E. (2024). Un análisis experimental de la relación entre las evaluaciones proporcionadas por la inteligencia artificial y las proporcionadas por los docentes en formación. *EduTec, Revista Electrónica de Tecnología Educativa*, 89, 84-104. <https://doi.org/10.21556/edutec.2024.89.3509>
- Gisbert-Caudeli, V., Fontes-Alayón, R., D Calderón-Garrido, D., & Anguita-Acero, J.M. (2023). Teacher Involvement in the Use of Digital Tools in Conservatory, Municipal Music School and University Classrooms. *Per Musi*, 24, 1-18. <https://doi.org/10.35699/2317-6377.2023.46441>

- Hallam, S., & Himonides, E. (2022). The power of music: An exploration of the evidence. Open Book Publishers.
- Hanrahan, F., Hughes, E., Banerjee, R., Eldridge, A., & Kiefer, C. (2019). Psychological benefits of networking technologies in children's experience of ensemble music making. *International Journal of Music Education*, 37(1), 59-77. <https://doi.org/10.1177/0255761418796864>
- Kata, A. (2016). The Structure and Development of Music Perception in Ages 5-17 Online Diagnostic Measurements in Kidergarten and School Environments. [Tesis de Maestría] Szegedi Tudományegyetem (Hungary).
- Kim, J. (2020). Learning and teaching online during Covid-19: Experiences of student teachers in an early childhood education practicum. *International journal of early childhood*, 52(2), 145-158. <https://doi.org/10.1007/s13158-020-00272-6>
- Kwon, H. M. (2019). A study of the trend and direction of keyboard instrument technology for education using patent information. *Journal of Next-generation Convergence Information Services Technology*, 8(3), 319-328. <http://doi.org/10.29056/jncist.2019.09.09>
- Lee, A., & Huai, Y. J. (2015). Interactive Whiteboard Integration into Music Teaching and Learning: Preschool Children as a Case Study. *Procedia - Social and Behavioral Sciences*, 177, 449-458. <https://doi.org/10.1016/j.sbspro.2015.02.394>
- Lee, L., & Ho, H.J. (2023). Effects of music technology on language comprehension and self-control in children with developmental delays. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2298. <https://doi.org/10.29333/ejmste/13343>
- Lee, Y., Kim, A.J., & Suk, K. (2016). Comparison of Perception between pre-service Teacher's and kindergarten Teachers' on Applications for Early Childhood education. *Journal of Educational Innovation Research*, 26(3), 267-288. <http://doi.org/10.21024/pnuedi.26.3.201612.267>
- Mosquera, C. F. B., Joaqui, M. F. S., Caicedo, M. I. V., & Velasco, E. P. (2022). Mapeo sistemático sobre la integración de tecnologías de inteligencia artificial y recursos TIC en aprendizaje-enseñanza musical. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação* 49, 479-490. <https://www.proquest.com/scholarly-journals/mapeo-sistemático-sobre-la-integración-de/docview/2714753112/se-2>
- Palazón-Herrera, J., & Soria-Vilchez, A. (2021). Students' perception and academic performance in a flipped classroom model within Early Childhood Education Degree. *Heliyon*, 7(4). <https://doi.org/10.1016/j.heliyon.2021.e06702>
- Paule-Ruiz, M., Álvarez-García, V., Pérez-Pérez, J. R., Álvarez-Sierra, M., & Trespalacios-Menéndez, F. (2016). Music learning in preschool with mobile devices. *Behaviour & Information Technology*, 36(1), 95-111. <https://doi.org/10.1080/0144929X.2016.1198421>

- Pérez-Moreno, J., & Reverté, L. (2019). Las actividades musicales preferidas de la voz de los propios niños y niñas de cuatro años. Un estudio exploratorio. *Revista Electrónica de Léeme*, 43, 19-34. <https://doi.org/10.7203/LEEME.43.13985>
- Preka, G. & Rangoussi, M. (2019) Augmented Reality and QR Codes for Teaching Music to Preschoolers and Kindergarteners: Educational Intervention and Evaluation. *SciTePress*, 113-123. <http://dx.doi.org/10.5220/0007682301130123>
- Quispe, A. M., Hinojosa-Ticona, Y., Miranda, H. A., & Sedano, C. A. (2021). Serie de Redacción Científica: Revisiones Sistemáticas. *Revista del Cuerpo Médico Hospital Nacional Almanzor Aguinaga Asenjo*, 14(1), 94-99. <https://dx.doi.org/10.35434/rcmhnaaa.2021.141.906>
- Rahiem, M.D.H. (2021) Storytelling in early childhood education: Time to go digital. *ICEP* 15. <https://doi.org/10.1186/s40723-021-00081-x>
- San Francisco Symphony. (s.f.). SFSKids: Fun with Music. <https://www.sfsymphony.org/education/community/sfskids>
- Strawhacker, A. L., & Bers, M. U. (2015). I want my robot to look for food: Comparing children's programming comprehension using tangible, graphical, and hybrid user interfaces. *International Journal of Technology and Design Education*, 25(3), 293–319. <https://doi.org/10.1007/s10798-014-9287-7>
- Sullivan, A., & Bers, M.U. (2018). Dancing robots: integrating art, music, and robotics in Singapore's early childhood centers. *International Journal Technology and Design Education*, 28, 325–346. <https://doi.org/10.1007/s10798-017-9397-0>
- Sverre, J., Sagmo, G., Danbolt, I., & Engesnes, N. (2019). Musical pathfinders of the Kindergarten. *Contemporary Issues in Early Childhood*, 20(2) 163–176. <https://doi.org/10.1177/1463949118756369>
- Valenzuela-Romero, R., Vera-Alcázar, M., & Herrera-Córdova, F. (2025). Realidad aumentada y educación inclusiva para la comunidad sorda: Una revisión sistemática. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, 57(3), 72-84. DOI: 10.17013/risti.57.72–84
- Vasil, M., Weiss, L., & Powell, B. (2018). Popular music pedagogies: An approach to teaching 21st-entury skills. *Journal of Music Teacher Education*, 28(3), 85-95. <https://doi.org/10.1177/1057083718814454>
- Young, P., & Suk, Y.J. (2022). Review on Digital Technology Convergence Studies for Therapeutic Use of Music. *Culture and Convergence*, 44(11), 953-975. <http://dx.doi.org/10.33645/cnc.2022.11.44.11.953>

