



**ESPE**  
UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

**DEPARTAMENTO DE ELÉCTRICA Y ELECTRONICA  
CARRERA DE INGENIERIA EN SOFTWARE**

**TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO  
DE INGENIERO EN SOFTWARE**

**TEMA: DESARROLLO DE UNA APLICACIÓN MÓVIL DE VISIÓN  
ARTIFICIAL QUE FORTALEZCA EL PROCESO DE ENSEÑANZA-  
APRENDIZAJE DE CONFIGURACIÓN MANUAL DEL LENGUAJE DE  
SEÑAS EN LOS ALUMNOS DE PRIMER Y SEGUNDO GRADO DE LA  
UNIDAD EDUCATIVA ESPECIALIZADA CAMILO GALLEGOS PARA  
PERSONAS CON DISCAPACIDAD AUDITIVA**

**AUTORES: RENGIFO TULCANAZ, PAOLA BELÉN  
CHICAIZA ROCHA, MAURO LENIN**

**DIRECTOR: ING. CARRILLO MEDINA, JOSÉ LUIS**

**LATACUNGA**

**2019**



**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA**  
**CARRERA DE INGENIERÍA EN SOFTWARE**

**CERTIFICADO DEL DIRECTOR**

Certifico que el trabajo de titulación, **“DESARROLLO DE UNA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL QUE FORTALEZCA EL PROCESO DE ENSEÑANZA-APRENDIZAJE DE CONFIGURACIÓN MANUAL DEL LENGUAJE DE SEÑAS EN LOS ALUMNOS DE PRIMER Y SEGUNDO GRADO DE LA UNIDAD EDUCATIVA ESPECIALIZADA CAMILO GALLEGOS PARA PERSONAS CON DISCAPACIDAD AUDITIVA”**, fue realizado por los señores: Chicaiza Rocha, Mauro Lenin y Rengifo Tulcanaz, Paola Belén, el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Latacunga, 5 de junio del 2019

---

Ing. José Luis Carrillo M.  
C.C: 0501553788  
DIRECTOR DEL PROYECTO




**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA  
CARRERA DE INGENIERÍA EN SOFTWARE**


**AUTORIA DE RESPONSABILIDAD**

Nosotros, **CHICAIZA ROCHA, MAURO LENIN y RENGIFO TULCANAZ, PAOLA BELÉN**, declaramos que el contenido, ideas y criterios del trabajo de titulación: **“DESARROLLO DE UNA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL QUE FORTALEZCA EL PROCESO DE ENSEÑANZA-APRENDIZAJE DE CONFIGURACIÓN MANUAL DEL LENGUAJE DE SEÑAS EN LOS ALUMNOS DE PRIMER Y SEGUNDO GRADO DE LA UNIDAD EDUCATIVA ESPECIALIZADA CAMILO GALLEGOS PARA PERSONAS CON DISCAPACIDAD AUDITIVA”**, es de nuestra autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

Latacunga, 5 de junio del 2019

  
Chicaiza Rocha, Mauro Lenin  
C.C: 1726849423

  
Rengifo Tulcanaz, Paola Belén  
C.C: 1724015506





**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA  
CARRERA DE INGENIERÍA EN SOFTWARE**

**AUTORIZACIÓN**

Nosotros, **CHICAIZA ROCHA, MAURO LENIN** y **RENGIFO TULCANAZ, PAOLA BELÉN**, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **“DESARROLLO DE UNA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL QUE FORTALEZCA EL PROCESO DE ENSEÑANZA-APRENDIZAJE DE CONFIGURACIÓN MANUAL DEL LENGUAJE DE SEÑAS EN LOS ALUMNOS DE PRIMER Y SEGUNDO GRADO DE LA UNIDAD EDUCATIVA ESPECIALIZADA CAMILO GALLEGOS PARA PERSONAS CON DISCAPACIDAD AUDITIVA”**, en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Latacunga, 6 de junio del 2019

  
Chicaiza Rocha, Mauro Lenin  
C.C: 1726849423

  
Rengifo Tulcanaz, Paola Belén  
C.C: 1724015506

## **DEDICATORIA**

Tras el esfuerzo realizado para culminar esta gran meta, quiero dedicar mi proyecto de Titulación a mi querida familia por la unión y el amor que me han brindado siempre, en especial a mi hermana a quien admiro y amo, pues ha sido mi guía y pilar durante toda mi vida. A mis padres por los consejos, el sacrificio, amor y apoyo que me dieron para ayudarme a alcanzar este triunfo. A mis hermanos, incluyendo a mi cuñado, quienes siempre están a mi lado cuando más los necesito.

Y sobre todo, se lo dedico a Dios por poner en mi camino a todas las personas que me han enseñado a seguir adelante para crecer y ser mejor día a día.

Paola Rengifo

## DEDICATORIA

Una vez finalizado el proyecto de titulación, la dedicatoria va dirigida a la memoria de mi abuela María Magdalena Chiquito, quien me brindo su fuerza y amor, para afrontar cualquier adversidad que la vida me ofreciera.

A mis padres que me brindaron su apoyo durante todo este proceso, con sus consejos, enseñanzas, sacrificio, amor y por la confianza depositada en mi persona y con el apoyo incondicional de mis hermanos me permitieron conseguir este logro.

Mauro Chicaiza

## **AGRADECIMIENTO**

Agradezco a mis padres por brindarme su amor y apoyo incondicional durante toda mi vida, a mi hermana por brindarme sus consejos y ayudar a que nunca me rinda; a mi familia gracias por estar pendiente de mí en todo momento; a mi compañero en el trabajo de titulación Mauro Chicaiza, por su esfuerzo, dedicación y colaboración para finiquitar exitosamente nuestra meta.

Agradezco a mi director del proyecto por su guía presentada durante el desarrollo del proyecto y a los profesores de carrera por el conocimiento compartido. Un agradecimiento especial a la Ing. Ximena López por sus constantes consejos, guía y apoyo durante toda la vida universitaria para desarrollarnos profesionalmente.

Finalmente agradezco a la Universidad de las Fuerzas Armadas-ESPE sede Latacunga, por ser mi segundo hogar y haberme permitido formarme profesionalmente en ella.

Paola Rengifo

## **AGRADECIMIENTO**

Por le presente trabajo de titulación quiero agradecer a mi abuela por cuidarme, ser el motivo y la fuerza para cumplir con una de mis metas.

Agradezco a mis padres y hermanos, que con su amor incondicional, supieron darme la fortaleza para conseguir la culminación de este proceso; a mi compañera de proyecto Paola Rengifo, por su esfuerzo y dedicación para concluir con nuestro trabajo de titulación.

Agradezco a mi director de tesis Ing. José Luis Carrillo, que fue una pieza fundamental para la culminación del presente trabajo. A mis compañeros, amigos quienes me dieron todo su apoyo en los buenos y malos momentos durante esta etapa.

Por ultimo quiero extender mi agradecimiento a la Universidad de la Fuerzas Armadas ESPE Extensión Latacunga, por los conocimientos impartidos por los docentes con sus respectivas cátedras, que me permitieron formarme como un profesional.

Mauro Chicaiza



## ÍNDICE DE CONTENIDO

### CARÁTULA

CERTIFICADO DEL DIRECTOR .....	ii
AUTORIA DE RESPONSABILIDAD.....	iii
AUTORIZACIÓN .....	iv
DEDICATORIA .....	v
AGRADECIMIENTO .....	vii
ÍNDICE DE CONTENIDO.....	ix
ÍNDICE DE FIGURAS .....	xiii
ÍNDICE DE TABLAS.....	xv
RESUMEN .....	xvii
ABSTRACT.....	xviii

### CAPÍTULO I

#### DEFINICIÓN DEL PROBLEMA

1.1. Planteamiento del problema .....	1
1.2. Formulación del problema .....	4
1.3. Justificación e Importancia.....	5
1.4. Objetivos.....	6
1.5. Hipótesis.....	7
1.6. Variables de la Investigación .....	7
1.6.1. Variable Independiente .....	7
1.6.2. Variable Dependiente .....	7
1.6.3. Indicadores .....	8

### CAPÍTULO II

#### MARCO TEÓRICO

2.1. Discapacidad Auditiva.....	9
---------------------------------	---

2.2.	Tipos de Discapacidad Auditiva .....	11
2.2.1.	Según el Momento de aparición .....	11
2.2.2.	Según la severidad .....	13
2.2.3.	Según el pronóstico .....	14
2.2.4.	Según el sitio de la lesión .....	15
2.3.	Métodos de apoyo para personas con discapacidad auditiva.....	16
2.3.1.	Métodos Tecnológicos .....	16
2.3.2.	Métodos Físicos.....	19
2.4.	Lenguaje de Señas .....	20
2.5.	Historia del Lenguaje de Señas .....	20
2.6.	Tipos de Lenguaje de Señas .....	25
2.6.1.	Sistema manual .....	25
2.6.2.	Sistema gestual expresivo .....	26
2.6.3.	Sistema corporal .....	26
2.7.	Texto y guías del Lenguaje de Señas Ecuatoriano.....	27
2.8.	Aprendizaje de máquina .....	28
2.8.1.	Tipos de aprendizaje de máquina .....	31
a)	Aprendizaje supervisado.....	32
b)	Aprendizaje no supervisado.....	32
c)	Aprendizaje por Refuerzo .....	33
d)	Redes neuronales artificiales (Aprendizaje supervisado) .....	34
e)	El Perceptrón .....	36
f)	El Perceptrón Multicapa.....	39
g)	Redes Neuronales Convolucionales.....	44
h)	Máquinas de vectores de soporte (aprendizaje supervisado) .....	53
i)	Regresión lineal (aprendizaje supervisado) .....	54
j)	Mezcla de Gaussianas (aprendizaje no supervisado).....	55
k)	Agrupamiento jerárquico (aprendizaje no supervisado).....	57
l)	Q-learning (aprendizaje por refuerzo).....	59

2.9. Visión por computador .....	60
2.10. Procesamiento de imagen .....	63
2.11. Herramientas de Visión por Computador .....	67
2.11.1. Cámaras digitales.....	67
2.11.2. Transferencia de Aprendizaje.....	69
a) MobileNets.....	71
b) Tensorflow .....	72
c) OpenCV .....	73
d) Python.....	74
e) Android Studio .....	74
2.12. Metodología MADAMDM .....	75

### **CAPÍTULO III**

#### **ANÁLISIS Y DISEÑO DE LA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL**

3.1. Introducción .....	78
3.2. Investigación y análisis del proceso enseñanza-aprendizaje del lenguaje de señas en la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.....	78
3.3. Recolección de palabras del lenguaje de señas para la formación del dataset.....	83
3.4. Modelamiento del conocimiento y estructura del modelo a usar .....	85
3.4.1. Análisis y Diseño .....	85
3.5. Diagrama de casos de uso .....	94

### **CAPÍTULO IV**

#### **IMPLEMENTACIÓN DE LA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL**

4.1. Introducción del capítulo .....	101
4.2. Análisis y procesamiento de imágenes.....	101

4.3. Programación de la red neuronal artificial y entrenamiento con el dataset de las imágenes.....	108
4.4. Desarrollo de la interfaz de usuario a utilizar en la aplicación móvil.....	120
4.5. Integración de la red neuronal artificial con la aplicación móvil.....	125
4.6. Pruebas e Implantación.....	127
4.6.1. Pruebas para validar el modelo y el dispositivo en ambientes controlados.....	128
4.6.2. Pruebas del refuerzo de enseñanza-aprendizaje con niños de primer y segundo grado de la Unidad Educativa Camilo Gallegos con la aplicación “En Señas”.....	132
4.6.3. Pruebas de Integración.....	136

## **CAPÍTULO V**

### **VALIDACIÓN Y ANÁLISIS**

5.1. Validación de la aplicación móvil para la enseñanza y aprendizaje del lenguaje de señas.....	138
5.1.1. Análisis y validación de la hipótesis.....	138
5.1.2. Validación de la aplicación.....	144

## **CAPÍTULO VI**

### **CONCLUSIONES Y RECOMENDACIONES**

6.1. Conclusiones.....	147
6.2. Recomendaciones.....	149

<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>151</b>
--	------------

<b>ANEXOS.....</b>	<b>160</b>
--------------------	------------

## ÍNDICE DE FIGURAS

<b>Figura 1</b>	Tipos de Discapacidad Auditiva.....	11
<b>Figura 2</b>	Implante coclear en oído interno .....	18
<b>Figura 3</b>	Tipos de Aprendizaje de Maquina .....	34
<b>Figura 4</b>	Diagrama de una Neurona Artificial (PE).....	37
<b>Figura 5</b>	Perceptrón Multicapa.....	40
<b>Figura 6</b>	Arquitectura de las redes convolucionales .....	45
<b>Figura 7</b>	Capa convolucional .....	46
<b>Figura 8</b>	Mapa de características .....	48
<b>Figura 9</b>	Capa de reducción o pooling.....	49
<b>Figura 10</b>	Capa clasificadora totalmente conectada.....	50
<b>Figura 11</b>	Visión por computador.....	61
<b>Figura 12</b>	Pixel de una imagen .....	64
<b>Figura 13</b>	Herramientas utilizadas para la Visión por Computador.....	67
<b>Figura 14</b>	Aprendizaje por Trasferencia.....	70
<b>Figura 15</b>	Arquitectura MobileNets .....	72
<b>Figura 16</b>	Arquitectura de la metodología MADAMDM.....	76
<b>Figura 17</b>	Imágenes de señas contenidas por carpetas con sus respectivos nombres.....	84
<b>Figura 18</b>	Diagrama de caso de uso general.....	95
<b>Figura 19</b>	Diagrama de caso de uso: Escoge seña .....	96

<b>Figura 20</b> Diagrama de caso de uso: Reconocer seña.....	98
<b>Figura 21</b> Diagrama de Secuencia: Reconocimiento de seña .....	100
<b>Figura 22</b> La Imagen (A) es la imagen original y a la derecha (B) la imagen procesada (Palabra lunes) .....	102
<b>Figura 23</b> Programación de la red neuronal artificial .....	108
<b>Figura 24</b> Preparación del modelo y extracción de las características .....	110
<b>Figura 25</b> Resultados obtenidos por archivos planos denominados Bottleneck de la seña pararse(A) y lunes (B) .....	111
<b>Figura 26</b> Proceso de predicción de la seña.....	114
<b>Figura 27</b> Entrenamiento de la red .....	116
<b>Figura 28</b> Localización del modelo y las etiquetas de las imágenes .....	125
<b>Figura 29</b> Proceso de reconocimiento de la seña por la cámara del móvil .....	126
<b>Figura 30</b> Detección y reconocimiento de la palabra “Color” .....	130
<b>Figura 31</b> Detección y reconocimiento de la palabra “Lunes” .....	131
<b>Figura 32</b> Análisis del refuerzo del conocimiento del lenguaje de señas .....	143

## ÍNDICE DE TABLAS

<b>Tabla 1</b>	<i>Requisito funcional del sistema. Consultas del abecedario .....</i>	85
<b>Tabla 2</b>	<i>Requisito funcional del sistema. Consultas del vocabulario.....</i>	86
<b>Tabla 3</b>	<i>Requisito funcional del sistema. Selección y consulta de datos.....</i>	87
<b>Tabla 4</b>	<i>Requisito funcional del sistema. Reconocimiento de los gestos del lenguaje de señas .....</i>	88
<b>Tabla 5</b>	<i>Requisito funcional del sistema. Validación del reconocimiento del gesto.....</i>	89
<b>Tabla 6</b>	<i>Requisito funcional del sistema. Ayuda del sistema.....</i>	90
<b>Tabla 7</b>	<i>Requisito no funcional del sistema. Rendimiento de la aplicación móvil .....</i>	91
<b>Tabla 8</b>	<i>Requisito no funcional del sistema. Interfaz de la aplicación móvil .</i>	91
<b>Tabla 9</b>	<i>Requisito interfaz de usuario. Pantalla inicio.....</i>	92
<b>Tabla 10</b>	<i>Requisito interfaz de usuario. Menú deslizante.....</i>	93
<b>Tabla 11</b>	<i>Caso de uso general.....</i>	95
<b>Tabla 12</b>	<i>Especificación de caso de uso: Escoger Señá .....</i>	97
<b>Tabla 13</b>	<i>Diagrama de Casos de uso: Reconocer seña.....</i>	98
<b>Tabla 14</b>	<i>Palabras con sus respectivas imágenes del lenguaje de señas: se ha clasificado las imágenes de cada seña como con fondo y sin fondo para probar la predicción del reconocimiento .....</i>	103

<b>Tabla 15</b> <i>Letra con su respectiva imagen del lenguaje de señas: se ha clasificado las imágenes de cada seña como con fondo y sin fondo para probar cuál de las dos formas es favorable al reconocimiento</i> .....	106
<b>Tabla 16</b> <i>Proceso de entrenamiento</i> .....	117
<b>Tabla 17</b> <i>Prueba del modelo predictivo</i> .....	118
<b>Tabla 18</b> <i>Pantalla de Inicio de la Aplicación</i> .....	120
<b>Tabla 19</b> <i>Menú de la Aplicación</i> .....	121
<b>Tabla 20</b> <i>Pantalla de vocabulario</i> .....	122
<b>Tabla 21</b> <i>Pantalla Muestra</i> .....	122
<b>Tabla 22</b> <i>Pantalla de Práctica</i> .....	123
<b>Tabla 23</b> <i>Pantalla Abecedario</i> .....	124
<b>Tabla 24</b> <i>Pantalla de Ayuda</i> .....	124
<b>Tabla 25</b> <i>Reconocimiento de imágenes. Pruebas al dispositivo móvil con el modelo predictivo en un ambiente controlado</i> .....	129
<b>Tabla 26</b> <i>Métricas de precisión para evaluar el refuerzo del aprendizaje con la aplicación "En Señas"</i> .....	134
<b>Tabla 27</b> <i>Pruebas de integración</i> .....	136
<b>Tabla 28</b> <i>Calculo de la media de las diferencias</i> .....	142



## **RESUMEN**

El presente proyecto brinda soporte al proceso de enseñanza-aprendizaje de la configuración manual del lenguaje de señas ecuatoriano, es dirigido a los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos, ubicado en la ciudad de Ambato. La aplicación “En Señas” está enfocada en funcionar como una herramienta iterativa e interactiva para la enseñanza-refuerzo del usuario y/o las personas con discapacidad auditiva, por lo que presenta una colección de 50 palabras del diccionario del lenguaje de señas publicado por la FENASEC.EC. Esta aplicación cuenta con tres segmentos, el primero de ellos es el vocabulario, el cual contiene la colección de palabras, que permite mostrar imágenes con el significado de la seña y videos que muestran la forma correcta de realizar la seña. El segundo segmento cuenta con el proceso práctico, en donde la cámara del dispositivo móvil detecta y reconoce la configuración manual realizada por el alumno. Por último el segmento de validación, una vez reconocida la seña, la aplicación determina si se realizó correcta o incorrectamente la configuración manual mostrando un mensaje de acierto o error. La aplicación se desarrolló mediante los frameworks Opencv y Python en conjunto con la librería de Tensorflow para el tratamiento de las imágenes y la construcción de la red neuronal. Fue realizada en Android Studio, utilizando la metodología MADAMDM para la interfaz de usuario y la unión de la red neuronal artificial.

### **CLAVE:**

- **DISCAPACIDADES AUDITIVAS**
- **LENGUAJE DE SEÑAS - ENSEÑANZA**
- **REDES NEURONALES**
- **VISIÓN ARTIFICIAL**

## **ABSTRACT**

This project provides support to the teaching-learning process of the manual configuration of the Ecuadorian sign language. It is aimed at first and second grade students of the Camilo Gallegos Specialized Educational Unit, located in the city of Ambato. The application "En Señas" is focused on functioning as an iterative and interactive tool for the teaching-reinforcement of the user and / or hearing impaired people, so it presents a collection of 50 words from the dictionary of sign language published by FENASEC.EC. This application has three segments, the first of them is the vocabulary, which contains the collection of words, which allows to show images with the meaning of the sign and videos that show the correct way to make the sign. The second segment has the practical process, where the camera of the mobile device detects and recognizes the manual configuration made by the student. Finally, the validation segment, once the password has been recognized, the application determines if the manual configuration was made correctly or incorrectly by displaying a message of success or error. The application was developed using Opencv and Python frameworks together with the Tensorflow library for image processing and neural network construction. It was done in Android Studio, using the MADAMDM methodology for the user interface and the union of the artificial neural network.

### **KEY:**

- **HEARING IMPAIRMENTS**
- **SIGN LANGUAGE - TEACHING**
- **NEURAL NETWORKS**
- **ARTIFICIAL VISION**

# CAPÍTULO I

## DEFINICIÓN DEL PROBLEMA

### 1.1. Planteamiento del problema

El ser humano al ser un animal racional, social desde la antigüedad ha sentido la necesidad de comunicarse, por tal necesidad, a través del tiempo creó muchos sistemas de comunicación como teléfono, radio, televisión, y por último el internet, pero todos estos inventos han sido desarrollados dando prioridad a personas sin discapacidad, excluyendo a una sociedad con discapacidad auditiva dado que la información que reciben estos es escasa.

La comunicación se da cuando existe interacción entre dos o más individuos que cumplen el proceso de transmitir y recibir información a través de un canal en específico y un código en común; pero si uno de estos elementos falla o no funciona, la comunicación no existe. Uno de los problemas que se han dado a lo largo de la historia es con personas que tienen discapacidad auditiva, ya que al ser una discapacidad sensorial que no se ve físicamente, recibe muy poca atención por parte de la sociedad, misma que desconoce el tipo de código o lenguaje que usan las personas con imposibilidad de oír para poder desarrollar una comunicación adecuada y natural (RENKEMA, 2011).

Para entender este tipo de discapacidad es menester señalar que la discapacidad auditiva en un ser humano puede ser de cualquier tipo (conductiva<sup>1</sup>, neurosensorial<sup>2</sup>, o mixta<sup>3</sup>), de cualquier grado (leve<sup>4</sup>, moderada<sup>5</sup>, severa<sup>6</sup>, profunda<sup>7</sup>) sea unilateral (de un oído) o bilateral (de dos oídos), que puede ser congénita (presente desde nacimiento), o adquirida (después de nacimiento) prelingue o postlingue (período en que adquirió el sistema del lenguaje), puede ser periférica (lesión en el oído externo, medio o interno) o central (relacionado a los desórdenes del nervio central o de la corteza cerebral); cualquiera de estos tipos generalmente utiliza y puede beneficiarse para identificar los sonidos con algún tipo de amplificación ya sean audífonos, implantes cocleares u otros. Se los llama sordos a personas con discapacidad auditiva que poseen un grado de deficiencia total o profunda. Cabe mencionar que las personas que necesitan aprender el lenguaje de señas son las que tienen un grado de sordera profunda de 91Db o más.

