



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2026,
Volumen 10, Número 4.

https://doi.org/10.37811/cl_rcm.v10i4

RECONOCIMIENTO DE TEXTO EN IMÁGENES USANDO APRENDIZAJE PROFUNDO PARA PERSONAS CON DISCAPACIDAD VISUAL

**TEXT RECOGNITION IN IMAGES USING DEEP LEARNING
FOR PEOPLE WITH VISUAL IMPAIRMENTS**

Mirko Martin Manrique Ronceros
Universidad Nacional del Santa

Juan Pablo Sánchez Chávez
Universidad Nacional del Santa

Carlos Alfredo Gil Narvaez
Universidad Nacional del Santa

Guillermo Edward Gil Albarrán
Universidad Nacional del Santa

Camilo Ernesto Suárez Rebaza
Universidad Nacional del Santa

Reconocimiento de texto en imágenes usando aprendizaje profundo para personas con discapacidad visual

Mirko Martin Manrique Ronceros¹
mmanrique@uns.edu.pe
<https://orcid.org/0000-0002-0364-4237>
Universidad Nacional del Santa
Perú

Juan Pablo Sánchez Chávez
jsanchez@uns.edu.pe
<https://orcid.org/0000-0002-3521-7037>
Universidad Nacional del Santa
Perú

Carlos Alfredo Gil Narvaez
cagil@uns.edu.pe
<https://orcid.org/0000-0003-0137-9545>
Universidad Nacional del Santa
Perú

Guillermo Edward Gil Albarrán
ggil@uns.edu.pe
<https://orcid.org/0000-0003-3782-6765>
Universidad Nacional del Santa
Perú

Camilo Ernesto Suárez Rebaza
csuarez@uns.edu.pe
<https://orcid.org/0000-0002-6870-4296>
Universidad Nacional del Santa
País Perú

RESUMEN

El acceso a la información escrita constituye una dificultad para las personas con discapacidad visual, debido a que los textos presentes en documentos, imágenes y señalizaciones pueden limitar su independencia y generar dependencia de terceros. La investigación tuvo como objetivo mejorar la accesibilidad y autonomía de las personas con discapacidad visual mediante el desarrollo de un sistema de reconocimiento de texto en imágenes basado en aprendizaje profundo. Se realizó una investigación aplicada, de nivel explicativo y diseño cuasiexperimental con pretest y posttest, en una muestra de 30 participantes seleccionados mediante muestreo por conveniencia. Para el desarrollo de la solución tecnológica se empleó la metodología CRISP-DM. Los resultados mostraron que la precisión del reconocimiento aumentó de 41.82 % a 80.19 %, el tiempo de procesamiento disminuyó de 15.39 a 2.98 segundos, el número promedio de tareas realizadas sin asistencia externa aumentó de 5.93 a 14.93 y la satisfacción del usuario se incrementó de 2.02 a 4.57 puntos. Se concluyó que el sistema basado en aprendizaje profundo mejoró significativamente la accesibilidad y autonomía de las personas con discapacidad visual, favoreciendo el reconocimiento de información escrita, la eficiencia del procesamiento y la realización independiente de tareas.

Palabras clave: Aprendizaje profundo, reconocimiento de texto, discapacidad visual, accesibilidad, autonomía.

¹ Autor principal.
Correspondencia: mmanrique@uns.edu.pe

Text Recognition in images using deep learning for people with Visual Impairments

ABSTRACT

Access to written information represents a significant challenge for people with visual impairments, as text contained in documents, images, and signs may limit their independence and increase their reliance on external assistance. This study aimed to improve accessibility and autonomy for people with visual impairments through the development of a deep learning-based text recognition system for images. An applied research approach with an explanatory level and a quasi-experimental pretest-posttest design was employed. The sample consisted of 30 participants selected through convenience sampling. The CRISP-DM methodology was used to develop the technological solution. The results showed that text recognition accuracy increased from 41.82% to 80.19%, processing time decreased from 15.39 to 2.98 seconds, the average number of tasks completed without external assistance increased from 5.93 to 14.93, and user satisfaction increased from 2.02 to 4.57 points. It was concluded that the deep learning-based system significantly improved accessibility and autonomy for people with visual impairments by enhancing text recognition, reducing processing time, and promoting independent task completion.