Caracterização espectral da cobertura do solo do Parque Estadual do Cocó a partir de imagens CBERS-4A

Mauricio Alejandro Perea-Ardila¹

mauricio.perea@alu.ufc.br

¹ Programa de Pós-Graduação em Geografia (PPGGeo) Universidade Federal do Ceará, 60440-554, Fortaleza, Brasil

DOI: 10.17013/risti.59.83-97

Resumo: Os parques urbanos são essenciais para a biodiversidade e o bem-estar humano. O Parque Estadual do Cocó (PEC), em Fortaleza, Brasil, sofre alta pressão antropogênica, exigindo monitoramento contínuo. Este estudo caracterizou espectralmente suas coberturas com imagens do sensor WPM do satélite CBERS-4A, usando uma cena de 24 de julho de 2024. Foram analisadas assinaturas espectrais das principais classes, com base na classificação do MapBiomas, e calculados os índices NDVI, SAVI e EVI para avaliar a vegetação. Os resultados mostraram que o sensor WPM discrimina eficazmente as coberturas, com assinaturas distintas para manguezais, restingas, água e áreas urbanas. O manguezal domina (50,78%), seguido pela restinga herbácea (27,24%) e áreas urbanizadas (10,78%), evidenciando pressão periférica. O EVI superou o NDVI em sensibilidade em áreas de alta biomassa, evitando a saturação. As imagens CBERS-4A provaram ser ferramentas eficazes e econômicas para monitorar ecossistemas urbanos complexos, fornecendo dados essenciais para a conservação e gestão sustentável do PEC.

Palavras-chave: Assinatura espectral; Índices de vegetação; Monitoramento ambiental.

Spectral characterization of land cover in Cocó State Park using CBERS-4A images

Abstract: Urban parks are essential for biodiversity and human well-being. The Cocó State Park (PEC) in Fortaleza, Brazil, faces significant anthropogenic pressure, necessitating continuous monitoring. This study spectrally characterized land cover using WPM sensor imagery from the CBERS-4A satellite, based on a scene from July 24, 2024. Spectral signatures of major land cover classes were analyzed, supported by MapBiomas classification, and vegetation indices (NDVI, SAVI, and EVI) were computed to assess vegetation condition. Results demonstrated that the WPM sensor effectively discriminates land cover types, with distinct spectral signatures for mangroves, restinga, water bodies, and urban areas. Mangroves dominate the park (50.78%), followed by herbaceous restinga (27.24%) and urbanized areas (10.78%), highlighting substantial peripheral pressure. EVI outperformed NDVI

in sensitivity to structural variations in high-biomass zones, avoiding saturation. CBERS-4A imagery proves to be a cost-effective and reliable tool for monitoring complex urban ecosystems, providing critical data for informed decision-making in the sustainable conservation and management of the PEC.

Keywords: Spectral signature; Vegetation indices; Environmental monitoring.

1. Introdução

Os parques urbanos e as áreas protegidas desempenham um papel crucial na conservação da biodiversidade, na prestação de serviços ecossistêmicos e na promoção do bem-estar humano em ambientes urbanos. No contexto de cidades em rápido crescimento, como Fortaleza, no Brasil, esses espaços verdes representam oásis de biodiversidade que mitigam os impactos da urbanização, como a perda de habitats e a degradação ambiental (Chiesura, 2004). O monitoramento constante e preciso da cobertura do solo em ambientes urbanos é essencial para o gerenciamento eficaz dos recursos naturais e a detecção oportuna de mudanças na paisagem. A dinâmica da cobertura do solo tem implicações significativas na sustentabilidade ambiental, no planejamento urbano e na conservação dos ecossistemas. De acordo com Hansen et al., (2013), as mudanças no uso do solo, como o desmatamento e a urbanização, aceleraram a perda de biodiversidade e alteraram os ciclos hidrológicos, o que ressalta a necessidade de ferramentas de monitoramento precisas para avaliar essas transformações em escala global e local.

Nesse contexto, a implementação de tecnologias avançadas, como a teledeteção, transformou a capacidade de monitoramento, permitindo uma avaliação mais detalhada das mudanças na paisagem (Robinson et al., 2022; Zhu et al., 2015). A teledeteção consolidou-se como uma ferramenta fundamental para o monitoramento de grandes áreas, oferecendo uma solução eficiente, de baixo custo e com alta capacidade de cobertura espacial e temporal para a observação da superfície terrestre (Bairwa et al., 2025). O satélite CBERS-4A, parte do Programa Sino-Brasileiro de Recursos Terrestres (CBERS), representa um avanço significativo na teledeteção moderna, graças à sua combinação de resolução espacial, temporal e espectral (Vieira et al., 2025). O estudo espectral da cobertura do solo é fundamental para compreender a dinâmica ambiental, avaliar a saúde dos ecossistemas e monitorar mudanças na vegetação (Sharma et al., 2019). Através da análise das assinaturas espectrais captadas pelos sensores, é possível identificar mudanças na cobertura vegetal, detectar o desmatamento, avaliar o impacto das atividades antropogênicas e monitorar a recuperação de áreas degradadas (Kumar et al., 2021).

No caso do Parque do Cocó (PEC), uma área de grande relevância ecológica dentro de um ambiente urbano em expansão, o uso da teledeteção fornece ferramentas essenciais para a tomada de decisões em matéria de conservação da biodiversidade, planejamento urbano sustentável e mitigação das pressões ambientais, contribuindo assim para a preservação de um dos últimos remanescentes de vegetação natural na região metropolitana de Fortaleza (Farias et al., 2018). A gestão eficiente desses ecossistemas requer informações essenciais que permitam monitorar de forma precisa e contínua a cobertura do solo, especialmente em áreas onde a pressão antropogênica ameaça a integridade ecológica.

O objetivo deste estudo é caracterizar as principais classes de cobertura do solo do PEC utilizando imagens do sensor WPM do satélite CBERS-4A. Através da análise de assinaturas espectrais e índices de vegetação, busca-se gerar informações espaciais atualizadas que apoiem a conservação do parque, facilitem a tomada de decisões no planejamento urbano e contribuam para o monitoramento dos ecossistemas naturais em um ambiente urbano em constante transformação.