---

<sup>1</sup>Todo inconveniente en el oído externo o medio que impida que el sonido se transmita adecuadamente, se conoce como pérdida auditiva conductiva o bien de transmisión. Las pérdidas auditivas conductivas son normalmente de grado leve o moderado, oscilando entre los 25 y los 65 decibelios.

<sup>2</sup>La pérdida auditiva neurosensorial es el resultado del deterioro o ausencia de células sensoriales (células ciliadas) en la cóclea y acostumbra a ser permanente.

<sup>3</sup>Una pérdida auditiva mixta es la combinación de una pérdida auditiva neurosensorial y conductiva. Se genera a consecuencia de inconvenientes tanto en el oído interno como medio.

<sup>4</sup>Entre 26-40 dB (Decibelios) Dificultad para escuchar un discurso y conversaciones suaves, especialmente en situaciones ruidosas, pero sin problemas en ambientes silenciosos.

<sup>5</sup>Entre 41-55 dB Dificultad para entender una conversación, especialmente en presencia de ruido de fondo. Se necesitan mayores niveles de volumen para escuchar la televisión o la radio.

<sup>6</sup>Entre 71-90 dB, la conversación regular es inaudible y hay dificultad incluso en voz alta. Frecuentemente la comprensión es solo posible a través de gritos o de amplificación.

<sup>7</sup>De 91 o más dB Incluso el sonido amplificado es difícil de comprender o es inaudible. (RENKEMA, 2011)

La gran mayoría de personas con discapacidades auditivas tienen una capacidad intelectual alta y hasta competitiva, pero se cree que hay una relación entre sordera y retraso mental, por el hecho de que se desconoce e ignora la manera en como ellos establecen una comunicación, haciendo sentir a esta minoría excluida, sin oportunidades académicas, laborales, sociales, etc. (Salicci, 2007).

Ante esta problemática que tiene el discapacitado de la audición para comunicarse, los métodos que se han desarrollado para entenderlos, son netamente físicos como por ejemplo la escritura, lectura de labios y/o el lenguaje de señas siendo este último en el que se va a profundizar en esta investigación ya que este ayuda más comúnmente a la comunicación entre estas personas por ser un lenguaje propio de personas con discapacidad auditiva.

En la actualidad, en el Ecuador, se están implementando planes de inclusión social, con alrededor 26 comisiones de diferentes partes del país con la Fenasec<sup>8</sup>; en donde se pretende ayudar a todas las personas, con discapacidad auditiva, a acceder a servicios, bienes y empleo, que sean ubicadas en áreas en donde son aptas para desarrollar su trabajo con total normalidad.

En algunas ciudades del país existen centros y escuelas de educación especial para personas con discapacidad auditiva, en donde se imparten asignaturas de educación

---

<sup>8</sup>FENASEC Es una organización nacional que aglutina a asociaciones en las diferentes provincias del país, en 1992 quedo legalmente constituida mediante acuerdo ministerial N°1333 en el Ministerio de Bienestar Social, está inscrita en el registro nacional de ONG's del Consejo Nacional de Igualdad de Discapacidades "CONADIS" y desde 1995 es filial a la World Federation of the Deaf "W.F.D", lo que le ha permitido relacionarse e interactuar con la Comunidad Mundial de Sordos (FENASEC, 2018).

básica mediante el lenguaje de señas. Sin embargo la mayor parte de personas con discapacidad auditiva son analfabetas no solo por falta de recursos, sino también, por la falta de información. Así mismo existen personas sin discapacidad auditiva que desean comunicarse con ellos por medio del lenguaje de señas sin tener éxito, por el hecho de no tener a una persona o sistema que les enseñe este lenguaje en cualquier momento y lugar.

Uno de estos centros es la Unidad Educativa Especializada Camilo Gallegos ubicada en la ciudad de Ambato, esta unidad educativa se especializa en la educación de personas con discapacidad auditiva la misma que no cuenta con un software específico (aplicación móvil) para ayudar al refuerzo de la enseñanza y aprendizaje de las palabras básicas del lenguaje de señas en los alumnos de primer y segundo grado, dentro y fuera de las aulas.

## **1.2. Formulación del problema**

Existen distintos factores que dificultan el aprendizaje del lenguaje de señas de los alumnos como la falta de concentración, falta de interés de los alumnos para reforzar el conocimiento adquirido en aulas, mala instrucción del lenguaje de señas por parte de los docentes, los cuales influyen directamente en el aprendizaje del alumno ocasionando una mala abstracción de conocimientos impartidos en clase.

Basándonos en estos inconvenientes se formula el siguiente problema: ¿Cómo fortalecer el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas para las personas con discapacidad auditiva?

### **1.3. Justificación e Importancia**

En el Ecuador según el Registro Nacional de Discapacidades y el Ministerio de Educación, existe a nivel nacional 50,580 personas con discapacidad auditiva. De estas, el 40 % tiene acceso a la educación básica, el 5 % al bachillerato y educación superior. La mayoría de discapacitados son analfabetos y sufren por discriminación ya sea académica, laboral o social, generando en ellos sensación de soledad, aislamiento y frustración, influyendo en el incremento de la pobreza, la desigualdad social y a la persistencia del subdesarrollo económico (eltelegrafo, 2017).

En la actualidad se hace necesario que las personas conozcan las expresiones básicas del lenguaje de señas, para poder interactuar entre personas con discapacidad auditiva y estas a su vez puedan fortalecer sus conocimientos sobre el lenguaje de señas, utilizando para ello tecnologías actuales como son los dispositivos móviles u ordenadores.

Por estos motivos antes mencionados y de acuerdo a las cifras dadas según el Registro Nacional de Discapacidades y el Ministerio de Salud Pública que al existir una población de 50,580 personas con discapacidad auditiva en distintas adversidades, se ha decidido realizar el presente proyecto que involucra la investigación sobre: el

lenguaje de señas, el empleo de técnicas de inteligencia artificial como es la visión artificial para analizar, diseñar y desarrollar una aplicación móvil que permita enseñar y retroalimentar los movimientos básicos de las manos del lenguaje de señas a las personas con discapacidad auditiva.

Mediante el proyecto de investigación se pretende proveer una aplicación para fortalecer la enseñanza-aprendizaje de configuración manual del lenguaje de señas mediante videos e imágenes y verificar si los alumnos de primer y segundo grado la Unidad Educativa Especializada Camilo Gallegos consigue imitar/simular los movimientos de las manos mediante un dispositivo móvil. Lo que permitirá fortalecer el aprendizaje del lenguaje de señas para personas con discapacidad auditiva.

#### **1.4. Objetivos**

##### **Objetivos generales**

Desarrollo de una aplicación móvil de visión artificial que fortalezca el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

##### **Objetivos específicos**

- Elaborar el marco teórico orientado a las técnicas de visión artificial.
- Identificar los movimientos básicos de rasgos manuales del lenguaje de señas.



- Elaborar una aplicación con visión artificial que permita reforzar el proceso de enseñanza-aprendizaje de rasgos manuales del lenguaje de señas para las personas con discapacidad auditiva.
- Validar los resultados obtenidos de la aplicación móvil realizada.

### **1.5. Hipótesis**

Si se desarrolla una aplicación móvil con visión artificial entonces se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

### **1.6. Variables de la Investigación**

#### **1.6.1. Variable Independiente**

Si se desarrolla una aplicación móvil con visión artificial

#### **1.6.2. Variable Dependiente**

Se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

### **1.6.3. Indicadores**

- Incremento del aprendizaje
- Reducción del tiempo de enseñanza
- Validación de la aproximación de la imitación/simulación de los rasgos manuales del lenguaje de señas.
- Porcentaje de precisión del algoritmo de aprendizaje superior al 70%

## CAPÍTULO II

### MARCO TEÓRICO

#### 2.1. Discapacidad Auditiva

La Organización Mundial de la Salud (OMS)<sup>9</sup> estima como una de las discapacidades más habituales en todo el mundo a la pérdida de audición, se estima que alrededor de 466 millones de personas, padece dicha discapacidad, esto es el 5% de la población mundial de los cuales: 34 millones son niños y 432 millones son adultos. (OMS, 2018)

La pérdida de la audición, una discapacidad sensorial, es la carencia o degradación en la aptitud para oír nítidamente los sonidos del ambiente, el lenguaje oral, ya sean estos débiles, distantes o en lugares que presentan excesivo ruido ambiental, depende de los contextos en el que se desenvuelve la persona, todo esto conllevado por un problema en algún sitio del aparato auditivo (OMS, 2018).

Conforme el Registro Nacional de Discapacidades y el Ministerio de Salud Pública del Ecuador, a nivel nacional existen cincuenta mil quinientas ochenta personas con discapacidad auditiva. De ellos, 54,34% son hombres, y el 45,65%, mujeres. De estas, las que tiene acceso a la educación básica son el 40 %, el 5 % a bachillerato y educación superior, mismos que tienen una capacidad intelectual alta, como cualquier ser humano, pero se cree que hay una relación entre sordera y retraso mental, por el

---

<sup>9</sup> OMS es la autoridad directiva y coordinadora en asuntos de sanidad internacional en el sistema de las Naciones Unidas (OMS, 2018)

hecho de desconocer e ignorar la manera en como ellos establecen una comunicación, haciendo sentir a esta minoría excluida, sin oportunidades académicas, laborales, sociales, etc. (Salicci, 2007)

Según la OMS, en el año 2017, las causas que inducen pérdida de la audición pueden ser: genéticas, dificultades al momento del nacimiento, padecimientos infecciosos crónicos del oído, el manejo de medicamentos inadecuados, el continuo riesgo al exponerse al ruido y la degeneración natural del oído con el paso del tiempo.

En países de Sudamérica, las causas principales de pérdida auditiva son las infecciones crónicas del oído. El predominio de la otitis media crónica o de las infecciones crónicas del oído varía entre el 1 y el 46% en los países desarrollados y los países en desarrollo. La otitis media crónica provoca pérdida de la audición, logrando ocasionar complicaciones potencialmente mortales e incluso de defunción. (OMS, 2018)

Esta discapacidad auditiva no se manifiesta, es invisible, ya que no expone características físicas evidentes. Se hace evidente principalmente por, el uso del audífono, las personas que han nacido sordas, haber adquirido la pérdida auditiva a temprana edad o por el tono de voz, el que en varios casos existe diferencia al común de la gente.

## 2.2. Tipos de Discapacidad Auditiva

Existen múltiples personas con distintos tipos de discapacidad auditiva, en función de cada particularidad que posee cada uno, que están clasificados según: el momento de aparición, la severidad, el pronóstico y el sitio de la lesión, tal como se presenta en **Figura 1**:



**Figura 1.** Tipos de Discapacidad Auditiva

Fuente: (ASPAS, 2018) (Sosa, 2016)

### 2.2.1. Según el Momento de aparición

La pérdida auditiva se puede dar en cualquier momento ya sea por malformaciones en el nacimiento (congénita) o en el transcurso del tiempo por diversos factores externos como infecciones en el oído (adquirida).

- **Congénita.-** Este tipo de discapacidad auditiva se origina por factores hereditarios que se pueden descubrir a lo largo del tiempo, son mutaciones genéticas en el ADN, estas se logran detectar en el nacimiento, exceptuando casos donde las mutaciones reaccionan con la edad. Esto se genera, ya sea por malformaciones en la estructura interna que compone el oído o por un anormal funcionamiento en el ámbito molecular e incluso celular a nivel neurosensorial, siendo este último resultado del deterioro o ausencia de células sensoriales (células ciliadas<sup>10</sup>) en la cóclea y acostumbra a ser permanente.

Por otra parte también pueden darse casos no hereditarios sino por complicaciones en el embarazo o en el parto como; rubeola materna, asfixia del bebé en el parto, bajo peso al nacer, etc. (OMS, 2018)

En este tipo de discapacidad se puede realizar pruebas para detectar la pérdida de audición a temprana edad, es decir, en los primeros meses del nacimiento o en el momento de nacer. (Sosa, 2016)

- **Adquirida.-** Se originan por agentes externos tóxicos, es decir, donde el individuo consume medicamentos sin prevención, drogas o se somete a

---

<sup>10</sup> Las células ciliadas son aquellas que poseen Cilios. Los cilios son prolongaciones cilíndricas delgadas que se proyectan desde la superficie de la célula, están formadas por micro túbulos del cito esqueleto envueltos por la membrana ciliar. (Lenoir, 2016)

ruidos extremos o constantes sin protección. En estos casos, la persona se puede realizar pruebas de respuesta auditiva en cualquier momento del transcurso de su vida, para comprobar si hay pérdida de audición, generalmente las industrias o médicos sugieren efectuar dichos exámenes de manera trimestral o mensual. (OMS, 2018)

### 2.2.2. Según la severidad

- **Hipoacusia.-** Es la pérdida o disminución de la sensibilidad auditiva que puede ser unilateral, es decir, que aquejan a un solo oído, o bilateral, si los dos oídos son afectados. Los sonidos se captan de acuerdo a rangos en dB<sup>11</sup> y se subdivide conforme a los niveles de audición que tenga la persona, estos pueden ser:
  - **Audición normal.-** Implica lograr percibir sonidos en el rango entre - 10dB a 25dB. (NIOSH, 2008)
  - **Pérdida auditiva leve.-** El rango de sonido mínimo detectable es entre 26 y 40 dB. (hear-it, 2017) El individuo tiene mayor dificultad para oír al entablar conversaciones, principalmente en ambientes con excesivo ruido ya que está incapacitado para captar los sonidos normalmente.

---

<sup>11</sup> dB.- Es una unidad relativa que se utiliza para medir los niveles del sonido (NOISSES, 2015)

- **Pérdida auditiva moderada.**- Los sonidos se detectan a partir de los 40-70 dB, generalmente la persona debe usar audífono, ya que sin él no podría escuchar.
- **Cofosis.**- Es también conocida como anacusia o sordera profunda. Es considerada una auditiva severa, el individuo detecta sonidos a más de 90 dB, es decir, que es incapaz de percibir algún tipo de sonido.
- **Sordera.**- Hay audición al percibir sonidos entre 75-90dB, por lo que técnicamente la persona es sorda, pero a diferencia de la sordera profunda en la cofosis, los sonidos son amplificados por medio de dispositivos electrónicos diseñados para que pueda lograr una audición relativamente aceptable. (contributors, 2013)

### 2.2.3. Según el pronóstico

Para verificar si la discapacidad auditiva será duradera o momentánea de acuerdo el daño que tenga el oído, estableciéndose una calificación que puede ser:

- **Permanente.**- En este primer caso, la audición se perderá para perpetuamente, es decir, no se restaura, hasta consigue deteriorarse con el paso del tiempo, llegando a la cofosis.



- **Transitoria.-** En el caso de la transitoria al contrario del primer caso, la incapacidad de escuchar es momentánea, ya que la persona afectada con el tiempo podría alcanzar el rango de pérdida de audición entre leve y moderada, sin que esto implique una discapacidad severa, pero ocasionalmente y/o en ciertos casos podría recuperar la audición normal. (Sosa, 2016).

#### 2.2.4. Según el sitio de la lesión

Como se ha mencionado anteriormente en algunos casos la pérdida de la audición suele suscitarse ya sea por problemas externos o internos que afectan a las partes o nervios del oído, estos pueden ser:

- **Sordera sensorial neural.-** En estos casos la lesión se centra en los nervios, ubicados en el oído interno y en la corteza auditiva, estos llevan el impulso auditivo al cerebro donde finalmente se interpretan los sonidos. (ASPAS, 2018)
- **Sordera de conducción.-** Son alteraciones de transmisión sonora en el oído externo y el oído medio, suelen tener un pronóstico favorable ya que son asequibles a tratamiento. (ASPAS, 2018)

Dependiendo de los tipos de discapacidad auditiva con posibilidad de tratarse, según las características que se han mencionado, existen algunos

métodos de apoyo que brindan facilidades a las personas para mejorar no solo su audición sino su calidad de vida.

### **2.3. Métodos de apoyo para personas con discapacidad auditiva**

Si bien es cierto se han creado métodos tanto tecnológicos como físicos para mejorar la escucha de personas con discapacidad auditiva, estos no responden únicamente al tema biológico, sino que también tienen un trasfondo interpersonal que les ayudará a obtener una mejor calidad de vida, tanto físico y mental como social. Donde los métodos pueden clasificarse de siguiente forma:

#### **2.3.1. Métodos Tecnológicos**

Entre las distintas formas para mejorar la audición de una persona se han creado dispositivos electrónicos, cuyo uso aumenta la amplificación del sonido, estos son:

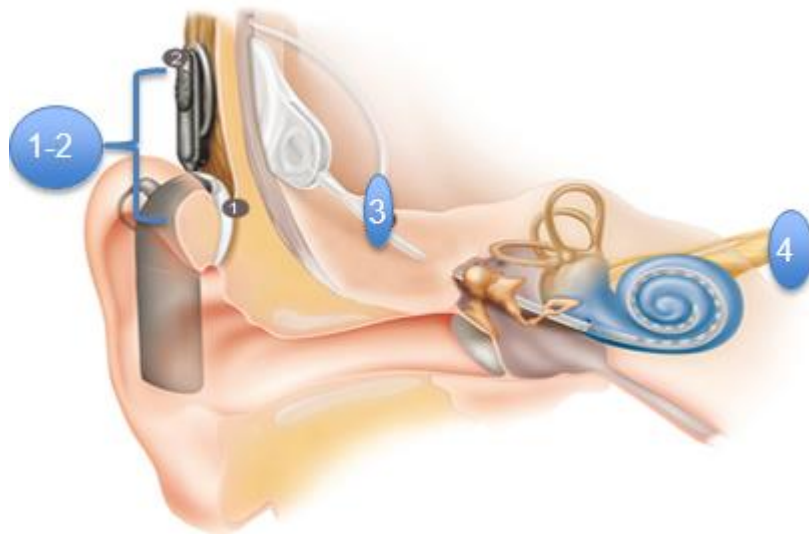
- **Audífonos o prótesis auditivas.-** Son dispositivos electrónicos especialmente diseñados, para tener la función de llevar y amplificar el sonido, sirve para que la persona pueda oír de mejor manera y así pueda comunicarse sin problema. El dispositivo funciona dependiendo del grado de sordera que se tenga, es decir que se pueden utilizar en casos como sordera, hipoacusia normal, hipoacusia leve, hipoacusia moderada.

- **Implantes cocleares.-** Es un mecanismo médico que sustituye las partes dañadas del oído interno con un transductor, es decir, que el implante recibe energía acústica o señales acústicas y las convierte en señales eléctricas, mismas que a su vez estimulan el nervio auditivo para captar el sonido.

Es menester señalar que las señales eléctricas son gestionadas mediante las partes que conforman dicho implante, las cuales se dividen en externas e internas:

- **Externas.-** Estos implantes son los que se utilizan para recolectar los sonidos, formados por:
  - **Micrófono:** Los sonidos captados son enviados al procesador.
  - **Procesador:** Distingue y recopila los sonidos para que llegue al individuo y pueda interpretar las ondas de audio receptadas. (AICE, 2015)
- **Internas.-** Implantes formados por piezas mecánicas en las distintas secciones del oído interno, cada una cumpliendo una función que ayuda a la audición, como se muestra en la Figura 2. Y se describe a continuación:
  - **1-2.- Receptor Estimulador:** Se implanta en el oído interno, su función es enviar señales eléctricas.

- **3.- Electroodos:** Son colocadas en el interior de la cóclea, es decir, en la parte del oído interno en forma de caracol, enviando estimulaciones a las células nerviosas que no están dañadas.
- **4.-Nervio Auditivo:** Son los nervios estimulados que aún trabajan llevando señales que llegan al cerebro para que las reconozca como sonidos.



**Figura 2** Implante coclear en oído interno

Fuente: (Arndt P, 1999)

- **Sistemas de Frecuencia Modular (FM).**- Son dispositivos que envían el sonido por medio de ondas de radio de alta frecuencia, desde el receptor hasta el audífono, creados para captar tanto la señal auditiva como el ruido de fondo, esto en razón de amplificar el sonido de acuerdo a la situación, logrando una comunicación transmisor-receptor.

### 2.3.2. Métodos Físicos

Los métodos físicos son netamente creados a nivel anatómico, puesto que tiene que ver con la comunicación a través de movimientos corporales que se llevan a cabo integrando gestos y rasgos manuales que utilizan las personas con discapacidad auditiva:

- **Lectura de labios.-** La lectura de labios es común en las personas sordas, en especial que no gozan de otro medio de comunicación puesto que es una técnica de intuición del habla a través de la que se interpreta y traduce el movimiento de los labios, la cara y la lengua, dependiendo del contexto o situación en que la persona se encuentre. Este método también lo puede usar cualquier sujeto interesado en la lectura de labios que desee comunicarse, ya sean discapacitados auditivos o personas normales.
- **Aprendizaje de lenguaje de signos.-** Se fundamenta en un pensamiento semántico de la comunicación, es decir, es un lenguaje no verbal que incluye expresiones faciales, movimientos gestuales del cuerpo como las manos, los brazos y los dedos, con el fin emitir y recibir información.

Con el transcurso del tiempo estos métodos han sido los medios de comunicación más importantes para que las personas con discapacidad auditiva, de las que hemos mencionado en este documento, se incluyan en varios ámbitos; social, cultural, político, laboral, entre otros. Esto, teniendo en cuenta que han

sobresalido en una sociedad indiferente a ellos, puesto que de una u otra forma han sido marginados y considerados como una minoría.

#### **2.4. Lenguaje de Señas**

Es el medio de comunicación que emplea expresiones corporales y faciales, cuenta con su propia estructura gramatical, relaciones y funciones de los componentes de la oración. Con este tipo de lenguaje las personas con discapacidad auditiva se expresan, obtienen conocimientos y, en último término, acceden a la cultura del lugar donde habitan.

