



UNIVERSIDAD DEL SINÚ

Elías Bechara Zainúm

Seccional Cartagena

**MODELO DE RED NEURONAL CONVOLUCIONAL PARA EL
RECONOCIMIENTO DEL ALFABETO EN LENGUAJE DE SEÑAS
COLOMBIANO**

PEDRO LUIS TORRES ÁLVAREZ

GUSTAVO CASTRO LOZANO

**UNIVERSIDAD DEL SINÚ ELÍAS BECHARÁ ZAINÚM
FACULTAD DE CIENCIAS EXACTAS E INGENIERÍAS
ESCUELA INGENIERÍA DE SISTEMAS
CARTAGENA, COLOMBIA**

2019



UNIVERSIDAD DEL SINÚ

Elías Bechara Zainúm
Seccional Cartagena

**MODELO DE RED NEURONAL CONVOLUCIONAL PARA EL
RECONOCIMIENTO DEL ALFABETO EN LENGUAJE DE SEÑAS
COLOMBIANO**

**Trabajo de grado presentado como requisito para optar el título de
INGENIERO DE SISTEMAS**

Autores

PEDRO LUIS TORRES ÁLVAREZ

GUSTAVO CASTRO LOZANO

Asesor

EUGENIA ARRIETA RODRÍGUEZ

**UNIVERSIDAD DEL SINÚ ELÍAS BECHARÁ ZAINÚM
FACULTAD DE CIENCIAS EXACTAS E INGENIERÍAS
ESCUELA INGENIERÍA DE SISTEMAS
CARTAGENA, COLOMBIA**

2019

CONTENIDO

ÍNDICE DE ILUSTRACIONES	5
INTRODUCCIÓN	6
1. DISEÑO METODOLÓGICO	8
Planteamiento del problema	8
Justificación	9
Formulación del problema.....	11
Alcance	11
Objetivos	11
Estado del arte.....	12
Marco de referencia	15
1.1.1. Marco teórico.....	15
1.1.2. Marco conceptual	17
1.1.3. Marco legal.....	18
Metodología	23
1.1.4. Línea de investigación.....	23
1.1.5. Metodología de desarrollo	24
2. ENTENDIMIENTO DEL NEGOCIO	27
3. ADQUISICIÓN DE LOS DATOS.....	30
Recolección de los datos	30
Construcción de la base de datos.....	30
4. Modelado.....	34
Diseño del modelo	34

4.1.1.	Cargue de datos	35
4.1.2.	Normalización de los datos.....	35
4.1.3.	División de los datos (Split)	35
	Diagrama del modelo.....	36
5.	ENTRENAMIENTO VALIDACIÓN Y PRUEBA	40
	Ambiente de entrenamiento.....	40
	Pruebas.....	40
	Métricas empleadas.....	42
5.1.1.	Exactitud o Accuracy	43
5.1.2.	Precisión.....	43
5.1.3.	Sensibilidad o Recall	44
5.1.4.	F1-score	44
	Matriz de confusión.....	45
6.	RESULTADO.....	47
7.	CONCLUSIONES Y TRABAJO EN EL FUTURO	50
8.	BIBLIOGRAFÍA.....	51
	ANEXO.....	53
	Cronograma.....	53

ÍNDICE DE ILUSTRACIONES

Ilustración 1 Abecedario lenguaje de señas -----	9
Ilustración 2 Redes neuronales [11]. -----	16
Ilustración 3 Categorización usando redes neuronales [12].-----	17
Ilustración 4 Proceso de ciencia de datos en equipo -----	24
Ilustración 5 Ciclo de vida del Proceso de ciencia de datos -----	25
Ilustración 6 Clasificación de los datos -----	31
Ilustración 7 Diseño del algoritmo de predicción de lenguaje de señas usando CNN. -----	34
Ilustración 8 División de los datos -----	36
Ilustración 9 Diagrama del modelo -----	39
Ilustración 10 Modelo de precisión -----	41
Ilustración 11 Modelo de pérdida -----	41
Ilustración 12 Representación genérica de la matriz de confusión. Fuente propia-----	42
Ilustración 13 Matriz de confusión -----	45
Ilustración 14 Clasificación de reporte-----	46

INTRODUCCIÓN

Hoy en día hay aplicaciones orientadas a personas discapacitadas que son limitadas y escasas, las personas con limitaciones para oír se relacionan con el mundo que los rodea a partir de información que les llega principalmente por la vista. El conocimiento del mundo, el establecimiento de conceptos y generalización, en los sordos, se hace muy necesaria la lengua de señas, elemento indispensable para la comunicación.

Según el Censo, en Colombia poco más del 6% de personas (2.624.898), tiene algún tipo de limitación. Esto quiere decir que, por cada 15 personas, hay 1 que tiene alguna discapacidad. Del total de personas con discapacidad, el 17% corresponde a personas con limitación para oír [1]. La educación se constituye en nuestra sociedad, en un factor determinante para el desarrollo integral de todas las personas, particularmente para la población sorda la educación formal se considera muy necesaria para acceder a toda información disponible en nuestro ambiente. La Lengua de Señas Colombiana o L.S.C. (ISO/DIS 639-3: csn), es la lengua de señas empleada por la comunidad sorda en Colombia. Está emparentada filogenéticamente con la lengua de señas francesa. Según Ethnologue, algunos signos son similares a aquellos de las lenguas de señas española, americana y salvadoreña.

La lengua de señas es el idioma natural con el que se comunican las personas sordas, cuenta con una gramática propia y difiere de otras lenguas. Además, es una lengua ágrafa, es decir no tiene escritura.

Por lo anteriormente mencionado, existe una necesidad muy marcada de contar con herramientas y estrategias que faciliten el proceso de enseñanza. Una de estas herramientas es, el Traductor Virtual de Lengua de Señas Colombiana creado por el ingeniero Colombiano Jorge Enrique Leal, quien es

el director de la Fundación HETAH (herramienta tecnológica para ayuda humanitaria) [2]. Esta herramienta virtual es gratuita, y consiste en un sistema unidireccional para traducir frases en español mediante inteligencia artificial, basado en un análisis gramatical que permite encontrar una secuencia de imágenes correspondientes, este va dirigido a personas oyentes, logrando incluir socialmente a la persona sorda en un solo sentido, debido a que la persona sorda no podrá expresarse en el mismo lenguaje que una persona oyente. Por lo cual se considera que este sistema no cumple con las normas gramaticales propias de la Lengua de Señas Colombiana LSC, y no es recomendado para la realización de comunicación con personas sordas.

En respuesta a este problema este proyecto plantea en su primera etapa un prototipo que sea de auto aprendizaje el cual identifique las letras del alfabeto de lengua de señas mediante de imágenes, con este pequeño paso esperamos influenciar en otros desarrolladores e implementar y darle avance a este proyecto.

1. DISEÑO METODOLÓGICO

En este capítulo se detallan los aspectos metodológicos necesarios para el desarrollo del proyecto, se describe la problemática, el alcance, los objetivos y la metodología de desarrollo de proyectos de ciencia de datos.

1.1. Planteamiento del problema

En la ciudad de Cartagena hay 6.783 personas con limitación para oír, es decir, una persona sorda por cada 132 habitantes. De este total, en cifras proyectadas, se tiene que el 8% corresponde a niños y niñas menores de 6 años; de manera acumulativa, el 19% se encuentran en el rango de 6 a 18 años; el 44% está entre 19 y 59 años; y el 29% se encuentra por encima de los 60 años, para en un futuro incluir a ese porcentaje de personas productivas a la sociedad [1].

Para ayudar a este porcentaje de la población colombiana, se ha proyectado construir un sistema de autoaprendizaje del abecedario de lengua de señas colombiano (Ver Ilustración 1). Esto Aportará a la población en la inclusión o aceptación de la persona con discapacidad como ente de desarrollo dentro de una sociedad, sustentado en los esfuerzos que están haciendo las naciones para cerrar la brecha abismal que existe hoy día en este aspecto, podemos ver este derecho sustentado por el documento emanado del Comité sobre los derechos de las personas con discapacidad realizada por las Naciones Unidas del 31 de marzo a 11 de abril de 2014, el mayor problema que se observa para que se dé la inclusión es la comunicación, ya que el discapacitado sensorial auditivo no tiene las suficientes herramientas para comunicarse dentro de una sociedad analfabeta en el lenguaje de señas, aunque existen pocos recursos,

aún no es suficiente y claro para la sociedad en general, porque hablando de apoyos, podemos apreciar que para muchos oyentes no tiene sentido el apoyo de lengua de señas en la telecomunicación, si no estamos educados o conocemos el lenguaje de señas, como hace un oyente para comentar y discernir con un discapacitado auditivo si no sabe cómo comunicarse en su lenguaje natural.

