

Strengthening communication channels for people with hearing and speech disabilities in basic education settings through motion capture using sign languages

INGENIERÍA AMBIENTAL

Fortalecimiento de canales de comunicación para personas con discapacidad fono-auditiva en ambientes de educación básica mediante captura de movimientos usando lenguajes de señas

Marco J. Suárez^{1§}, Juan S González¹, Julián D. González¹ , Sergio A. Rojas¹ 

¹*Universidad Pedagógica y Tecnológica de Colombia, Escuela de Ingeniería de Sistemas y Computación, Sogamoso, Colombia*

§ marco.suarez@uptc.edu.co, juansebastian.gonzalez@uptc.edu.co

julian.gonzalez0@uptc.edu.co, sergio.rojas04@uptc.edu.co

Suarez MJ, González JS, González JD, Rojas S.A. Fortalecimiento de canales de comunicación para personas con discapacidad fono-auditiva en ambientes de educación básica mediante captura de movimientos usando lenguajes de señas. Ingeniería y Competitividad, 2023, 25(1); e-21212066 <https://doi.org/10.25100/iytc.v25i1.12066>

Recibido: 1 de abril de 2022 – **Aceptado:** 19 de julio de 2022

Resumen

Se presenta la implementación de un modelo deductivo dirigido al reconocimiento de vocabulario basado en los gestos de Lenguaje de Señas Colombiano (LSC) que enfoca una solución a la falta de conocimiento y acompañamiento en su aprendizaje en personas que están continuamente relacionadas con esta población que presentan discapacidad fono-auditiva. Estos gestos con las manos son utilizados debido a su gran nivel de expresividad y siendo la principal fuente de comunicación para personas con este tipo de discapacidad.

Dentro del reconocimiento de patrones/gestos de movimiento en LSC, es necesario percibir y reconocer la ubicación, orientación, lugar de articulación y punto de contacto de las manos. Conociendo acerca de las tecnologías y las investigaciones sobre algoritmos de reconocimiento gestual, análisis de patrones y redes neuronales que ayudaron a la selección correcta del modelo deductivo implementado.

De tal manera que la implementación del modelo de reconocimiento de imágenes permitió analizar fotogramas y/o imágenes reales. Analizando la información importante y resolviendo problemas específicos. Estas características se encuentran enfocadas dentro de esta investigación, logrando acompañar a la población en ambientes educativos, mediante la detección de objetos en una imagen en tiempo real usando parte de la inteligencia artificial expuesta en el modelo deductivo.

Palabras Clave: Ambiente educativo, Inteligencia artificial, Lenguaje de señas colombiana (LSC), Modelo de aprendizaje deductivo, Reconocimiento de imágenes.

Abstract

The implementation of a deductive model aimed at vocabulary recognition based on Colombian Sign Language (CL) gestures is presented, which focuses on a solution to the lack of knowledge and accompaniment in its learning in people who are continuously related to this population with a speech and hearing disability. These hand gestures are used due to their high level of expressiveness and are the main source of communication for people with this type of disability.

Within the recognition of movement patterns and gestures in LSC, it is necessary to perceive and recognize the location, orientation, point of articulation and contact of the hands. Knowing about the technologies and research on gesture recognition algorithms, pattern analysis and neural networks that helped the correct selection of the implemented deductive model.

In such a way that the implementation of the image recognition model allowed to analyze frames and/or real images, analyzing important information and solving specific problems. These characteristics are focused within this research, managing to accompany the population in educational environments, by detecting objects in an image in real time using part of the artificial intelligence exposed in the deductive model.

Keywords: Educational environment, Artificial intelligence, Colombian Sing Language (LSC), Deductive learning model, Image recognition.

1. Introducción

La inclusión social y el lenguaje de señas han abarcado un papel importante en la mayoría de aspectos de la vida cotidiana, favoreciendo tanto a niños como adultos que presentan limitaciones auditivas y por ende se les dificulta establecer una comunicación en la sociedad y hacer parte de ella (1). Gracias a la constante evolución tecnológica se pueden capturar los gestos/patrones de movimiento del lenguaje de señas y mostrar de forma deductiva la correcta traducción de cada seña favoreciendo el aprendizaje de este medio de comunicación.

La disminución de este aprendizaje se ve afectado dentro del caso de estudio en ambientes

tempranos a la educación ya que dejan a un lado este aprendizaje por su complejidad o por el tiempo al que le tienen que dedicar para poder obtener buenos resultados. Según Natalia Martínez, directora del Instituto Nacional para Sordos (INSOR) el 90% de los niños sordos nacen en ambientes con personas oyentes y así mismo en su mayoría están dentro de los estratos 1 y 2, donde los familiares no conocen el lenguaje de señas(2).