Keywords: deep learning, text recognition, visual impairment, accessibility, autonomy.

*Artículo recibido 20 julio 2026
Aceptado para publicación: 20 agosto 2026*



INTRODUCCIÓN

El acceso a la información escrita constituye un componente esencial para la participación educativa, laboral y social; sin embargo, para las personas con discapacidad visual, los textos contenidos en documentos, imágenes, etiquetas, carteles y señalizaciones continúan representando una barrera para el desenvolvimiento independiente. En el contexto peruano, esta situación se agrava cuando el acceso a tecnologías adaptadas es limitado y las personas deben recurrir al apoyo de terceros para interpretar información visual o escrita. El Instituto Nacional de Estadística e Informática (INEI, 2022) ha documentado las condiciones de vida de la población con discapacidad en el país, mientras que la Comisión Económica para América Latina y el Caribe (CEPAL, 2022) ha señalado que persisten desafíos de inclusión y acceso a oportunidades para esta población.

Frente a estas limitaciones, el desarrollo de tecnologías de asistencia constituye una alternativa para mejorar la accesibilidad y favorecer una mayor autonomía. Yataco Marín (2022) sostiene que la tiflotecnología facilita el acceso de las personas con discapacidad visual a distintos formatos de información, mientras que Sebastiani Ronquillo et al. (2023) destacan la importancia de las competencias y herramientas digitales para favorecer procesos de inclusión. Del mismo modo, Kumar et al. (2024) plantean soluciones auditivas orientadas a ampliar la accesibilidad de las personas con discapacidad visual, lo que evidencia que el valor de la tecnología no depende únicamente de su capacidad computacional, sino también de su utilidad para facilitar la interacción del usuario con el entorno.

En este escenario, la inteligencia artificial y, particularmente, el aprendizaje profundo ofrece posibilidades relevantes para procesar imágenes y reconocer automáticamente patrones visuales. Los modelos de aprendizaje profundo emplean arquitecturas neuronales capaces de extraer características de los datos y resolver tareas complejas de visión por computadora. Kantipudi et al. (2021), mediante una combinación de redes neuronales profundas y redes BiLSTM, alcanzaron una precisión de 96.69 % en reconocimiento de texto en escenas, evidenciando el potencial de estas técnicas para procesar información textual contenida en imágenes. De manera complementaria, Dasari y Mehta (2022) abordaron la detección y reconocimiento de texto mediante una arquitectura neuronal de fusión, mientras

que Feilong (2021) desarrolló un sistema de reconocimiento de texto en escenas naturales basado en aprendizaje profundo.

La aplicación del aprendizaje profundo adquiere especial relevancia cuando se orienta a tecnologías de asistencia. Ganesan et al. (2022) desarrollaron un lector basado en aprendizaje profundo dirigido a personas con discapacidad visual, mientras que Hussan et al. (2022) abordaron el reconocimiento de objetos en tiempo real para esta misma población. Asimismo, Alsultan y Mohammad (2023) plantearon un sistema de asistencia basado en YOLO-V7, y Ben Atitallah et al. (2023) desarrollaron un sistema eficiente de reconocimiento de texto en imágenes complejas orientado a personas con discapacidad visual. Estos trabajos muestran que la combinación de visión por computadora, reconocimiento automatizado y mecanismos de accesibilidad constituye una línea de desarrollo con aplicaciones directas en la vida cotidiana.

Los antecedentes específicamente relacionados con reconocimiento y lectura de texto muestran resultados favorables. Langde et al. (2024) desarrollaron un sistema de narración automática de documentos basado en aprendizaje profundo que integró OCR, redes neuronales convolucionales, ResNet50, LSTM y conversión de texto a voz, alcanzando 94 % de precisión en reconocimiento textual. Ouali et al. (2023) aplicaron una red neuronal convolucional al reconocimiento de texto manuscrito árabe y obtuvieron una precisión de 95 %, mientras que Rafael Rosales et al. (2025) implementaron un dispositivo de asistencia para lectura de textos impresos mediante visión computacional, obteniendo una precisión de reconocimiento de 97.79 %. En conjunto, estos estudios evidencian la capacidad de los sistemas inteligentes para transformar contenido visual en información accesible.