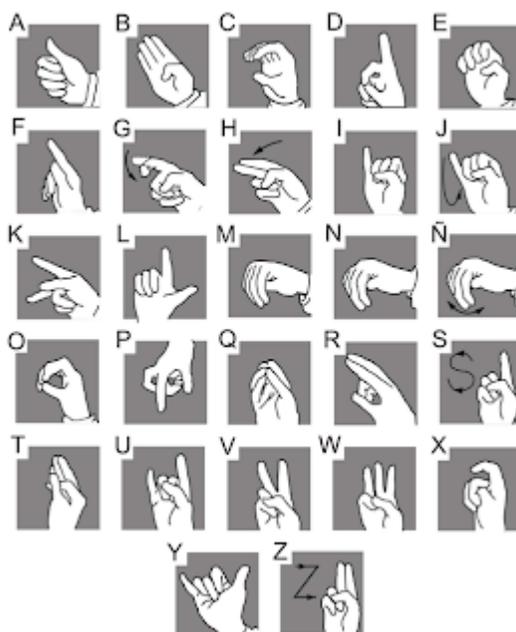


Ilustración 1 Abecedario lenguaje de señas

1.2. Justificación

Es importante cerrar la brecha de exclusión que existe en nuestra sociedad, dar inicio al proceso de reconocimiento de derechos que tienen las personas audio impedidas es sumamente importante, ya que a pesar que las leyes [1] y los gobiernos abren un pequeño espacio para la inclusión, la falta, desconocimiento o interés de aprender a comunicarnos con este grupo de

personas ha impedido que el proceso de inclusión sea exitoso, incluso dentro de los mismos medios de comunicación convencionales, se hace difícil entender que es importante el apoyo de lengua de señas para que esta población se mantenga informada y actualizada dentro de lo que sucede a su alrededor.

Cabe también hacer alusión a las pocas oportunidades que este grupo de personas audio impedidas tienen para sentirse plenamente desarrolladas como profesionales, visto que el problema de comunicación se ha convertido en el talón de Aquiles, para que logren ser empleados o puedan expresarse libremente frente a cualquier situación que se les presente, es tan grave la necesidad que tienen, porque los profesionales de salud o funcionarios de cualquier ente en los que ellos tengan que realizar algún trámite no los pueden atender por no saber lengua de señas, por lo que siempre están necesitando el acompañamiento de un familiar o de un traductor de lengua de señas para que puedan ser atendidos como debe ser.

El propósito del proyecto es aplicar un prototipo de inteligencia artificial, basado en el aprendizaje de lenguaje de señas colombiana para el reconocimiento de imágenes del alfabeto, para esto se piensa implementar en Python un lenguaje de autoaprendizaje que reconozca dichas imágenes. El impacto esperado para este proyecto es socio - económico el cual se espera que las personas sordas sean incluidas tanto en la vida escolar, como oportunidades de empleo. Para estas personas sería una nueva forma de comunicarse con el resto del mundo, para sentirse incluidas en la sociedad.

1.3. Formulación del problema

¿Como aplicar redes neuronales para el reconocimiento del alfabeto en lenguaje de señas colombiano?

1.4. Alcance

Lograr con un set de imágenes el aprendizaje de un prototipo para distinguir las letras del alfabeto de lenguaje de señas colombiano. Esto es un pequeño paso para estrechar la brecha comunicativa que existe en la sociedad, por falta de conocimiento de la lengua natural de las personas audio impedidas, muy a pesar de que es tan natural porque desde niños lo primero que aprendemos antes de empezar a articular palabras es el lenguaje de gestos. Por hecho el prototipo que esperamos alcanzar en esta primera etapa es alimentar nuestro programa para que este logre identificar las letras del abecedario en un medio controlado. El prototipo va a tener un algoritmo de autoaprendizaje, una base de dato con un set de imágenes para el entrenamiento del modelo. Su funcionalidad será reconocer a través de imágenes el alfabeto del lenguaje de señas colombiano.

1.5. Objetivos

Objetivo general

Desarrollar un modelo de predicción de lenguaje de señas usando redes neuronales convolucionales (CNN) que permita el reconocimiento del alfabeto del lenguaje de señas colombiano por imagen.

Objetivos específicos

- Construir un conjunto de datos con imágenes que representen el abecedario en lenguaje de señas a nivel nacional (Colombia) con la intención de entrenar y validar la red neuronal.
- Entrenar la red neuronal convolucional con el 80% del conjunto de datos, capaz de interpretar el alfabeto del lenguaje de señas.
- Validar el modelo de lenguaje de señas con ejemplos que no fueron usados durante los entrenamientos de la red neuronales, correspondiente al 20% del conjunto de datos.

1.6. Estado del arte

Se han realizado numerosas investigaciones sobre la incorporación del lenguaje de señas en recursos informáticos en el ámbito educativo, tanto en nuestro país como en el mundo, este apartado se centra fundamentalmente en los estudios y experiencias nacionales y latinoamericanas, debido a la importancia que se le adjudica en esta tesina al contexto social que caracteriza el ámbito educativo donde se incorporan las TI y Cs.

El proyecto “*Herramienta de Apoyo para la Interpretación de Lenguaje de Señas mexicano (HILSEM)*” desarrollado por Castro Estévez, Luis Alberto y Cruz García, Katia Karen, junio 2014. propone el desarrollo de un sistema utilizando MonoDevelop enfocado a personas sordomudas que, en el momento en que el usuario esté “ejecutando” una palabra en lenguaje de señas mexicana, sea capaz de detectar dichas señas e interpretarlas para dar como

resultado una salida de sonido, comprensible para un usuario secundario, el cual es una persona oyente [3].

Encontramos que en la ciudad México, D.F., en la INSTITUTO POLITÉCNICO NACIONAL en el centro de investigación en computación, el redactor Fausto Pável Priego Pérez, realizó un proyecto con el nombre de *“Reconocimiento de imágenes del lenguaje de señas mexicano.”* [4]. Este trabajo propone un sistema de dos módulos principales los cuales constan de aprendizaje y reconocimiento. Usa un dispositivo de cámara de video (dispositivo Kinect), para detectar el patrón o seña de cada palabra; se usan dos métodos para el reconocimiento de seña, en el primero se requiere algún tipo de guantes en específico para reconocer la seña. En el segundo se usa el dispositivo Kinect para su reconocimiento.

El proyecto *“Clasificador difuso para el reconocimiento continuo del lenguaje de señas a partir de las características de seguimiento y forma”* desarrollado por MVD Prasad, PVV Kishore, D. Anil Kumar, Cap. Raghava Prasad, agosto 2016. Propone el seguimiento y la captura de vectores de posición de la mano es la obra de arte del algoritmo de flujo óptico horn schunck. Los contornos activos extraen características de forma de marcos de signos en la secuencia de video. Las dos características más dominantes del lenguaje de señas se combinan para construir características de señas [5].

El proyecto *“Reconocimiento del lenguaje de señas usando procesamiento de imágenes”* desarrollado por Ananya Roy y Sandhya Arora, enero 2018. Propone el sistema reconoce el alfabeto del lenguaje de señas calculando la métrica de distancia de Bhattacharyya entre los histogramas de la imagen capturada, procesada y la imagen almacenada y dando ese alfabeto como

salida cuyo histograma de imagen tiene una menor distancia de Bhattacharyya con el histograma de la imagen capturada [6].

El proyecto “*Método automático de reconocimiento de voz para la clasificación de vocales al lenguaje de señas colombiano.*” desarrollado por Andres Arias y David Rubiano, 2018. Propone el sistema reconoce las vocales del abecedario colombiano por medio de audio con el fin de que reconozca la letra por audio y de vuelta la vocal por lenguaje de señas colombiano [7].

Creado en el 2012 en la Facultad Córdoba de la Universidad Tecnológica – Argentina como proyecto de tesis por los entonces estudiantes, Nacional Matías Reggi, Fabián Bertetto y Esteban Rodríguez, basados en el sistema de Kinect de la consola de juegos Xbox se produjo un software intérprete y traductor tanto del lenguaje de señas como del lenguaje natural, por lo que permite a los usuarios interactuar de diferentes maneras, el programa es muy complejo, tiene una señal para cada palabra y también para cada letra por si se quiere deletrear una palabra inexistente. (Araujo, 2019) [8].

Investigadores de Microsoft en Asia, en colaboración con el Instituto de la tecnología en computación de la academia china de la ciencia, utilizando la tecnología del Kinect se dieron a la tarea de desarrollar un software para la interpretación del lenguaje de señas chinas y traducirlo a su lenguaje, convirtiéndolo en texto o audio, también emplea un avatar en el que se apoya para las traducción a lengua de seña. (Saavedra, 2013) [9].

En Honduras los investigadores de la facultad de ingeniería de la Universidad Nacional Autónoma de Honduras han desarrollado una plataforma utilizando el principio tecnológico del avatar tridimensional que realiza la función de interprete virtual, esta plataforma seria estudiada en el 2015, en las aulas de

clase en la cual se encargaría de captar la voz del profesor he interpretada en las pantallas por medio del interprete virtual. El ingeniero Raúl Palma uno de los creadores de la plataforma, afirma tener alrededor de 300 señas para su implementación. (Andrade, 2014) [10].

1.7. Marco de referencia

1.7.1. Marco teórico

Machine learning

Es una disciplina científica del ámbito de la Inteligencia Artificial que crea sistemas que aprenden automáticamente. Aprender en este contexto quiere decir identificar patrones complejos en millones de datos.

Aprendizaje supervisado

Uno de los métodos del Machine Learning es el del aprendizaje supervisado, este algoritmo es entrenado entregándole las características que serían las preguntas, y la etiquetas denominadas respuestas, esto con el fin de que el algoritmo mezcle las dos y lograr tener predicciones.