2. Área do estudo

O PEC, localizado em Fortaleza, Ceará, Brasil, representa uma das maiores áreas verdes urbanas protegidas do norte e nordeste do país, com uma extensão de aproximadamente 1080,73 hectares (trecho 1) e está localizado entre as coordenadas geográficas $3^{\circ} 44' 37,44''$ e $3^{\circ} 47' 15,41''$ S e $38^{\circ} 30' 45,96''$ e $38^{\circ} 25' 52,36''$ W (Perea-Ardila & Muñoz, 2024). O clima predominante na região é tropical semiárido, com temperaturas médias mensais que oscilam entre 24°C e 30°C e uma precipitação média anual próxima a 1300 mm, concentrada principalmente entre os meses de fevereiro e maio (Price et al., 2024). De acordo com a classificação climática de Köppen-Geiger, está dentro da unidade climática “Aw” (Pasquini, 2020). Este regime climático influencia diretamente a dinâmica ecológica do parque, que abriga ecossistemas como manguezais, dunas, restingas e mata de tabuleiros (Sales & De Araújo, 2023).

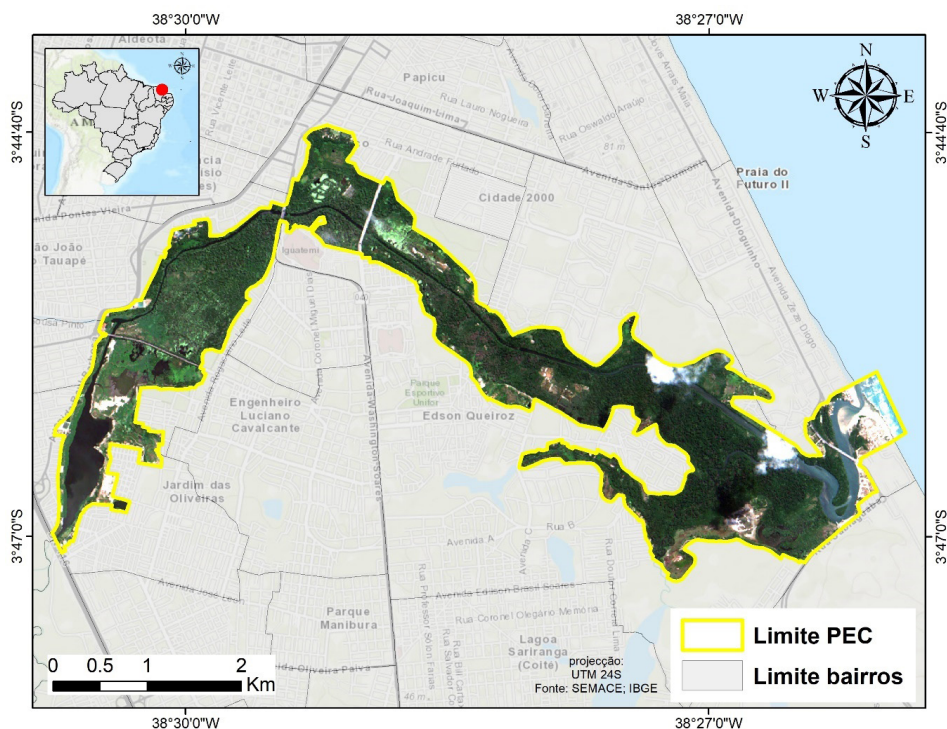


Figura 1 – Localização do Parque Estadual do Cocó-PEC.

2.1. Satélite e bandas espectrais utilizadas

O satélite CBERS-4A (China-Brazil Earth Resources Satellite-4A) é uma missão conjunta entre o Brasil e a China projetada para a observação da Terra e o monitoramento de recursos naturais. Lançado em dezembro de 2019, ele incorpora sensores avançados que melhoram significativamente a qualidade e a resolução das imagens. Um de seus principais instrumentos é o sensor WPM (Wide Field Imager), que oferece ampla cobertura com resolução espacial de 8 metros e revisita o mesmo ponto em 31 dias, ideal para o monitoramento regional da cobertura do solo, uso do território e mudanças ambientais (Tabela 1). Além disso, conta com bandas espectrais no visível e no infravermelho próximo, o que permite a diferenciação eficaz de tipos de vegetação, áreas urbanas, corpos d’água e zonas degradadas. A imagem utilizada neste estudo foi da data de 24/07/2024, pois estava livre de nuvens em mais de 95% da cena.

Sensor	Bandas espectrais	Resolução espectral	Resolução espacial
WPM	Pancromática (0)	0,45 – 0,90 µm	2 m
	Azul (1)	0,45 – 0,52 µm	8 m
	Verde (2)	0,52 – 0,59 µm	
	Vermelho (3)	0,63 – 0,69 µm	
	NIR (4)	0,77 – 0,89 µm	

Tabela 1 – Características do satélite CBERS-4A.

2.2. Processamento de imagens

Foi realizado um processo de ajuste radiométrico para as bandas espectrais do CBERS-4A, a fim de garantir a precisão e a consistência dos valores de refletância registrados, corrigindo as distorções causadas por fatores como iluminação variável, atmosfera e características próprias do sensor (Polidorio et al., 2005). Para tal, foi necessário converter os valores dos números digitais (ND) das bandas espectrais em radiação TOA (Equação 1), utilizando os parâmetros estabelecidos por (Pinto et al., 2016; Ponzoni et al., 2005).

$$L_{\lambda} = G_n \times DN_n + offset \tag{1}$$

Onde: L_{λ} = radiação espectral aparente expressa em (W/(m2” sr” µm));

G= coeficiente de ganho na banda n

ND= número digital extraído da imagem na banda n (n=1,2,3,4).