En el Ecuador existe la Federación Nacional de Personas Sordas (FENASEC) cuya comunidad agrupa y ayuda a personas con discapacidad auditiva de todo el país, para el reconocimiento de sus derechos, identidad cultural y lingüística. Posee cursos gratuitos con los cuales se puede aprender el lenguaje de señas de una manera visogestual, mediante prácticas con movimiento de manos, gestual y corporal (FENASEC, 2018)

#### **2.5. Historia del Lenguaje de Señas**

En el siglo XV las personas que nacían con una discapacidad auditiva severa no podían hablar normalmente, por lo que, se expresaban por medio de señas y gestos. Creándose el lenguaje de señas cuando este grupo de personas sordas se reúnen y comparten su conocimiento.

En el siglo XVI, el monje Benedicto español, Pedro Ponce de León, empezó a enseñar a los niños, con deficiencia auditiva, a desarrollar las habilidades de leer y escribir, considerado como el primer maestro de las personas sordas, para enseñar a los niños sordos el monje Pedro Ponce de León recurría a graficas en las cuales representaba objetos reales y palabras, estas representaciones eran vinculas con el sistema de señas manuales. (Sykes, 2009)

En 1620 Juan Pablo de Bonet, sordo de nacimiento, pedagogo, versado en sociología, historia, filosofía recopiló la información divulgada por el monje Pedro Ponce de León, redactó y publicó el primer alfabeto manual del lenguaje de señas para personas con discapacidad auditiva en la obra denominada “Reducción de las letras y arte para enseñar a hablar a los mudos” (Sykes, 2009), esta obra fue retomada por L'Epée, para describir los signos utilizados por sordos, en París. El francés Charles Michel de L'Epée, en 1712 estudió el lenguaje de señas y se esforzó por difundirlo, durante la segunda mitad del siglo XVII funda la “Institution Nationale des Sourds-Muets” en París que es la primera escuela pública para sordos, utilizando un idioma de gestos y configuración de manos, para referirse a determinadas palabras y conceptos. Los profesores aplicaban estos métodos para enseñar a sus alumnos, esto mejoró la educación de los niños con discapacidad auditiva. Con estos conocimientos las personas sordas eran capaces de establecer una comunicación por medio de un sistema de gestos con funciones y estructuras lingüísticas propias de una lengua. (Venegas, 2016)

L'Épée recopiló las señas lingüísticas en un diccionario de lenguaje de señas parisina titulado "Dictionarie de Sourd-muets" (diccionario de sordomudos), el cual contenía adjetivos, pronombres, preposiciones, conjunciones y más de doscientos verbos con su conjugación, persona, tiempo, y modo de expresión (Oviedo, cultura-sorda, 2007). Esta información fue compartida con personas interesadas en esta lengua. L'Épée consideraba al lenguaje de señas como el único vehículo adecuado para desarrollar el pensamiento y la comunicación entre las personas con discapacidad auditiva. (Rodríguez M. I., 2000). Mejorando la estructura gramatical y lingüística del lenguaje de señas para ello elaboró "Los Signos Metódicos", es un recopilado signos que facilita la enseñanza-aprendizaje del lenguaje de señas francés, contenían gestos que no existían en este lenguaje, como artículos y preposiciones uso las señas empleadas por sus alumnos como base y agrego gestos para nombrar objetos, hechos y actividades. (Oviedo, cultura-sorda, 2007)

En el año 1816, se divulga la metodología de la enseñanza del lenguaje de señas en Estados Unidos. En este país se fundó la escuela para personas con discapacidad auditiva llamada "La escuela de Hartford" usando la metodológica concebida por L'Épée, teniendo como primer maestro a Laurent Clerc, destacado profesor y pedagogo francés. Quien estudio en la escuela pública para sordos dirigida en París por el Abad Michel de l'Épée. En él años 1864, en Washington DC, se funda la Universidad Gallaudet, como la única universidad, para sordos, en todo el mundo; en la que se fusionaron: el método francés y el alemán. Este método se enfoca en la comprensión



del significado de una palabra utilizando objetos, señas y el alfabeto manual para su debida representación. Hasta el año 1867 se crearon alrededor de 20 escuelas, para niños sordos con el apoyo del gobierno.

En el Congreso de Milán (1880), en el cual participaron países como: Bélgica, Francia, Alemania, Inglaterra, Italia, Sueca, Rusia, Estados Unidos y Canadá, entre otros. Thomas Gallaudet, docente de ciencias morales y políticas en el Consejo Nacional para Sordos de Washington, en donde enfatizo la importancia de los métodos mixtos que comprendían los gestos naturales y el alfabeto manual para la enseñanza del lenguaje de señas. (Ignacio, 2009)

En el año de 1897 se creó en Argentina el Instituto Nacional en donde se separó a los niños y a las niñas para dictar clase, esto influyo en el proceso de desarrollo del lenguaje de señas argentina (LSA), que al día de hoy se evidencia en las señas que solamente son utilizadas por las mujeres y otras son exclusivamente usadas por hombres, como por ejemplo: las diferentes señas para referirse a los números, colores, meses, entre otros.

A partir de 1960 se realizó un análisis más profundo sobre el lenguaje de señas utilizando para ello categorías gramaticales universales (fonética, semántica, morfosintaxis) con las que se estudian diversas lenguas. (Rodriguez M. I., 2000). William Stokoe estudio la estructura gramatical y fonológica de la lengua de señas y determino que esta es similar a las demás lenguas. En 1980 se difundieron los estudios

llevados a cabo por Stokoe, que reconocían a las personas sordas como una comunidad minoritaria que habla su propia lengua con una estructura gramatical y lingüística. (Rodríguez M. I., 2000)

En el Ecuador en el año 1940, se abrió en Quito la primera escuela del país para niños con necesidades educativas especiales por las hermanas Enriqueta y Teresa Santillán. Teniendo como concurrencia de manera particular niños con discapacidad auditiva creándose una sección solo para ellos. En la actualidad la institución lleva el nombre de “Enriqueta Santillán Cepeda”. Una segunda institución se crea en el año de 1952 en Quito, La escuela Mariana de Jesús, que atendía a niños sordos y ciegos. En donde, la lengua de señas española ejerció una importante influencia en la comunidad sorda ecuatoriana. En el Ecuador actualmente existen un poco más de 100 escuelas en el país que atienden a niños sordos, bien sea en aulas especiales o integrados junto con otros niños.

En la actualidad, se reconoce que el lenguaje de señas es el idioma natural de los sordos. Hoy, se defiende el derecho que tienen, para educarse en su propia lengua, textos y guías. En donde el lenguaje de señas se puede clasificar en tipos de señas, dependiendo en gran medida a la parte o partes del cuerpo humano que es empleado por las personas con discapacidad auditiva para comunicarse con otras personas

## **2.6. Tipos de Lenguaje de Señas**

El lenguaje de señas se organiza en tres aspectos fundamentales, estos son: sistema manual, sistema gestual, sistema corporal.

### **2.6.1. Sistema manual**

Es el empleo de una o dos manos para realizar un conjunto de señas que representan una forma de expresión lingüística, constituyen un estimable apoyo en el lenguaje de señas. El sistema manual tiene las siguientes características:

- Formas de la mano
- Ubicaciones espaciales que la mano presenta en las señas
- Tipo de movimientos de las manos que pueden ser observados en las señas.

Cabe mencionar que, los niños sordos, hijos de padres con discapacidad auditiva, repiten los movimientos básicos del lenguaje de señas realizado por sus padres en sus primeros meses (ocho o nueve); en consecuencia los niños sordos educados en ese tipo de entorno, pronuncia las primeras señas significativas. (Biblioteca virtual Miguel de Cervantes, 2016).

### **2.6.2. Sistema gestual expresivo**

Es el empleo de expresiones y gestos de la cara como los ojos, boca y lengua para representar un conjunto de señales. Algunos de estos gestos son expresiones de orgullo, triunfo, poder, tristeza, alegría, entre otros. El sistema gestual se clasifican en:

- Emblemáticos.- tienen su propio significado sin necesidad de palabras.
- Adaptadores.- manipulaciones de nuestro propio cuerpo u objetos para canalizar las emociones.
- Manifestadores de afecto.- con los que transmitimos nuestros sentimientos.

### **2.6.3. Sistema corporal**

Es la utilización de diversas partes del cuerpo (hombros, el tronco, los pies y el cuello) para expresar y comunicar ideas. En el sistema corporal no se emplea manos y expresiones faciales. (Alberto, 2009). En el lenguaje de signos los gestos corporales son elementos importantes donde las manos cobran la mayor importancia, son la entonación que en la lengua oral complementa a los gestos, es decir, si un intérprete de signos no maneja bien su expresión corporal no podrá comunicarse correctamente por ese canal, ya que el receptor no sabrá si le está realizando una afirmación, una pregunta o una exclamación.

Los distintos significados que tienen en el lenguaje de signos la expresión corporal y facial son:

- Ceño fruncido, músculos de la cara tensos expresa enfado, tristeza o miedo, dependiendo del matiz que se le dé.
- Músculos de la cara relajados, presencia de sonrisa: Significa que el mensaje es positivo. Expresa amor, alegría, amistad, etc.
- Movimientos amplios y fluidos: Nos expresan alegría, amor, amistad o cualquier otro mensaje positivo.
- Movimientos secos y cortos: Expresan miedo, tristeza, odio, entre otros.

## **2.7. Texto y guías del Lenguaje de Señas Ecuatoriano**

En la actualidad en Ecuador existen dos diccionarios del lenguaje de señas, mismos que han sido divulgados por las asociaciones de personas sordas: Consejo Nacional de Igualdad de Discapacitados y la Federación Nacional de Personas Sordas del Ecuador. (FENASEC, 2018)

- En 1983, en Quito, se inició el Proyecto “Mano a Mano” por la asociación de sordos Ponce de León, con el objetivo de registrar las señas locales en un texto sobre el lenguaje de señas ecuatoriano (LSE), llamado “Lenguaje de Señas:

guía básica sobre una comunicación especial”. La obra ilustra, con gráficos y dibujos acompañados por su significado en español, cerca de doscientas cincuenta señas usadas en Quito y Guayaquil. (sordosecuador, 2017)

- El Diccionario oficial de lengua de señas ecuatoriana se publicó en el año 2012, una obra impresa en 2 tomos, que contiene 4.363 entradas (señas ilustradas con dibujos y acompañadas de glosa española).
- En junio del 2014 el Consejo Nacional de Igualdad de Discapacitados (CONADIS) junto con la Federación Nacional de Personas Sordas del Ecuador (FENASEC) presentaron el diccionario virtual de lenguaje de señas ecuatoriano “Gabriel Román” (Sordos Ecuador, 2017). Este proyecto reúne aproximadamente 5000 palabras del Diccionario oficial de la Lengua de señas ecuatoriana, que consta de imágenes y videos donde se muestra la forma adecuada de realizar una determinada seña (Sordos Ecuador, 2017). Para más información ver la página web: <http://fenasec.ec/diccionario-lsec.html>

## **2.8. APRENDIZAJE DE MAQUINA**

El aprendizaje es el proceso de obtener conocimiento, destrezas o actitudes, a través de distintas maneras, como por ejemplo la investigación, la experiencia, la reproducción continua, etc.

El ser humano es capaz de utilizar todos sus sentidos para aprender, pero siendo que en el caso de estos individuos no es posible utilizar el sentido auditivo, inmediatamente se recurre al de la vista primordialmente, mismo que envía información al cerebro para realizar el aprendizaje.

El proceso de aprendizaje produce un cambio constante, medible y específico en la conducta de un individuo y hace que el mismo manifieste un constructo mental nuevo o que examine uno previo y lo almacene en su cerebro llegando a concretarse en un conocimiento permanente (Gonzalez, 2007)

El Aprendizaje de Máquina (ML por sus siglas en inglés, Machine Learning) es el desarrollo de sistemas que usan patrones y/o algoritmos para lograr que la máquina aprenda a solucionar problemas específicos y de esta manera cambie el comportamiento de la misma, respondiendo de manera autónoma en base a su experiencia.

El ML es una de las técnicas de aprendizaje de la inteligencia artificial, que se basa en el conocimiento previo de la máquina, esto debido a que encuentra patrones en grandes volúmenes de datos (Gonzales, 2014)

La razón por la cual ha llegado a ser reconocido a nivel mundial, es su efectividad y que el ML juega un importante papel dentro de una gran gama de sistemas cruciales, como son; minería de datos (deduce los patrones y tendencias que existen en grandes

conjuntos de información); procesador de idioma natural (pretende conseguir que una máquina comprenda lo que expresa una persona mediante el uso de una lengua natural o un idioma hablado); reconocimiento de imagen (donde se reconocen objetos específicos de una imagen); sistemas de expertos (estos emulan el comportamiento humano para la solución de un problema), entre otros.

Además, también existen varias aplicaciones que las personas usan en la vida cotidiana por medio de su tecnología inteligente, como los asistentes personales virtuales que hoy en día empresas como; Amazon y Google, los han hecho en forma de parlantes inteligentes. Otras industrias los desarrollan incluyendo asistentes virtuales en los teléfonos celulares (por ejemplo SIRI, asistente de Apple) o como aplicaciones que se pueden descargar para computadoras (como es el caso de CORTANA, asistente de Windows).

Otras de las aplicaciones son; **los mensajes spam en correos electrónicos**, es decir, mensajes o correos no deseados de cadenas de publicidad u ofertas de dudosa procedencia; **los motores de búsqueda**, utilizados para realizar búsquedas de información en la web, un claro ejemplo es Google, mismo que emplea ML para filtrar las peticiones del usuario y darle opciones apegadas a lo que este busca; **recomendaciones de productos en línea**, este realiza sugerencias o recomendaciones de productos o servicios a la venta basados en las compras realizadas anteriormente por el usuario; **videos de vigilancia**, en los cuales el ML



monitorea todas las cámaras de vigilancia en busca de comportamientos inusuales como: objetos olvidados, personas perdidas, robos, etc., es decir, cosas fuera de lo común en las áreas de vigilancia; **predicciones durante desplazamientos** al reservar un taxi, como Uber, o simplemente predicciones de tráfico con navegación de GPS, entre muchas más. (Software Narciso, 2017)

Una de las aplicaciones de ML más populares y que muchas de las personas lo usan inconscientemente, son los servicios de personalización de las redes sociales como; personas que quizás conozcas, reconocimiento facial en fotos (etiquetas), páginas de interés, eventos recomendados, entre otros. En todos estos casos se utiliza el aprendizaje continuo de gustos, sitios web recurrentes, búsquedas frecuentes, perfiles visitados, para extraer información, ya sean imágenes, audios, videos, textos y datos de todo tipo.

En conclusión el ML, es una técnica que le permite a la máquina utilizar conceptos de aprendizaje del ser humano para que el ordenador sea capaz de aprender, entender y ejecutar reglas u órdenes, basado en la experiencia y con la ayuda de aplicaciones que seguirán abriendo puertas a un nuevo futuro tecnológico (Software Narciso, 2017).

### **2.8.1. Tipos de aprendizaje de máquina**

El aprendizaje de maquina se clasifica dependiendo del algoritmo que se utilice para resolver un determinado problema, en el cual influye el ambiente

en donde se desarrolla y los factores que afectan en la toma de decisiones. Dependiendo del tipo de aprendizaje se clasifican en: supervisado, no supervisado y por refuerzo.

#### **a) Aprendizaje supervisado**

El aprendizaje supervisado utiliza algoritmos que trabajan con datos clasificados y etiquetados, para construir modelos predictivos basados en los patrones encontrados a partir de estos datos.

El algoritmo se entrena con un conjunto de datos denominado datos de entrenamiento, que deben estar especificado en su valor, grupo o categoría al que pertenece, es decir, los datos deben estar asociados a un conjunto de etiquetas. Esto permite que el algoritmo construya un modelo que asigne etiquetas de salida adecuadas a un nuevo valor, predice el valor de salida (John, 2018).

#### **b) Aprendizaje no supervisado**

Este tipo de aprendizaje utiliza datos que no están: clasificados, etiquetados y categorizados para el proceso de entrenamiento. Se conoce los datos de entrada, pero no existen valores de respuesta que correspondan a estos datos de entrenamiento. Por tanto, sólo podemos describir la estructura de los datos, para intentar encontrar algún tipo de organización

que simplifique el análisis. Por ello, tienen un carácter exploratorio (Gonzales, 2014).

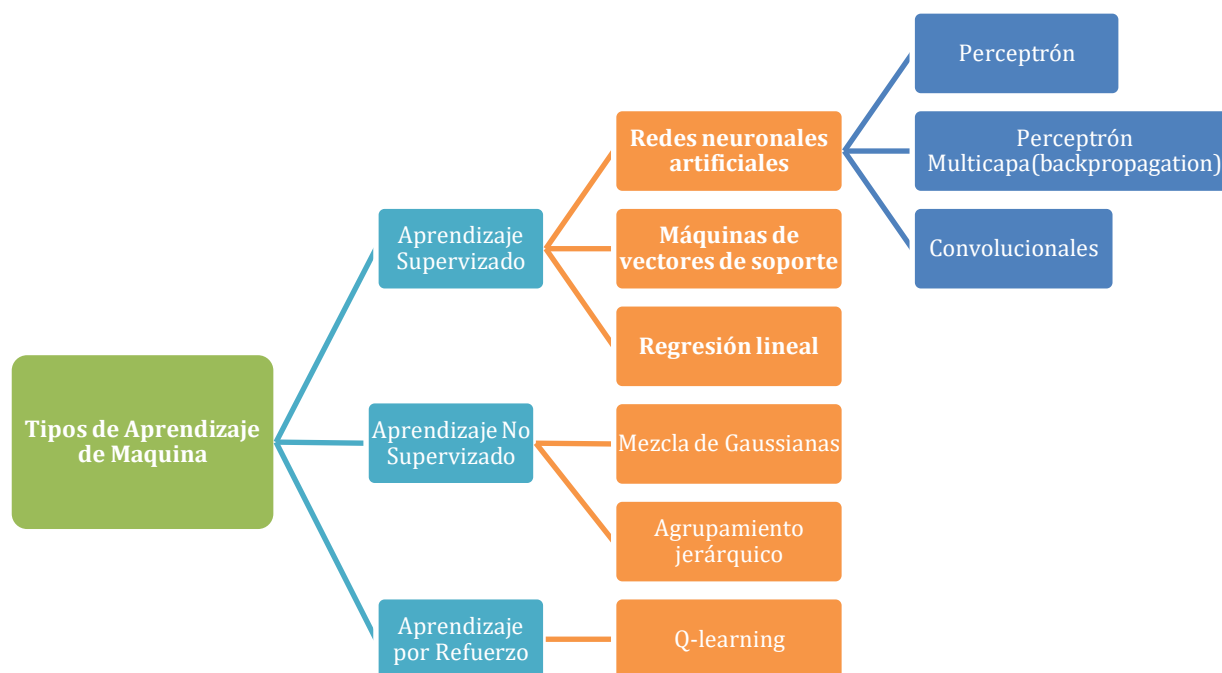
### **c) Aprendizaje por Refuerzo**

En el aprendizaje por refuerzo un agente, que puede ser un sistema, interactúa con el entorno para conseguir un objetivo. Mediante la propia actividad de interacción, el agente debe aprender una política de interacción; es decir, una estrategia de elección de posibles acciones para llegar a la solución del problema (Nolla, 2014).

La realización de una acción cambia el estado del entorno y las consecuencias de ello son sancionadas por el entorno mediante una señal de recompensa (premio o castigo).

El aprendizaje por refuerzo tiene por objetivo aprender como mapear situaciones a acciones para maximizar una cierta señal de recompensa.

En la Figura 3 se puede observar los tipos de aprendizaje y los algoritmos para realizar aprendizaje de máquina.



**Figura 3 .** Tipos de Aprendizaje de Máquina

#### **d) Redes neuronales artificiales (Aprendizaje supervisado)**

Una RNA (red neuronal artificial) es un arquetipo matemático inspirado en el proceder de las redes neuronales de modelos biológicos, ya que consta de elementos que se comportan de forma similar a la neurona biológica en sus funciones más comunes. Las redes neuronales artificiales aprenden de la experiencia y abstraen las características principales de una serie de datos (Claudio Javier Tablada, 2009).

Las características fundamentales de las RNA son:

- **Aprenden de la experiencia:** Las redes neuronales artificiales modifican su comportamiento como respuesta a su entorno. Se presenta un conjunto de

datos en la entrada de la red, esta aprende y se ajusta para producir respuestas consistentes a nuevos valores. Existe una amplia variedad de algoritmos de entrenamiento que se han desarrollado para representar este proceso, cada uno con sus propias ventajas y desventajas

- **Generalizan de ejemplos anteriores a los ejemplos nuevos:** una vez que la RNA fue entrenada con datos de entrenamiento, la respuesta de la red para nuevos valores pueden ser insensibles a pequeñas variaciones, lo que las convierte en la mejor opción para el reconocimiento de patrones.
- La unidad básica de una red neurona artificial es el procesador elemental denominado neurona, la capacidad de este elemento es la de calcular una suma ponderada de sus entradas y luego aplica una función de activación para obtener una señal que será transmitida a la próxima neurona. Estas redes neuronales artificiales se agrupan en capas con un alto grado de conectividad entre las neuronas de diferentes capas, conectividad que es representada por pesos. Mediante algoritmos de aprendizaje supervisado y no supervisado, las RNA ajustan su arquitectura y parámetros para minimizar errores en su modelo de predicción y la capacidad de generalización en nuevos ejemplos.

La red neuronal asimila los datos de un determinado problema, estos datos denominado conjunto de entrenamiento, debe ser amplio y variado, para que la

red pueda hallar las características relevantes y generalizarlos en un conjunto adecuado de pesos y sesgo para que la red pueda resolver correctamente un problema (Matich, 2001).

Se puede distinguir varios tipos de arquitecturas neuronales en función de la forma en que se organizan las neuronas y por el número de capas, como: la red perceptrón, la red perceptrón multicapa y redes convolucionales.

### e) El Perceptrón

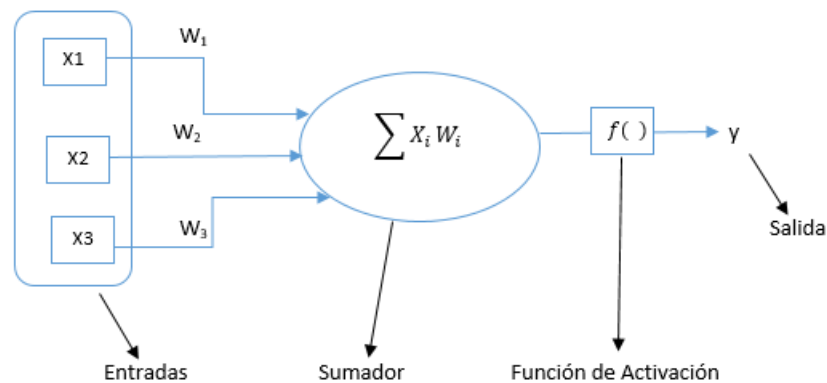
Es la representación más sencilla de una red neuronal, y es utilizado para la categorización de ejemplos, es decir características que se localizan en los dos lados del hiperplano (datos linealmente separables). Fue desarrollado por el psicólogo Frank Rosenblatt en el año 1957, basándose en la regla de aprendizaje de hebb<sup>12</sup> y de los modelos de neuronas biológicas (Anguas Pérez, 2005).