Redes Neuronales Artificiales (RNA)

Inspirado en el comportamiento observado en el proceso biológico, es un modelo computacional que consiste en un conjunto de unidades llamado neuronas artificiales conectadas entre sí para transmitirse señales, la entrada de información atraviesa la red neuronal en este paso se aplican diversas operaciones dando resultados unos valores de salida. En las redes neuronales el objetivo es llegar a la capacidad de resolver problemas de la misma manera

que el cerebro humano, estos sistemas se entrenan y se encargan del auto aprendizaje.

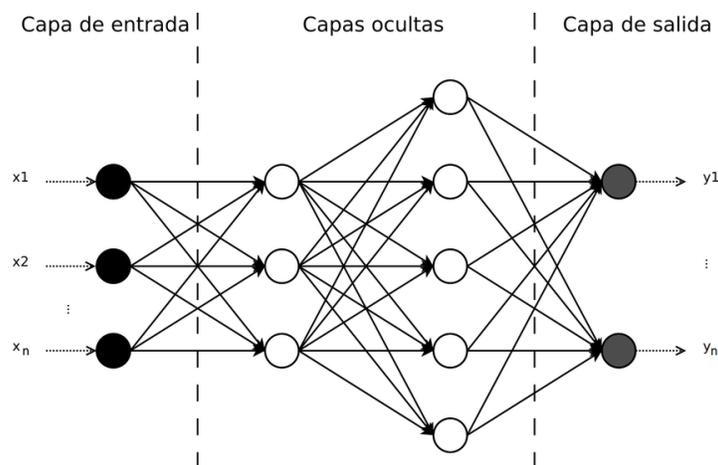


Ilustración 2 Redes neuronales [11].

Red neuronal Convolutiva (CNN)

Las redes neuronales convolucionales son semejantes a las neuronas receptoras de la corteza visual primaria de un cerebro. Estas neuronas son capaces de una visión artificial para aplicaciones de clasificación y segmentación de imágenes entre otras utilidades, estas actividades se pueden llevar a cabo gracias a las matrices bidimensionales, gracias a estas matrices es que se pueden desarrollar las distintas aplicaciones.

Las CNN son especialmente útiles para localizar patrones en imágenes con el objetivo de reconocer objetos, caras y escenas. Aprenden directamente a partir de los datos de imágenes, utilizando patrones para clasificar las imágenes y eliminar la necesidad de una extracción manual de características.

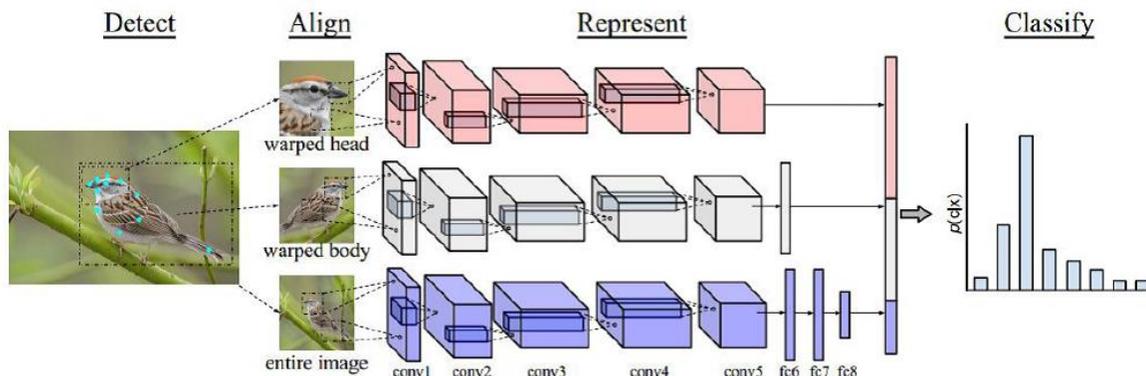


Ilustración 3 Categorización usando redes neuronales [12].

1.7.2. Marco conceptual

- **Python**

Es un lenguaje de programación versátil multiplataforma y multiparadigma, destaca por su código legible y limpio. Obtiene una gran ventaja al ser un lenguaje de código abierto que le permite ser utilizado en diferentes plataformas de alto tráfico, como lo son Google, YouTube, etc. El principal objetivo de este lenguaje es la automatización de los procesos para reducir dificultades y tiempos. Esto lo puede realizar en reducidas líneas de código que puedes implementar en distintas plataformas y sistemas operativos

- **TensorFlow**

Es una biblioteca de software libre que se utiliza para realizar cálculos numéricos mediante diagramas de flujo de datos. Los nodos de las matrices de datos multidimensionales (tensores) comunicadas entre ellas.

- **Keras**

Keras es un framework de alto nivel para el aprendizaje, escrito en Python y capaz de correr sobre los frameworks TensorFlow, CNTK, o Theano. Fue desarrollado con el objeto de facilitar un proceso de experimentación rápida. Lo que haremos en este experimento es entrenar modelos de clasificación de imágenes. Esto consiste en dada una serie de imágenes etiquetadas, reconocer una imagen y asignarle dicha etiqueta (por ejemplo, si es la foto de un gato, el modelo reconocerá que hay un gato).

- **Google colabatory**

Esta herramienta es un entorno gratuito de Jupyter Notebook el cual no requiere configuración y no es necesario contar con una maquina física con grandes capacidades, ya que esta herramienta es totalmente ejecutable en la nube. Nos permite escribir, guardar, ejecutar código y hasta compartir el análisis de datos y tienes recursos de información muy potentes todo desde tu navegador y de forma gratuita. Se recomienda usar el navegador de Google Chrome.

1.7.3. Marco legal

En el ámbito legal la emisión de la ley General de Educación 115/94, y su decreto reglamentario el 2082 de 1996, que establecen la integración de las poblaciones especiales al sistema educativo regular, legislando sobre la transformación gradual de las instituciones actuales de educación especial:

- a) Transformándose en Unidades de Atención Integral que presten servicios tecnológicos, terapéuticos y pedagógicos a las escuelas integradoras.

- b) Ofreciendo un servicio educativo formal a la población especial, en la que se cuentan las personas sordas.
- c) Convirtiéndose en una oferta de educación no formal.

En 1996, se aprueba en el congreso de la república la ley 324 por la cual se crean algunas normas a favor de la población sorda. En esta ley, se reconoce a la Lengua de Señas Colombiana (en adelante LSC) como la lengua de la comunidad sorda del país, y en su decreto reglamentario 2369/97, se incluye la necesidad de una educación bilingüe para los sordos y de los servicios de interpretación.

A pesar de que las anteriores condiciones favorecen el cambio hacia una educación verdaderamente significativa para los sordos, se puede anticipar que dicha transformación será a largo plazo, pues implica una reestructuración profunda del servicio educativo que se ofrece a la población sorda colombiana y exige el tratamiento de aspectos tales como: re significación de la escuela y de los procesos formativos que en ella se desarrollan; definición de la política lingüística al interior del espacio escolar; la formación de docentes; la formación y preparación de adultos sordos; la formación de intérpretes; los procesos de enseñanza de la LSC para los oyentes involucrados en los procesos educativos; la participación de las organizaciones de padres de sordos; el trabajo colaborativo con la comunidad sorda; la re conceptualización del significado de la integración social y educativa de los educandos sordos. (Ramírez, et al. 2003)

Se puede deducir entonces, que la historia de la educación de los sordos en Colombia ha estado más vinculada a la controversia sobre la enseñanza formal de la lengua oral que a los procesos pedagógicos que

favorecen su desempeño académico. Esto explicaría en parte, el escaso desarrollo de innovaciones pedagógicas que ha caracterizado la educación especial impartida a la población sorda y su sistemático distanciamiento de las reflexiones pedagógicas que en el plano de la educación en general se han presentado a lo largo de la historia de la educación en Colombia.

Todo este estudio y bases legales es lo que nos lleva a desarrollar la posibilidad de crear este prototipo, ya que la educación o cierre de la brecha educacional para las personas con deficiencias auditivas, también tiene el limitante del desconocimiento de la sociedad auditiva del lenguaje de señas, que se hace imprescindible al momento de integrar a estas personas ya educadas dentro de cualquier ciencia del saber, para que se desempeñen como miembros activos y productivos de la sociedad, por esto se hace necesario que también como oyentes nos eduquemos y conozcamos este lenguaje, por lo que es necesario una herramienta que pueda leer la palabra hablada y la pueda reproducir en una seña perceptible para los audio impedido, también podemos ver este derecho sustentado por el documento emanado del **Comité sobre los derechos de las personas con discapacidad** realizada por las Naciones Unidas del 31 de marzo a 11 de abril de 2014.

Comité sobre los Derechos de las Personas con Discapacidad
11º período de sesiones.