No obstante, el problema no se limita exclusivamente a alcanzar elevados niveles de precisión. Una tecnología dirigida a personas con discapacidad visual debe responder también a criterios relacionados con rapidez, accesibilidad, facilidad de interacción y autonomía. Nasir et al. (2021) desarrollaron una aplicación móvil basada en aprendizaje profundo capaz de reconocer elementos del entorno y proporcionar información auditiva, concluyendo que la solución favorecía la independencia de sus usuarios.

Asimismo, Hussan et al. (2022) señalan que las soluciones de reconocimiento destinadas a esta población deben orientarse a incrementar su independencia, mientras que los planteamientos sobre



accesibilidad tecnológica resaltan la necesidad de transformar la información visual en formatos que puedan ser utilizados de manera inclusiva.

A pesar de estos avances, persiste la necesidad de estudiar soluciones que articulen el desempeño técnico con los efectos concretos producidos en los usuarios. Muchos trabajos previos centran su evaluación en métricas como exactitud o precisión, pero resulta igualmente relevante analizar cuánto tiempo requiere el procesamiento, cómo valoran los usuarios la solución y en qué medida esta les permite realizar tareas sin asistencia externa. En tal sentido, la presente investigación aborda el reconocimiento de texto en imágenes desde una perspectiva integral, considerando simultáneamente precisión, eficiencia, satisfacción y autonomía, de modo que la evaluación trascienda el funcionamiento computacional y se relacione directamente con la utilidad de la tecnología para las personas con discapacidad visual. Esta orientación coincide con la relevancia científica y social planteada en el informe de investigación.

En consecuencia, el problema de investigación se centró en determinar cómo un sistema de reconocimiento de texto en imágenes basado en aprendizaje profundo podía mejorar la accesibilidad y autonomía de las personas con discapacidad visual. Esta relación consideró como variable independiente el sistema de reconocimiento de texto basado en aprendizaje profundo y como variable dependiente la accesibilidad y autonomía de los usuarios. Se planteó como hipótesis que el uso de dicho sistema mejora significativamente ambas condiciones.

Por tanto, el objetivo general de la investigación fue mejorar la accesibilidad y la autonomía de las personas con discapacidad visual mediante un sistema de reconocimiento de texto en imágenes basado en aprendizaje profundo. Para alcanzar este propósito se establecieron cuatro objetivos específicos: evaluar la precisión del reconocimiento de texto, analizar el tiempo de procesamiento desde la captura de la imagen hasta la generación del resultado, determinar el nivel de satisfacción de los usuarios y medir la autonomía mediante el número de tareas realizadas sin asistencia externa. De esta manera, el estudio buscó aportar evidencia sobre la aplicación práctica del aprendizaje profundo en tecnologías de asistencia, considerando conjuntamente el desempeño del sistema y los beneficios alcanzados por sus usuarios.

METODOLOGÍA

La investigación se desarrolló bajo un enfoque cuantitativo, debido a que se analizaron mediciones numéricas obtenidas antes y después de la utilización de la aplicación web de reconocimiento de texto en imágenes. El estudio fue de tipo aplicado y nivel explicativo, puesto que se orientó al desarrollo de una solución tecnológica destinada a atender una problemática concreta y a determinar el efecto de su implementación sobre la accesibilidad y autonomía de las personas con discapacidad visual. En este sentido, la investigación aplicada busca trasladar el conocimiento científico hacia soluciones prácticas, mientras que el nivel explicativo permite examinar relaciones de causa y efecto derivadas de una intervención.