Los elementos que forman el perceptrón son: entrada, sumador, activador y elemento de salida (ver Figura 4), las operaciones realizadas por cada elemento son:

---

<sup>12</sup> **Aprendizaje de hebb** es conectar dos neuronas sincrónicamente y asincrónicamente, la conexión sincrónica en este vínculo se produce un incremento de la energía al conectar simultáneamente las neuronas, entre los dos lados de la sinapsis, por el contrario al ser conectadas asincrónicamente, proporciona una unión débil de la sinapsis (Pérez, 2018).

- **Entradas (receptor)**, donde llegan una o varias señales de entrada ( $X_i$ ) que proviene de otra neurona y que son amplificadas cada una de ellas con un factor de peso  $W_i$  que constituye la conectividad entre la neurona fuente de donde provienen y la neurona de destino.
- **Sumador**, efectúa la suma algebraica ponderada de las señales de entrada, ponderándolas de acuerdo con su peso.
- **Función de activación**, una función no lineal de umbral (que es una función escalón o una curva logística) a la salida del sumador para decidir si la neurona se activa, disparando una salida o no.
- **Elemento de salida**, produce la señal, de acuerdo con el elemento anterior, que constituye la salida de la neurona.



**Figura 4.** Diagrama de una Neurona Artificial (PE)

Fuente: (CAMELO, 2009)

El entrenamiento del perceptrón se lo realiza mostrándole un conjunto de datos de entrenamiento en forma repetida. Cada dato de entrenamiento es una pareja formada por un vector de entrada  $x$  y su vector de salida  $y$  deseable. La dimensión del vector de entrada es igual al número de neuronas de la capa de entrada, en tanto que la dimensión del vector de salida es igual al número de neuronas de la capa de salida.

La ley de aprendizaje del perceptrón es ajustar los pesos para obtener con mayor probabilidad la salida deseada correspondiente aun cierto conjunto de datos.

Para el reajuste de los pesos se realizan los siguientes pasos:

- Inicialización de los pesos, se asignan valores aleatorios a cada uno de los pesos ( $W_i$ ) de las conexiones.
- Se presenta los datos de entrada y la salida deseada, el patrón de entrada  $X=(x_1, x_2, x_3, \dots, x_n)$ , salida esperada  $y$
- Calculo de la salida actual

$$y(t) = f\left(\sum_{i=1}^n w_i x_i\right)$$

Siendo  $f$  la función de transferencia escalón



- El reajuste de los peso es igual al peso anterior  $w_i$  mas una cantidad proporcional a la diferencia entre la salda deseada ( $t$ ) y la salida real ( $y$ )

$$w_{i+1} = w_i + n(t - y)$$

Donde  $n$  es una constante de proporcionalidad menor que la unidad que se llama razón de aprendizaje. Se vuelve al paso dos hasta ajustar los pesos de la red perceptrón para obtener la salida deseada con nuevas entradas.

El perceptrón simple sólo sirve para clasificar problemas linealmente separables, cosa que ya se podía hacer mediante métodos estadísticos, y de una forma mucho más eficiente.

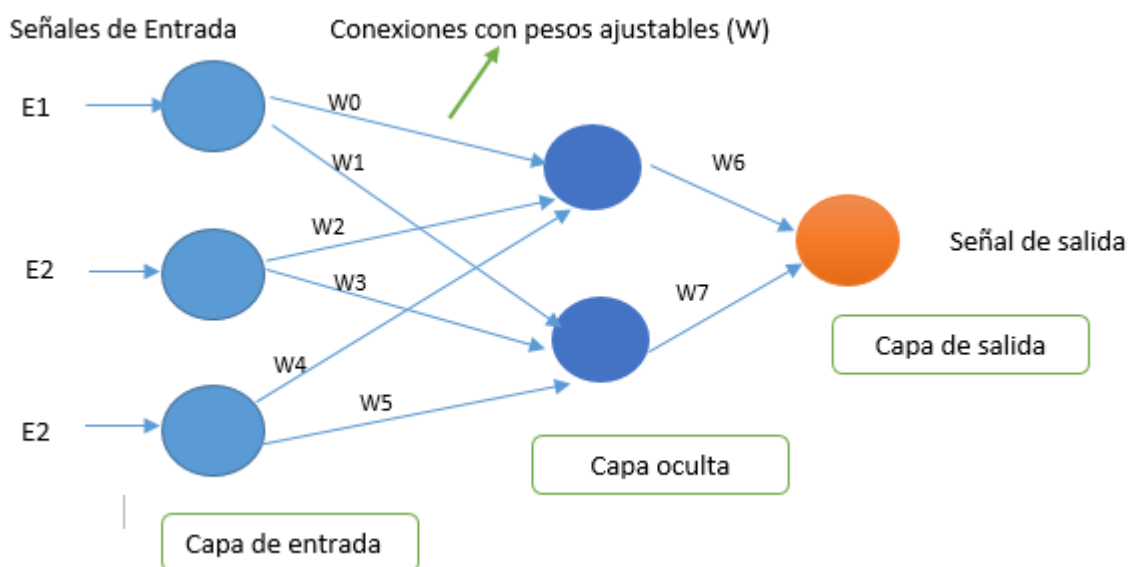
#### **f) El Perceptrón Multicapa**

Es una red neuronal artificial formada por múltiples capas, esto le permite resolver problemas que no son linealmente separables, lo cual es la principal limitación del perceptrón. El perceptrón multicapa tiene la siguiente estructura:

- Capa de entrada. Está conformada por neuronas, que introducen variables de entrada a la red.
- Capas ocultas. Son conjuntos de capas, que reciben los datos de entrada de las neuronas de la capa anterior y envían su salida a la capa

posterior. Su salida representa la propagación de la suma ponderada de las entradas (sesgo y pesos) provenientes de la capa anterior.

- Capa de salida: conecta las capas ocultas con la salida de la red, en esta capa se calcula los valores de salida que produce la red, es decir, proporciona los resultados de predicción.



**Figura 5.** Perceptrón Multicapa

*Fuente:* (Calvo, Perceptrón Multicapa – Red Neuronal, 2018)

Esta red multicapa está totalmente conectada, las neuronas de la capa de entrada están conectadas a cada neurona de la capa oculta y estas a su vez están conectadas a las neuronas de la capa de salida, es decir, todas las neuronas de una capa están conectadas a todas las neuronas de la siguiente capa (ver Figura 5).

La red neuronal perceptrón multicapa tiene dos fases, la primera fase se denominada propagación hacia adelante, que es la propagación de los datos de la capa de entrada hasta la capa de salida, se obtienen los valores de salida. Para producir esta propagación es necesario que las neuronas de las capas ocultas procesen la información recibida por sus entradas y produzcan una respuesta que se propaga, a través de las conexiones, hacia las neuronas de la siguiente capa oculta hasta llegar a la capa de salida y comparar los resultados que se obtiene con los que se desea tener (Calvo, Perceptrón Multicapa – Red Neuronal, 2018).

La segunda fase comienza comparando los valores de salida de la red con la salida esperada, para obtener el error. Se ajustan los pesos de la capa de salida proporcionalmente a este error, se propaga este valor a la capa anterior, se ajustan los pesos y se continua a con este proceso hasta llegar a la primera capa, este proceso se lo denomina retropropagación del error.

El algoritmo de aprendizaje, Backpropagation (propagación hacia atrás) modifica los pesos de las conexiones de la red para cada patrón de aprendizaje del problema. Como el error se propaga de forma inversa, al que comúnmente funciona una red neuronal artificial; este método también es lo denomina descenso del gradiente, por la manera en cómo se consigue el error (García, 2006).

A continuación se muestra como se minimiza el error que está contenido en en los pesos de las neuronas que integran la capa de salida.

En la capa de entrada compuesta por neuronal (k) y la salida (p) se representa en la siguiente ecuación:

$$entrada_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + w_{k0}^o$$

En esta ecuación la letra  $L$  es el número total de entradas (i) de la red neurona artificial, misma que se encuentra equilibrada por sus pesos (w). La salida se determina mediante los cálculos de sus entradas y la función de activación.

La ecuación:  $Error_{pk} = (y_{pk} - o_{pk})$  determina el error en la salida que se obtiene al realizar la operación de la ecuación anterior.

En donde se calcula el error entre la salida deseada ( $y_{pk}$ ) de los datos de entrada y la salida resultante ( $o_{pk}$ ) mediante el ajuste de pesos y sesgo.

$$E_p = \frac{1}{2} \sum_{k=1}^M (y_{pk} - o_{pk})^2$$

Donde M es el número total de salidas que posee la red.

El error que se produce entre la salida deseada con la salida de red neuronal se define como:

$$\delta_{pk}^o = (y_{pk} - o_{pk}) \times f^o(entrada_{pk}^o)$$

Mediante el cálculo del error en la capa de salida, los pesos son ajustados mediante la siguiente ecuación:

$$W_{kj}^o_{nuevo} = W_{kj}^o_{actual} + \eta * \delta_{pk}^o * i_{pj}$$

El coeficiente de velocidad de aprendizaje ( $\eta$ ) en la ecuación anterior, tiene valores entre 0 y 1, para asegurar que la red perceptrón multicapa obtenga una respuesta, para conseguir dicha respuesta es necesario que la red realice una gran cantidad de iteraciones con valores altos para  $\eta$  la red podría alejarse de la solución. (García, 2006)

Los errores de la última capa (salida) se propagan hacia atrás durante el proceso de entrenamiento, hasta llegar a la capa oculta. Para conseguir el error en la última capa, se utiliza el método descenso del gradiente, con el cual se calcula el error medio, en base a los pesos de la capa de salida. Este método también es utilizado en las capas ocultas, pero utilizando sus respectivos pesos.

Las principales ventajas del perceptrón multicapa es el aprendizaje con datos históricos, disminución del ruido y generalización en la resolución de problemas.

Esta red es utilizada en diversas aplicaciones en donde se procesa información compleja, entre se destaca la codificación de datos, reconocimiento del habla e imágenes visuales, control de robots, pronóstico de eventos y control automático de procesos (García, 2006).

Una aplicación que utiliza la red perceptrón multicapa es el NETtalk, que es capaz de transformar texto en inglés en sonido individual (representaciones

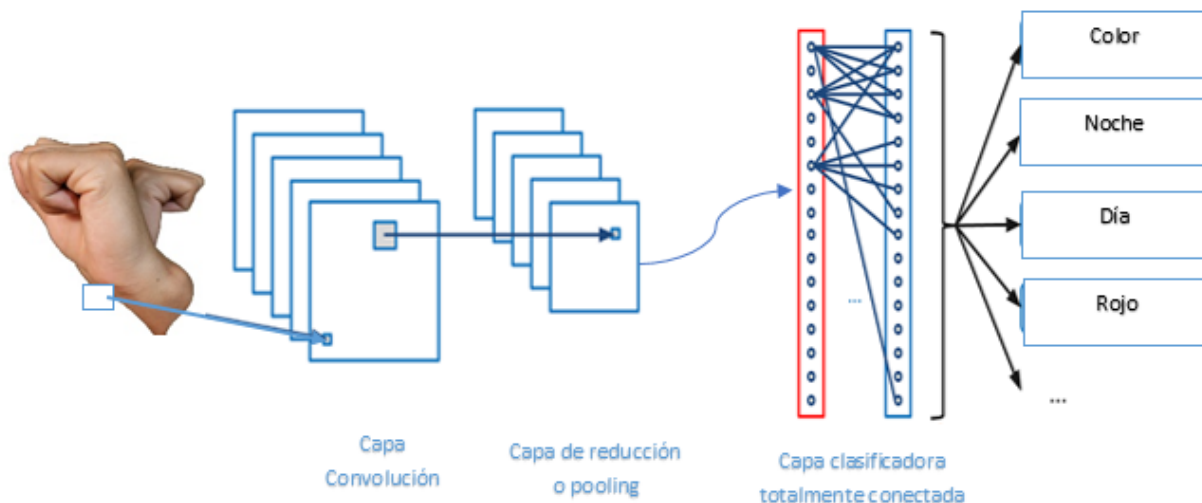
fonéticas) y la pronunciación con la utilización de un sintetizador de voz; cuenta con aproximadamente 300 nodos de neuronas (80 neuronas en la capa oculta) y 20,000 conexiones individuales.

### **g) Redes Neuronales Convolucionales.**

Las redes neuronales ordinarias vistas en la sección anterior presentan problemas de carga computacional al realizar procesos complejos de reconocimiento de imágenes. Por ejemplo se dispone de una imagen con unas dimensiones de 640 x 640 píxeles, esto produciría un valor de 409.600 datos de entrada que estarán conectados a cada una de las neuronas de la primera capa. Esto ralentizaría el proceso de entrenamiento por la gran cantidad de cálculos que se realizaría. Para solucionar este problema se presentan las redes neuronales convolucionales ya usadas a finales de los años 90 pero no muy difundidas, sin embargo, en los últimos años se han popularizado enormemente al lograr resultados muy impresionantes en el reconocimiento de imagen, impactando profundamente en el área de visión por computador (Torres, 2018).

Las redes neuronales convolucionales son redes que están formadas por diferentes capas (ver figura 6), que consiguen identificar patrones más complejos en los datos de entrada. Tratan de buscar características en pequeñas regiones de las imágenes, como pueden ser bordes, formas, colores homogéneos, en los grupos de datos de entrenamiento. Las últimas

capas están totalmente conectadas y son las encargadas de dar una predicción a partir de las características extraídas.

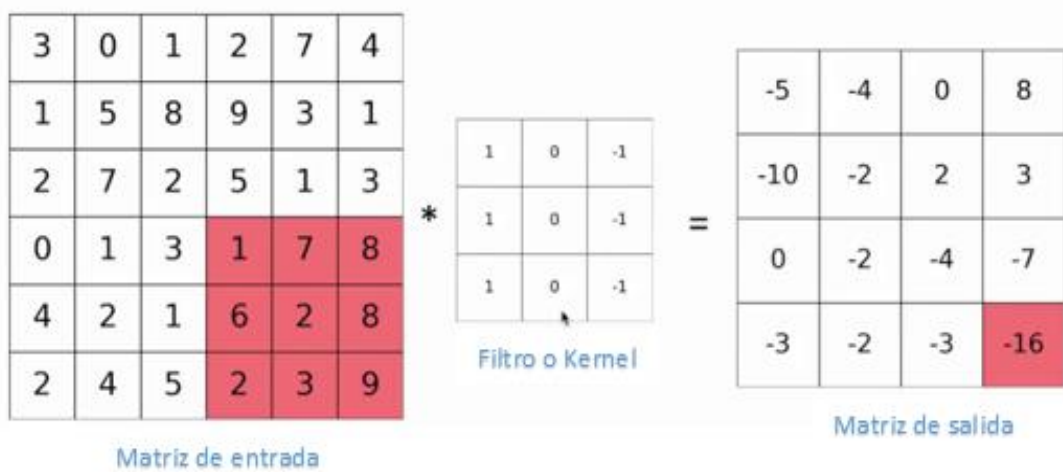


**Figura 6.** Arquitectura de las redes convolucionales

**Fuente:** (Bendemra, 2018)

La arquitectura de las redes neuronales convolucionales está formada de tres tipos de capas: capa convolucional, capa de reducción o de pooling y una capa clasificadora totalmente conectada (ver Figura 6).

- **Capa convolucional.** Es la encargada de realizar operaciones llamadas convolución, estas operaciones recibe como entrada una imagen a la cual se le aplica un filtro o kernel que nos devuelve una nueva matriz con las principales características de la imagen original (ver Figura 7), con esto se logra reducir el tamaño de los parámetros.



**Figura 7.** *Capa convolucional*

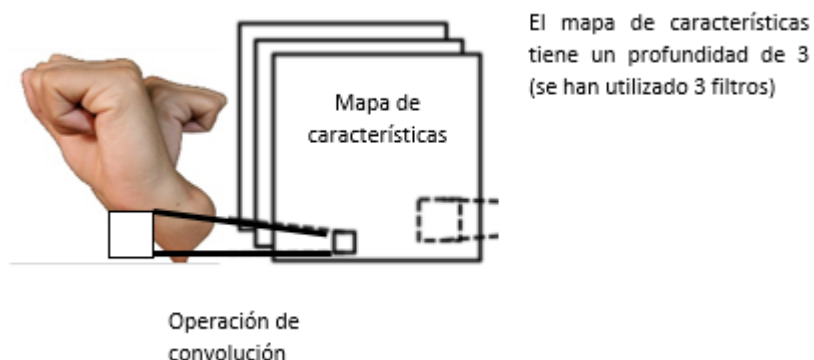
Para representar una imagen en formato digital se utilizan las matrices, donde cada celda representa un pixel, un pixel es la menor unidad de color que conforma una imagen digital y estará representado por un valor numérico (ver figura 7). Se utiliza una matriz denominada filtro (ver Figura 7), que actúa como detector de características de la imagen de entrada.

La operación de convolución consiste en desplazar de izquierda a derecha la matriz filtro sobre la imagen de entrada con un salto de un pixel y en cada movimiento se realiza una multiplicación de matrices y se suman los elementos de esta matriz para obtener un valor, que formara parte de la matriz de salida, denominada mapa de características (ver Figura 7).



El tamaño del mapa de características es controlado por tres parámetros: profundidad, zancada más conocida como en inglés stride y relleno de cero (Bendemra, 2018).

- La profundidad corresponde al número de filtros que se utiliza para la operación de convolución. En la figura 8 se utiliza tres filtros distintos, produciendo tres mapas de características diferentes, estos mapas no son más que tres matrices de dos dimensiones apiladas.
- Zancada (stride). Es el número de píxeles por los que deslizamos la matriz de filtro sobre la matriz de entrada. Cuando el stride es uno, la el filtro se mueve un píxel a la vez. Al ser dos el stride, esta matriz se mueve saltando 2 píxeles a la vez, a medida que se desliza sobre la matriz de entrada.
- Relleno cero. En algunas ocasiones es conveniente rellenar la matriz de entrada con ceros alrededor del borde, para aplicar el filtro a los elementos adyacentes de la matriz de entrada.



**Figura 8.** *Mapa de características*

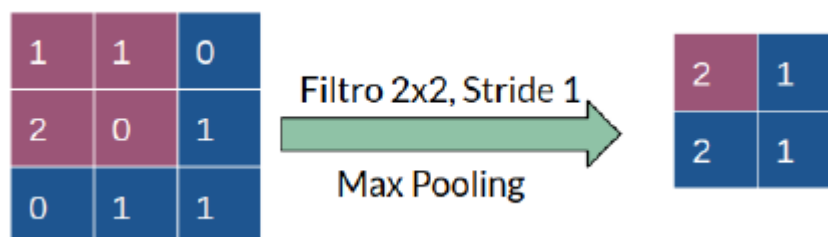
**Fuente:** (Bendemra, 2018)

Las redes neuronales convolucionales aprenden los valores de estos filtros durante el proceso de entrenamiento. Cuando más número de filtros se utilizó, más características de la imagen se extraen, mejor será el reconocimiento de patrones en la red.

Una operación adicional llamada ReLu (Unidad Lineal Rectificada) es utilizada después de cada operación de convolución, es una función de activación que se encarga de reemplazar todos los valores negativos en el mapa de características por cero.

- **Capa de reducción o pooling.** Esta capa se ubica después de la capa convolucional, se utiliza para reducir las dimensiones de cada mapa de características, sin afectar su profundidad, reteniendo la información más importante. Se utiliza diferentes maneras para realizar esta reducción:

- **Max pooling:** se divide en áreas o regiones a la matriz que contiene el mapa de características, estas áreas pueden ser de 2x2 como se ve en la figura 10, se toma el elemento más grande dentro de las áreas y se forman una matriz
- **Average pooling:** tiene la misma función que el anterior, pero en vez de tomar los valores altos para formar la nueva matriz, se realiza el promedio de los elementos que comprenden las áreas del mapa de características.

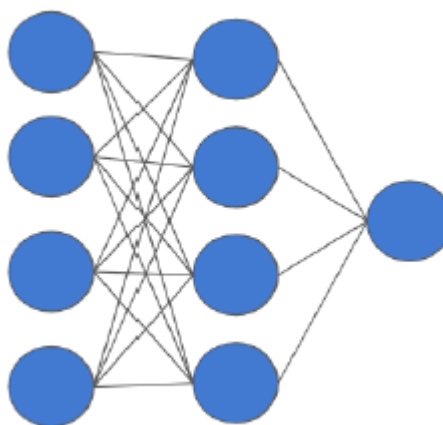


**Figura 9.** Capa de reducción o pooling

Fuente: (Calvo, Red Neuronal Convolutiva CNN, 2017)

- **Capa clasificadora totalmente conectada.** Es un perceptrón multicapa que utiliza una función softmax en la capa de salida. Al final de las capas convolutiva y de pooling, se utilizan generalmente capas completamente conectadas, esto implica que cada neurona de la capa anterior está conectada a cada neurona de la siguiente capa.

La salida de las capas convolucional y de pooling representan características de alto nivel de la imagen de entrada. El propósito de la capa totalmente conectada es utilizar estas características para clasificar la imagen de entrada en varias clases basadas en el conjunto de datos de entrenamiento



**Figura 10.** *Capa clasificadora totalmente conectada*

*Fuente:* (Perez, 2018)

La suma de las probabilidades de salida de la capa totalmente conectada es 1. Esto se garantiza utilizando el softmax como función de activación en la capa de salida de la capa totalmente conectada.

La **función softmax** toma un vector de salida con valores arbitrarios y lo reduce a un vector con valores entre uno y cero, es decir, divide cada salida de tal manera que la suma total de las salidas es igual a 1, esta salida es

equivalente a una distribución de probabilidad categórica, indica la probabilidad de que cualquiera de las clases sea verdadera.