31 de marzo a 11 de abril de 2014

Observación general N.º 1 (2014)

Artículo 12: Igual reconocimiento como persona ante la ley

I. Introducción

1. La igualdad ante la ley es un principio básico general de la protección de los derechos humanos y es indispensable para el ejercicio de otros derechos humanos. La Declaración Universal de Derechos Humanos y el Pacto Internacional de Derechos Civiles y Políticos garantizan específicamente el derecho a la igualdad ante la ley. En el artículo 12 de la Convención sobre los derechos de las personas con discapacidad se describe en mayor profundidad el contenido de ese derecho civil, centrándose en las esferas en que tradicionalmente se les ha denegado a las personas con discapacidad. En el artículo 12 no se establecen derechos adicionales para las personas con discapacidad; simplemente se describen los elementos específicos que los Estados partes deben tener en cuenta para garantizar a las personas con discapacidad el derecho a la igualdad ante la ley, en igualdad de condiciones con las demás.

II. Contenido normativo del artículo 12

Artículo 12, párrafo 1

11. En el artículo 12, párrafo 1, se reafirma que las personas con discapacidad tienen derecho al reconocimiento de su personalidad jurídica. Esto garantiza que todo ser humano sea respetado como una persona titular de personalidad jurídica, lo que es un requisito indispensable para que se reconozca la capacidad jurídica de la persona.

Artículo 12, párrafo 2

A continuación, se detalla uno de los apartes del artículo 12 en el que claramente expresa “El apoyo también puede consistir en la elaboración y el reconocimiento de métodos de comunicación distintos y no convencionales, especialmente para quienes utilizan formas de comunicación no verbales.

17. El apoyo en el ejercicio de la capacidad jurídica debe respetar los derechos, la voluntad y las preferencias de las personas con discapacidad y nunca debe consistir en decidir por ellas. En el artículo 12, párrafo 3, no se especifica cómo debe ser el apoyo. "Apoyo" es un término amplio que engloba arreglos oficiales y oficiosos, de distintos tipos e intensidades. Por ejemplo, las personas con discapacidad pueden escoger a una o más personas de apoyo en las que confíen para que les ayuden a ejercer su capacidad jurídica respecto de determinados tipos de decisiones, o pueden recurrir a otras formas de apoyo, como el apoyo entre pares, la defensa de sus intereses (incluido el apoyo para la defensa de los intereses propios) o la asistencia para comunicarse. El apoyo a las personas con discapacidad en el ejercicio de su capacidad jurídica puede incluir medidas relacionadas con el diseño universal y la accesibilidad —por ejemplo, la exigencia de que las entidades privadas y públicas, como los bancos y las instituciones financieras, proporcionen información en un formato que sea comprensible u ofrezcan interpretación profesional en la lengua de señas, a fin de que las personas con discapacidad puedan realizar los actos jurídicos necesarios para abrir una cuenta bancaria, celebrar contratos o llevar a cabo otras transacciones sociales. El apoyo también puede consistir en la elaboración y el reconocimiento de métodos de

comunicaciones distintas y no convencionales, especialmente para quienes utilizan formas de comunicación no verbales para expresar su voluntad y sus preferencias. Para muchas personas con discapacidad, la posibilidad de planificar anticipadamente es una forma importante de apoyo por la que pueden expresar su voluntad y sus preferencias, que deben respetarse si llegan a encontrarse en la imposibilidad de comunicar sus deseos a los demás. Todas las personas con discapacidad tienen el derecho de planificar anticipadamente, y se les debe dar la oportunidad de hacerlo en condiciones de igualdad con las demás. Los Estados partes pueden ofrecer diversas formas de mecanismos de planificación anticipada para tener en cuenta las distintas preferencias, pero todas las opciones deben estar exentas de discriminación. Debe prestarse apoyo a la persona que así lo desee para llevar a cabo un proceso de planificación anticipada. El momento en que una directiva dada por anticipado entra en vigor (y deja de tener efecto) debe ser decidido por la persona e indicado en el texto de la directiva; no debe basarse en una evaluación de que la persona carece de capacidad mental.

1.8. Metodología

1.8.1. Línea de investigación

Este proyecto es un modelo que busca aportar a la línea de investigación de inteligencia artificial del grupo DEARTICA, para la implementación de reconocimiento de imagen mediante una red neuronal convolucional.

1.8.2. Metodología de desarrollo

La metodología empleada para este proyecto es “Team Data Science Process” que cuenta con un ciclo de vida para proyectos de ciencia de datos, es una metodología ágil e iterativa para ofrecer soluciones de análisis predictivo y aplicaciones inteligentes de manera eficiente, ayuda a mejorar la colaboración y el aprendizaje del equipo.

El ciclo de vida original contempla 5 pasos, como se muestra en la (ilustración 4). Para efectos de este proyecto se usarán tan sólo tres pasos, esto debido a que se plantea un alcance que permita el acercamiento a este tipo de técnicas y como punto de partida o línea de base para continuar con la investigación. En la (Ilustración 5) se plantea el ciclo de vida con los pasos a seguir en este trabajo para los objetivos.

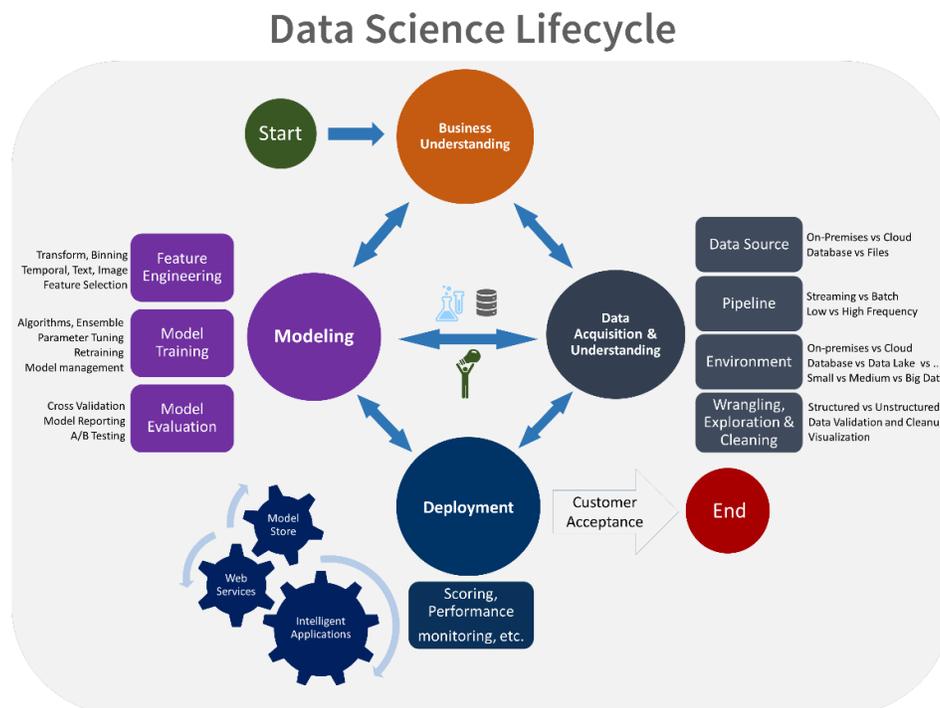


Ilustración 4 Proceso de ciencia de datos en equipo

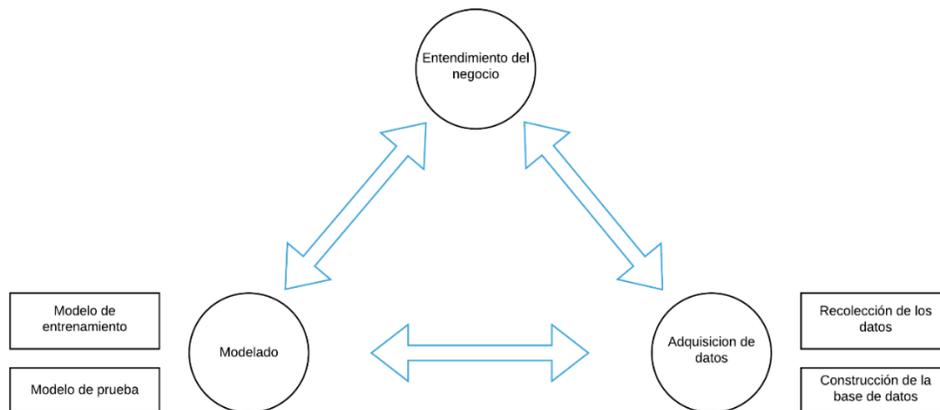


Ilustración 5 Ciclo de vida del Proceso de ciencia de datos

Entendimiento del negocio.

Para el entendimiento del negocio lo primero que se debe tener en cuentas son los objetivos del modelo y cuya implementación se determinan para establecer el éxito del proyecto. Lo segundo es buscar la fuente de datos importantes de la cual se tenga acceso o donde se va a obtener.

Definir objetivos; trabajar con las partes interesadas para establecer e identificar los problemas. Formular preguntas para encontrar objetivos a los cuales se puedan apuntar con ayuda de las técnicas de ciencia de datos. Identifique fuentes de datos; Encontrar datos importantes para ayudar a la solución de las preguntas de los objetivos del proyecto.

Adquisición de los datos.

En este proceso tiene como objetivo producir un conjunto de datos limpio y de alta calidad con el que se pueda usar en el entrenamiento del modelo. En este proyecto el conjunto de datos se dividió en carpeta por cada letra del abecedario en lenguaje de señas, al ser datos de diferentes rasgos característicos.

Modelado.