Esto permite inferir que son pocas las personas que están preparadas para afrontar o acompañar a este tipo de población dentro de Colombia y en especial del departamento de Boyacá. Por lo anterior, en este documento se presenta la implementación de un modelo deductivo que

combinado con otras técnicas de aprendizaje de máquina junto con herramientas de captura y reconcomiendo de señas en tiempo de real, que permite aumentar de manera positiva la inclusión de personas a los ambientes educativos fortaleciendo la práctica del lenguaje de señas en personas sordas y oyentes desde edades tempranas con el fin de garantizar su derecho a la educación, dando como resultado un conjunto de datos para alimentar un modelo deductivo dentro de una estructura de reconocimiento de patrones de señas.

Junto a lo anterior, un problema común es la falta de conocimiento y acompañamiento de personas que están inmersas en ambientes educativos en donde la comunicación con el tipo de población que presentan discapacidad fono-auditiva es de vital importancia. Esto debido a que no se ven influenciadas en tener que aprender este tipo de lenguaje, lo cual restringe la participación de la población que padece limitación auditiva a pertenecer en una sociedad cambiante que gira en torno a un nivel de comunicación oral y poder acceder una educación de calidad.

De otro lado, la falta de información e iniciativa de aprendizaje por parte de docentes, el mismo núcleo familiar y personas allegadas que no manejan lenguaje de señas colombiano presentan dificultades ya que al ser un lenguaje propio posee un léxico y estructura en oración únicos, no obstante la mayoría de la población sorda nace en familias donde sus integrantes son oyentes, y dentro del mismo núcleo familiar no dominan la lengua de señas, logrando generar entornos aislados en donde los niños no consiguen comunicarse con la misma familia y es un factor negativo ya que el espacio natural en donde los estudiantes o niños desarrollan y apropian su lengua en un ambiente educativo como las escuelas y se pierde un correcto nivel de comunicación.

El fortalecimiento de la inclusión de personas al aprendizaje de lenguaje de señas por medio de aprendizaje de máquina junto con herramientas de captura de movimiento han estado en constante desarrollo en Latinoamérica en los últimos años, ejemplos son los casos de Chile (3), Ecuador (4) en donde la ayuda para este tiempo de población está presente. No obstante, dentro de Colombia el desarrollo de ayudas aumenta el fortalecimiento de su inclusión en la educación básica (5–7) pero sin tener en cuenta la captura de movimiento en tiempo real y sin la clasificación correspondiente por cada imagen en tiempo real de forma que el modelo está dentro un ambiente centrado a nivel de desarrollo.

En la Universidad Pedagógica y Tecnológica de Colombia – Sogamoso (5), se implementó un sistema de reconocimiento del alfabeto de la lengua de señas colombiana de manera estática. Esta solución empleada tiene la limitación estandarizada en la toma de información; el sistema desarrollado está diseñado para funcionar bajo restricciones de fondo de color negro para simplificar la detección de las señas, y el uso de iluminación uniforme para eliminar la aparición de sombras al momento de tomar la imagen.

Esto da un gran punto de diferencia, logrando enmarcar y mejorar un medio de ayuda dentro de una población y ambiente social en específico. Es bueno aclarar que al hacer correcciones a tiempo de los problemas del lenguaje de señas permite al individuo mejorar no sólo la capacidad para expresar todo su mundo interior, sino ayuda a mejorar la capacidad de auto modulación de conductas, así como la organización del pensamiento (8).

2. Metodología

Asegurando la construcción del modelo deductivo, se estableció una metodología que proporcionó un acercamiento detallado

dentro de 5 etapas (Figura 1), cada una de las cuales cumplió una característica primordial para el desarrollo del modelo deductivo: Recolección de información, procesamiento de información, estudio de modelo de aprendizaje, evaluación y entrenamiento, y, diseño y codificación.



Figura 1 Etapas de la metodología de investigación.
Fuente: Elaboración propia

2.1 Recolección y captura de información

Se realizó la toma y estudio de las imágenes correspondientes a cada patrón y gesto del lenguaje de señas colombiano, para así formar un sistema homogéneo de información el cual formó un conjunto de datos, en este conjunto de datos se encuentran estipuladas 64 palabras básicas de comunicación, letras del alfabeto y números del 1 al 10 (Figura 2). Cada una de estas imágenes adquiridas están dentro del ambiente de estudio educativo básico que conformaron el conjunto de entrenamiento y evaluación del modelo deductivo implementado.