Além disso, para determinar o valor da refletância TOA das bandas espectrais, foi necessário usar a equação 2, com o objetivo de transformar os valores de radiação em uma medida padronizada de refletância planetária, que permite comparar diferentes

coberturas do solo independentemente das condições atmosféricas, da geometria solar ou das características do sensor. Foram utilizados os parâmetros estabelecidos por (Pinto et al., 2016).

$$P_{\lambda} = \frac{\pi \cdot L_{\lambda} \cdot d^2}{ESUN \cdot \cos(\theta_{sz})} \quad (2)$$

Onde: P_{λ} = é a refletância planetária TOA (sem unidades);

L_{λ} = é a radiação espectral na banda n ;

d = é a distância Terra-Sol (unidades astronômicas);

$ESUN$ = é a irradiância solar exoatmosférica média

θ_{sz} = é o ângulo cenital solar (graus ou radianos).

Além disso, foi realizada a correção atmosférica das bandas utilizando parâmetros extraídos da própria imagem, através do método DOS (Dark Object Subtraction) proposto por Chávez, (1996) implementado no QGIS através do módulo Semi-Automatic Classification Plugin (SCP) (Congedo, 2021). O objetivo foi minimizar os efeitos da atmosfera e obter valores de refletância mais próximos dos reais da superfície terrestre.

2.3. Caracterização espectral

Foi baixada uma camada de coberturas e uso do solo do MapBiomass do ano de 2023, que foi recortada até o limite do PEC. Isso foi feito com o objetivo de conhecer a distribuição das coberturas e analisar os padrões de ocupação territorial, identificando áreas naturais, agrícolas, urbanas e corpos d'água dentro do perímetro do projeto. A Coleção 9 do MapBiomass fornece dados anuais de 1985 a 2023, com mais de 30 classes temáticas que permitem uma avaliação detalhada das mudanças na paisagem brasileira (MapBiomass, 2024). Foram criados 50 pontos aleatórios por cobertura, totalizando 500 pontos de controle que serviram para extrair os valores de TOA e realizar uma assinatura espectral das coberturas.

Para caracterizar o vigor, cobertura e estrutura da vegetação, foram determinados três índices de vegetação (NDVI, SAVI e EVI) utilizando imagens multiespectrais do satélite CBERS-4A. Os valores médios, mínimos e máximos desses índices foram analisados para identificar padrões. O NDVI é comum para estimar biomassa, mas satura em áreas densas. O SAVI corrige a influência do solo, sendo mais eficaz em regiões com vegetação escassa. O EVI (Enhanced Vegetation Index) oferece a maior sensibilidade em áreas de alta densidade e corrige interferências atmosféricas pela banda azul (Leivas et al., 2022). O cálculo e as estatísticas descritivas (Tabela 2) permitem uma análise detalhada das condições da vegetação na área de estudo.

Índice de vegetação	Formula	Fonte
Índice de vegetação de diferença normalizada (NDVI)	$NDVI = \frac{NIR - Vermelho}{NIR + Vermelho}$	(Rouse et al., 1974)
Índice de vegetação ajustado ao solo (SAVI)	$SAVI = \frac{NIR - Vermelho}{NIR + Vermelho + L} * 1$	(A. R. Huete, 1988)
Índice de Vegetação Melhorado (EVI)	$EVI = 2.5 \frac{NIR - Vermelho}{NIR + (2.4 * Vermelho) + 1}$	(A. Huete et al., 2002)

Tabela 2 – Índices de vegetação determinados para o PEC.

3. Resultados

A cor vermelha intensa representou vegetação saudável e densa. No contexto do rio Cocó, essas áreas correspondem principalmente às florestas de manguezais e à vegetação ribeirinha. A intensidade do vermelho é diretamente proporcional à atividade fotossintética e à biomassa; portanto, as áreas com um vermelho mais brilhante indicam um ecossistema vegetal mais vigoroso e exuberante (Figura 2).

Os tons em preto/azul escuro representam corpos de água profundos e/ou limpos. Correspondem ao canal principal do rio Cocó e às lagoas associadas. A água absorve fortemente a radiação infravermelha próxima, pelo que aparece escura nesta composição. Os tons ciano/azul claro pode indicar dois fenômenos principais: água com alta carga de sedimentos ou turbidez ou águas muito rasas ou bancos de areia/lama úmida (planícies de maré). Na imagem, essa tonalidade é proeminente perto da foz do rio (extremo direito, em direção à Praia do Futuro), o que sugere fortemente que se trata de planícies de maré e águas turvas onde o rio interage com o oceano.

Os tons em branco/cinza/roxo claro representam superfícies com alta refletividade no espectro visível, como áreas urbanizadas (edifícios, ruas, asfalto) e solo exposto ou areia seca. A trama quadriculada que circunda o parque, com essas cores, confirma que se trata de setores da cidade de Fortaleza. Qualquer mancha dessa cor dentro do limite amarelo do parque pode indicar infraestrutura, estradas ou áreas degradadas.

O PEC revelou um ecossistema dominado por formações vegetais nativas. A classe Mangue é a mais representativa, abrangendo 50,78% (540,54 ha) da superfície total, com uma distribuição espacial concentrada ao longo do curso principal do rio, formando o núcleo ecológico do parque (Figura 3). A segunda classe mais extensa é a Restinga Herbácea, que corresponde a 27,24% da área e geralmente margeia as zonas de manguezais. Uma

descoberta significativa é a proporção considerável de Área Urbanizada, que ocupa 10,78% da área delimitada e está localizada principalmente na periferia do parque, o que evidencia a forte pressão urbana nos limites da unidade de conservação.