Después de la creación del conjunto de datos se cargan los datos al modelo, para obtener las características de los datos para entrenar el conjunto de datos, en el modelo se genera en base a plantillas que proporcionan detalles para la implementación de él.

2. ENTENDIMIENTO DEL NEGOCIO

Para comprender el contexto del problema se hace necesario explicar cuatro aristas:

Contexto familiar

Para iniciar el proceso de comprensión y entendimiento del contexto del problema y las necesidades de los interesados, se entrevistó a una persona que tiene a un familiar con discapacidad auditiva, quien relató cómo fue crecer con esa persona a su lado y cómo fue el crecimiento de esa persona en forma académica y personal, a continuación, se presenta la historia.

“Al crecer en una familia estable y con mucho amor, todo es fácil y divertido, pero como dice un dicho *no todo es color de rosas*. Mi hermana mayor nació con la discapacidad de ser sorda, al no poder escuchar no pudo desarrollar su capacidad de habla al 100%. Ella siempre fue muy atenta conmigo y me apoyaba en muchas cosas, incluso en las tareas. Cuando ella se graduó del bachillerato, quiso ser alguien en la vida, y estudió una carrera “Estética y Cosmetología” y llegó la hora de ayudar a mi hermana en sus estudios, yo era su “conejillo de india”, ella practicaba sus masajes conmigo, y cuando en sus clases necesitaba llevar a una persona para que fuera su modelo (adivinen quien la acompañaba), yo me ofrecía en acompañarla, se graduó ella y 2 de sus compañeras de colegio con honores y aún no termina la historia, siendo la primera promoción de Esteticistas y Cosmetólogas sordas, las apoyaron para que realizaran el curso de docencia en el Sena, trabajó como docente en la misma institución donde se graduó, llevando a la segunda promoción de Esteticistas y Cosmetólogas sordas, ahora trabaja de independiente en un SPA que tiene en casa, y en estos momentos que tengo la capacidad y el

conocimiento de devolverle los esfuerzos a esa persona que estuvo tantos años conmigo porque no hacerlo, y mis queridos lectores por ese motivo estoy llevando a cabo este proyecto de grado, más que proyecto de grado, *proyecto de vida.*”

Contexto educativo

Al ser un lenguaje que solo se enseñan en instituciones educativas especiales, el mayor porcentaje de los individuos en el planeta no saben este idioma, este abismo de conocimiento se cerraría un poco teniendo a la mano una herramienta que facilitaría la comunicación entre las personas oyentes y las personas sordos muda, haciendo a estas últimas personas incluidas.

Contexto social y afectivo

En el ámbito social el lenguaje de señas representa un gran reto dado que causa restricciones y limitaciones que impiden el desarrollo normal de algunas actividades, por ejemplo, si llegas a un centro comercial y necesitas ubicar un producto en específico, comunicarte con los asesores puede ser un gran reto dado que la mayoría de las personas no manejan este tipo de lenguaje por tanto se les hace difícil dar la información que solicitas.

La interacción con otras personas cuando se tienen limitaciones del habla usualmente crea muchos obstáculos a los individuos con esta condición, las relaciones amorosas e interpersonales se dificultan al tratar de expresarte y no ser entendido. Un individuo sordo mudo tratando de pretender de forma amorosa a otra persona que no sabe ni entiende su lenguaje, no puede expresar sus ideas y sentimientos de la misma forma que los demás, esto crea desventajas y podría.

Contexto laboral

Las empresas ven una gran limitante en contratar personas audio impedidas en la colocación de estas en un puesto de trabajo, ya que las demás personas no manejan el lenguaje de señas, lo que dificulta establecer un puesto de trabajo. Para las empresas contratar a una persona audio impedida implica capacitar a parte o a su totalidad de sus empleados en este lenguaje, lo que generan costos y tiempos. Con una herramienta que facilite la conversación entre personas oyentes y sorda mudas, ya estas personas serian útiles y se podrían aprovechar en diversas áreas laborales.

3. ADQUISICIÓN DE LOS DATOS

En este capítulo se describe el proceso de recolección, organización y finalmente construcción de la base de datos del alfabeto del lenguaje de señas, mediante registro fotográfico.

3.1. Recolección de los datos

Teniendo en cuenta que no se encontró disponible una base de datos de imágenes con el alfabeto de las señas en un contexto local, se decidió construirla de forma manual. Para esto se hizo la toma de fotos, usando cámaras de celulares, esto fue necesario para cumplir con la cantidad mínima requeridas para el entrenamiento del modelo. Los dos Smartphone que se usaron fueron el ZenFone Max (M2) y el Neffos y5s. La recolección fue hecha por los dos investigadores, y los participantes que sirvieron como modelos para tomarles las fotos de la mano indicando cada letra del lenguaje de señas. Se tomaron 4043 fotos para el entrenamiento y prueba del modelo.

3.2. Construcción de la base de datos

Seguido de esto se realizó el proceso de segmentación por seña; a este proceso también se le denomina proceso de etiquetado, donde los nombres de cada carpeta representen una letra del alfabeto colombiano en lenguaje de señas (ver ilustración 6). Se tomaron 4043 fotos de las 26 letras del abecedario en lenguaje de señas (ver tabla 1), en el conjunto de datos se encuentran dos tipos de imágenes por la diferencia en dimensiones que se ocasionó por el uso de los dos dispositivos para la toma de las imágenes. Para solucionar este problema, en el modelo se estandariza las dimensiones de los datos.

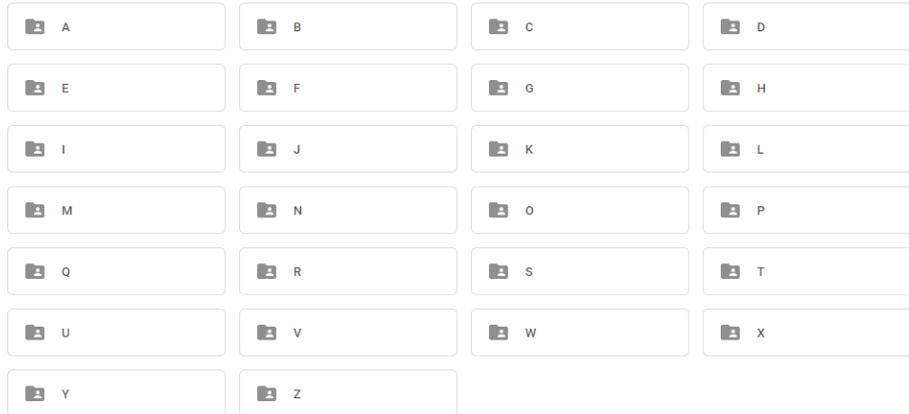


Ilustración 6 Clasificación de los datos

Una vez organizado el conjunto de datos, en la tabla 1, se muestran las letras del abecedario, la letra en lenguaje de señas y el número que la representa en el algoritmo.

Tabla 1 Representación del lenguaje de señas y alfabético

Letra LSC	Letra ABC	N° representativo	Letra LSC	Letra ABC	N° representativo
	A	0		D	3
	B	1		E	4
	C	2		F	5

Letra LSC	Letra ABC	N° Rep.
	G	6
	H	7
	I	8
	J	9
	K	10
	L	11
	M	12
	N	13
	O	14

Letra LSC	Letra ABC	N° representativo
	P	15
	Q	16
	R	17
	S	18
	T	19
	U	20
	V	21
	W	22

Letra LSC	Letra ABC	N° representativo
 A hand sign for the letter X, formed by crossing the index and middle fingers of the right hand.	X	23
 A hand sign for the letter Y, formed by crossing the index and middle fingers of the right hand, with the thumb pointing up.	Y	24
 A hand sign for the letter Z, formed by pointing the index finger of the right hand to the left.	Z	25

4. Modelado

En este capítulo se habla de la construcción del modelo, su diseño, la carga, la normalización, la división de los datos y el diagrama de los datos.

4.1. Diseño del modelo

En la ilustración 7 se presenta el diseño del algoritmo para la detección de lenguaje de señas, esta plantea el desarrollo segmentado en dos fases, la primera es la creación del conjunto de datos y la segunda corresponde a la etapa de construcción del modelo de red neuronal convolucional.