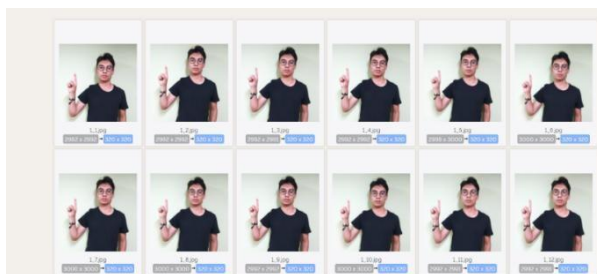


Figura 2 Parte del conjunto de imágenes para entrenamiento y evaluación del modelo. Fuente: Elaboración propia

2.2 Procesamiento de información

Seguido de la recolección de la información se realiza un estudio y procesamiento detallado de la información e imágenes generadas dentro del conjunto de datos ya establecido, validando y clasificando cada una de las imágenes adquiridas. La Figura 3 muestra como las imágenes capturadas pasan a un procesamiento y análisis para la eficacia del modelo de entrenamiento, extractando automáticamente de cada una los datos como: La ubicación, orientación, lugar de articulación y punto de contacto de las manos, por medio de una herramienta de clasificación de imágenes llamada LabelImg (9,10).

Al tener listas las imágenes se le indica a la herramienta de LabelImg en que coordenadas de imagen se encuentra el objeto o seña que se quiere detectar. Es necesario que cada una de estas imágenes se basen en la ruta directa del lugar de almacenamiento local, para poder exportar el archivo XML con las características previas a la clasificación, Este proceso se presenta en la Figura 4, donde las anotaciones son almacenadas en el formato Pascal Voc (VisualObject Classes)(11).

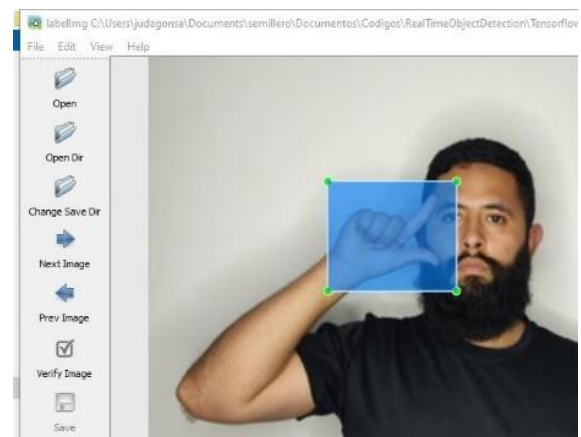


Figura 3 Clasificación de imágenes y tomas de información mediante LabelImg. Fuente: Elaboración propia

Es importante señalar que a la hora de tener y/o construir el conjunto de datos, es bueno contar con un buen número de imágenes desde varios tipos de ángulos, posiciones y que no sean de gran tamaño o variante dentro de cada una de las imágenes (12,13). Por motivos de peso en los archivos puede tardar el procesamiento de imágenes a la hora de ejecutar el entrenamiento del modelo. Obteniendo de cada una de las imágenes como resultado un archivo en formato

XML, en este archivo se encuentra, posición, tamaño, ruta de la etiqueta y nombre exacto de la sección a la cual el modelo debe basarse para el conjunto de entrenamiento y prueba (Figura 5). Cada parte de estas características están descritas dentro de etiquetas XML. De esta forma el modelo toma a la hora del procesamiento cada uno de estos archivos y los valida con cada clase suministrada.



Figura 4 Proceso de clasificación de imágenes por etiqueta dentro de LabelImg. Fuente: Elaboración propia



Figura 5 Composición del archivo XML al etiquetar la imagen. Fuente: Elaboración propia

2.3 Modelo de aprendizaje

Se realizó el estudio e implementación de un modelo de aprendizaje que realice la detección de objetos y esté acorde a las necesidades y requerimientos del modelo deductivo de la sección 2.5, donde ya generado un conjunto de imágenes tratadas para el proyecto logre de forma correcta generar un conocimiento y clasificación de objetos en tiempo real para deducir y clasificar cada imagen capturada mediante la cámara del dispositivo de captura de movimiento, de la mano de una red neuronal convolucional (14) y modelos de reconocimiento de imágenes (15).

Dentro de este paso se utiliza el método de validación cruzada para evaluar los resultados cuando el conjunto de datos se ha segmentado en dos partes. La primera muestra es utilizada para el entrenamiento y la segunda o la parte restante es utilizada para pruebas (14,16), con el fin de lograr identificar qué distribución del conjunto de datos es la más adecuada para obtener los mejores resultados en la ejecución del modelo.