Se empleó un diseño cuasiexperimental de un solo grupo con pretest y posttest, debido a que se evaluaron los mismos participantes antes y después de implementar la solución, sin realizar asignación aleatoria. Este tipo de diseño resulta pertinente cuando una intervención se aplica en condiciones reales y no es factible conformar grupos mediante aleatorización, permitiendo valorar los cambios asociados a la intervención (Gopalan et al., 2020; Nianogo et al., 2023). El informe final estableció precisamente que el sistema fue evaluado mediante un diseño cuasiexperimental sin asignación aleatoria de los participantes.

La población estuvo conformada por personas con discapacidad visual, incluyendo usuarios con diferentes grados de pérdida de visión, desde baja visión hasta ceguera total. La muestra estuvo constituida por 30 participantes, seleccionados mediante muestreo no probabilístico por conveniencia, de acuerdo con su disponibilidad y accesibilidad para participar en el estudio. Como criterios de inclusión se consideraron personas que aceptaron participar voluntariamente, que podían desarrollar las actividades de manera presencial o remota y que disponían de acceso básico a un teléfono inteligente o computadora. Se excluyeron aquellos participantes que no completaron las mediciones pretest y posttest o cuyos registros resultaron incompletos para alguno de los indicadores evaluados.

La intervención consistió en la utilización de una aplicación web de reconocimiento de texto en imágenes basada en aprendizaje profundo, integrada con un modelo multimodal preentrenado para procesar las imágenes y proporcionar el contenido textual reconocido. La evaluación contempló cuatro indicadores: precisión del reconocimiento de texto, tiempo de procesamiento, número de tareas

realizadas sin asistencia externa y nivel de satisfacción del usuario. El pretest permitió establecer la condición inicial sin la aplicación, mientras que el posttest se realizó después de que los participantes utilizaron la solución tecnológica. Esta estructura permitió determinar los cambios producidos sobre los mismos sujetos y reducir la variabilidad asociada a diferencias individuales.

Para la recolección de información se emplearon la observación directa, pruebas funcionales, encuesta y análisis de registros del sistema. La observación directa permitió registrar el número de actividades realizadas sin ayuda de terceros; las pruebas funcionales permitieron medir la precisión del reconocimiento y el tiempo de procesamiento; la encuesta se utilizó para evaluar la percepción y satisfacción de los usuarios; y los registros del sistema proporcionaron información generada automáticamente durante las pruebas. Como instrumentos se utilizaron fichas de observación y registro para los indicadores técnicos y de autonomía, además de un cuestionario estructurado de satisfacción basado en escala Likert de cinco alternativas.

Para el desarrollo tecnológico se utilizó la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining). Chapman et al. (2000) organizan este proceso en seis fases: comprensión del problema, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. En el presente estudio, estas fases permitieron identificar las necesidades de los usuarios, preparar las imágenes con contenido textual, integrar el modelo de aprendizaje profundo, evaluar el reconocimiento obtenido y desplegar la aplicación web. Debido a que la solución utilizó un modelo multimodal preentrenado, las imágenes fueron empleadas para las pruebas y evaluación funcional y no para entrenar desde cero una arquitectura neuronal propia.

El procedimiento experimental comprendió tres momentos principales. Primero, se realizó el pretest, registrándose el desempeño de los 30 participantes sin utilizar la aplicación. Posteriormente, se efectuó la interacción con la solución web y se desarrollaron las actividades establecidas para evaluar su funcionamiento. Finalmente, se aplicó el posttest, registrándose nuevamente los cuatro indicadores bajo las mismas condiciones generales de evaluación. La información obtenida fue revisada y organizada antes de efectuar el análisis estadístico, verificándose previamente su integridad y consistencia.

El procesamiento estadístico se realizó mediante SPSS. Inicialmente se calcularon estadísticos descriptivos, principalmente media y desviación estándar. Debido a que las mediciones pretest y posttest

correspondieron a los mismos participantes, se evaluó la normalidad de las diferencias mediante la prueba de Shapiro-Wilk, considerando normalidad cuando $p > 0.05$. Cumplido este supuesto, se empleó la *t* de Student para muestras relacionadas con un nivel de significancia de $\alpha=0.05$. Asimismo, se calculó la *d* de Cohen para estimar la magnitud práctica del efecto producido por la intervención. Se aplicó este procedimiento estadístico a los indicadores evaluados, justificando la utilización de la prueba paramétrica después de comprobar la normalidad de las diferencias.