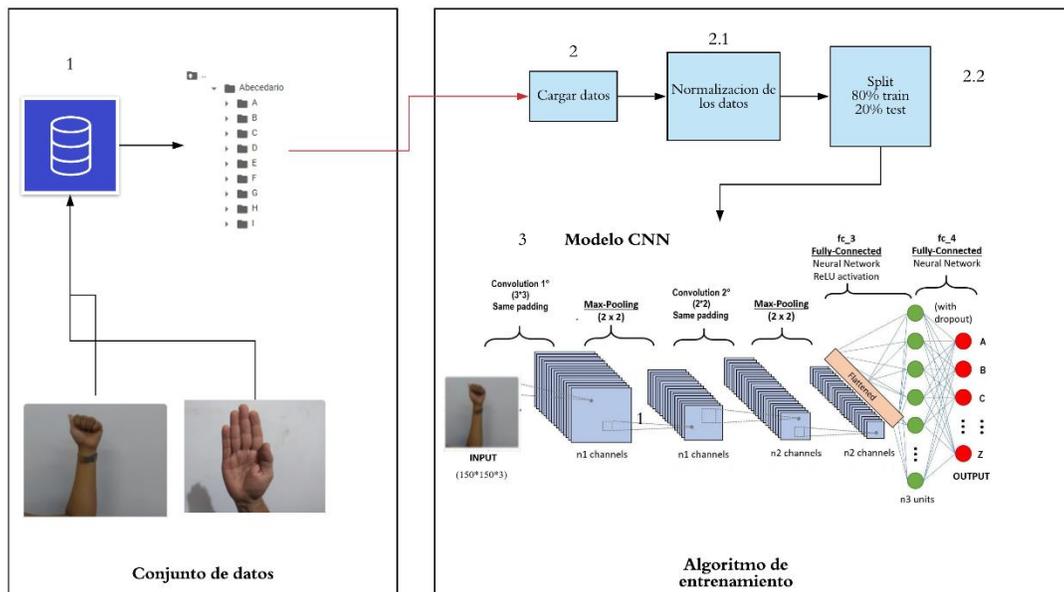


Ilustración 7 Diseño del algoritmo de predicción de lenguaje de señas usando CNN.

4.1.1. Cargue de datos

Con el conjunto de datos ya construido y guardados en la herramienta Google drive, usando unas las librerías únicas de Google Colaboratory para cargar los datos en ella. En la misma se ejecutó el código Python con las librerías de TensorFlow y Keras para la construcción y evaluación del modelo CNN (Convolutional Neural Network).

4.1.2. Normalización de los datos

Es útil normalizar los datos para hacerle más fácil a la técnica de Machine Learning el aprendizaje. Por normalizar nos referimos a poner a todos los datos en una escala similar.

Fue necesario tomar la matriz de datos y realizar una normalización aplicando una división de la matriz que contiene los datos de entrenamiento entre 255 (datos de entrenamiento/255 el valor máximo de un pixel), Se utilizó Keras para procesar las imágenes a la escala de los pixeles de datos $1/255$, esta normalización permitió que las operaciones matriciales que se realizaron en entrenamiento sean mucho más rápidas y eficientes tanto para la multiplicación que se realiza con los pesos que se inicializan y para la actualización de pesos con respecto a nivel de error que se debe disminuir.

4.1.3. División de los datos (Split)

En la ilustración 8, representa la técnica Split, que opera con una división del total de datos que se tiene cargados en memoria en una distribución de 80%

para datos de entrenamiento y un 20% para evaluación, el objetivo de tener estos datos divididos nos fue pertinente para corroborar con los datos de evaluación, si nuestros modelos generalizados con datos que desconocía.

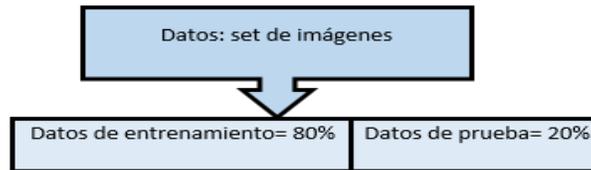


Ilustración 8 División de los datos

4.2. Diagrama del modelo

En la arquitectura del modelo (Ilustración 9), se definen las funciones que ayudaron a la construcción de la red neuronal convolucional, a continuación, se dará a explicar de forma breve cómo funcionan estos fragmentos de código.

Sequential ()

Esta función es una pila de capas lineales, que se crean pasando una lista de instancia de capas al constructor.

add(convolution2D ())

Esta capa crea un núcleo de convolución que este entrelazado con la entrada de capas para producir un tensor de salidas. Cuando se usa esta capa como la primera de un modelo, se proporciona ciertas palabras claves cuales son: filters, kernel_size, strides, padding, input_shape y activation.

- **Filters:** es la dimensión de la salida de la convolucion.
- **Kernel_size:** especifica las dimensiones de la imagen de entrada.
- **Strides:** zancadas de la convolución a lo largo y ancho.

- **Padding:** "same" proporciona como resultado el relleno de la entrada de modo que la salida tenga la misma longitud que la entrada original.
- **Input_shape:** indica que la entrada esperada serán lotes de vectores de N° de dimensiones.
- **Activation:** las activaciones se pueden usar a través de una capa de activación o puede ser admitido por todas las capas.

MaxPooling2D

Operación de agrupación máxima para datos espaciales. MaxPooling2D es un tensor 3D de forma (peso, altura, canal) las dimensiones altura y peso tienden a reducirse a medida que nos adentramos en las capas ocultas de la red neuronal.

- **Pool_size:** tupla de 2 enteros que reduce la escala (vertical, horizontal). (2, 2) reducirá a la mitad la entrada en ambas dimensiones espaciales. Si solo se especifica un número entero, se utilizará la misma longitud de ventana para ambas dimensiones.

Flatten

Convierte los elementos de la matriz de imágenes de entrada en un array plano.

Dense

Con esta instrucción añadimos una capa oculta (hidden layer) de la red neuronal.

- **Units:** dimensionalidad del espacio de salida.

Dropout

La deserción es establecer aleatoriamente una fracción rate de las unidades de entrada en 0 o 1 en cada actualización durante el tiempo de entrenamiento, lo que ayuda a evitar el sobreajuste.

Compile

Configura el modelo para el entrenamiento.

- **Optimizer:** instancia del optimizador.
- **Loss:** función de pérdida de instancia, si el modelo tiene múltiples salidas, puede usar una pérdida diferente en cada salida pasando un diccionario o una lista de pérdidas.
- **Metrics:** Lista de métricas que el modelo evaluará durante el entrenamiento y las pruebas.

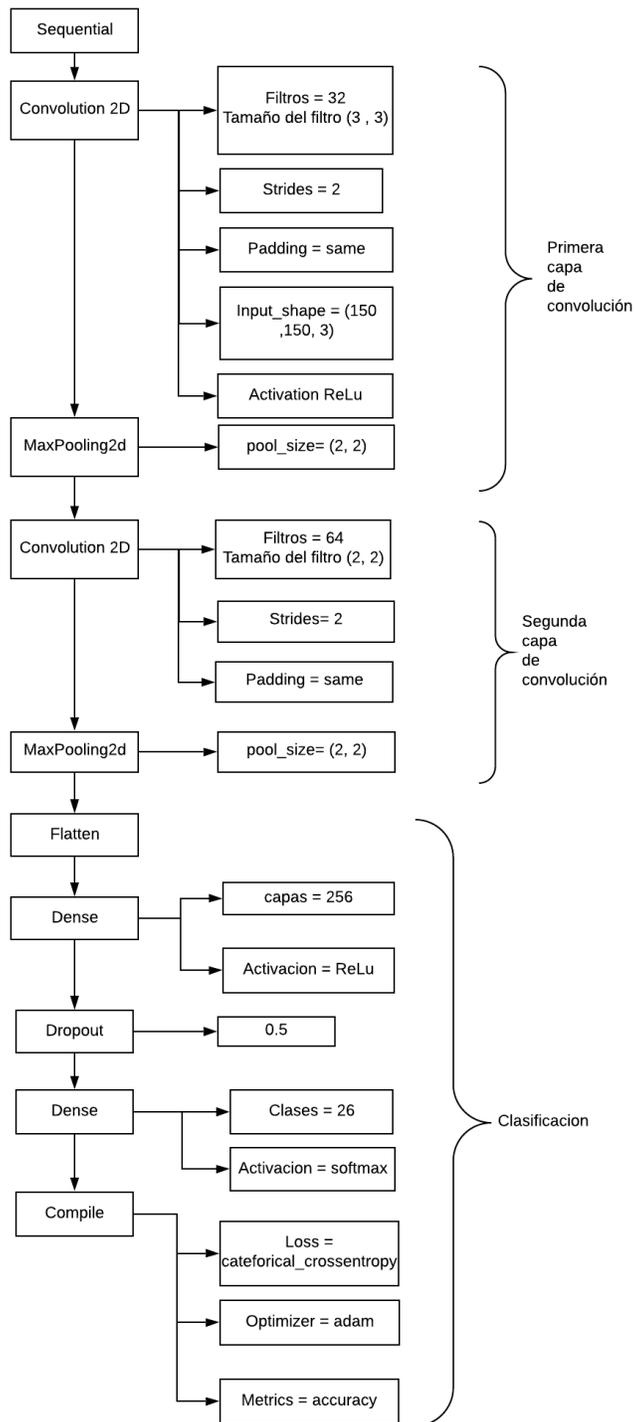


Ilustración 9 Diagrama del modelo

5. ENTRENAMIENTO VALIDACIÓN Y PRUEBA

Una vez obtenido el conjunto de datos apto y también la técnica de Machine Learning que se usara para desarrollar el clasificador de letras para el lenguaje de señas.

5.1. Ambiente de entrenamiento

El entrenamiento y pruebas fueron realizados en Google Colaboratory es un entorno de máquinas basado en Jupyter Notebooks, se corre en la nube puede correr nuestro notebook de forma gratuita, tienes la opción de guardar en la aplicación de Google Drive los notebooks creados en él, tiene como alternativa poder guardar los conjuntos de datos que son usados en Machine Learning y utilizarlos en Google Colaboratory.