Una vez normalizado el conjunto de datos en las partes que se desean entrenar y realizando el entrenamiento, se hizo una simple comparación del resultado expuesto por el modelo, el cual permite validar la óptima segmentación de datos dentro 70% para entrenamiento y 30% para pruebas como se presenta en la Figura 6.

Dentro de la búsqueda del modelo de aprendizaje e implementación de la detección de objeto, se basan en Tensorflow, la cual presenta una serie de modelos a disposición, cada uno con características sobresalientes en cuanto a tiempo de procesamiento como de efectividad al reconocimiento. Cada característica se encuentra dentro de la biblioteca TensorFlow Model Garden (17), usando el modelo de detección llamado Detection Zoo.

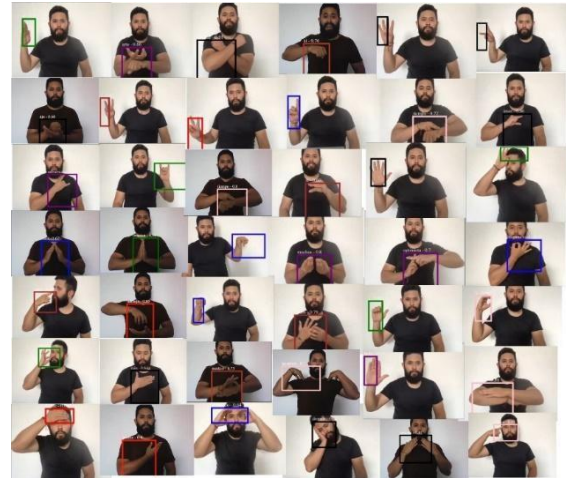


Figura 6 Conjunto de datos de entrenamiento y test para la detección de objetos. Fuente: Elaboración propia s

El modelo pertenece a un Red Neuronal Convolucional (CNN) (18,19), para el reconocimiento de objetos, ya que procesa sus capas imitando a la capacidad del ojo humano en lograr identificar y/o distinguir características que definen la detección o clasificación de un objeto. En base a los datos suministrados dentro del conjunto de datos, la red neuronal extrae automáticamente características que luego se utilizaran para la clasificación de objetos.

Este tipo de detección se maneja internamente dentro de una jerarquía, lo que esto infiere es que las primeras capas pueden detectar características como: curvas, líneas hasta llegar a la capa más profunda que reconoce formas complejas como un gesto/seña (20).

En las capas convolucionales las neuronas no están totalmente interconectadas, ya que estas lo hacen con una pequeña región de la capa anterior. Las neuronas de esta capa comparten los mismos pesos o vías, lo que reduce en una buena parte el número de parámetros a la red. Lo cual ayuda a construir un sistema que permita “ver” características propias de cada imagen/objeto y de forma correcta obtener la respuesta de dicha detección y reconocimiento.

Dando como salida una región rectangular enmarcada dentro de una imagen establecida y conocida como Región-CNN (R-CNN) (21), la cual es una técnica que consta de varias etapas establecidas para realizar la detección de objetos.

De forma general los pasos de cómo funciona la detección de dicha imagen, es mediante el uso de una técnica dentro del modelo seleccionado llamada Búsqueda selectiva (22,23), el cual pasa a escanear la imagen de entrada y de esa forma completar la consulta de posibles objetos dentro

de la imagen y posiblemente general alrededor de 2000 propuestas por región (23) Luego se procesa esta región deformada en la red neuronal convolucional (CNN), donde al final se obtendrá un vector de 3 características de la región deformada como se presenta en la Figura 7.

2.4 Evaluación y entrenamiento

En esta fase o proceso se establecerá la medición y control de efectividad de la respuesta del modelo deductivo escogido, pasando parte del conjunto de datos para la evaluación y

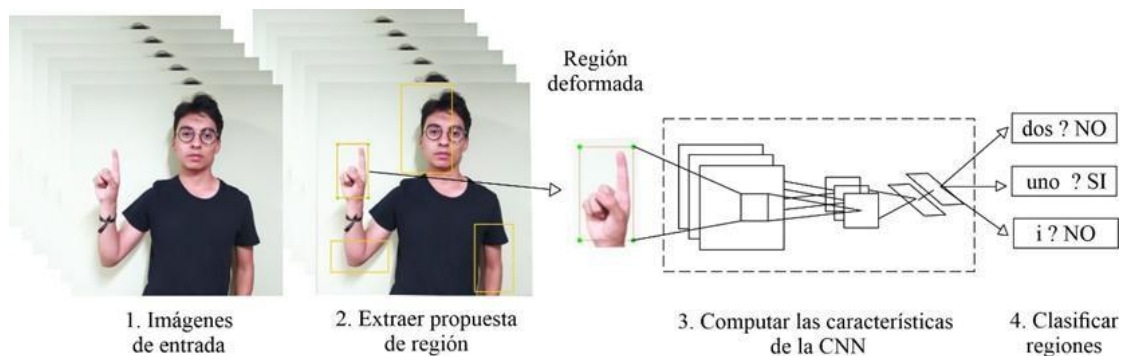


Figura 7 Proceso de detección y reconocimiento de imagen en CNN. Fuente: Elaboración propia