RESULTADOS Y DISCUSIÓN

Los resultados obtenidos evidenciaron cambios favorables en los cuatro indicadores considerados para evaluar el efecto de la aplicación web basada en aprendizaje profundo. Para evitar la reiteración de información y facilitar una lectura comparativa, los principales estadísticos descriptivos e inferenciales se presentan de manera consolidada en la Tabla 1.

Tabla 1:
Resultados estadísticos de los indicadores evaluados

Indicador		Pretest Media ± DE	Posttest Media ± DE	Shapiro- Wilk de las diferencias W (p)	t de Student, gl = 29	p	Cohen d	Mejora promedio (%)
Precisión del reconocimiento de texto (%)		41.82 ± 2.79	80.19 ± 5.54	0.968 (0.498)	-65.81	< 0.001	12.02	91.81
Tiempo de procesamiento (s)		15.39 ± 2.02	2.98 ± 0.49	0.975 (0.689)	41.37	< 0.001	7.55	80.66
Tareas realizadas sin asistencia externa		5.93 ± 1.70	14.93 ± 2.24	0.947 (0.143)	27.68	< 0.001	5.05	168.12
Nivel de satisfacción del usuario ¹		2.02 ± 0.36	4.57 ± 0.33	0.974 (0.657)	35.84	< 0.001	6.54	135.18

Nota. DE = desviación estándar; gl = grados de libertad. La normalidad fue evaluada sobre las diferencias pretest-posttest mediante Shapiro-Wilk. Para el tiempo de procesamiento, una reducción representa una mejora. ¹La satisfacción se expresó mediante la puntuación media por ítem en una escala Likert de 1 a 5

Los cuatro indicadores presentaron valores de significancia superiores a 0.05 en la prueba de Shapiro-Wilk, lo que permitió asumir normalidad en las diferencias y justificar el uso de la prueba t de Student para muestras relacionadas. En todos los casos se obtuvieron diferencias estadísticamente significativas ($p < .001$), acompañadas por tamaños del efecto elevados, lo que indicó que los cambios observados no solo fueron estadísticamente significativos, sino también relevantes en términos prácticos.

Respecto al primer objetivo específico, la precisión del reconocimiento de texto mostró un incremento significativo entre el pretest y el posttest. La media pasó de 41.82 % a 80.19 %, evidenciándose una mejora considerable después de utilizar la aplicación. Estos resultados confirmaron que la solución incrementó la capacidad para identificar correctamente el contenido textual presente en las imágenes.

Este hallazgo fue consistente con Kantipudi et al. (2021), quienes desarrollaron un sistema de reconocimiento de texto basado en redes neuronales profundas y BiLSTM, alcanzando una precisión de 96.69 %. Asimismo, Langde et al. (2024) reportaron una precisión de 94 % mediante una arquitectura que integró OCR, redes convolucionales, ResNet50 y LSTM, mientras que Ouali et al. (2023) alcanzaron 95 % en reconocimiento de texto manuscrito árabe mediante CNN. Aunque los porcentajes obtenidos en dichos estudios fueron superiores, coincidieron en demostrar la efectividad del aprendizaje profundo para procesar información textual contenida en imágenes. Las diferencias pueden asociarse a las características de los datos, condiciones de captura, tipo de texto y arquitectura tecnológica utilizada. Los resultados también guardaron relación con Ben Atitallah et al. (2023), quienes desarrollaron un sistema de reconocimiento de texto para imágenes complejas orientado a personas con discapacidad visual, y con Dasari y Mehta (2022), quienes propusieron una arquitectura neuronal de fusión para tareas de detección y reconocimiento textual. Estos antecedentes respaldaron la capacidad de las técnicas basadas en aprendizaje profundo para mejorar el reconocimiento en escenarios visuales diversos.