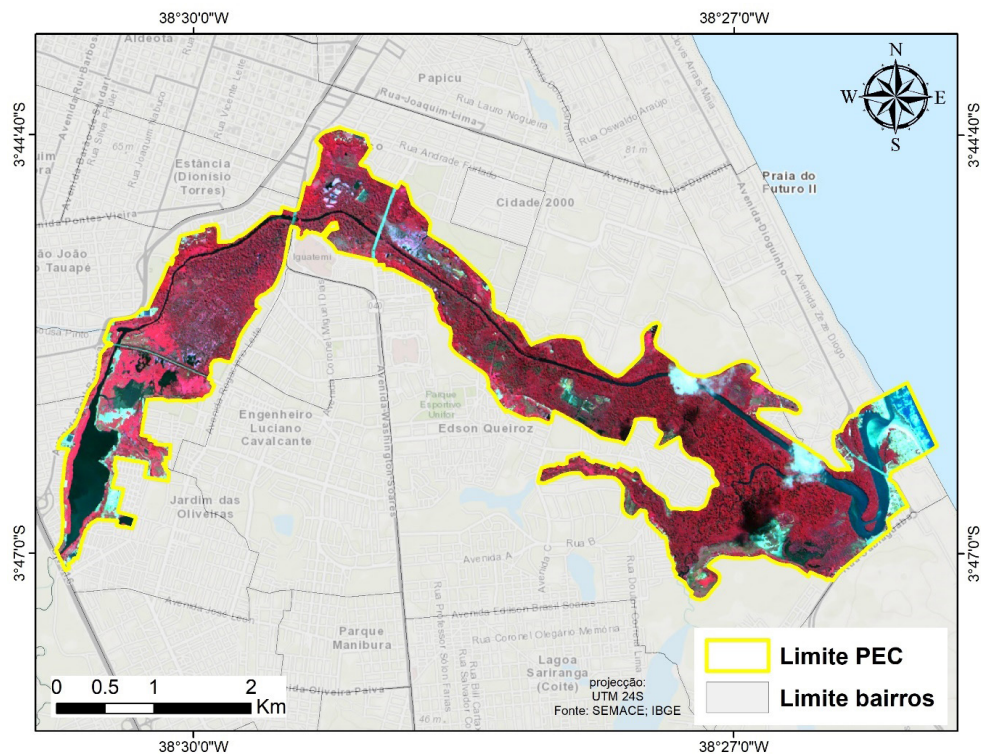


Figura 2 – Imagem CBERS-4A composta em falsa cor do PEC.

O gráfico de assinaturas espectrais ilustra o comportamento da reflectância média para cada classe de cobertura do solo nas quatro bandas espectrais do sensor (Figura 4). A análise das curvas permite confirmar a separabilidade entre as classes e compreender sua resposta biofísica. As classes de vegetação de Mangue, Restinga Herbácea, Formação Savânica e Pastagem exibem a assinatura característica da vegetação saudável. Elas apresentam uma absorção de energia notável na banda do Vermelho (B3), seguida por um aumento acentuado da reflectância na banda do Infravermelho Próximo (NIR, B4), resultado da estrutura celular das folhas. A Restinga Herbácea e a Pastagem demonstram os maiores valores de reflectância no NIR, indicando um alto vigor vegetativo.

As análises espectrais revelaram assinaturas distintas para as diferentes classes de cobertura do solo. A classe Rio, Lago e Oceano apresentou um comportamento típico de corpos d'água, com a reflectância diminuindo progressivamente do azul para o infravermelho devido à alta absorção de energia. Em contraste, superfícies com

pouca ou nenhuma vegetação, como Área Urbanizada, Apicum e Praia, Duna e Areal, mostraram um aumento gradual da reflectância do visível para o infravermelho, sem o “salto” abrupto da vegetação. A classe Praia, Duna e Areal destacou-se com a reflectância mais elevada em todas as bandas, consistente com a resposta da areia. A sobreposição parcial entre Área Urbanizada e Formação Savânica nas bandas visíveis sugere um potencial dificuldade de separação em áreas de baixa densidade. No geral, a distinção das assinaturas espectrais valida a capacidade do sensor em discriminar as principais coberturas.

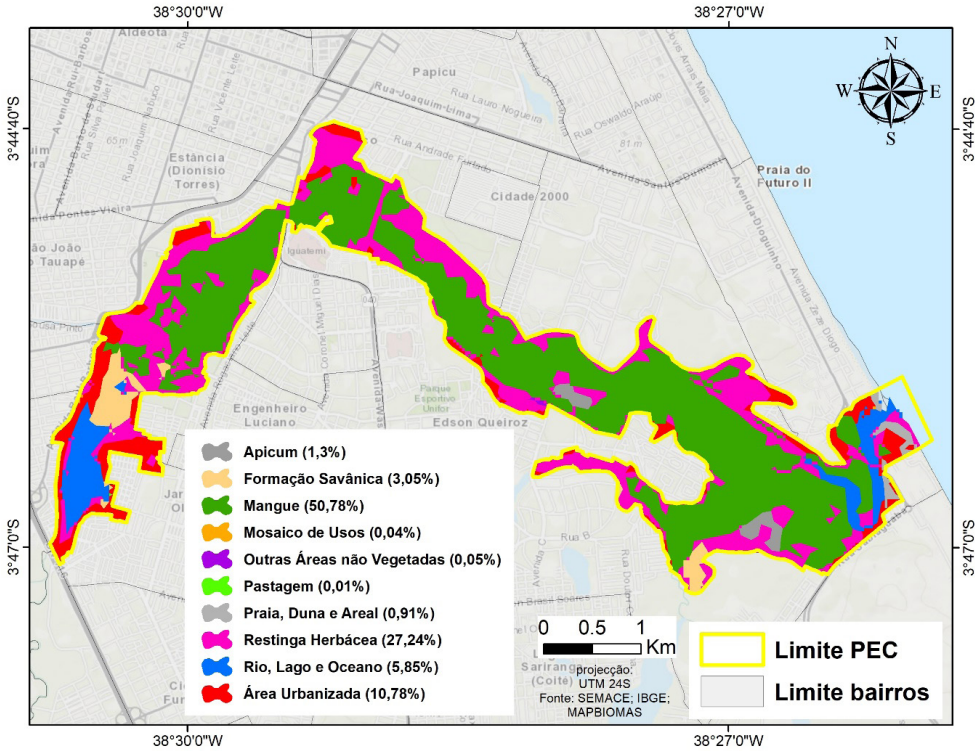


Figura 3 – Cobertura e uso do solo do PEC.

A análise estatística dos índices de vegetação (Tabela 3) fornece uma avaliação quantitativa da resposta espectral de cada classe de cobertura do solo. Os resultados demonstram uma clara diferenciação entre as classes com base na densidade e no vigor da vegetação. Conforme esperado, as classes de vegetação densa apresentaram os maiores valores médios. O Mangue se destacou com o valor médio mais alto de NDVI (0,654), seguido pela Restinga Herbácea (0,579) e pelo Mosaico de Usos (0,589), o que reflete a alta atividade fotossintética e biomassa dessas coberturas.