Matemáticamente la función softmax, se muestra a continuación, donde:

$z$  es un vector de las entradas a la capa de salida

$j$  indexa las unidades de salida, así que  $j = 1, 2, 3, \dots, K$

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

Este tipo de redes se las utiliza en el “Desarrollo e implementación de una aplicación que convierta el abecedario y los números del uno al diez del lenguaje de señas a texto para ayuda de personas con discapacidad auditiva, mediante dispositivos móviles” (Sarmiento, 2016). La aplicación utiliza redes neuronales artificiales para aprender y reconocer una letra o número del lenguaje de señas, la aplicación indicara si la imagen tomada corresponde a la letra que se desea reconocer (Sarmiento, 2016).

- **Convolución separable en profundidad**

La capa convolucional mencionada en la sección g), para cada pixel de entrada, sin considerar el número de canales, genera un nuevo pixel de salida que consta de un solo canal. Porque en esta capa el filtro se desliza

a través de la imagen y en cada paso realiza una suma ponderada de los píxeles de entrada cubiertos por el filtro en todos los canales de entrada.

La convolución separable en profundidad se compone de dos capas: las convoluciones en profundidad y las convoluciones puntuales. Se utiliza convoluciones en profundidad para aplicar un solo filtro por cada canal de entrada (profundidad de entrada). Por ejemplo en una imagen de tres canales, la convolución profunda genera una salida que tienen 3 canales. Cada canal tiene su propio conjunto de pesos. Esta operación tiene como propósito filtrar los canales de entrada.

Después de realizar la convolución profunda, se realiza la convolución puntual, que es como realizar una convolución regular con un filtro de  $1 \times 1$ . En esta operación se suman todos los canales como una suma ponderada con el propósito de combinar los canales de salida de la convolución profunda para crear un mapa de características.

Al unir estas dos operaciones, el resultado se denomina convolución separable en profundidad.

Los resultados para los dos tipos de convoluciones son similares, ambos filtran los datos y crean el mapa de características, pero la convolución regular tiene que hacer mucho más tareas computacionales

para llegar a generar el mapa de características y necesita aprender más ponderaciones. Entonces las redes convolucionales separable en profundidad son más rápidas.

#### **h) Máquinas de vectores de soporte (aprendizaje supervisado)**

Es una técnica de categorización utilizada para la resolución de problemas linealmente separables, permite obtener la separación más óptima entre dos conjuntos de datos, que son casos sencillos y fáciles de solucionar. Por el contrario, los problemas linealmente no separable, los datos se superponen espacialmente siendo insostenible la dispersión por medio de un hiperplano (Vargas Jose, 2012).

Una alternativa es cambiar la frontera lineal del hiper plano por una no lineal. MVS lo hace de una manera indirecta que resulta computacionalmente eficiente: se apela a una transformación no lineal del espacio en otro de mucha mayor dimensión en el cual los dos grupos transformados quedan separados como en el primer caso, es decir por una distancia positiva. Luego, es posible proceder como el caso linealmente separable. La transformación inversa deforma dicho hiper plano en una frontera no lineal que separa a los dos grupos en el espacio original (Vargas Jose, 2012).

#### **Caso Linealmente separable**

Sea  $(X_1, Y_1), \dots, (X_1, Y_1)$  dos grupos separables por un hiperplano, donde  $X_1$  corresponde a las coordenadas espaciales de cada punto y  $Y_1$  toma los valores

{1,-1}, estableciendo los dos grupos. La ecuación del hiperplano que separa a los dos grupos tiene la forma:

$$\phi(x) = \beta'x + b_0$$

Donde  $\beta$  es un vector normal al hiperplano. Los coeficientes  $\beta$  y  $b_0$  son seleccionados para aumentar el margen  $M$ , el cual mide la distancia entre los datos de cada grupo.

Aplicación MVS otorga diagnóstico clínico de la enfermedad de párkinson y el temblor, este trabajo utiliza clasificadores binarios ajustados en las Máquinas Vectoriales para conseguir un diagnóstico diferencial entre las dos patologías (RobertoGonzález, 2017).

### **i) Regresión lineal (aprendizaje supervisado)**

La regresión es una forma estadística de establecer una relación entre una variable dependiente y un conjunto de variables independientes. En donde la variable independiente es una variable que se manipula para determinar el valor de una variable dependiente. Son las características que se usa para predecir algún valor dado de  $y$ . La variable dependiente depende de los valores de la variable independiente, la característica que se trata de predecir, que se conoce comúnmente como una variable de respuesta (Orellana, 2008).

La regresión lineal es un método para predecir la variable dependiente ( $y$ ) en función de los valores de las variables independientes ( $x$ ). Se tienen las parejas de datos  $(x,y)$  para definir una función  $f(x)$  donde los valores sean muy



cercanos a  $y$ . En el caso de la regresión lineal, la función  $f(x) = ax + b$  es la que se considera mejor para describir la relación de las variables  $x$  y  $y$ .

El algoritmo consiste en hallar una línea recta que mejor encaje en un conjunto de datos dados, este conjunto de datos comprende las variables independientes e dependientes. Para esto se puede utilizar métodos matemáticos como el de los mínimos cuadrados, para buscar minimizar la distancia vertical de todos los puntos a la línea recta (Orellana, 2008).

El proceso de pigmentación de una empresa del sector de la automoción. En un caso industrial se puede saber que la pintura, para partes automotrices, está relacionada con la cantidad de pigmentación con la que se lleva a cabo. Se desarrolla un método de predicción para estimar el contenido de pigmentación que deben de tener las pinturas para cumplir con las especificaciones de las armadoras de tal manera que el problema consiste en lograr la mejor estimación de la relación entre las variables.

#### **j) Mezcla de Gaussianas (aprendizaje no supervisado)**

Un modelo de mezcla es definido como un modelo probabilístico que representa la presencia de subpoblaciones dentro de una misma población. Estos modelos son usados para crear inferencia estadística, aproximaciones y predicciones acerca de las propiedades de las subpoblaciones a partir de las observaciones o datos adquiridos de la población estudiada sin necesidad de información que identifique a la subpoblación.

La mezcla de Gaussianas es un algoritmo basado en la distribución de probabilidad para variables continuas: la distribución Gaussiana o distribución Normal. La distribución gaussiana depende de dos parámetros: la media y la varianza. En la mezcla gaussianas se utiliza el algoritmo Expectation maximization (EM) que encuentra la máxima probabilidad estimada para los parámetros del modelo

**El algoritmo Esperanza-Maximización (EM).** Es un método ampliamente aplicado en la computación iterativa de la estimación de máxima verosimilitud (ML), útil en problemas de datos incompletos. En cada iteración del algoritmo EM, hay dos pasos llamados el paso de esperanza (expectación) o paso-E y el paso de maximización o paso-M.

Los problemas de datos con un nivel alto de complejidad son generados por el paso-E, mediante la utilización de problemas con datos fragmentados o incompletos y el valor de las medidas actuales, se produce el cálculo del Paso-M, que por la ejecución del paso anterior se vuelve más simple, al momento de aplicar este cálculo a los datos complejos. Al estar basados en estos datos, sus correspondiente esperanzas condicionales son perjudicados en al paso-E por el ajuste de los parámetros. Se debe inicial con valores bajos, para inicializar los parámetros, e iterar los pasos E y M hasta la convergencia. El inconveniente con este algoritmo es su lentitud en converger para algunos escenarios. Por este problema, se producen versiones del algoritmo que intentan disminuir esta debilidad (Jiménez, 2006).

La definición de este tipo de algoritmo es vincular un problema de datos complejos e incompletos, a otro que o tenga estos inconvenientes, para que la apreciación de máxima verosimilitud pueda tener una solución (Jiménez, 2006).

Los modelos de mezcla gaussiana tienen como objetivo encontrar una aproximación o estimación a partir de sus componentes encontrando un acomodamiento de los datos que contienen las componentes (Ramírez Aldama, 2015).

Modelo de mezcla de gaussianas generalizadas para segmentación de melanomas es un sistema de diagnóstico asistido por ordenador, en este trabajo, tras corregir la iluminación de la fotografía original mediante un filtro de transformación homomórfica, utilizamos un modelo de mezcla de Gaussianas Generalizadas, donde los parámetros de la mezcla se obtienen mediante el algoritmo EM (Expectation-Maximization). Este modelo de selección basado en las distribuciones Gaussianas generalizadas (GGD) es capaz de cubrir un amplio rango de datos debido a su versatilidad.

#### **k) Agrupamiento jerárquico (aprendizaje no supervisado)**

Los algoritmos de agrupamiento jerárquicos se utilizan para el agrupamiento de patrones de los cuales se desconoce la organización interna que tienen, es decir, no existe conocimiento acerca de la etiqueta de clase a la que pertenecen, se divide los datos sin etiquetas en grupos (clusters) de tal forma que datos que pertenecen al mismo grupo son similares, y datos que

pertenecen a otros grupos son diferentes. Para conseguirlo se puede hacer mediante las siguientes estrategias: aglomerativa y divisiva.

- Las estrategias aglomerativas parten de un conjunto de elementos individuales y van juntando los elementos que más se parezcan hasta quedarse con un número de clusters que se considere óptimo. Su complejidad computacional es del orden de  $n^3$ .
- Las estrategias divisivas parten del conjunto de elementos completos y se van separando los grupos que más diferentes sean entre ellos hasta quedarse con un número de clusters que se considere óptimo. Su complejidad computacional es del orden de  $2^n$ .

El agrupamiento jerárquico suele representarse a través de un dendograma, que muestra en qué orden se han unido los cluster y cuál es el grado de proximidad que tienen los clusters que se unen. Su estructura es la siguiente:

- Los nodos hojas del dendograma se corresponden con los elementos individuales.
- En el nodo raíz se representa el cluster al que pertenecen todos los elementos.

- El resto de nodos se corresponde con los clusters que se van formando.

Análisis del clúster jerárquico como técnica análisis exploratorio de registros múltiples en datos meteorológicos, el proyecto analizo y demostró la hipótesis de las épocas atípicas, es decir, el cambio en la media y la varianza (asociación individual) pertenecen o son del tipo homogéneo (Castro Heredia, 2012).

### I) Q-learning (aprendizaje por refuerzo)

Es un algoritmo de aprendizaje reforzado basado en el cambio de estados y la retroalimentación, su ecuación de actualización es la siguiente:

$$Q_{t+1}(s_t, a_t) \leftarrow r_t + \gamma \max_a Q_t(s_{t+1}, a)$$

El algoritmo de aprendiza por refuerzo Q-Learning actualiza sus valores mediante la estimulación del estado-acción que mejor se acople en la siguiente etapa. Esta acción realizada en el estado  $S_{t+1}$  no influye en el valor que se tomó para llegar a dicho estado. El valor de estas acciones es la suma de las recompensas del estado  $S_{t+1}$ .

La matriz Q toma estos valores, sin consideran las acciones que se tomaron, esta matriz asume que siempre se realizó una elección correcta y óptima.

La elección de acciones para algunos casos no es importante, pero por otro lado, esta elección al ser influenciado por una determinada política, se obtiene

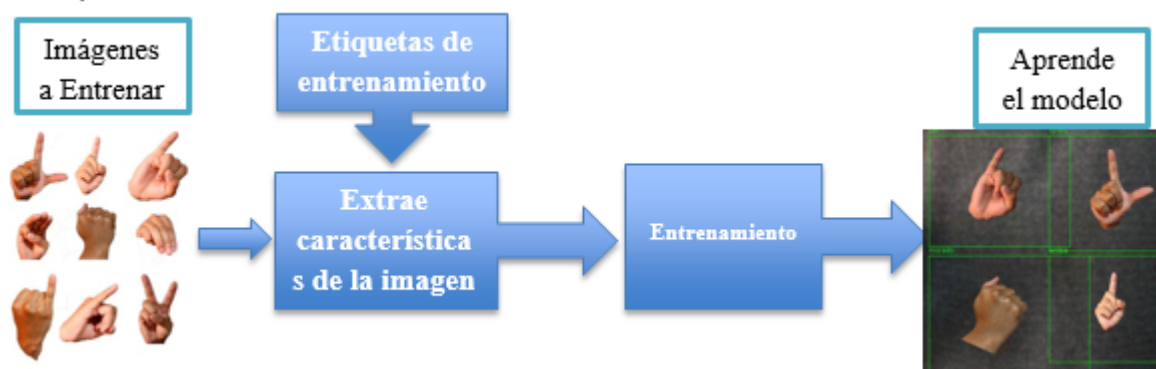
resultados que nos lleven a estados que influyen en la pérdida de la partida, como se asume la elección correcta. La política óptima es la consideración de buenas acciones a dichos estados (camino útil). Si no se tiene una política y acción óptima se podría perder la partida (Sarrió, 2015).

Q-Learning es usado para la resolución de laberintos en su más básica implementación. Uno de sus principales problemas es que es un método lento, ya que el agente requiere un tiempo considerable de entrenamiento para aprender una política aceptable (Sarrió, 2015).

Seguimiento visual en mecanismos robóticos con Q-Learning. El proceso del sistema de control inicia en la obtención de las características de significativas de la imagen. Con dichas características se obtiene un error de la posición medida con respecto a la posición deseada. Este error sirve como parámetro de entrada para el controlador Q-learning, que su función es aprender el movimiento para seguir el objetivo y controlar el robot. El ciclo de control se cierra cuando la planta controlada es modificada por el controlador Q-learning.

## **2.9. Visión por computador**

La visión por computador llamada también visión artificial es un conjunto de algoritmos o métodos computacionales que tienen el fin de originar y ordenar reglas para la mejora de la realización de distintas tareas o toma de decisiones (E.H.U, 2018)



**Figura 11.** Visión por computador

La visión por computador se basa en el sentido de la vista parecido al del ser humano y se lo toma como principal sentido, ya que es más importante y complejo que los otros, es un medio por el que el hombre interacciona con el mundo que nos rodea.

La mayoría de la información (los objetos de un entorno) procesada en el cerebro ingresa por la vista, por ejemplo si un individuo se encuentra en una situación extrema como en un incendio, la primera reacción será buscar la salida de emergencia, para ello se necesita saber la posición de una posible puerta o ventana de salida, el tamaño, la forma, la trayectoria para llegar, si existen obstáculos, etc., es por ello que todo este proceso que el cerebro humano realiza sirve para tomar acciones sobre cómo reaccionar por alertas ante posibles peligros (HUESO, 2018).

Los objetos físicos obtienen una descripción específica o detallada al ser captados por la cámara de un computador que ha sido programado para dicho fin, una de las ventajas de la visión por computador es que se lo ha implementado para la reducción de mano de obra en varias áreas, por ejemplo; existen aplicaciones que han simulado el

“sentido de la vista” en robots industriales, y otras que lo emplean para realizan trabajos como control de calidad de productos, inspecciones, clasificación, empaquetado, etc.

Aplicaciones significativas en las que se emplea visión por computador es la captura de imágenes en tiempo real, puesto que existen, por ejemplo micro-cámaras que están integradas en todos los dispositivos electrónicos inteligentes de uso cotidiano, en otros casos se conjuga la visión de la cámara del dispositivo con videojuegos en 3D, realidad virtual, realidad aumentada, u otros para de esta manera reconocer imágenes en tiempo real; en la industria automotriz la visión por computador se utiliza tanto en la fabricación y ensamblaje como en inspecciones de calidad, aunque también en los últimos años se ha venido investigando e implementando el uso de visión artificial para que sirvan en sistemas autónomos de conducción; en la alimentación realiza control de calidad de los productos en especial en fases finales; entre otras aplicaciones (para más detalle ver *Introducción a la Visión por Computador*, 2018) (INFAIMON, 2018).

En el caso de las computadoras su vista es a través de cámaras digitales integradas o externas, mismas que pueden obtener fotografías, videos o imágenes y de este modo conseguir una o varias imágenes que son almacenadas en una base de datos y que posteriormente servirá para que la computadora realice el procesamiento de la imagen y se pueda utilizar para ejecutar el reconocimiento de imágenes<sup>13</sup>.

---

<sup>13</sup> **Reconocimiento de Imagen:** El reconocimiento de imágenes es un proceso complejo que requiere una serie de pasos que sucesivamente transforman los datos icónicos a información que la computadora pueda reconocer (*Reconocimiento de Imagen*, 2018).



## 2.10. Procesamiento de imagen

Una imagen, fotografía, animación o video en su interior se representan como una matriz de números, esto se logra de acuerdo a la codificación de la combinación de los colores RGB (rojo, verde y azul) que lo compone.

Las imágenes capturadas por la cámara son procesadas, tratadas y convertidas en una matriz para extraer información importante o lo que se desee obtener de ellas, llegando así a traducir la matriz de números en objetos para utilizarlas al solucionar algún problema, este proceso es conocido también como análisis de imágenes.

- **Análisis de imágenes.-** Para llegar al reconocimiento de una imagen digital en el que una maquina pueda hacer el trabajo del ojo humano, es necesario realizar análisis de imágenes.

En el análisis de imágenes se toman en cuenta diversas características de la óptica<sup>14</sup> como son: rotación, translación, escala, movimiento del objeto y/o de la cámara, iluminación, o coexistencia con otros objetos llamados oclusiones, esto dependiendo de los grados de libertad<sup>15</sup> que se le dé al objeto, aunque también se necesita información limitada del objeto; restricción temporal, exactitud, etc.

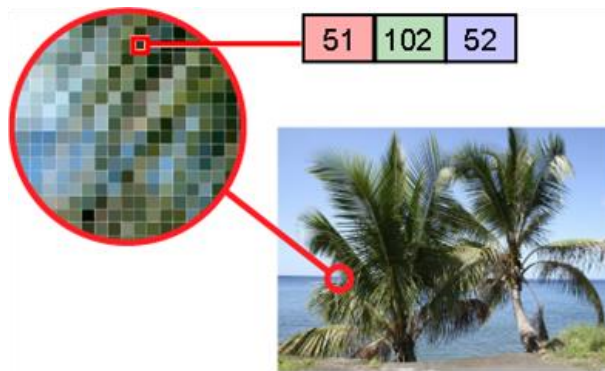
---

<sup>14</sup> **La óptica** se refiere al aparato compuesto de lentes que sirve o se utiliza para ver imágenes, escritos agrandados (Óptica, 2017).

<sup>15</sup> **Grados de Libertad.-** Expresión que dice que están dados por el número de valores que pueden ser asignados pde forma arbitraria, antes de que el resto de las variables tomen un valor automáticamente (Glosario, 2018).

en este sentido las imágenes presentarán apariencias distintas de acuerdo al punto de vista de la óptica(González, 2007)(*Introducción a la Visión por Computador*, 2018).

- **Traslación, Rotación, escala.-** Esto sucede en la imagen digital y son transformaciones que alteran los píxeles de la imagen, representados en matriz RGB o escala de grises, de acuerdo al tamaño que esta tenga, la posición y el color de cada píxel que se obtiene determinando la columna y la fila donde este se encuentre, construyendo así una matriz donde cada entrada es un vector bidimensional (x, y) que detalla la posición de los píxeles mencionados en la matriz, esto se debe a que un píxel es considerado como un punto  $P(a,b)$  en un plano x,y como se muestra en la **Figura 12** (Rodríguez A. , 2016).



**Figura 12.** Pixel de una imagen

Fuente: (UOC, 2012)

Este grupo de transformaciones recibe coordenadas de un píxel como entrada, y devuelve a la matriz las coordenadas en las que será colocado el píxel de acuerdo a la traslación, rotación o escala que se efectuó; de forma tal que, al transformar cada uno de los píxeles de una imagen digital, se obtiene una nueva imagen con el cambio deseado (Rodríguez A. , 2016).

- **Movimiento de cámara y/o de objeto.-** Esto se refiere tanto a un objeto específico dentro del encuadre de la imagen que no se encuentra de forma fija (por ejemplo un animal, personas caminando, medios de transporte, etc.), como cuando la cámara no se mantiene estática y por lo tanto el reconocimiento de la imagen se vuelve más complejo.
- **Iluminación.-** Puede ser artificial como lámparas, focos, luces, rebotadores, etc., o natural, provoca que una imagen se vean más clara o más oscura, al mismo tiempo produce sombras, modifica el contraste<sup>16</sup>, etc. Para el reconocimiento de imagen la iluminación juega un papel muy importante puesto que se necesita que el objeto a reconocer este esclarecido.

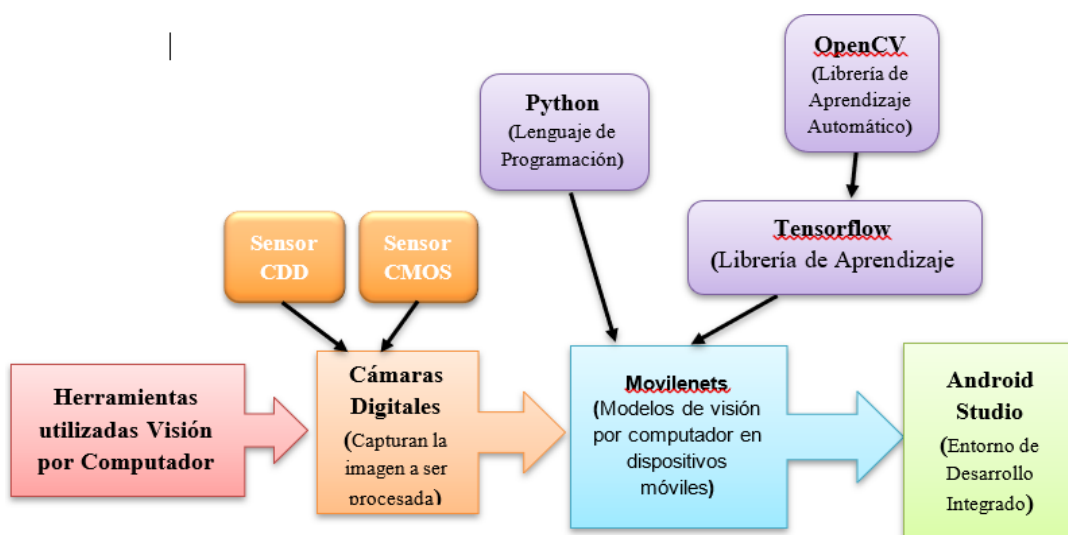
---

<sup>16</sup> **El contraste** incrementa el cambio de luminosidad entre las zonas más oscuras o más claras de una fotografía, simulando a su vez, un mejor enfoque y claridad de imagen.