En relación con el segundo objetivo específico, el tiempo de procesamiento disminuyó considerablemente después de la intervención. El promedio se redujo de 15.39 segundos a 2.98 segundos, equivalente a una mejora de 80.66 %. La diferencia fue estadísticamente significativa y presentó un tamaño del efecto elevado. Este resultado mostró que la aplicación permitió entregar el texto reconocido en un periodo sustancialmente menor, aspecto especialmente relevante en tecnologías de asistencia que requieren una respuesta rápida durante actividades cotidianas.



El hallazgo fue concordante con Hussan et al. (2022), quienes desarrollaron un sistema de detección y reconocimiento en tiempo real para personas con discapacidad visual, destacando la importancia de una respuesta eficiente durante la interacción con el entorno. De igual manera, Ganesan et al. (2022) plantearon un lector basado en aprendizaje profundo orientado a facilitar el acceso inmediato a información visual, mientras que Feilong (2021) evidenció la aplicabilidad del aprendizaje profundo en reconocimiento de texto de escenas naturales. En conjunto, estos estudios sugieren que el desempeño de una solución de reconocimiento no depende únicamente de su exactitud, sino también del tiempo requerido para generar una respuesta útil.

Respecto al tercer objetivo específico, referido a la satisfacción del usuario, se obtuvo un incremento estadísticamente significativo después de utilizar la aplicación. La puntuación media por ítem pasó de 2.02 a 4.60 puntos en una escala de 1 a 5. La prueba de Shapiro-Wilk confirmó la normalidad de las diferencias ($W=.974$, $p=.657$) y la t de Student mostró una diferencia significativa, $t(29)=35.84$, $p<.001$, con un tamaño del efecto de $d=6.54$.

Este resultado evidenció una valoración favorable de aspectos relacionados con facilidad de uso, eficiencia percibida, utilidad y satisfacción general. El hallazgo coincidió con Yataco Marín (2022), quien destacó que las tecnologías de asistencia contribuyen al acceso a diferentes formatos de información y pueden favorecer una mayor autonomía informativa. Asimismo, Kumar et al. (2024) resaltaron la importancia de desarrollar interfaces y mecanismos auditivos que permitan una interacción accesible con la información, mientras que Sebastiani Ronquillo et al. (2023) subrayaron la relevancia de las herramientas digitales para favorecer procesos de inclusión de personas con discapacidad visual. En cuanto al cuarto objetivo específico, relacionado con la autonomía, el número promedio de tareas realizadas sin asistencia externa aumentó de 5.93 a 14.93. La mejora alcanzó 168.12 %, mientras que la t de Student evidenció una diferencia estadísticamente significativa, $t(29)=27.68$, $p<.001$, con un tamaño del efecto de $d=5.05$. La prueba de Shapiro-Wilk confirmó la normalidad de las diferencias ($W=.947$, $p=.143$). Estos resultados mostraron que, después de utilizar la aplicación, los participantes pudieron realizar una mayor cantidad de actividades relacionadas con el acceso a información textual sin depender de terceros.

El hallazgo fue coherente con Nasir et al. (2021), quienes desarrollaron una aplicación móvil basada en aprendizaje profundo orientada a reconocer elementos cotidianos y proporcionar información auditiva, concluyendo que la tecnología podía favorecer una mayor independencia de las personas con discapacidad visual. Asimismo, Alsultan y Mohammad (2023) propusieron un sistema de asistencia basado en YOLO-V7, mientras que Yousif y Al-Jammas (2024) desarrollaron una solución de comprensión visual destinada a mejorar la asistencia a personas con discapacidad visual. Estos estudios coincidieron en que el valor de la inteligencia artificial aplicada a tecnologías de apoyo no se limita a la automatización del reconocimiento, sino que se extiende a su capacidad para fortalecer la autonomía del usuario.