A classe Rio, Lago e Oceano apresentou os menores valores médios para os índices de vegetação, incluindo mínimos negativos de NDVI (-0,232) e EVI (-0,045), comportamento

típico de corpos d’água. Outras classes com pouca vegetação, como Apicum, Praia, Duna e Areal e Pastagem, também exibiram valores médios consistentemente baixos. A dispersão dos dados revela diferenças importantes: classes heterogêneas como Área Urbanizada e Mosaico de Usos mostram uma ampla faixa de valores, refletindo a mistura de alvos (vegetação, solo, construções). Por outro lado, o Mangue possui uma faixa de valores mais concentrada em níveis altos, indicando uma cobertura vegetal mais homogênea e densa

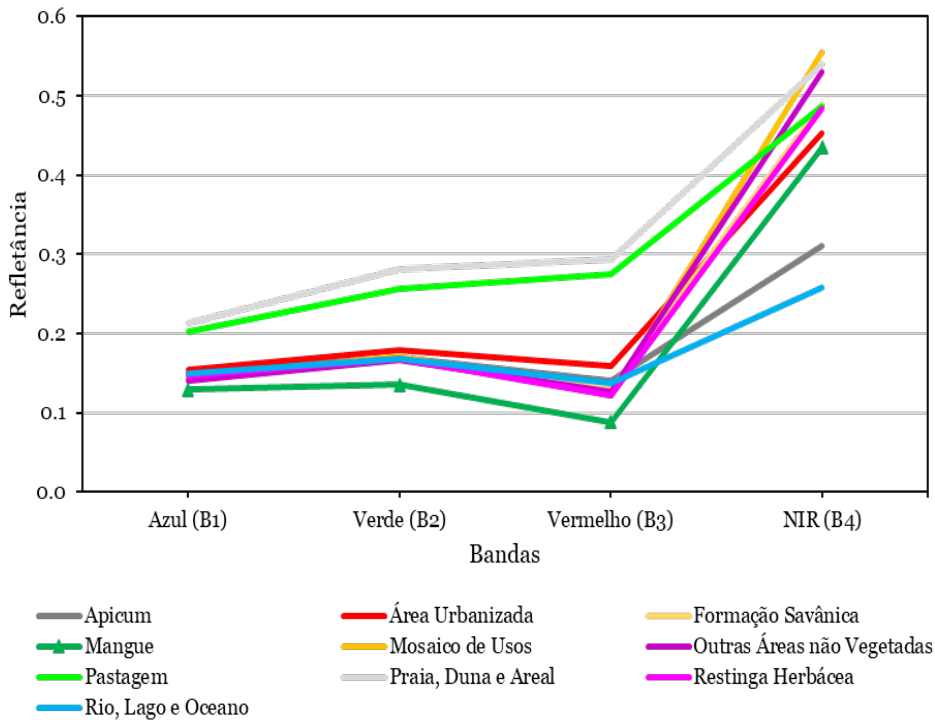


Figura 4 – Assinaturas espectrais médias dos uso e coberturas identificadas do PEC.

Cobertura	NDVI			SAVI			EVI		
	Min	Max	Media	Min	Max	Media	Min	Max	Media
Apicum	0.128	0.624	0.379	0.071	0.461	0.265	0.023	0.420	0.180
Área Urbanizada	0.115	0.720	0.456	0.049	0.625	0.381	0.016	0.634	0.373
Formação Savânica	0.214	0.811	0.540	0.108	0.800	0.452	0.043	1.111	0.458
Mangue	0.401	0.759	0.654	0.221	0.618	0.500	0.077	0.507	0.353
Mosaico de Usos	0.027	0.786	0.589	0.015	0.753	0.523	0.007	0.946	0.553
Outras Áreas não Vegetadas	0.420	0.719	0.616	0.372	0.639	0.522	0.290	0.690	0.496

Cobertura	NDVI			SAVI			EVI		
	Min	Max	Media	Min	Max	Media	Min	Max	Media
Pastagem	0.203	0.388	0.277	0.179	0.343	0.251	0.231	0.443	0.335
Praia, Duna e Areal	0.158	0.592	0.309	0.151	0.522	0.283	0.231	0.663	0.410
Restinga Herbácea	0.219	0.788	0.579	0.174	0.757	0.476	0.079	0.966	0.415
Rio, Lago e Oceano	-0.232	0.810	0.172	-0.118	0.802	0.143	-0.045	1.149	0.170

Tabela 3 – Estatísticas dos índices de vegetação por classe de cobertura do PEC.

A Figura 5 ilustra a distribuição espacial dos índices de vegetação (NDVI, SAVI e EVI) no Parque Estadual do Cocó (PEC), mostrando uma alta coerência espacial. Áreas de vegetação densa (manguezais e restingas) são consistentemente representadas com valores elevados (tons verdes), enquanto superfícies de água, solo e áreas urbanizadas têm valores baixos (tons laranja/vermelho). Comparativamente, o mapa NDVI exibe o maior contraste e mais zonas de verde intenso, sugerindo saturação em copas mais densas. O SAVI, projetado para mitigar a influência do solo, segue um padrão similar ao NDVI, mas com uma atenuação geral dos valores.

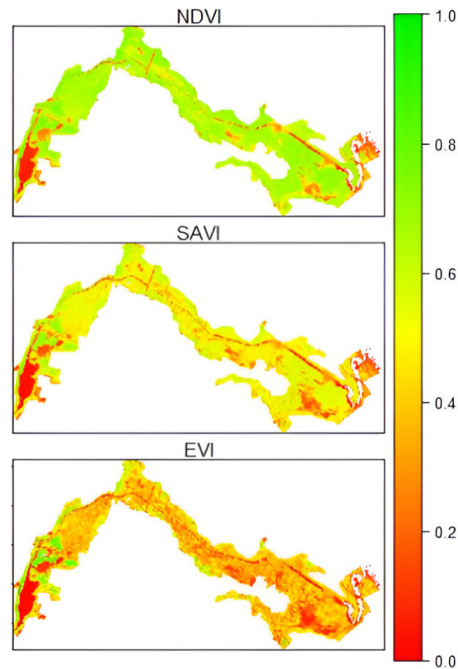


Figura 5 – Caracterização dos índices de vegetação do PEC.