- **Información limitada del objeto.-** Son las características específicas para buscar un objeto en la imagen, aunque en ciertas ocasiones es difícil explicar qué es lo que define a los objetos, es decir, no es fácil delimitar concretamente las particularidades que queremos buscar en la imagen para determinar si un objeto está presente o no en ella (*Introducción a la Visión por Computador, 2018*).
- **Restricción temporal.-** Es el tiempo de respuesta en el que el sistema logró el procesamiento de imagen capturada.
- **Exactitud.-** En todo este proceso se examinan miles de posibilidades de acierto en cada imagen y en consecuencia el sistema puede funcionar y reconocer el objeto sin problema, pero también puede suceder lo que es conocido como falsos positivos y falso negativos, en el primer caso si el algoritmo funciona 99.9% de acierto, se tendrá miles de errores, es decir, que el sistema indicará que un objeto se encuentra presente en la imagen lo cual será incorrecto, en el segundo caso por el contrario, el algoritmo indicará que no se encuentra presente el objeto indicando cuando si lo está (*Introducción a la Visión por Computador, 2018*).

## 2.11. Herramientas de Visión por Computador

En el mundo de la visión por computador, existen variedad de herramientas que permiten la realización de esta, como YOLO, librerías (opencv, tensorflow, sherlok, halcon, PCL (Point Cloud Library), etc.), también se usan lenguajes de programación, así como Python, Cython, C++, C#, java, etc., o dispositivos electrónicos como cámaras digitales, micrófonos, etc., como se ve en la **Figura 13**.



**Figura 13.** Herramientas utilizadas para la Visión por Computador

Fuente: (Tensorflow TM, 2018) (Developers Android Studio, 2018)(Introducción a la Visión por Computador, 2018)

### 2.11.1. Cámaras digitales

Las cámaras digitales son las encargadas de capturar imágenes digitales, se pueden diferenciar de acuerdo al tipo de sensor: CCD y CMOS.

- **Sensor CCD.-** Estos son sensores que se basan en elementos sensibles a la luz, que al ser activados por un reloj, crean un pozo de potencia. De

esta forma, se consigue que el valor de intensidad lumínica de cada punto del sensor, corresponda con cada píxel de la imagen.

**Sensor CMOS.-** Las cámaras con CMOS son más modernas que las cámaras CCD, En estos casos, la conversión de luz nos da un valor digital, mediante el uso de varios transistores en el propio elemento sensible. Por lo tanto, y a diferencia de las cámaras CCD, no necesitan registrar desplazamiento, ni conversor analógico-digital. Se puede acceder a cada píxel de forma individual, por ello, es frecuente que, en las cámaras de tipo CMOS, se defina lo que se conoce como una región de interés. Se consiguen así velocidades de transferencia superiores, al centrarnos en una región más pequeña de la imagen, que es sobre la que vamos a trabajar (Liarte Diego, 2018).

El objetivo, de lograr una buena óptica<sup>17</sup> de la imagen es captar los rayos luminosos y agruparlos sobre el elemento sensible de la cámara, los cuales deben tener parámetros de luminosidad, distancia focal que son los rayos paralelos que pasan por una lente convexa, convergen hacia un

---

<sup>17</sup> **Óptica.-** involucra el estudio del comportamiento y las propiedades de la luz, incluidas sus interacciones con la materia (UNIVERSAL., 1986)

punto (punto focal<sup>18</sup>). Es el parámetro principal a la hora de calcular el tamaño de los objetos en la imagen.

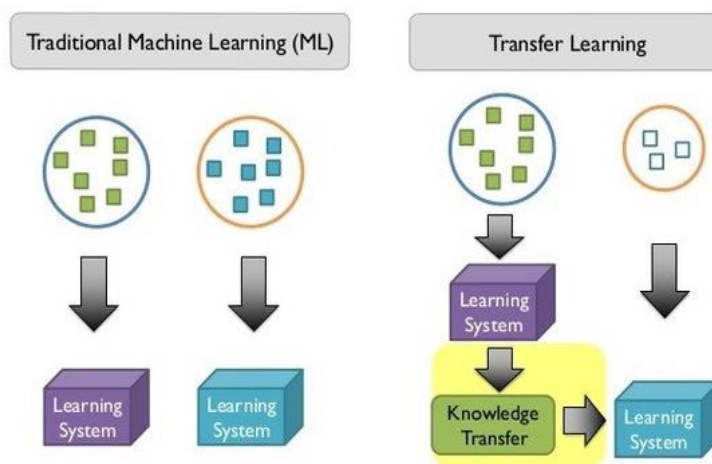
### **2.11.2. Transferencia de Aprendizaje**

Los modelos de reconocimiento de imágenes modernos tienen millones de parámetros, entrenarlos desde cero requiere una gran cantidad de datos de entrenamiento y potencia de computo (cientos de horas de CPU). El aprendizaje por transferencia es una técnica que reduce gran parte de estos costos computacionales, tomando una parte de un modelo que ya ha sido entrenado en una tarea relacionada y reutilizándolo en un nuevo modelo, como se ve en la Figura 14 Aprendizaje por Trasterencia.

---

<sup>18</sup> **Punto Focal** se denomina al espacio en el que chocan en algún punto, tras haber atravesado el objetivo que se está enfocando (Foto Nostra, 2018)

## Transfer Learning



**Figura 14** Aprendizaje por Trasterencia

Fuente: (Jain, 2018)

Tensorflow Hub es una biblioteca para fomentar la publicación, el descubrimiento y el consumo de partes reutilizables de modelos de aprendizaje automático. En particular, proporciona módulos, que son piezas previamente entrenadas e independientes de modelos de tensorflow que se pueden reutilizar en nuevas tareas, este procesos se lo conoce como aprendizaje por transferencia (tensorflow, 2019)

Sus principales ventajas son:

- Entrenar un modelo con un conjunto de datos más pequeño
- Mejorar la generalización
- Acelerar la formación del modelo



## a) MobileNets

MobileNets son redes neuronales convoluciones separables en profundidad, para construir redes ligeras y profundas, pertenecientes a la familia de modelos de visión por computadora para tecnologías móviles de Tensorflow, diseñado para maximizar efectivamente la precisión, considerando los recursos limitados de un dispositivo móvil. Este modelo es utilizado para: la clasificación, detección de objetos y segmentación de manera similar a cómo se usan otros modelos populares de gran escala, como Inception<sup>19</sup> (Howard, 2017).

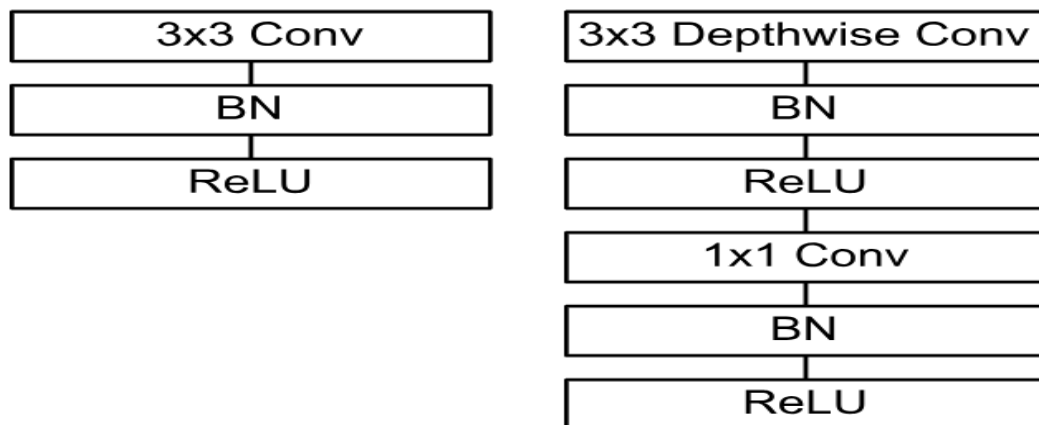
La arquitectura de MobileNet también utiliza la convolución estándar vista en la sección (Redes Neuronales Convolucionales). Pero solo en la primera capa. Todas las demás capas tienen una convolución “separable en profundidad”. La arquitectura del modelo es organizada en capas que realizan convolución en profundidad y puntual, que van seguidas de una normalización de lotes <sup>20</sup>(batchnorm) y de una función ReLU, la Figura 15 muestra esta organización de capas.

La capa final completamente conectada se utiliza la función softmax para su clasificación. Contando con las capas convoluciones profundas y puntuales como capas separadas, MobileNet consta de 30 capas.

---

<sup>19</sup> **Inception** es un modelo de aprendizaje profundo, desarrollado por Google. Tiene un alto rendimiento de clasificación de detección de objetos. (Quiza, 2018)

<sup>20</sup> **La normalización de lotes**, normaliza la salida de una capa de activación previa restando la media del lote y dividiendo por la desviación estándar del lote.



**Figura 15.** *Arquitectura MobileNets*

Fuente: (Andrew G. Howard, 2017)

## b) Tensorflow

Es una biblioteca software de aprendizaje automático de código abierto lanzada por google, tanto para la investigación como producción, desarrolladas en aplicaciones de escritorio, móvil, web y nube.

TensorFlow incluye XLA (Accelerated Linear Algebra), un potente compilador de álgebra lineal que optimiza la velocidad de ejecución del código de TensorFlow en procesadores, CPUs, GPUs, TPUs y otras plataformas de hardware, mejora el uso de la memoria y la portabilidad en servidores y aplicaciones móviles, esto se experimenta especialmente en compilaciones de tiempo de ejecución o avance del tiempo (AOT).

TensorFlow ofrece APIs de alto nivel para facilitar el desarrollo y la preparación de los modelos. Además, proporciona un nivel de control bajo para maximizar la flexibilidad y el rendimiento, este ha popularizado en el mercado siendo utilizado en empresas como twitter, ebay, snapchat, google, uber, inte, IBM, qualcomm, entre otras (tensorflow, 2019).

### **c) OpenCV**

Es una librería libre (para uso investigativo y social) hecha en lenguaje de programación C/C++ tiene interfaces C++, python y java, fue planteado para la eficacia computacional para aplicaciones de visión artificial con más 2500 algoritmos de machine learning tratándose de enfocar al tiempo real, ya que aprovecha la aceleración del hardware con procesamiento multi-core, permitiendo reconocer objetos, caras, clasificar acciones humanas, movimiento de objetos, eliminar ojos rojos de fotografías, etc (OpenCV, 2018).

OpenCV ha sido implementada por empresas como Google, Sony, Honda, Yahoo, Toyota, Microsoft, IBM, Intel, entre otras que hacen aplicaciones como monitorización de equipaje, ayuda a navegación de robots, detección de intrusos en video, inspección de etiquetas en productos, monitoreando el equipo de minas, entre otros.

**d) Python**

Se trata de un lenguaje de programación orientado a objetos, de código abierto y multiplataforma, por lo cual existe extensa documentación en la web de este, en los últimos tiempos se ha hecho muy útil para el área de inteligencia artificial ya que posee integrada el framework keras específicamente para redes neuronales es una API (Valarezo Pardo, 2017).

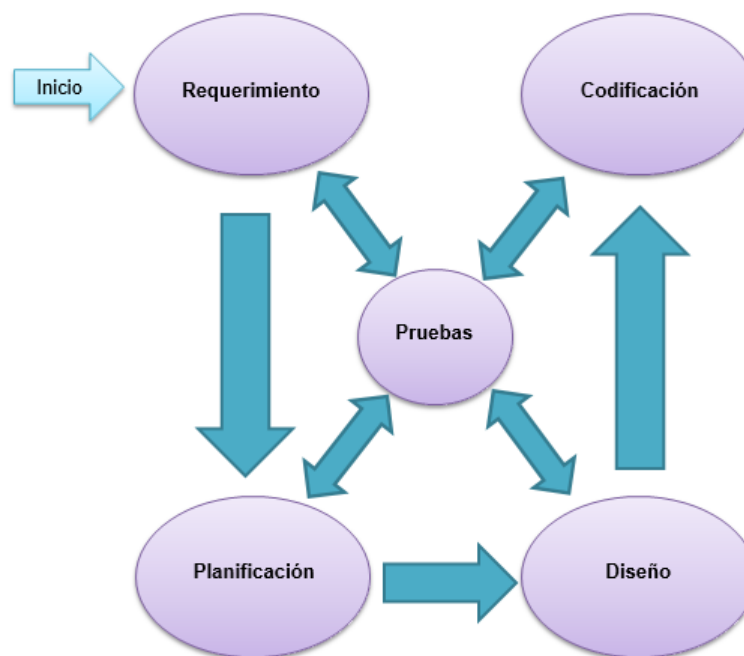
**e) Android Studio**

Entorno (IDE) gratuito para el desarrollo y creación de aplicaciones para el sistema operativo Android, se basa en IntelliJ IDEA, es decir, es un dominante editor de códigos de herramientas para desarrolladores, es robusto compatible con Windows, Linux, Mac OS, brinda funciones que amplifican el rendimiento, eficiencia y producción en la compilación de apps para Android como: sistema de compilación basado en Gradle flexible, un emulador veloz con múltiples funciones, un entorno integrado en el que se realiza desarrollos de app para todos los dispositivos Android, instant Run para emplear cambios durante el tiempo en que se ejecuta la aplicación sin que se requiera compilar un nuevo APK, contiene herramientas y framework de prueba, tiene compatibilidad con C++ (Developers, 2018).

## 2.12. Metodología MADAMDM

Es una metodología ágil orientada a desarrolladores de aplicaciones móviles, tiene una estructura obtenida de características importantes de cuatro metodologías distintas entre ellas XP (Programación Externa), OOHDM (Metodología de diseño hipertexto orientado a objetos), UWE (UML-Based Web) y MOOMH (Metodologías orientadas a objetos para la producción de software multimedia e hipertexto), uniendo estas particularidades ayudan a dar factibilidad al desarrollo de una aplicación móvil. (Jorge Adelmo Hernández Pardo Mg Tomas Antonio Vázquez Arrieta José Vicente Reyes Mozo Roberto Salas Ruiz Mariluz Romero García Miguel Ángel Leguizamón Páez Jorge Enrique Rodríguez Rodríguez, Reynaldo Becerra Correa, Alejandra Quintero, & Morales Editorial FABBECORONG, 2017)

Sus fases son requisitos, diseño, planificación, codificación y pruebas, cabe mencionar que después de terminar cada una de estas fases pasa por la fase de pruebas para determinar su correcta realización. (Ver Figura 16)



**Figura 16** Arquitectura de la metodología MADAMDM

Fuente: (GARCÍA, 2016) (Fuzi, 2013)

**Fase de Requisitos.-** Basada en las prácticas de la metodología XP, en esta fase el usuario muestra las características deseadas y necesarias para la creación de la aplicación móvil, el compromiso entre el usuario y el desarrollador debe estar enlazado robustamente, teniendo como objetivo la retroalimentación continua ya que será importante para obtener información adecuada que será analizada adquiriendo la idea clara de lo que se va a desarrollar.

**Fase de Planificación.-** Basada en la metodología UWE, en esta etapa se planifica y organiza el desarrollo de los eventos entre el usuario y la aplicación móvil, modelando y mostrando las características que tendrá al momento de codificarla, de esta fase se basa todo el proyecto.

**Fase de Diseño.-** Una vez planificada la estructura del proyecto, se empieza a diseñar toda la aplicación construyendo la conceptualización de las entidades con sus respectivos atributos, donde se puede utilizar los diferentes diagramas UML.

**Fase de Codificación.-** En esta fase y una vez analizada y realizada correctamente las fases anteriores, el desarrollador tiene la idea clara y ajustada al tipo de aplicación que va a realizar ya sea nativa (se realiza para un sistema operativo específico), web (se ejecutan dentro del navegador) o híbrida (Combinación de nativa y web). Cabe notar que esta fase es la más extensa de la metodología.

**Fase de Pruebas.-** Esta fase hace que esta metodología sea ágil en el uso de aplicaciones móviles ya que se puede llegar a esta fase desde cualquiera de las etapas de esta metodología, haciendo que el desarrollador pueda evaluar y conocer un resultado del funcionamiento de la aplicación, ya sea del diseño o su modelo para verificarlo y en caso de ser necesario corregirlo, teniendo así un resultado satisfactorio y óptimo terminada cualquier fase (GARCÍA, 2016).

## **CAPÍTULO III**

### **ANÁLISIS Y DISEÑO DE LA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL**

#### **3.1. Introducción**

En el presente capítulo se estudia el proceso de enseñanza-aprendizaje del lenguaje de señas en los estudiantes de la Unidad Educativa Camilo Gallegos, recolectando palabras simples y complejas del lenguaje, para el diseño de la aplicación móvil que permitirá fortalecer el proceso de enseñanza-aprendizaje del lenguaje de señas.

#### **3.2. Investigación y análisis del proceso enseñanza-aprendizaje del lenguaje de señas en la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.**

El modelo pedagógico implementado por Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva para la enseñanza y aprendizaje del lenguaje de señas, se enfoca en el desarrollo de destrezas visuales y memorización visual. En donde, la enseñanza y aprendizaje del lenguaje de señas se fundamenta en modelos pedagógicos, centrados en el análisis de la lengua<sup>21</sup> desde la perspectiva de la comunicación. Estos modelos otorgan una visión amplia del aprendizaje y el uso del lenguaje de señas y preparan a las personas que han

---

<sup>21</sup>Análisis de la lengua: conjunto o sistema de sonidos, gestos, formas o signos orales y escritos, que sirven para la comunicación entre las personas de una misma comunidad lingüística. (Chacoa, 2015)



nacido sordas, o sufrieron pérdida de la audición para que sean capaces de desenvolverse en la sociedad (Cesar Hernández, 2015).

El lenguaje de señas es un lenguaje visual-gestual y espacial que requiere el desarrollo previo de destrezas básicas, que las personas con discapacidad auditiva no tienen desarrollado, estas destrezas se consideran fundamentales en el aprendizaje del lenguaje de señas y está enfocado en mejorar las habilidades de los alumnos con discapacidad auditiva de primer y segundo grado de la unidad educativa y se mencionan a continuación:

- La atención y discriminación visual son tan importantes en el lenguaje de señas, ya que existen señas que se realizan de forma idéntica, pero se diferencian en un solo aspecto, como puede ser la expresión facial o la configuración de la mano. Por eso es necesario que los alumnos con discapacidad auditiva aprendan a centrar la atención visual en brazos, manos, expresiones faciales y movimientos del cuerpo.
- Para aprender el lenguaje de señas se hace necesario que el alumno memorice cómo se realizan las señas o expresiones faciales, es decir, aquí trabaja la memoria visual.

- La expresión facial es fundamental en el lenguaje de señas puesto que ésta ofrece mucha información sobre el mensaje que se quiere transmitir.
- Además, la percepción y el uso correcto del espacio es un recurso gramatical fundamental.
- El modelo pedagógico que maneja la unidad educativa, establece cuatro ideas básicas para la enseñanza del lenguaje de señas para los alumnos:
- Comprensión del lenguaje de señas. Son actividades receptoras orientadas a la comprensión, los alumnos comienzan a entender gestos y movimientos simples del lenguaje de señas.
- Es necesario tener un intervalo de tiempo entre el lenguaje recibido y la producción, es decir, que no podemos pedir al alumno produzca todo aquello que entiende, debemos dejar transcurrir un tiempo para que el alumno pueda procesar el lenguaje recibido.
- Motivación y variables afectivas. Debemos crear un clima adecuado para el aprendizaje, de tal modo que el alumno esté motivado

- Práctica del lenguaje de señas. El trabajo en parejas y en grupos es el medio más adecuado para adquirir conocimiento y destrezas comunicativas del lenguaje de señas (Santiago Prado, 2003).

Desarrollo de habilidades en los alumnos para el proceso de enseñanza-aprendizaje del lenguaje de señas. Dentro de las cuales podemos mencionar:

- Desarrollar la expresión facial
- Aprender a controlar los movimientos del cuerpo
- Trabajar la simetría, la coordinación de movimientos.
- Tomar conciencia del espacio
- Aprender a “narrar” una secuencia de acontecimientos con gestos.
- Ubicar varios objetos o personas en el espacio al mismo tiempo.
- Mejorar la agilidad y la precisión en el movimiento de las manos a través de la práctica del Alfabeto dactilológico<sup>22</sup>.
- Mejorar la agilidad de dedos y manos de los números ordinales.

---

<sup>22</sup> **Alfabeto dactilológico**, medio de comunicación utilizado por las personas con discapacidad auditiva severa, es un tipo de alfabeto, en el cual las letras del abecedario son representadas manualmente (USEFEDORA, 2017).

Los docentes de la unidad educativa aplican tres principios metodológicos en el proceso de enseñanza y aprendizaje del lenguaje de señas, los cuales son:

- Los estudiantes deben tener el deseo y la oportunidad de comunicarse mediante el lenguaje de señas. Siendo responsabilidad del profesor transformar la simple curiosidad en deseo de comunicarse
- Selección de los elementos de motivación. Estos elementos se acomodarán a las condiciones de los alumnos y su entorno social y son los siguientes: diálogos, ejercicios que desarrollen habilidades, actividades comunicativas y juegos. Son esenciales para aprender el lenguaje de señas. La práctica y la repetición ayudan y son necesarias.
- Aumentar progresivamente el nivel de complejidad; ir de lo más fácil a lo más difícil, de lo más simple a lo más complejo, de lo conocido a lo desconocido.

Las estrategias de enseñanza empleadas por los docentes de la unidad educativa, son los procedimientos y recursos que promueven la enseñanza y el aprendizaje del lenguaje de señas en la formación básica de un niño con discapacidad auditiva. Existen momentos en donde se aplican estas estrategias: pre-instruccionales (antes), co-instrucciones (durante) y pos-instruccionales (después) (Chacoa, 2015).

- Pre-instruccionales (antes) preparan al alumno para preguntas y como va aprender.

- Co-instrucciones (durante), estas apoyan los contenidos curriculares durante el proceso de la enseñanza.
- Pos-instruccionales (después); permite al alumno valorar su propio aprendizaje.