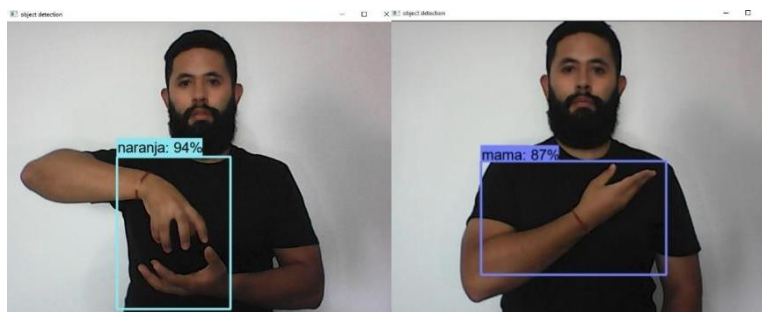


Figura 8 Respuesta del modelo con región delimitada por imagen. Fuente: Elaboración propia

entrenamiento, y de esta manera poder llevarlo a la correcta implementación dentro del proyecto, evaluando cada nivel de acierto dentro del ambiente de estudio, asegurando que cada deducción dentro de la ejecución sea la correcta y la esperada. Dentro de esta fase Tensorflow ofrece una herramienta llamada Tensorboard para poder llevar de una manera más controlada y fácil el tema de evaluación, estos resultados se

encuentran expresados en la sección de resultados y discusiones.

2.5 Diseño y codificación del modelo

Por último, se planea el diseño metodológico de desarrollo, generación de esquemas de acciones, paquetes y demás correspondiente a un desarrollo de software, en donde ya establecido el modelo deductivo (ver Figura 8), se integra dentro de la codificación para obtener como

resultado esperado un sistema en donde se pueda observar la captura de movimiento, nivel de asertividad de la seña lingüística del modelo expuesto.

Es por esto que se implementa y desarrolla el despliegue dentro de IBM Cloud (24), donde presenta un servicio de Kubernetes (25,26), de almacenamiento de datos conocido como IBM Cloud Object Storage, la cual ofrece una característica gratuita con limitación de almacenamiento persistente de alta disponibilidad. La información que se almacena con IBM Cloud Object Storage está cifrada y

distribuida en varias ubicaciones geográficas. Como resultado del despliegue del modelo dentro de IBM Cloud da el link de accesos al modelo de forma pública.

La ruta nombrada anteriormente, asegura que el acceso y uso al modelo queda ajeno a cualquier plataforma tecnológica, con lo que es pertinente recalcar que puede hacer uso del modelo tanto plataformas web como móviles o tal así dentro de la compilación de TensorFlow. Para la prueba de acceso al modelo se desarrolló una plataforma web mediante el uso del Framework de Facebook conocido como React JS al cual

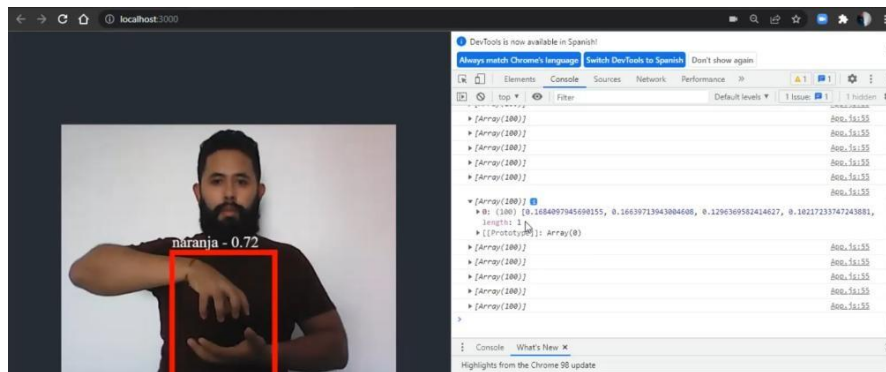


Figura 9 Uso del modelo de reconocimiento de señas. Fuente: Elaboración propia

accede a la cámara del dispositivo y hace el llamado del modelo en tiempo real para lograr la detección el cual se detalla en la Figura 9, código expuesto en el siguiente enlace <https://github.com/Sar123/ReactComputerVision>.