5.2. Pruebas

En el desarrollo de este proyecto, se realizó el entrenamiento con un algoritmo de aprendizaje supervisado. En este entrenamiento se utilizó 80% del conjunto de datos y el 20% del conjunto se usó para la validación del modelo.

En el entrenamiento se hizo en quince epochs (época) que son las veces que se van a entrenar el modelo, con cien steps_per_epoch (lotes de muestra) es la que dice cuando una época se acaba y comienza la otra.

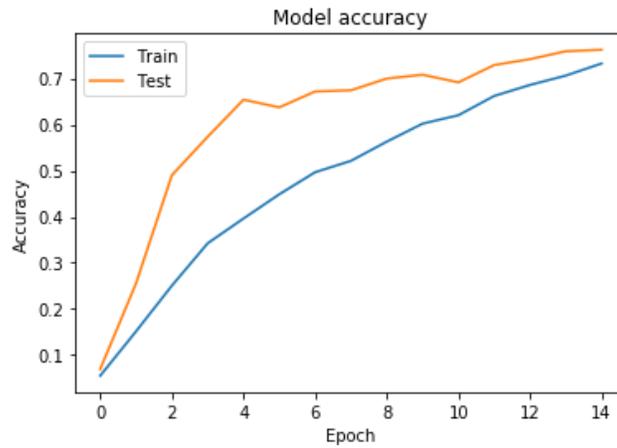


Ilustración 10 Modelo de precisión

En la ilustración anterior se destaca el paso de la época en la precisión del modelo, que se obtiene entre el entrenamiento (train) y la validación (test) del modelo.

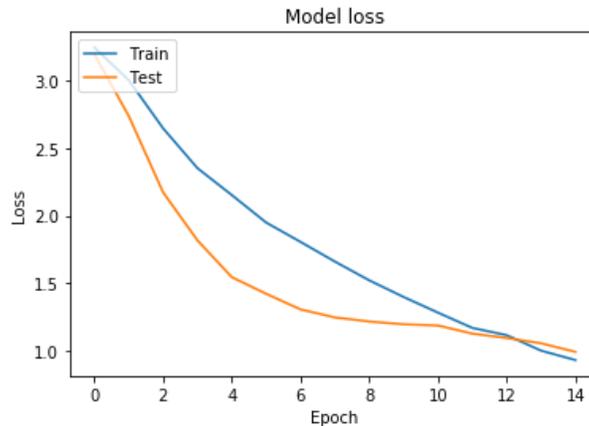


Ilustración 11 Modelo de perdida

En la ilustración anterior se destaca el paso de la época en la pérdida del modelo, que se obtiene entre el entrenamiento (train) y la validación (test) del modelo.

Observando los valores del diagrama se esperó que el porcentaje de perdida en la prueba no se fuera tan desviada del entrenamiento.

5.3. Métricas empleadas

Luego de entrenar el modelo de Machine Learning es necesario evaluar el rendimiento del clasificador para determinar la viabilidad del mismo; una de las métricas más usadas a la hora de evaluar un algoritmo es la matriz de confusión. La matriz de confusión por si sola nos da información relevante acerca del rendimiento del clasificador, pero a partir de ella se extraen otras métricas que dan información veraz del mismo.

La matriz de confusión consta de cuatro partes fundamentales como se muestra en la ilustración 12, donde en las columnas se ubican las clases ya establecidas y en las filas las clases de prueba.

	Positivos	Negativos
Positivos	TP	FP
Negativos	FN	TN

Ilustración 12 Representación genérica de la matriz de confusión. Fuente propia

- **Verdaderos positivos (TP por sus siglas en inglés):** Es el número de clasificaciones verdaderas correctas.
- **Verdaderos negativos (TN por sus siglas en inglés):** Es el número de clasificaciones falsas correctas.
- **Falsos positivos (FP por sus siglas en inglés):** Es el número de clasificaciones verdaderas incorrectas.
- **falsos negativos (FN por sus siglas en inglés):** Es el número de clasificaciones falsas incorrectas.

A partir de las cuatro partes fundamentales de la matriz de confusión es posible extraer la exactitud, la precisión, la sensibilidad, entre otras métricas importantes.

5.3.1. Exactitud o Accuracy

Evalúa el número total de clasificaciones correctas frente al total de muestras clasificadas, esto en la matriz de confusión se describe en la ecuación 1. La exactitud toma valores de cero a uno (o cero a cien si se representa como porcentaje) siendo cero una exactitud muy baja y uno una exactitud muy alta.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad [1]$$

La exactitud únicamente debe ser utilizada cuando el número de muestra tanto de entrenamiento de como prueba es similar para cada clase, de lo contrario la métrica siempre tendrá un sesgo hacia la clase que tenga mayor número de muestra.

5.3.2. Precisión

Evalúa el número de clasificaciones correctas para cada clase de prueba; es decir cuántas clasificaciones verdaderas son realmente verdaderas para cada clase, esto en la matriz de confusión se describe en la Ecuación 2. La precisión al igual que la exactitud, toma valores de cero a uno, siendo cero una precisión muy baja y uno una precisión muy alta.

$$Precision = \frac{TP}{TP + FP} \quad [2]$$

5.3.3. Sensibilidad o Recall

Indica que número de muestras que corresponden a una clase fueron clasificadas dentro de ella para cada muestra evaluada; es decir, cuántas clasificaciones que debían ser verdaderas fueron clasificadas como verdaderas, esto en la matriz de confusión se describe en la ecuación 3. La sensibilidad toma valores de cero a uno, siendo cero una sensibilidad muy baja y uno una sensibilidad muy alta.

$$Recall = \frac{TP}{TP + FN} \quad [3]$$

A primera vista la precisión y la sensibilidad parecen ser iguales pero la diferencia radica en que la precisión tiene en cuenta los falsos positivos mientras que la sensibilidad tiene en cuenta los falsos negativos.

5.3.4. F1-score

Para conocer el rendimiento general del modelo habitualmente se calcula la relación entre las medidas de precisión y sensibilidad mediante una media armónica, este indicador es conocido como F1-Score y es calculado mediante la ecuación 4.

$$F_1 \text{ Score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad [4]$$

5.4. Matriz de confusión

En el campo del aprendizaje automático la matriz de confusión es una tabla específico que permite la visualización del rendimiento de un algoritmo.

La matriz de confusión nos ayuda a la validación de los aciertos por cada variable y con cual otra se está equivocando y con qué frecuencia. Como se muestra en la ilustración 13, Las variables H, N y Z se presentan bajos aciertos ya que en el lenguaje de señas tiene similitudes con otras.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	[24	0	0	1	0	5	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0	0	1	2	0	0]	
B	[0	30	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
C	[0	0	11	3	0	2	1	0	1	0	0	0	0	0	1	0	0	0	4	0	0	1	0	0	3	0]	
D	[0	0	1	20	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0]
E	[0	1	0	0	22	0	0	0	0	0	5	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0]	
F	[0	1	0	0	0	29	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
G	[0	0	0	0	0	0	18	2	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
H	[1	0	0	2	0	0	0	8	0	2	1	0	1	0	1	2	2	1	5	2	0	2	0	0	6	0]	
I	[0	0	0	0	0	0	0	0	23	0	0	1	0	0	0	0	0	0	0	0	3	4	0	0	0	0]	
J	[0	0	0	0	0	0	0	0	0	28	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
K	[0	0	0	0	0	0	0	1	7	22	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
L	[0	0	0	0	0	0	0	0	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0]	
M	[0	0	0	0	0	0	0	0	0	0	0	0	24	1	0	0	0	0	0	0	3	0	0	0	0	0]	
N	[0	0	0	0	0	0	0	0	7	2	0	0	14	8	0	1	0	0	1	0	0	0	0	0	1	0]	
O	[0	0	0	0	0	1	0	0	0	1	0	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0]	
P	[0	0	0	0	0	0	0	0	0	0	0	0	8	5	0	16	0	0	0	0	0	0	0	0	0	0]	
Q	[0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	27	0	1	1	0	0	0	0	0	0	0]	
R	[0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	26	3	2	0	1	0	0	0	0	0]	
S	[0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	29	0	0	0	0	0	1	0	0]	
T	[0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	27	0	1	0	0	0	0	0	0]	
U	[0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	1	1	0	0	0	0	0]	
V	[0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	30	0	0	0	0	0]	
W	[0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	28	0	0	0	0]	
X	[0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	21	1	0	0]	
Y	[0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	34	0]	
Z	[0	0	0	0	0	1	0	4	0	2	0	0	0	1	0	0	0	1	2	1	0	0	0	2	0	12]	

Ilustración 13 Matriz de confusión

Classification Report				
	precision	recall	f1-score	support
A	0.96	0.67	0.79	36
B	0.94	0.86	0.90	35
C	0.92	0.41	0.56	27
D	0.77	0.65	0.70	31
E	0.92	0.71	0.80	31
F	0.66	0.94	0.77	31
G	0.90	0.60	0.72	30
H	0.50	0.22	0.31	36
I	0.66	0.74	0.70	31
J	0.68	0.97	0.80	29
K	0.65	0.69	0.67	32
L	0.76	1.00	0.86	28
M	0.50	0.86	0.63	28
N	0.53	0.24	0.33	34
O	0.90	0.93	0.92	30
P	0.84	0.55	0.67	29
Q	0.90	0.90	0.90	30
R	0.79	0.79	0.79	33
S	0.62	0.94	0.74	31
T	0.73	0.93	0.82	29
U	0.80	0.93	0.86	30
V	0.81	0.97	0.88	31
W	0.88	0.93	0.90	30
X	0.78	0.78	0.78	27
Y	0.65	0.94	0.77	36
Z	1.00	0.46	0.63	26
accuracy			0.75	801
macro avg	0.77	0.75	0.74	801
weighted avg	0.77	0.75	0.74	801

Ilustración 14 Clasificación de reporte

En la ilustración 14, se muestra en detalle de precisión, recall, f1-score, support. La precisión proporciona información sobre su rendimiento de cuantos han capturado, recall da información de un clasificador de cuantos han fallado, f1-score es el promedio ponderado de precisión y recuperación, y support fueron la cantidad de imágenes que se usaron para cada letra para hacer la matriz. La exactitud del modelo es de 0.75 por ciento de precisión fallando más donde el f1-score tenga menor valor. Eso da a entender que el modelo tiene que ser entrenado con una cantidad más de imágenes para obtener un resultado mayor en la exactitud.