Notavelmente, o mapa EVI revela uma maior textura e variabilidade dentro das massas florestais. Ao contrário do NDVI, o EVI parece ser mais sensível às variações estruturais do dossel, mostrando uma menor saturação em áreas de alta biomassa. Esta característica permite diferenciar com maior detalhe a heterogeneidade interna da vegetação densa. Em conjunto, a representação espacial dos índices confirma sua eficácia para mapear a vegetação e destaca a capacidade do EVI de fornecer uma visão mais matizada da estrutura do dossel em ecossistemas complexos como o manguezal.

4. Discussão

Os resultados confirmam a utilidade da teledeteção multiespectral e do satélite CBERS-4A como ferramenta eficaz para distinguir tipos de cobertura do solo em ambientes urbanos complexos como o PEC. A capacidade do CBERS-4A, via seu sensor WPM e bandas visíveis/infravermelho próximo, de capturar assinaturas espectrais diferenciadas entre vegetação, corpos d'água e áreas urbanizadas demonstra seu potencial para monitoramento ambiental (Dias dos Santos et al., 2023; Vieira et al., 2025). Sua aplicação valida-se em contextos de conservação, planejamento urbano e avaliação de impactos (Choi et al., 2011). Por outro lado, uma resolução temporária de 31 dias, como a oferecida pelo sensor WPM do satélite CBERS-4A, pode ser um obstáculo significativo para o monitoramento de mudanças rápidas, pois eventos importantes podem ocorrer e desaparecer completamente no intervalo entre duas aquisições de imagens. Por exemplo, incêndios florestais que ocorrem no PEC podem acontecer em questão de dias ou semanas (Perea-Ardila et al., 2023).

Os índices NDVI, SAVI e EVI distinguem bem as coberturas; contudo, o NDVI satura em vegetação densa. O EVI se destaca, mostrando maior sensibilidade estrutural do dossel e precisão em ecossistemas complexos (A. Huete et al., 2002), sendo recomendado em áreas de alta biomassa. Os tipos de água (rio, lago e oceano) apresentaram assinaturas espectrais típicas, validando o uso do NDWI para identificação (Yang et al., 2017). Zonas de transição, como planícies de maré, requerem a combinação de múltiplos índices para caracterização precisa. A presença de 10,78% de infraestrutura urbana dentro do parque evidencia uma pressão antropogênica significativa. Essa fragmentação da paisagem urbana pode comprometer a conectividade e resiliência dos ecossistemas (Seto et al., 2012). Pesquisas destacam que a expansão urbana descontrolada é chave na perda de biodiversidade, recomendando a integração da biodiversidade como eixo central no planejamento urbano (Elmqvist et al., 2021).

5. Conclusões

O estudo confirma que o satélite CBERS-4A é uma ferramenta eficaz para mapear e monitorar coberturas em ecossistemas urbanos complexos. As assinaturas espectrais e índices de vegetação permitiram uma clara discriminação, sendo o índice EVI mais sensível que o NDVI para caracterizar o vigor da vegetação densa.

Os resultados revelam que o manguezal é o ecossistema predominante, cobrindo 50,78% da área e consolidando-se como o núcleo ecológico. Críticamente, a análise também quantificou a forte pressão antropogênica, com 10,78% de área urbanizada detectada

dentro do perímetro do parque. Em termos práticos, a pesquisa fornece uma base de dados espaciais atualizada, essencial para a gestão ambiental do PEC e o planejamento urbano de Fortaleza. O mapeamento preciso da pressão urbana e dos ecossistemas nativos oferece insumos diretos para que gestores públicos formulem políticas de conservação mais eficazes, fiscalizem os limites e promovam um desenvolvimento sustentável. A pesquisa sugere a expansão para um sistema de monitoramento contínuo com imagens CBERS-4A para gestão adaptativa futura. Outros estudos podem explorar séries temporais de imagens multiespectrais de diferentes sensores (como Sentinel-2 e Landsat) para avaliar a dinâmica espaço-temporal da cobertura do solo no PEC, ampliando a compreensão sobre processos de degradação e regeneração da vegetação. Além disso, a integração de dados hiperespectrais, LiDAR e drones poderia fornecer informações mais detalhadas sobre a estrutura do dossel e a conectividade dos ecossistemas. Estudos também podem incluir a modelagem preditiva dos impactos das mudanças climáticas e da expansão urbana sobre a biodiversidade local, fortalecendo estratégias de conservação e gestão adaptativa do parque.

Referências

- Bairwa, B., Sharma, R., Kundu, A., Sammen, S. S., Alshehri, F., Pande, C. B., Orban, Z., & Salem, A. (2025). Predicting changes in land use and land cover using remote sensing and land change modeler. *Frontiers in Environmental Science*, 13, 1–17. <https://doi.org/10.3389/fenvs.2025.1540140>
- Chávez, P. S. J. (1996). Image-based atmospheric corrections - revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62(9), 1025–1036. <https://doi.org/10.1112/96/6209-1025>
- Chiesura, A. (2004). The role of urban parks for the sustainable city. *Landscape and Urban Planning*, 68(1), 129–138. <https://doi.org/10.1016/j.landurbplan.2003.08.003>
- Choi, J., Khanal, S., & Ambinakudige, S. (2011). Análise Comparativa do CBERS-2 e Imagens de Satélite Landsat-TM em Cartografia da Vegetação. *Revista Brasileira de Cartografia*, 63(1), 115–122. <https://doi.org/10.14393/rbcv63n1-43723>
- Congedo, L. (2021). Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. *Journal of Open Source Software*, 6(64), 3172. <https://doi.org/10.21105/joss.03172>
- Dias dos Santos, B., Neves de Azevedo, M., Bernardino Kanzato, L.A., Duque de Pinho, C. M., Paez, A., & Amaral, S. (2023). Cbers-4A Imagery for Mapping Urban Land Cover in the Amazon. In XX Simpósio Brasileiro de Sensoriamento Remoto (pp. 2169–2172). <https://proceedings.science/p/164620?lang=pt-br>
- Elmqvist, T., Andersson, E., McPhearson, T., Bai, X., Bettencourt, L., Brondizio, E., Colding, J., Daily, G., Folke, C., Grimm, N., Haase, D., Ospina, D., Parnell, S., Polasky, S., Seto, K. C., & Van Der Leeuw, S. (2021). Urbanization in and for the Anthropocene. *Npj Urban Sustainability*, 1(1), 6. <https://doi.org/10.1038/s42949-021-00018-w>