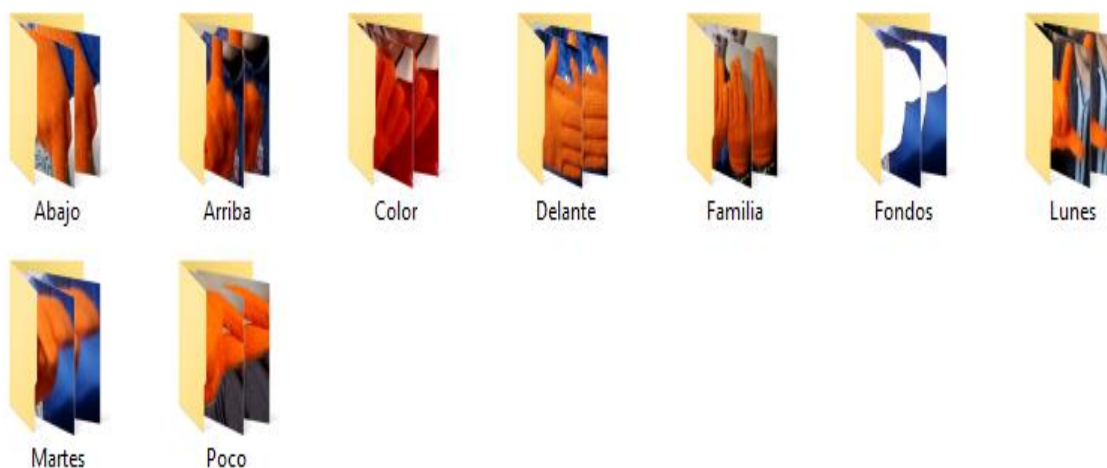
En donde, los docentes promueven la creación de ambientes de aprendizaje con orientación visual, el cual será desarrollado considerando los aprendizajes esperados, el contexto del alumno, los materiales educativos ya sean impresos o digitales.

### **3.3. Recolección de palabras del lenguaje de señas para la formación del dataset**

Para la recolección y clasificación de palabras del lenguaje de señas(dataset) se estableció en forma conjunta (tesistas y docentes de la Unidad Educativa Camilo Gallegos) la elección de letras y palabras simples y complejas que permita establecer un amplio vocabulario, inicialmente se contó con 30 palabras como nociones, colores, días de la semana y familia, con las que las personas discapacitadas auditivas comienzan su educación poder comunicarse, se toman en cuenta principalmente estas palabras porque son fáciles de aprender y realizar, las cuales son estáticas, diferentes al resto de palabras y lentas en su movimiento.

Una vez elegidos las palabras y letras que van a formar parte del dataset se realizaron fotografías a través de una cámara Cannon 70D con: sensor CMOS de

20,2 megapíxeles, disparos a alta velocidad, tecnología dual pixel CMOS AF; etiquetando las palabras del lenguaje de señas, conformando un dataset inicialmente de 90 imágenes por seña, debidamente clasificadas o agrupadas en carpetas especificadas con el nombre de cada seña como se muestra en la Figura 17.



**Figura 17** Imágenes de señas contenidas por carpetas con sus respectivos nombres

Las imágenes obtenidas tienen dimensiones superiores a 300 píxeles para posteriormente ser tratadas o procesadas, descritas en la sección 4.2 Análisis y procesamiento de imágenes.

Además se realizó un video con la Sra. Rectora de la Unidad Educativa Especializada Camilo Gallegos con las palabras elegidas del cual, se extrajeron pequeñas secuencias de este para posteriormente situarlas en la aplicación móvil como una guía para que el practicante sepa cómo realizar la seña, de la misma forma se extrajeron imágenes con el abecedario del lenguaje de señas.

### 3.4. Modelamiento del conocimiento y estructura del modelo a usar

Para el modelamiento se basó en la metodología MADAMDM descrita en la sección 2.12, por ello se recolectó los requerimientos pertinentes para la realización de la aplicación y se analizó las funcionalidades del sistema través de casos de uso y diagramas de secuencia para un mejor entendimiento de la aplicación.

#### 3.4.1. Análisis y Diseño

##### Análisis Requerimientos del Proyecto

En este punto se especifican los requisitos funcionales, los no funcionales, es decir, requisitos que permiten darle funcionalidad al sistema:

##### Requerimientos Específicos Funcionales

###### Tabla 1

*Requisito funcional del sistema. Consultas del abecedario*

Identificación del requerimiento:	<b>RF01</b>
Nombre del Requerimiento:	Consulta del abecedario del lenguaje de señas.
Descripción del requerimiento:	La aplicación móvil permitirá al usuario visualizar el abecedario del lenguaje de señas (ver tabla 11), en donde cada letra del abecedario estará representada por dos imágenes, las cuales son:  Representación simbólica (letra)

**CONTINÚA** 

---

Representación manual (gesto con la mano)

**Prioridad del requerimiento:**

**Alta**

**Tabla 2**

*Requisito funcional del sistema. Consultas del vocabulario*

Identificación del requerimiento:	<b>RF02</b>
Nombre del Requerimiento:	Consulta del vocabulario del lenguaje de señas.
Descripción del requerimiento:	<p>La aplicación móvil permitirá al usuario visualizar un listado de 50 palabras que corresponden al vocabulario del lenguaje de señas, (ver tabla 8), cada ítem del listado mostrara:</p> <ul style="list-style-type: none"> <li>• La palabra</li> <li>• Una imagen que simbolice a la palabra</li> <li>• La imagen de una letra del abecedario, con la que empieza esa la palabra.</li> </ul>

Prioridad del requerimiento:

**Alta**



**Tabla 3***Requisito funcional del sistema. Selección y consulta de datos*

Identificación del requerimiento:	<b>RF03</b>
Nombre del Requerimiento:	Seleccionar y consultar datos
Descripción del requerimiento:	<p>La aplicación móvil permitirá al usuario seleccionar cualquier palabra del listado de palabras del lenguaje de señas (ver tabla 9), que muestra archivos multimedia sobre:</p> <ul style="list-style-type: none"> <li>• Los gestos faciales y movimientos corporales de la palabra del lenguaje de señas mostradas mediante archivo gif.</li> <li>• Representación de la palabra mediante imagen</li> <li>• La descripción de la palabra(etiquetas)</li> </ul> <p>La pantalla también tendrá los botones: practicar el gesto, volver al listado, inicio y ayuda</p>
Prioridad del requerimiento:	<b>Alta</b>

**Tabla 4**

*Requisito funcional del sistema. Reconocimiento de los gestos del lenguaje de señas*

Identificación del requerimiento:	<b>RF04</b>
Nombre del Requerimiento:	Detección y Reconocimiento de los gestos manuales del lenguaje de señas
Descripción del requerimiento:	<p>La aplicación móvil contará con la opción de práctica de gestos, mostrado en la pantalla de contenido multimedia (requisito RF03). Para realizar el reconocimiento de gestos se utilizará la cámara frontal del dispositivo móvil y estará enfocado en los siguientes parámetros:</p> <ul style="list-style-type: none"> <li>• La configuración manual de la seña (movimiento de las manos)</li> <li>• La distancia de las manos con respecto a la cámara frontal del dispositivo</li> </ul> <p>La cámara del dispositivo comenzara a detectar los movimientos, que se realicen con las manos y reconocerá en un intervalo de tiempo si dicho movimiento es parecido al que fue previamente mostrado en la pantalla de contenido multimedia. Dependiendo de la respuesta que</p>

se consigue al realizar el reconocimiento de la seña (configuración manual), la aplicación mostrar una pantalla en la cual se evalué este reconocimiento (requisito RF05).

Prioridad del requerimiento:

**Alta**

### **Tabla 5**

*Requisito funcional del sistema. Validación del reconocimiento del gesto*

Identificación del **RF05**

requerimiento:

Nombre del Validación de la precisión del reconocimiento

Requerimiento:

Descripción del requerimiento: La aplicación móvil permitirá al usuario visualizar en una pantalla si el movimiento fue realizado correctamente o no, mediante la precisión del reconocimiento. Para esto se utilizaran las siguientes imágenes:

- Una X si no se realizó correctamente el movimiento y por el contrario de un visto si se lo realizo correctamente.
- La imagen de la seña.

Prioridad del requerimiento:

**CONTINÚA** 

**Alta**

**Tabla 6**

*Requisito funcional del sistema. Ayuda del sistema*

Identificación del requerimiento:	<b>RF06</b>
Nombre del Requerimiento:	Ayuda en el uso de la aplicación móvil
Descripción del requerimiento:	La aplicación móvil presentara una pantalla de soporte a los usuarios del sistema, con la finalidad de conocer el funcionamiento de la aplicación móvil a través de imágenes sobre: el abecedario, vocabulario y el proceso de reconocimiento de un gesto manual de una seña del lenguaje de señas.
Prioridad del requerimiento:	<b>Alta</b>

## Requerimientos No Funcionales

**Tabla 7**

*Requisito no funcional del sistema. Rendimiento de la aplicación móvil*

Identificación del **RNF01**

requerimiento:

Nombre del Rendimiento de la aplicación móvil

Requerimiento:

Descripción del requerimiento: La aplicación móvil garantiza al usuario que el diseño de las pantallas u otros procesos no afecte el desempeño de la base de datos, ni el proceso de reconocimiento y validación de la configuración manual de una determinada seña perteneciente al lenguaje de señas.

Prioridad del requerimiento:

**Alta**

**Tabla 8**

*Requisito no funcional del sistema. Interfaz de la aplicación móvil*

Identificación del **RNF02**

requerimiento:

Nombre del Interfaz de la aplicación móvil

Requerimiento:

**CONTINÚA** 

---

Descripción del requerimiento: La aplicación móvil presentara una interfaz de usuario sencilla para que sea fácil manejo a los usuarios del sistema. Siendo intuitiva, sencilla y manejable.

Prioridad del requerimiento: **Alta**

---

### Requisitos interfaces

- **Interfaces de usuario**

La interfaz con el usuario consistirá en un conjunto de pantallas con menús, botones, listas y campos de textos. Ésta deberá ser construida específicamente para el sistema propuesto y, será visualizada desde un dispositivo móvil, a continuación se detalla cada una de estas:

#### Tabla 9

*Requisito interfaz de usuario. Pantalla inicio*

---

Identificación del requerimiento:	<b>RIU01</b>
Nombre del Requerimiento:	Pantalla inicio
Descripción del requerimiento:	La aplicación móvil tendrá cuatro iconos en su pantalla principal, con sus respectivas etiquetas (ver tabla6), para acceder a las siguientes pantallas: <ul style="list-style-type: none"> <li>• Inicio</li> </ul>

---

**CONTINÚA** 

- Vocabulario
- Abecedario
- Ayuda

Prioridad del requerimiento:

**Alta**

### Tabla 10

*Requisito interfaz de usuario. Menú deslizable*

Identificación del **RIU02**

requerimiento:

Nombre del Menú deslizable.

Requerimiento:

Descripción del requerimiento: La aplicación móvil tendrá un menú deslizable en la parte izquierda de la pantalla del dispositivo (ver tabla 7), para acceder desde cualquier parte de la aplicación a las siguientes pantallas:

- Inicio
- Vocabulario
- Abecedario
- Ayuda

Prioridad del requerimiento:

**Alta**

- **Interfaces de hardware**

Se utiliza un dispositivo móvil que tiene las siguientes características:

- Dispositivo móvil compatible con Android 6.0
- Memoria 2 GB de RAM - Mínimo
- Almacenamiento interno 1 GB Mínimo
- Procesador Quad-Core 2.5 GHz

- **Interfaces de software**

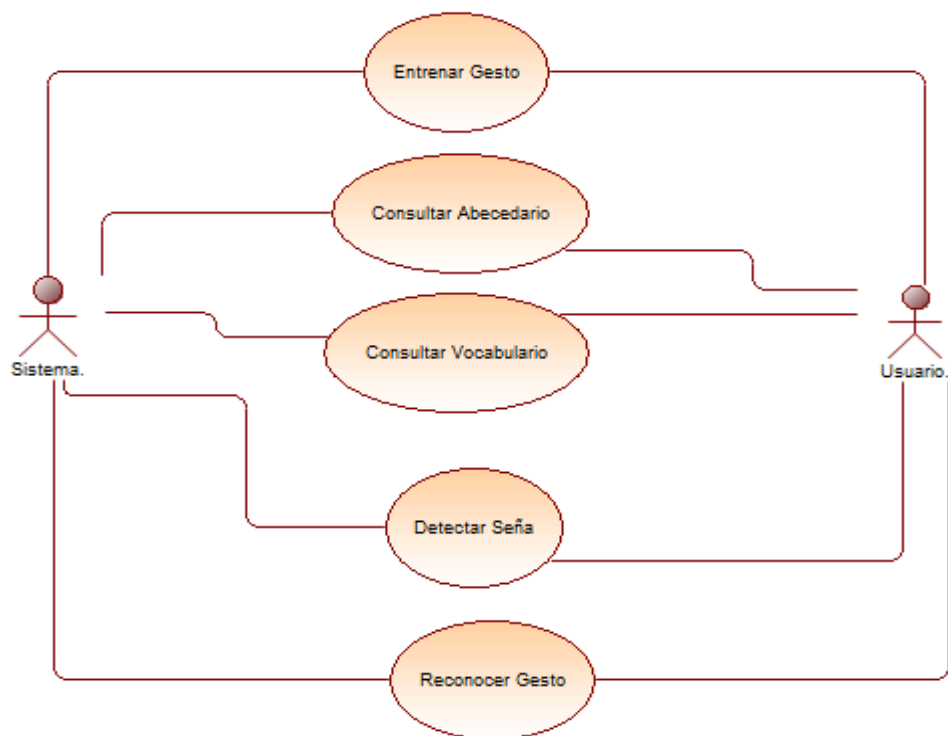
Sistema operativo: Android 6.0 o superior.

### **3.5. Diagrama de casos de uso**

#### **Caso de uso general**

Muestra la generalidad del desenvolvimiento sobre el uso que tienen los actores hacia el sistema:





**Figura 18** Diagrama de caso de uso general

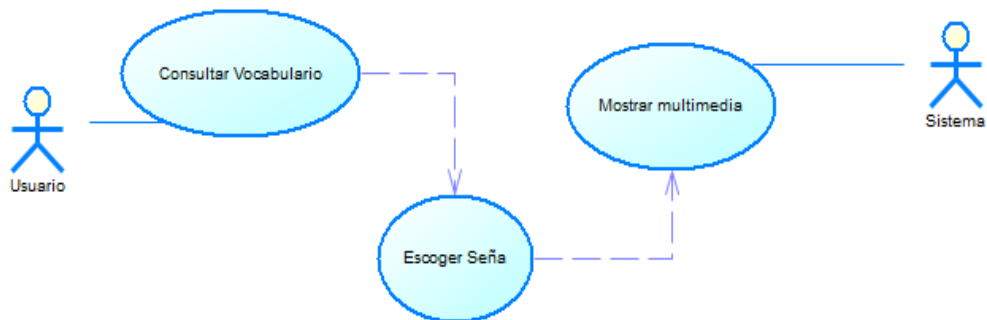
**Tabla 11**

*Caso de uso general*

<b>Caso de Uso General</b>	
<b>Caso de uso</b>	General
<b>Actores</b>	Usuario, Sistema
<b>Objetivo</b>	Realizar acciones en el sistema que se desee para interactuar.
<b>Pre-Requisito</b>	<ul style="list-style-type: none"> <li>El usuario ingresa a la aplicación</li> </ul>
<b>Curso Normal de los Eventos</b>	

- 
- **El usuario ingresa a la aplicación**
  - **El sistema permitirá realizar las siguientes acciones:**
  - **Consultar el vocabulario:** contará con un listado de las palabras del vocabulario.
  - **Consultar abecedario:** contará con el abecedario del lenguaje de señas
  - **Detectar seña:** se podrá detectar la seña una vez realizada por el usuario.
  - **Reconocer gesto:** se podrá reconocer el gesto una vez detectada y verificada por el sistema
- 

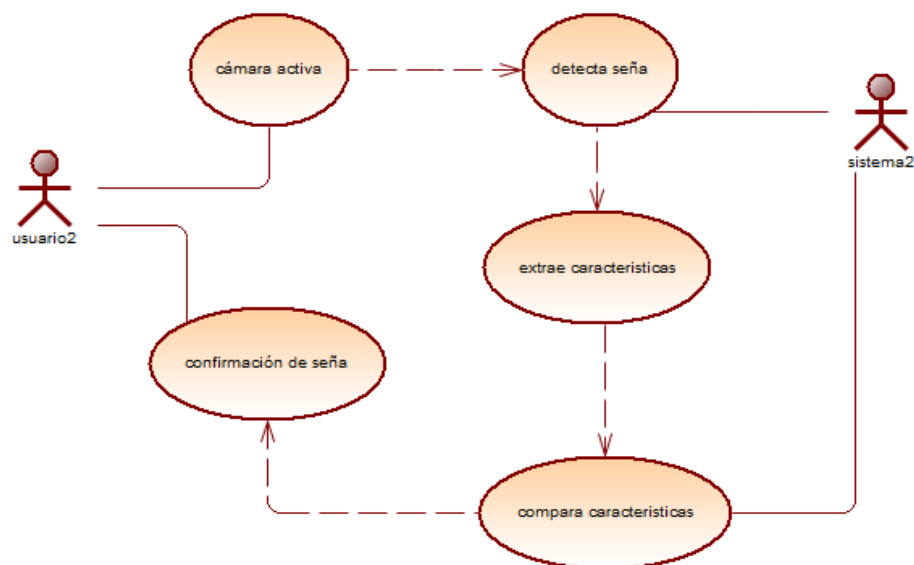
### Casos de uso específicos



**Figura 19** Diagrama de caso de uso: Escoge seña

**Tabla 12***Especificación de caso de uso: Escoger Señal*

<b>Caso de Uso Escoger Señal</b>	
<b>Caso de uso</b>	Escoger señal
<b>Actores</b>	Usuario
<b>Objetivo</b>	Seleccionar una de las señales de la lista del vocabulario que se desee aprender.
<b>Pre-Requisito</b>	<ul style="list-style-type: none"> <li>El usuario ingresa a la aplicación en la sección vocabulario</li> </ul>
<b>Curso Normal de los Eventos</b>	
<ul style="list-style-type: none"> <li><b>El usuario ingresa a la aplicación</b></li> <li><b>El usuario consulta el vocabulario</b></li> <li><b>El sistema muestra un listado de las señales que tiene en su base de datos</b></li> <li><b>El usuario escoge una señal</b></li> <li><b>Se muestra contenido multimedia</b></li> </ul>	



**Figura 20** Diagrama de caso de uso: Reconocer seña

**Tabla 13**

*Diagrama de Casos de uso: Reconocer seña*

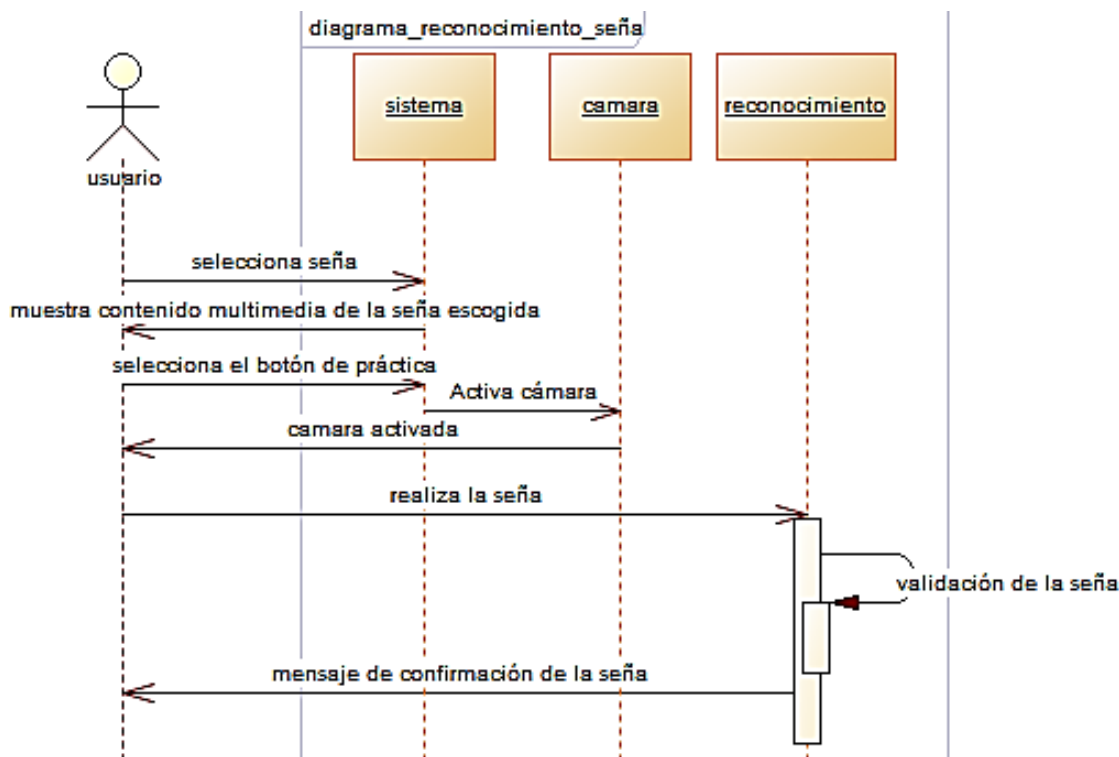
Caso de Uso Reconocer Seña	
<b>Caso de uso</b>	Reconocer seña
<b>Actores</b>	Sistema
<b>Objetivo</b>	Detecta y Reconoce la seña para posteriormente identificar su etiqueta
<b>Pre-Requisito</b>	-El usuario selecciona reconocer gesto. - Realizar Seña
<b>Curso Normal de los Eventos</b>	
<ul style="list-style-type: none"> <li>• <b>El usuario ingresa a la aplicación</b></li> <li>• <b>El usuario consulta el vocabulario</b></li> </ul>	

**CONTINÚA** 

- 
- **El sistema muestra un listado de las señas que tiene en su base de datos**
  - **El usuario escoge una seña**
  - **El sistema muestra el contenido multimedia**
  - **Se activa la cámara**
  - **Se realiza la seña**
  - **El sistema detecta la seña**
  - **Se extrae características**
  - **Se Valida y compara las características**
  - **Se Confirmar y etiqueta la seña**
- 

### **Diagrama de Secuencia**

En esta sección, se muestra la interacción entre el usuario y el reconocimiento de señas con la aplicación móvil, es decir, el orden de comportamiento de la aplicación conforme el usuario utilice el reconocimiento.



**Figura 21.** Diagrama de Secuencia: Reconocimiento de seña

Como se muestra en la *Figura 21*. Diagrama de Secuencia: Reconocimiento de seña el usuario selecciona la seña el sistema le muestra contenido multimedia de la seña escogida misma que una vez aprendida por el usuario seleccionará el botón de práctica mismo que, activa la cámara donde, el usuario posteriormente realiza la seña aprendida frente a la cámara y el reconocimiento valida la seña enviando al usuario un mensaje de confirmación de la misma.