En conjunto, los resultados mostraron que la aplicación produjo mejoras simultáneas en dimensiones técnicas y funcionales. La precisión del reconocimiento y el tiempo de procesamiento reflejaron una mejora del desempeño computacional, mientras que la satisfacción y el aumento de tareas realizadas sin asistencia evidenciaron beneficios relacionados con la experiencia del usuario y la autonomía. Esta correspondencia entre desempeño técnico y utilidad práctica constituyó uno de los principales aportes de la investigación, ya que permitió evaluar la solución desde una perspectiva más amplia que la utilizada en estudios centrados exclusivamente en métricas algorítmicas.

Desde el punto de vista científico, la novedad del trabajo radicó en integrar un sistema web basado en aprendizaje profundo con una evaluación simultánea de precisión, eficiencia, satisfacción y autonomía. Ello permitió evidenciar que una solución tecnológica destinada a personas con discapacidad visual debe valorarse no solo por su capacidad de reconocer texto, sino también por la rapidez de respuesta, la percepción del usuario y su contribución a reducir la dependencia de asistencia externa. Este enfoque resulta pertinente con la línea de investigación en tecnologías de información, inteligencia artificial y accesibilidad, al vincular métricas computacionales con efectos concretos sobre los usuarios. No obstante, los resultados deben interpretarse considerando las limitaciones metodológicas del estudio, particularmente el tamaño de la muestra y el empleo de muestreo por conveniencia. Asimismo, el desempeño del reconocimiento pudo verse condicionado por características propias de las imágenes, como iluminación, nitidez, orientación y tipo de texto. Por ello, futuras investigaciones deberían

considerar muestras más amplias y heterogéneas, así como escenarios de uso más diversos, con la finalidad de determinar la generalización, robustez y estabilidad de los resultados obtenidos.

CONCLUSIONES

Los hallazgos permitieron sostener que la aplicación web basada en aprendizaje profundo constituyó una alternativa tecnológica pertinente para favorecer el acceso autónomo a información textual por parte de personas con discapacidad visual. La evidencia obtenida mostró que el beneficio de la solución no se limitó al reconocimiento automatizado del texto, sino que también se manifestó en una interacción más eficiente, una valoración favorable de la herramienta y una mayor capacidad para desarrollar actividades sin apoyo de terceros. En ese sentido, la investigación confirmó que la utilidad de una tecnología de asistencia debe valorarse de manera integral, considerando tanto su desempeño técnico como el efecto funcional que produce en sus usuarios.

Desde una perspectiva aplicada, se concluyó que la incorporación del aprendizaje profundo dentro de una solución accesible puede contribuir a disminuir barreras vinculadas con la lectura e interpretación de información escrita en imágenes. La significancia estadística obtenida en los indicadores evaluados y la elevada magnitud de los tamaños del efecto respaldaron la existencia de cambios relevantes después de la intervención, particularmente en precisión, eficiencia, satisfacción y autonomía. En consecuencia, la solución desarrollada presentó potencial para su utilización como herramienta de apoyo en actividades cotidianas que requieran identificar y comprender contenido textual.

REFERENCIAS BIBLIOGRÁFICAS

- Alsultan, O. K. T., & Mohammad, M. T. (2023). A deep learning-based assistive system for the visually impaired using YOLO-V7. *Revue d'Intelligence Artificielle*, 37(4), 901–906. <https://doi.org/10.18280/ria.370409>
- Ben Atitallah, A., Ben Atitallah, M. A., Said, Y., Albekairi, M., Boudabous, A., Alanazi, T. M., Kaaniche, K., & Atri, M. (2023). An efficient text recognition system from complex color image for helping the visually impaired persons. *Computer Systems Science and Engineering*, 46(1), 701–717. <https://doi.org/10.32604/csse.2023.035871>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS.