6. RESULTADO

En este capítulo se exponen los resultados obtenidos después de realizar todas las etapas de construcción del modelo de red neuronal para el reconocimiento de lenguaje de señas por imágenes, este resultado fue del mejor ajuste del modelo dado, con el conjunto de datos que se le fue suministrado para entrenarlo y validarlo.

Los resultados obtenidos en nuestro modelo de Redes Neuronales Convolucionales para el reconocimiento del alfabeto de lenguaje de señas se muestran en la tabla 2.

Tabla 2 Reporte de clasificación

	Precision	Recall	F1-score
A	0.96	0.67	0.79
B	0.94	0.86	0.90
C	0.92	0.41	0.56
D	0.77	0.65	0.70
E	0.92	0.72	0.80
F	0.66	0.94	0.77
G	0.90	0.60	0.72
H	0.50	0.22	0.31
I	0.66	0.74	0.77
J	0.68	0.97	0.80
K	0.65	0.69	0.67
L	0.76	1.00	0.86
M	0.50	0.86	0.63
N	0.53	0.24	0.33
O	0.90	0.93	0.92
P	0.84	0.55	0.67
Q	0.90	0.90	0.90
R	0.79	0.79	0.79
S	0.62	0.94	0.74
T	0.73	0.93	0.82

U	0.80	0.93	0.86
V	0.81	0.97	0.88
W	0.88	0.93	0.90
X	0.78	0.78	0.78
Y	0.65	0.94	0.77
Z	1.00	0.46	0.63

En la tabla anterior se muestra una precisión, sensibilidad y la puntuación F1 por cada letra entrenada y validada por el conjunto de datos donde se nota que la mayoría de las letras del abecedario muestran una precisión y sensibilidad más alta, dando como resultado una predicción del modelo muy bueno, esto sale de una teoría de detección de señales la cual es la curva ROC. El análisis de la curva ROC proporciona herramientas para seleccionar los modelos posiblemente óptimos, la curva ROC se usa para pruebas de diagnóstico, la elección se realiza mediante la comparación del área bajo la curva, esta área posee un valor comprendido entre 0,5 y 1 donde 1 es un valor perfecto y 0,5 es una prueba sin capacidad discriminativa diagnóstica.

Hay una forma de interpretar la curva ROC se ha establecido los siguientes intervalos para los valores de AUC:

[0.5]: Es como lanzar una moneda.

[0.5, 0.6): Test malo.

[0.6, 0.75): Test regular.

[0.75, 0.9): Test bueno.

[0.9, 0.97): Test muy bueno.

[0.97, 1): Test excelente.

Las características de la red neuronal se muestran a continuación:

- Capas: 2 capas convolucionales una de ellas con una activación relu, 2 maxpooling, 1 capa flatten, 2 capas dense con 2 activaciones diferentes relu y softmax, 1 capa dropout.
- Optimizador: Adam.
- Función de activación: Relu y Softmax.
- Cantidad de entrenamiento: 3242
- Tamaño de imagen: 150*150
- Stride de filtro de convolucion: 2
- Padding: Same.
- Flatten: Predeterminado.
- Tamaño de filtro MaxPooling: (3*3) y (2*2)
- Dropout: 0.5
- Clases: 26
- Función de pérdida: categorical_crossentropy.
- Métrica: accuracy
- Batch size: 32
- Epochs: 15
- Pasos: 100
- Learning rate: 0.0004

Cuando se entrena y se valida el modelo se usan los parámetros anteriores dados para eso da un margen de error de 0.25 y una exactitud de 0.75. Dado a que son 26 clases que tiene que entrenar y validar hay unas letras del abecedario se predicen mejor que otra, pasa cuando las señas se parecen y eso tiende a confundir al modelo cuando es tanto entrenado como validado.

7. CONCLUSIONES Y TRABAJO EN EL FUTURO

Con la construcción de este proyecto se logró cumplir el propósito general que fue el desarrollar un modelo de predicción de lenguaje de señas usando redes neuronales convolucionales (CNN) que permita el reconocimiento del alfabeto del lenguaje de señas colombiano por imagen, a través del cumplimiento de cada uno de los objetivos específicos planteados, para lo cual se inició construyendo un set de datos de imágenes por medio de celulares móviles capturando fotos de cada letra del abecedario del lenguaje de señas colombiano, teniendo diferentes personas como modelos para lograr una variedad por cada letra.

El siguiente paso fue crear una red neuronal convolucional para entrenarla con el 80% del set de datos que obtuvimos y así lograr una Exactitud del 75% de aciertos. Para lograr esta estabilidad se validó el modelo utilizando el 20% restante del set de datos para su entrenamiento y concluir logrando el propósito del proyecto.

Recomendaciones

Este proyecto se desarrolló como una base científica, para darle continuidad a este tipo de proyecto se espera que se pueda aportar significativamente cambios a la población, para ello se hacen las siguientes recomendaciones.

1. Aumentar el set de datos de cada letra para tener un modelo de entrenamiento mucho más eficiente.
2. Retirar el fondo de las imágenes para lograr una imagen más estable y menos ruido de fondo con el cual se pueda confundir el modelo.
3. Bien pueda seguir con imágenes o avanzar y realizar las capturas en tiempo real con video ya que algunas de las letras tienen movimiento.

8. BIBLIOGRAFÍA

[1] Observatorio Social Población Sorda Colombiana. Instituto Nacional para Sordos - INSOR www.insor.gov.co/observatorio/

[2] Colombiano desarrolla primer traductor online de lenguas a señas - Colombia Digital colombiadigital.net/actualidad/noticias/item/5452-colombiano-desarrolla-primer-traductor-online-de-lenguas-a-senas.html

[3] Herramienta de Apoyo para la Interpretación de Lenguaje de Señas Mexicano (HILSEM). <https://tesis.ipn.mx/bitstream/handle/123456789/15434/Reporte%20t%C3%A9cnico%20final%202014-A041.pdf?sequence=1>

[4] Reconocimiento de imágenes del lenguaje de señas mexicano. <https://tesis.ipn.mx/bitstream/handle/123456789/15991/XM%2013.05.pdf?sequence=1&isAllowed=y>

[5] Clasificador difuso para el reconocimiento continuo del lenguaje de señas a partir de las características de seguimiento y forma. <http://www.indjst.org/index.php/indjst/article/view/98726>

[6] Reconocimiento del lenguaje de señas usando procesamiento de imágenes. https://www.researchgate.net/publication/322186019_Recognition_of_sign_language_using_image_processing

[7] Método automático de reconocimiento de voz para la clasificación de vocales al lenguaje de señas colombiano. <https://repository.ucatolica.edu.co/bitstream/10983/22809/1/M%C3%89TODO%20AUTOM%C3%81TICO%20DE%20RECONOCIMIENTO%20DE%20VOZ%20PARA%20LA%20CLASIFICACI%C3%93N%20DE%20VOCALES%20AL%20L>

ENGUAJE%20DE%20SE%C3%91AS%20COLOMBIANO.pdf#page=55&zoom=100,0,114

[8] Inventaron “software” que interpreta lenguaje de señas - <http://www.buendiario.com/inventaron-software-interpreta-lenguaje-senas/>

[9] Microsoft planea usar Kinect para interpretar el lenguaje de señas - <https://hipertextual.com/2013/07/kinect-lenguaje-senas>

[10] Software convierte voz en lenguaje de señas para sordos - <https://www.scidev.net/america-latina/comunicacion/noticias/software-convierte-voz-en-lenguaje-de-se-as-para-sordos.html>

[11] Pronóstico del tipo de cambio USD/MXN con redes neuronales de retropropagación - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Red-neuronal-artificial-de-cuatro-capas_fig1_323985249

[12] Categorización usando redes neuronales convolucionales. Available from: <https://medium.com/espanol/avances-en-redes-neuronales-705c2efe53d2>

[13] «Especificación de Requisitos según el estándar,» [En línea]. Available: <https://www.fdi.ucm.es/profesor/gmendez/docs/is0809/ieee830.pdf>. [Último acceso: 10 09 2018]