3. Resultados

El entrenamiento y fase de resultados del modelo marco un número de iteraciones dentro de 5000 pasos por clasificación de imágenes de esta manera fue necesario crear un conjunto de datos de buena calidad y acorde con el ambiente del modelo en desarrollo. Haciendo uso de un conjunto de datos creado para este modelo expuesto en el enlace https://drive.google.com/drive/folders/1FWbtbzgV3_PTbFYD4t0C7vmNFwvZK781?usp=sharing

Dentro de esta ruta se encuentran los archivos e imágenes tomadas y utilizadas para el entrenamiento y test del modelo junto con sus respectivas anotaciones y cuadros delimitadores de acuerdo a su clase mediante la herramienta LabelImg.

El análisis por pasos en calidad de tiempo (Figura 10), demuestra que en cada iteración el modelo se aproxima al punto objetivo disminuyendo el ruido que generan las imágenes al momento de entrenar. A mayor cantidad de pasos se genera menor cantidad de ruido, de esta forma se lleva un control sobre el entrenamiento del modelo respecto al tiempo con la cantidad de iteraciones que se tomaron para el modelo, dando en totalidad como base 5.000 pasos, ya que al aumentar el número de iteraciones excede

limitación de hardware y se aumenta el nivel de tiempo de respuesta.



Figura 10 Número de pasos de entrenamiento por tiempo en segundos. Fuente: Elaboración propia

En la Figura 11 se determina qué durante cada iteración del entrenamiento, el modelo gana un nivel de aprendizaje disminuyendo la cantidad de pérdida de información, en donde el valor final de iteraciones es 0,14, ganando un nivel de asertividad para cada clasificación, puesto que a mayor nivel de pasos de entrenamiento menor ruido.

De esta manera se disminuye el nivel de pérdida en la localización de cajas o región delimitada dentro de la imagen de la Figura 12, en donde al tener menos fugas de información que causan ineficiencia a la hora de la localización, se gana una estabilidad periódica sin sobrepasar valores de pérdida mayores al 0,15, asegurando que el modelo es capaz de deducir y detectar en tiempo real cada señal o clase.



Figura 11 Pérdida de información por clasificación por cantidad de pasos. Fuente: Elaboración propia

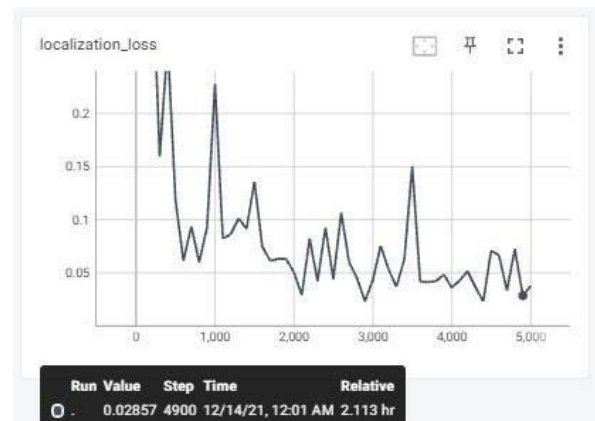


Figura 12 Pérdida de localización por cantidad de pasos. Fuente: Elaboración propia

La función de pérdida consta de estas dos partes, que se visualizan en Figura 11 y Figura 12, donde ambas partes se calculan como la suma de los errores al cuadrado. Se utilizan los parámetros de escala y predicción para controlar cuánto incrementar la pérdida de coordenadas del cuadro delimitador y cuánto se quiere disminuir las predicciones de puntuación de pérdida de confianza para cajas sin objetos, es importante reducir la pérdida aportada por los cuadros de fondo, ya que la mayoría de los cuadros delimitadores no implican ninguna instancia. En cuanto menor sea la pérdida, mejor será el modelo, es la suma de errores cometidos para cada ejemplo en los conjuntos de entrenamiento o validación.