- Farias, I. F., Silva, C. R. M. da, Lima, D. S. V. R., Oliveira, L. V. C., & Fontenele, R. E. S. (2018). Valoração Ambiental do Parque Ecológico do Rio Cocó. *Desenvolvimento Em Questão*, 16(45), 191–213. <https://doi.org/10.21527/2237-6453.2018.45.191-213>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342(6160), 850–853. <https://doi.org/10.1126/science.1244693>
- Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Huete, A. R. (1988). Soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Kumar, S., Shwetank, S., & Jain, K. (2021). Development of spectral signature of land cover and feature extraction using artificial neural network model. 2021. In International Conference on Computing, Communication, and Intelligent Systems, (pp. 113–118). <https://doi.org/10.1109/ICCCIS51004.2021.9397172>
- Leivas, J. L., Teixeira, A. H. C., Takemura, C. M., & Garçon, E. A. (2022). Análise da dinâmica temporal de índices de vegetação NDVI, EVI, SAVI e IRECI através de imagens Sentinel-2a e MODIS. In J. O. F. Melo (Ed.), *Ciências Agrárias: o avanço da ciência no Brasil* (Vol. 4, pp. 332–343). Editora Científica Digital. <https://doi.org/10.37885/220709585>
- MapBiomass. (2024). *MapBiomass General “Handbook”: Algorithm Theoretical Basis Document (ATBD) Collection 9*. MapBiomass. <https://brasil.mapbiomas.org/wp-content/uploads/sites/4/2025/02/ATBD-Collection-9-versao2-v2.pdf>
- Pasquini, B. (2020). Plano de manejo do Parque Estadual do Cocó. Arcadis. https://www.sema.ce.gov.br/wp-content/uploads/sites/36/2021/03/PMPC_01.pdf
- Perea-Ardila, M. A., & Muñoz, S. I. (2024). Caracterização de uma nova queimada utilizando sensoriamento remoto do Parque Estadual do Cocó, Região Metropolitana de Fortaleza/CE, Brasil. *Journal of Hyperspectral Remote Sensing*, 14(1), 858–869. <https://doi.org/10.29150/jhrs.v14i1.261692>
- Perea-Ardila, M. A., Muñoz, S. I., & Sopchaki, C. H. (2023). Análise de áreas queimadas utilizando imagens Sentinel-2 no Parque Estadual do Cocó, Região Metropolitana de Fortaleza (Ceará). *Revista Brasileira de Sensoriamento Remoto*, 4(3), 67–83. <https://doi.org/10.5281/zenodo.10440566>
- Pinto, C., Ponzoni, F., Castro, R., Leigh, L., Mishra, N., Aaron, D., & Helder, D. (2016). First in-Flight Radiometric Calibration of MUX and WFI on-Board CBERS-4. *Remote Sensing*, 8(5), 405. <https://doi.org/10.3390/rs8050405>

- Polidorio, A., Clélia, F., Imai, N., Tommaselli, A., & Galo, M. de L. (2005). Correção radiométrica de imagens multiespectrais CBERS e Landsat ETM usando atributos de reflectância e de cor. In XII Simpósio Brasileiro de Sensoriamento Remoto (pp. 4241–4248). <http://marte.sid.inpe.br/col/ltid.inpe.br/sbsr/2004/11.19.19.46/doc/thisInformationItemHomePage.html>
- Ponzone, F. J., Junior, J., & Lamparelli, R. A. (2005). Calibração absoluta da câmera CCD/CBERS-2. In XII Simpósio Brasileiro de Sensoriamento Remoto (pp. 1067–1074). <http://marte.dpi.inpe.br/col/ltid.inpe.br/sbsr/2004/11.09.15.05/doc/1067.pdf>
- Price, J., Forstnerhäusler, N., Graham, E., Osborn, T. J., & Warren, R. (2024). Report on the observed climate, projected climate, and projected biodiversity changes for Parque Estadual Do Cocó under differing levels of warming. Centre for Climate Change Research. <https://wallaceparcs.uea.ac.uk/Brazil/Tembe.pdf>
- Robinson, J. M., Harrison, P. A., Mavoa, S., & Breed, M. F. (2022). Existing and emerging uses of drones in restoration ecology. *Methods in Ecology and Evolution*, 13(9), 1899–1911. <https://doi.org/10.1111/2041-210X.13912>
- Rouse, J. W., Haas, D. W., Deering, J. A., Schell, J. A., & Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. NASA. <https://ntrs.nasa.gov/citations/19740022555>
- Sales, F. R., & De Araújo, R. C. P. (2023). Mudanças Temporais de Uso e Cobertura da Terra na Foz do Rio Cocó, Fortaleza-Ceará. *Ciência & Trópico*, 47(2), 239–258. [https://doi.org/10.33148/cetropv47n2\(2023\)art12](https://doi.org/10.33148/cetropv47n2(2023)art12)
- Seto, K. C., Güneralp, B., & Hutyrá, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109(40), 16083–16088. <https://doi.org/10.1073/pnas.1211658109>
- Sharma, R. C., Nguyen, H. T., Gharechelou, S., Bai, X., Nguyen, L. V., & Tateishi, R. (2019). Spectral Features for the Detection of Land Cover Changes. *Journal of Geoscience and Environment Protection*, 7, 81–93. <https://doi.org/10.4236/gep.2019.75009>
- Vieira, C. F. M., Guerra, F. C., Targino, A., & Gonçalves, R. D. (2025). Application of an Orbital Remote Sensing Vegetation Index for Urban Tree Cover Mapping to Support the Tree Census. *Earth*, 6(3), 87. <https://doi.org/10.3390/earth6030087>
- Yang, X., Zhao, S., Qin, X., Zhao, N., & Liang, L. (2017). Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening. *Remote Sensing*, 9(6), 1–19. <https://doi.org/10.3390/rs9060596>
- Zhu, Z., Woodcock, C. E., Holden, C., & Yang, Z. (2015). Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sensing of Environment*, 162, 67–83. <https://doi.org/10.1016/j.rse.2015.02.009>

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 2025 <http://www.risti.xyz>