- Comisión Económica para América Latina y el Caribe. (2022). Personas con discapacidad e inclusión laboral en América Latina y el Caribe: Principales desafíos y respuestas de política pública. CEPAL. <https://repositorio.cepal.org/handle/11362/48735>
- Dasari, S. K., & Mehta, S. (2022). Text detection and recognition using fusion neural network architecture. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE. <https://doi.org/10.1109/ICACCS54159.2022.9785137>
- Feilong, L. (2021). Design and implementation of natural scene text recognition system based on deep learning. In 2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture. ACM. <https://doi.org/10.1145/3495018.3495323>
- Ganesan, J., Azar, A. T., Alsenan, S., Kamal, N. A., Qureshi, B., & Hassanien, A. E. (2022). Deep learning reader for visually impaired. Electronics, 11(20), 3335. <https://doi.org/10.3390/electronics11203335>
- Gopalan, M., Rosinger, K., & Ahn, J. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. Review of Research in Education, 44, 218–243. <https://doi.org/10.3102/0091732X20903302>
- Hussan, M. I. T., Saidulu, D., Anitha, P. T., Manikandan, A., & Naresh, P. (2022). Object detection and recognition in real time using deep learning for visually impaired people. International Journal of Electrical and Electronics Research, 10(2), 80–86. <https://doi.org/10.37391/IJEER.100205>
- Instituto Nacional de Estadística e Informática. (2022). Perú: Caracterización de las condiciones de vida de la población con discapacidad, 2020. INEI. <https://www.gob.pe/institucion/inei/informes-publicaciones/6025488-peru-caracterizacion-de-las-condiciones-de-vida-de-la-poblacion-con-discapacidad-2020>
- Kantipudi, M. P., Kumar, S., & Jha, A. (2021). Scene text recognition based on bidirectional LSTM and deep neural network. Computational Intelligence and Neuroscience, 2021, Article 2676780. <https://doi.org/10.1155/2021/2676780>
- Kumar, L., Kushwaha, A. S., & Jain, A. (2024). An auditory system interface for augmented accessibility: Empowering the visually impaired. Engineering Proceedings, 62(1), 8.



<https://doi.org/10.3390/engproc2024062008>

- Langde, P., Patil, S., & Langde, P. (2024). Automating document narration: A deep learning-based speech captioning system for visually impaired person. *International Journal of Intelligent Systems and Applications in Engineering*, 12(23s), 2794–2810.
- Nasir, H. M., Brahni, N. M. A., Aminuddin, M. M. M., Mispan, M. S., & Zulkifli, M. F. B. (2021). Android-based application for visually impaired using deep learning approach. *IAES International Journal of Artificial Intelligence*, 10(4), 879–888. <https://doi.org/10.11591/ijai.v10.i4.pp879-888>
- Nianogo, R., Benmarhnia, T., & O'Neill, S. (2023). A comparison of quasi-experimental methods with data before and after an intervention: An introduction for epidemiologists and a simulation study. *International Journal of Epidemiology*. <https://doi.org/10.1093/ije/dyad032>
- Ouali, I., Ben Halima, M., & Wali, A. (2023). An augmented reality for an Arabic text reading and visualization assistant for the visually impaired. *Multimedia Tools and Applications*, 82, 43569–43597. <https://doi.org/10.1007/s11042-023-14880-6>
- Rafael Rosales, C. K., Salazar Minchan, J. E., & Ortiz Porras, J. E. (2025). Diseño e implementación de un dispositivo electrónico para asistencia de lectura de textos impresos mediante visión computacional para personas con discapacidad visual. *Ingeniería Industrial*, 48, 241–261. <https://doi.org/10.26439/ing.ind2025.n48.7607>
- Sebastiani Ronquillo, M. R., Bazualdo Fiorini, E. R., & Sánchez Soto, J. M. (2023). Competencias digitales en escolares peruanos con discapacidad visual: Análisis y perspectivas. *Horizontes. Revista de Investigación en Ciencias de la Educación*, 7(28), 942–957. <https://doi.org/10.33996/revistahorizontes.v7i28.565>
- Yataco Marín, R. M. (2022). Tiflotecnología y el acceso a la información de las personas con discapacidad visual. *Fénix*, 50, 76–90. <https://doi.org/10.51433/fenix-bnp.2022.n50.p76-90>
- Yousif, A. J., & Al-Jammas, M. H. (2024). A lightweight visual understanding system for enhanced assistance to the visually impaired using an embedded platform. *Diyala Journal of Engineering Sciences*. <https://doi.org/10.24237/djes.2024.1731071>