Figura 13 Tasa de aprendizaje por cantidad de iteración.
Fuente Elaboración propia:

Dentro de la tasa de aprendizaje de la Figura 13, se percibe que en las actualizaciones de ponderaciones del modelo que hace referencia a la predicción del desplazamiento del cuadro delimitador, se obtiene una tasa de aprendizaje baja, indicando que la probabilidad de pasar del punto objetivo sea menor, pero tardará en llegar a este ya que requiere una mayor cantidad de repeticiones dentro del modelo

4. Conclusiones

En este proyecto se logra determinar que el uso de una red neuronal convolucional, por medio de búsqueda de selección, permite mediante algoritmos de reconocimiento gestual, análisis de patrones y redes neuronales, una eficiente elaboración del modelo deductivo. Mismo que toma como base el modelo SSD MobileNet V23 FPNlite 320x320 de clasificación de imágenes proporcionado por Tensorflow, que ha demostrado brindar una eficiente tasa de detección y un muy bajo error de ubicación de los objetos quién brindó la respectiva supervisión mediante procesos de inferencia para detectar patrones en el LSC entrenando y evaluando 64 palabras, letras y números de manera eficiente en cuanto a su precisión, exhaustividad y pérdida de información.

Se aplica entonces una primera etapa en el desarrollo de este prototipo. Plasmando estas características en los resultados anteriores y garantizando que el modelo deductivo junto con el conjunto de datos se encuentran en óptimas condiciones consolidando la base para ejecutarse en la segunda etapa de pruebas en ambientes reales como: Hogares y escuelas y al encontrarse almacenado en IBM Cloud lo vuelve totalmente accesible al público facilitando su análisis e interactuando con el usuario final ya sea implementado en un página web o dispositivo móvil para capturar movimientos en tiempo real visualizando la información ya analizada, sin utilizar guantes o dispositivos especializados.

Estos objetivos cumplidos, son solo la base de muchos otros para lograr el cometido final, se requiere aplicar diferentes métricas para evitar el sobreajuste al modelo. También explorar, descubrir y reinventar a medida que surjan nuevas herramientas tecnológicas mitigando un ambiente controlado para poder llevar esta herramienta a otros ambientes, optimizando procesos técnicos de reconocimiento y tiempos de entrenamientos.

De esta manera se estará logrando que este modelo sea parte de una herramienta que logre incentivar la sensibilización de los usuarios para poder involucrarse en este ambiente, ya que permitirá a medida que aumente su conjunto de datos unificando este lenguaje, promoviendo la enseñanza de lenguaje de señas desde ambientes pequeños a grandes logrando la inclusión de personas sordomudas a todos los ambientes educativos.

5. Declaración de financiamiento y/o agradecimientos

Los autores agradecen la colaboración de la Facultad seccional Sogamoso y la escuela de Ingeniería de sistemas y computación de Universidad Pedagógica y Tecnológica de Colombia-UPTC

6. Referencias

1. Muñoz I, Ruiz M, Darbet C, Lago E, Fernández E. Comunidades sordas: ¿pacientes o ciudadanas? Deaf communities: patients or citizens? ScienceDirect [Internet]. 2011;25(1):72–8. Available from: <https://doi.org/10.1016/j.gaceta.2010.09.020>
2. Espectador E. Inclusión en la educación: estudiantes sordos en Colombia. In: Espectador E, editor. Maleducados. Bogotá, Colombia; 2021.
3. González C, Yimes F. Sistema de reconocimiento gestual de lengua de señas Chilena mediante cámara digital [Internet]. PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO; 2016. Available from: http://opac.pucv.cl/pucv_txt/txt-0500/UCC0990_01.pdf
4. Pichuho J, Constante P, Gordón A, Mendoza D. Interpretación de lenguaje de señas ecuatoriano empleando visión por computador. Ibérica Sist e Tecnol Inf [Internet]. 2019;(960–971):13. Available from: <https://www.proquest.com/openview/84337a90042da817ad3453747249970a/1?pq-origsite=gscholar&cbl=1006393>
5. Guerrero J. Algoritmos de procesamiento de imágenes y redes neuronales artificiales para el reconocimiento de la lengua de señas colombiana (LSC). Rev Colomb Tecnol Av [Internet]. 2016;2(28):1–8. Available from: https://www.unipamplona.edu.co/unipamplona/portallG/home_40/recursos/05_v25_30/revista_28/20052017/01.pdf
6. Muñoz Y, Moreno L. Implementación de un algoritmo para a clasificación automática de lenguaje de señas colombiano en video usando aprendizaje profundo [Internet]. Universidad Católica de Colombia; 2020. Available from: [https://repository.ucatolica.edu.co/bitstream/10983/24980/1/Proyecto de grado.pdf](https://repository.ucatolica.edu.co/bitstream/10983/24980/1/Proyecto%20de%20grado.pdf)
7. Alvarez P, Castro G. Modelo de red neuronal convolucional para el reconocimiento del alfabeto en lenguaje de señas colombiano [Internet]. Universidad del Sinú; 2019. Available from: <http://repositorio.unisinucartagena.edu.co:8080/xmlui/handle/123456789/56>
8. Barragán E, Lozano S. Identificación temprana de trastornos del lenguaje Early identification of language disorders. ScienceDirect [Internet]. 2011;22:227–32. Available from: <https://www.sciencedirect.com/science/article/pii/S0716864011704175?via%3Dihub>
9. Tzuta L. LabelImg [Internet]. Github. 2020. Available from: <https://github.com/tzutalin/labelImg#labelimg>
10. Nabende J, Tusubira J, Babirye C, Nsumba S, Omongo Christopher. A dataset of cassava whitefly count images. ScienceDirect [Internet]. 2022;41. Available from: https://www.sciencedirect.com/science/article/pii/S2352340922001238#cit_1

11. Everingham M, Eslami A, Gool L, Williams C, Winn J, Zisserman A. The PASCAL Visual Object Classes Challenge: A Retrospective. *Int J Comput Vis Vol* [Internet]. 2015;111:98–136. Available from: <https://link.springer.com/article/10.1007/s11263-014-0733-5>
12. Martinez E, Morillas F. Deep Learning Techniques for Spanish Sign Language Interpretation. *Hindawi* [Internet]. 2021;2021:10. Available from: <https://doi.org/10.1155/2021/5532580>
13. Brownlee J. How Do Convolutional Layers Work in Deep Learning Neural Networks? [Internet]. *Machine Learning Mastery*. 2019. Available from: <https://machinelearningmastery.com/deep-learning-for-computer-vision/>
14. Adeyanju I, Bello O, Adegboye M. Machine learning methods for sign language recognition: A critical review and analysis. *Intell Syst with Appl* [Internet]. 2021; Available from: <https://doi.org/10.1016/j.iswa.2021.200056>
15. Liang T, Bao H, Pan W, Pan F. Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles. *Hindawi* [Internet]. 2022;2022:16. Available from: <https://doi.org/10.1155/2022/3825532>
16. Farfán O, Camargo J. Modelo computacional para reconocimiento de lenguaje de señas en un contexto colombiano. *TecnoLógicas* [Internet]. 2020;23(48):197–232. Available from: <https://revistas.itm.edu.co/index.php/tecno logicas/article/view/1585/1637>
17. Tensorflow. TensorFlow JS [Internet]. npm. 2022. Available from: <https://www.npmjs.com/package/@tensorflow/tfjs>
18. Anirudh R. Convolutional Neural Network Tutorial (CNN) – Developing An Image Classifier In Python Using TensorFlow. *edureka* [Internet]. 2020;7. Available from: <https://www.edureka.co/blog/convolutional-neural-network/>
19. Lu, Jiaqi Liu, Ruiqing Yuejuan, Zhang Zhang, Xiaxiang Zheng, Longbo Zhang, Chao Zhang, Kaiming Li, Shuan Lu Y. Research on the development and application of a detection platform for colorectal cancer tumor sprouting pathological characteristics based on artificial intelligence. *ScienceDirect* [Internet]. 221AD;6. Available from: <https://doi.org/10.1016/j.imed.2021.08.003>
20. Cabrera J, Cervantes J, Lamont G, Ruiz J, Jalili L. Mexican sign language segmentation using color based neuronal networks to detect the individual skin color. *Expert Syst Appl* [Internet]. 2021; Available from: <https://doi.org/10.1016/j.eswa.2021.115295>
21. Wei X, Chen X, Lai C, Zhu Y, Yang H, Du Y. Automatic Liver Segmentation in CT Images with Enhanced GAN and Mask Region-Based CNN Architectures. *Hindawi* [Internet]. 2021;2021:11. Available from: <https://doi.org/10.1155/2021/9956983>
22. Guo X, Liu S. A Scatter Search Approach for Multiobjective Selective Disassembly Sequence Problem. *Hindawi* [Internet]. 2014;9. Available from: <https://doi.org/10.1155/2014/756891>

23. Programmer Click. SSD para detección de objetos en tiempo real [Internet]. 2020. Available from: <https://programmerclick.com/article/91891967042/>
24. Cloud I. IBM Cloud Object Storage [Internet]. IBM. 2022. Available from: <https://www.ibm.com/co-es/cloud/object-storage>
25. Vayghan L, Saied M, Toeroe M, Ferhat K. A Kubernetes controller for managing the availability of elastic microservice based stateful applications. *J Syst Softw* [Internet]. 2021; Available from: <https://doi.org/10.1016/j.jss.2021.110924>
26. Kubernetes. ¿Qué es Kubernetes? [Internet]. Documentación distribuida Kubernetes. 2022. Available from: <https://kubernetes.io/es/docs/concepts/overview/what-is-kubernetes/>