## **CAPÍTULO IV**

### **IMPLEMENTACIÓN DE LA APLICACIÓN MÓVIL DE VISIÓN ARTIFICIAL**

#### **4.1. Introducción del capítulo**

En el presente capítulo se realiza el análisis y procesamiento de imágenes para el reconocimiento de señas, utilizando el modelo de red convolucional Mobilenet v2, (ver sección a) MobileNets), en donde se describe el algoritmo para la recolección del dataset (procesamiento de imágenes), entrenamiento y ajuste de la red neuronal para mejorar la precisión en el reconocimiento de la configuración manual por parte del modelo predictivo.

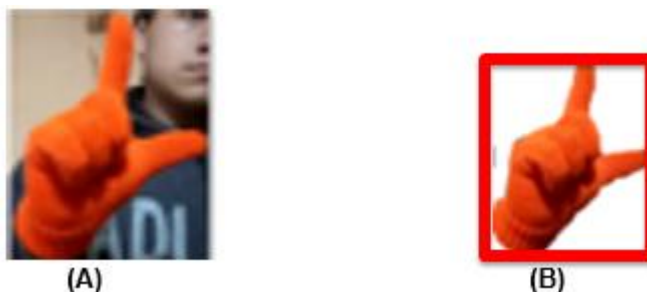
#### **4.2. Análisis y procesamiento de imágenes**

Con la recolección de las imágenes debidamente clasificadas (ver sección 3.3 Recolección de palabras del lenguaje de señas para la formación del dataset) se da inicio al procesamiento de imágenes.

El dataset de la red contó con 90 imágenes de señas (sin guantes), mismas que no dieron buenos resultados en la etapa de pruebas, por ello se incrementó el número de imágenes a 250 por cada palabra, así como se propuso el uso de guantes de lana tomatos, todo esto para la mejora de precisión del reconocimiento, como se puede observar en la sección 4.5 Integración de la red neuronal artificial

con la aplicación móvil, en donde se comprobó que con un dataset menor a 200 imágenes, el modelo predictivo no da los resultados esperados.

Para precisar el modelo en el reconocimiento de la configuración manual, al grupo de 250 imágenes capturadas de cada seña se dividió en dos grupos, la primera mitad se les extrajo el fondo y la otra mitad se le dejó con fondo. El proceso de eliminación de fondos se realizó mediante un programa desarrollado en Python con la librería de openCV inRange, en donde se implanta filtros para que solo se detecte el color de los guantes que estén en el rango de valores mínimos y máximos de HSV (sus siglas en inglés Hue saturation value), una transformación no lineal del espacio RGB, del naranja que es nuestra área de interés, se corta la forma de la seña haciéndola una máscara quitando todo el fondo y colocando el color negro o blanco al contorno de la seña, ya que estos colores son neutrales y no difieren en el área de interés antes mencionado, como se muestra en la *Figura 22*.



**Figura 22** La Imagen (A) es la imagen original y a la derecha (B) la imagen procesada (Palabra lunes)










En la Tabla 14 se muestra un pequeño ejemplo de las palabras que van a hacer el reconocimiento en las cuales se tiene en la primera fila las categorías de las

















señas, en la tercera fila las imágenes sin fondo que en algunas señas como martes, abajo o delante no se le dejó el fondo para verificar con cuál de las dos formas (fondo o sin fondo) es más favorable para el reconocimiento y en la cuarta fila las imágenes con fondo variado.

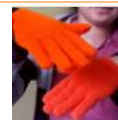
**Tabla 14**

*Palabras con sus respectivas imágenes del lenguaje de señas: se ha clasificado las imágenes de cada seña como con fondo y sin fondo para probar la predicción del reconocimiento*

Categoría	Seña	Imagen Sin Fondo	Imagen Con Fondo
Colores	<b>Color</b>		 
	<b>Rojo</b>		
Días de la semana	<b>Lunes</b>		
	<b>Martes</b>	<b>No Tiene</b>	 

CONTINÚA 

	<b>Viernes</b>		
			
<b>Nociones</b>	<b>Arriba</b>		
			
	<b>Abajo</b>	<b>No Tiene</b>	
			
	<b>Casa</b>		
			
	<b>Delante</b>	<b>No Tiene</b>	
			
	<b>Encima</b>		
			



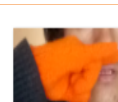
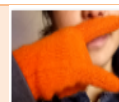
**Familia**



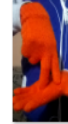
**Mama**



**Niño**



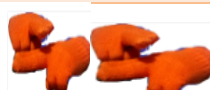
**Pararse**



**Poco**








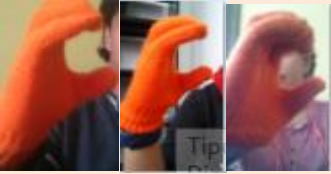



**Sentarse**

















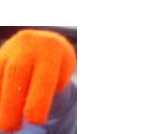





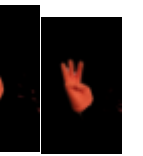
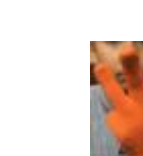




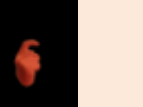

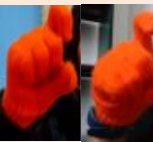



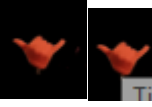





De la misma forma para el abecedario en la Tabla 15 se muestran la categoría de la palabra, la seña, la imagen sin fondo y las imágenes con fondos variados.

**Tabla 15**

*Letra con su respectiva imagen del lenguaje de señas: se ha clasificado las imágenes de cada seña como con fondo y sin fondo para probar cuál de las dos formas es favorable al reconocimiento*

Categoría	Seña	Imagen Sin Fondo	Imagen Con Fondo
Abecedario	<b>A</b>		
	<b>B</b>		
	<b>C</b>		
	<b>CH</b>	<b>NO TIENE</b>	
	<b>F</b>	<b>NO TIENE</b>	
	<b>G</b>	<b>NO TIENE</b>	

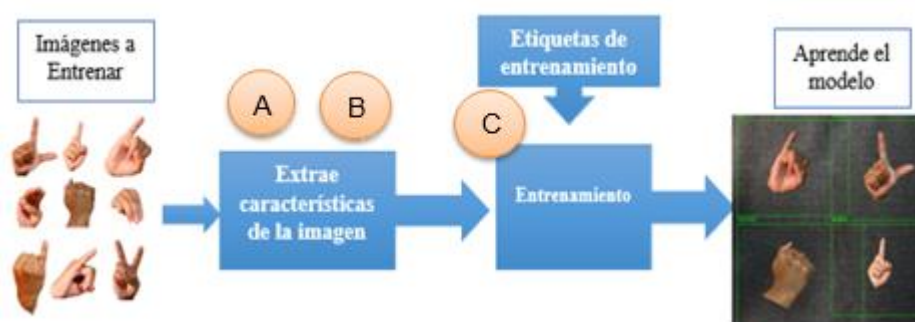
CONTINÚA 

		
<b>H</b>	<b>NO TIENE</b>	   <small>Tipo de eleme</small>
<b>K</b>		    
<b>L</b>	<b>NO TIENE</b>	  
<b>M</b>	<b>NO TIENE</b>	   
<b>U</b>	<b>NO TIENE</b>	  
<b>W</b>		     
<b>X</b>		       <small>Tipo</small>
<b>Y</b>		     

### 4.3. Programación de la red neuronal artificial y entrenamiento con el dataset de las imágenes.

Para la programación de la red neuronal y su entrenamiento con el dataset de imágenes, se utiliza el modelo MobileNet, publicado en la biblioteca de Tensorflow Hub<sup>23</sup>. En particular se emplea los módulos de extracción de características de las imágenes del dataset, para facilitar la construcción de la red y teniendo la particularidad de reducir el tiempo de entrenamiento, este proceso de desarrollo cuenta con los siguientes pasos (Figura 23):

- A. Preparación del modelo y extracción de las características de las imágenes de entrenamiento en archivos planos (bottleneck).
- B. Desarrollo de la capa de entrenamiento de la red.
- C. Entrenamiento del modelo (red neuronal).
- D. Pruebas al modelo.



**Figura 23.** Programación de la red neuronal artificial

<sup>23</sup> Tensorflow Hub, es una biblioteca para la publicación, descubrimiento y consumo de partes reutilizables de modelos de aprendizaje automático

En cada uno de estos pasos se describe de forma parcial el algoritmo utilizado para implantar la red neuronal

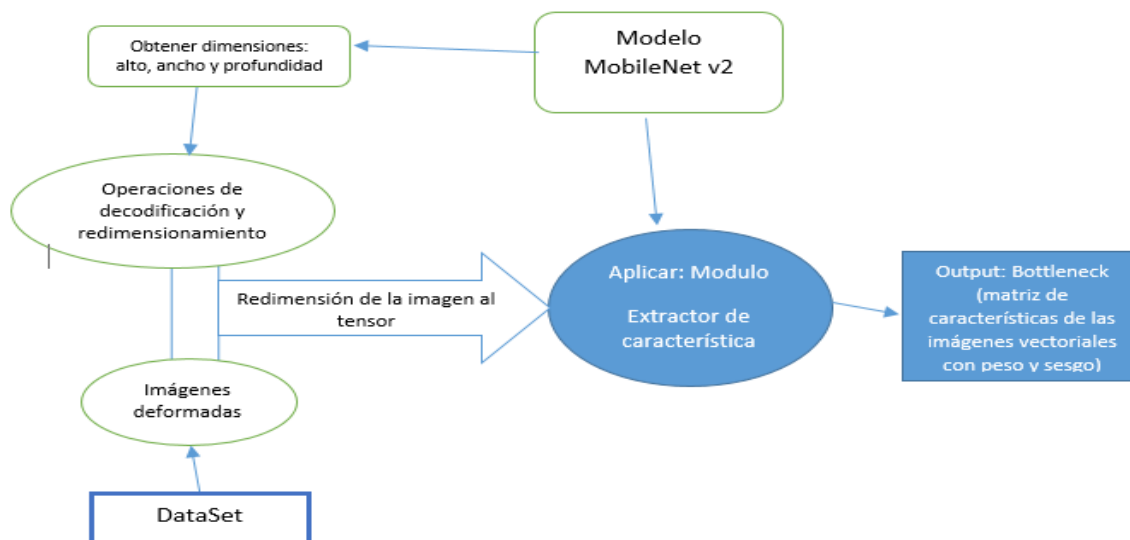
#### **A. Preparación del modelo y extracción de las características de las imágenes de entrenamiento en archivos planos (bottleneck)**

La preparación comienza por el desarrollo de la función *“create\_module”*, donde se instancia funciones que se encargan de importar y carga el modelo MobileNet, desde la biblioteca de Tensorflow-Hub, para realizar las operaciones de extracción de características de las imágenes de entrenamiento y almacenar estos datos en archivos planos denominados bottleneck, como se presenta en la Figura 24.

Con el objetivo de conseguir estos archivos, primero se debe redimensionar las imágenes, con el fin de ajustarse a los parámetros de entrada del modelo. Para realizar estas operaciones se desarrolló la función *“add\_img\_decoding”*, que instancia las siguientes funciones:

- Decodificación de imagen, el modelo trabaja con el formato uint8, (enteros positivo de 8 bits), que es la representación de una imagen en escala de grises de un solo canal (uniovied, 2012). Este proceso transforma imágenes con formato jpg y png al formato uint8.
- Redimensionamiento del alto y ancho de las imágenes en escala de grises y asignarlo al tensor de imagen decodificado, para posteriormente ser, enviado al grafo de operaciones computacionales, el cual realizara la extracción de características de estas imágenes.

Al realizar las operaciones de decodificación y redimensionamiento se agrega imágenes deformadas (recortar, escalar, ajustar el brillo y voltear las imágenes), esto permite simular situaciones más reales, mejorando el modelo predictivo.

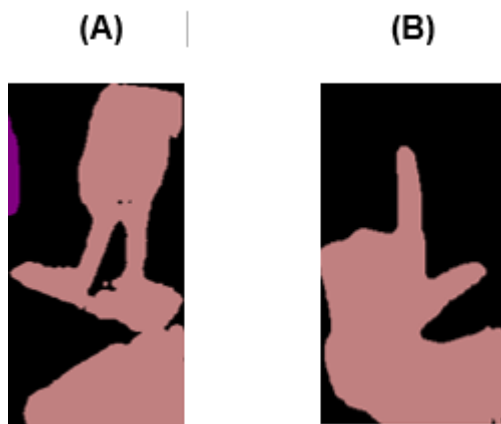


**Figura 24.** Preparación del modelo y extracción de las características

Al ejecutar las operaciones de: importación, carga y de extracción de características para preparar el modelo MobileNet, se crean los archivos planos denominados Bottleneck, a partir del dataset, en el directorio *bottlenecks*, por carpetas etiquetadas con el nombre que fueron clasificados los datos de entrenamiento. Estos archivos planos contienen matrices con decimales, que son una recopilación de la información importante (pesos), en donde toman mayor importancia a los valores mayores a uno ya que reflejan la intensidad de los píxeles, permitiendo obtener la forma de la seña, como por ejemplo de los archivos bottleneck pararse y lunes, se puede visualizar la forma de



la configuración manual (ver Figura 25. *Resultados obtenidos por archivos planos denominados Bottleneck de la seña pararse(A) y lunes (B)*)



**Figura 25.** *Resultados obtenidos por archivos planos denominados Bottleneck de la seña pararse(A) y lunes (B)*

Para realizar las operaciones de extracción de características de las imágenes de entrenamiento (7.500 por las 15 palabras y 15 letras del abecedario del lenguaje de señas, vistas en la tabla 13), se desarrolló los siguientes métodos: “*get\_or\_create\_bottleneck*” y “*cache\_bottlenecks*”, que leen los archivos almacenados en el directorio *bottlenecks* y los guarda en cache, con estos métodos se verifica que cada imagen del dataset tenga su correspondiente archivo *bottleneck*, al no existir este archivo, estos métodos los crean. Esta funcionalidad permite que podamos agregar más imágenes al dataset y que solo se creen los archivos *bottleneck* de las imágenes agregadas, permitiendo minimizar el tiempo de entrenamiento de modelo con nuevas imágenes.

Estos archivos serán utilizados en el desarrollo de la capa de entrenamiento de la red, más específicamente los datos serán almacenados en un tensor bottleneck para ser utilizados en la capa de entrenamiento.

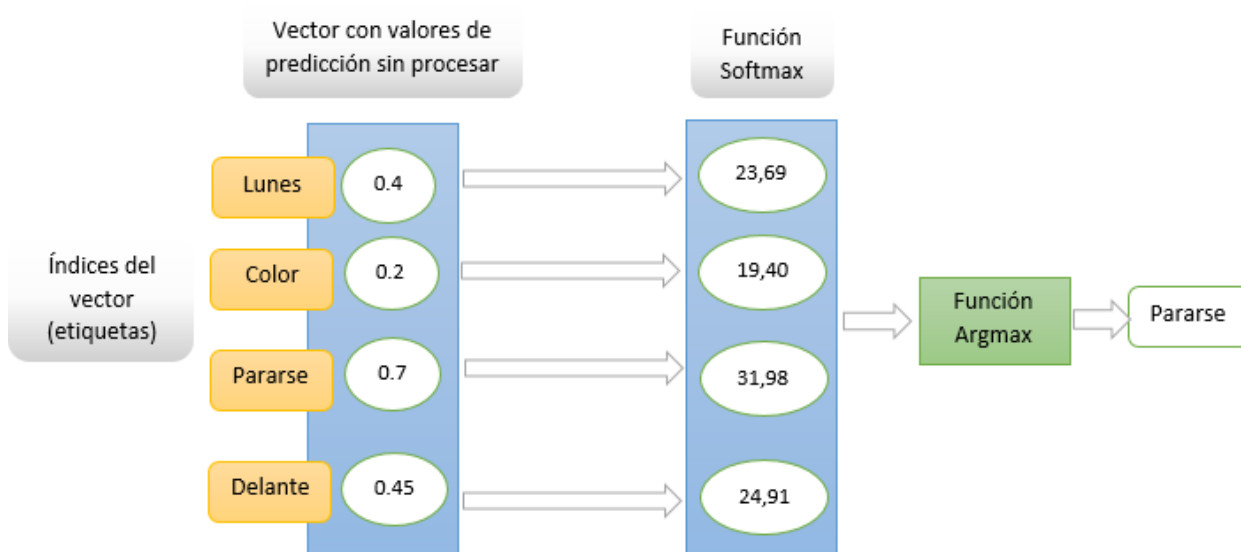
### **B. Desarrollo de la capa de entrenamiento de la red.**

Para la etapa de desarrollo de la red, primero se define las operaciones de entrenamiento, las cuales son:

- La variable `bottleneck_input`, está diseñada para contener los valores del tensor de características de las imágenes, denominado bottleneck.
- Se declara las variables de pesos y sesgo, las cuales toman la forma de la matriz del tensor bottleneck, esto genera el espacio necesario para contener los valores que inicializaran estas variables.
- Para la inicialización de la variable peso se utiliza la media de los datos del tensor bottleneck y su magnitud de separación está en función a la desviación estándar, esto datos son aleatorios y mayores a 0. La variable sesgo comienza con una matriz formada por ceros
- La variable `Wx_plus_b` almacena los datos de la multiplicación entre la matriz de pesos por la variable `bottleneck_input` y la suma del sesgo. Los valores que se obtienen al realizar esta operación se denominan valores de predicción sin procesar (resultados de la capa final o capa totalmente conectada), estos datos

son enviados a las funciones softmax y argmax (ver Figura 26), las cuales se detallan a continuación:

- La función softmax se encarga de asignar probabilidades decimales a cada valor del vector que nos proporciona la capa final, estos valores de probabilidad suman 1. Esta función proporciona porcentajes de probabilidad de las etiquetas a las cuales podría pertenecer el objeto (seña) que reconoce la red neuronal, obteniendo un vector de porcentajes de probabilidad.
- La función Argmax se encarga de devolver el índice más alto del vector de porcentajes de probabilidad entregado por softmax, este índice es una referencia de las etiquetas del dataset. Mediante esta función se asignan las categorías a las que podría pertenecer la seña. La etiqueta con el porcentaje más alto se la considera como la correcta predicción del reconocimiento de la seña.



**Figura 26.** Proceso de predicción de la señal

Para verificar que la red se haya entrenado adecuadamente, se implementa la función entropía cruzada (cross-entropy) con la finalidad de medir el rendimiento en la clasificación para cualquier tipo de señal (imágenes y audio), esta función continua siempre es positiva. Si la predicción del modelo coincide con la salida deseada, entonces la entropía cruzada tiende a cero. Pero si la salida del modelo no coincide con la salida deseada, se produce un valor de error (medida de error o perdida).

En este caso de clasificación de señas es necesario minimizar la entropía cruzada para que se acerque lo más posible a cero modificando peso y sesgo de la red.

Para minimizar esta media de error se realiza las siguientes operaciones:

- Obtención del valor de perdida en cada iteración de la red (mediante la función “*sparse\_softmax\_cross\_entropy*”), que permite obtener el error que se produce

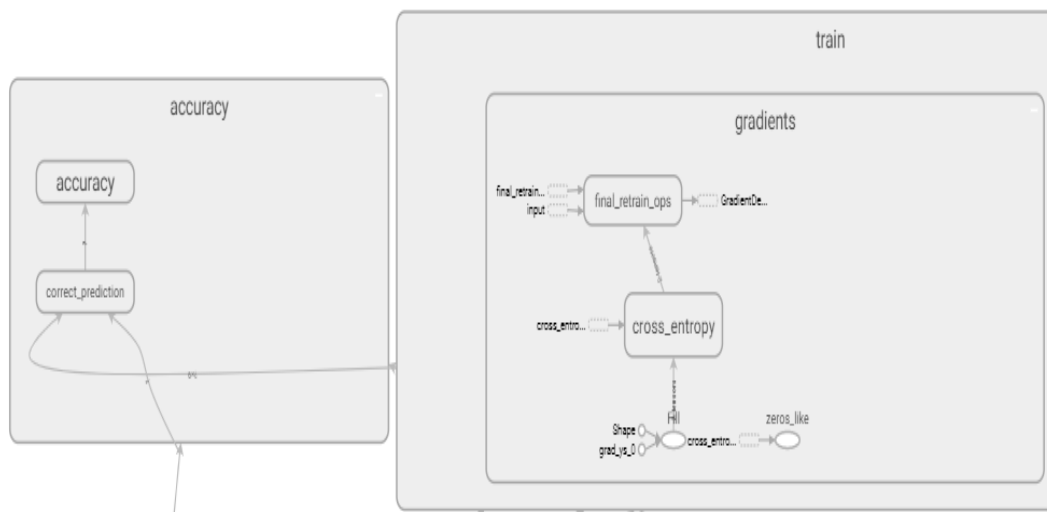
entre la salida predicha y la deseada que es entregada por cada iteración del entrenamiento de la red.

- Al tener el valor medida de error, se procede a utilizar un optimizador. En este caso se utiliza el método conocido como Gradient Descent, para esto se instancia la función “*tf.train.GradientDescentOptimizer*”, con el objetivo de minimizar este error por cada iteración del entrenamiento de la red.

Definidas las operaciones en la capa de entrenamiento de la red con los datos almacenados en tensor `bottleneck` (características de las imágenes) se procede a realizar el entrenamiento del modelo.

### **C. Entrenamiento del modelo (red neuronal).**

En esta etapa se ejecuta las operaciones definidas en la sección B, y se asigna memoria a la variable `bottleneck_input`, para almacenar los datos de los archivos `bottlenecks`, para el entrenamiento, tal y como se muestra en la Figura 27.



**Figura 27.** Entrenamiento de la red

El modelo fue entrenado con diferentes números de iteraciones las cuales comenzaron con 1000, con la finalidad de alcanzar el 100% de precisión del modelo con las imágenes de entrenamiento y minimizar el error de predicción para lo cual se midió:

- La precisión del entrenamiento (train accuracy). Que es la precisión que tiene un modelo entrenado con ejemplos clasificados y etiquetados, en este caso con imágenes de señas.
- Entropía cruzada (cross entropy). Medida de rendimiento utilizada en la clasificación. En este caso de clasificación de señas es necesario minimizar la entropía cruzada para que se acerque lo más posible a cero modificando peso y sesgos del modelo.

- Precisión de la validación (validation accuracy). El porcentaje de clasificaciones correctas / la cantidad total de clasificaciones.

La precisión del modelo con las imágenes del dataset alcanzaron el 100% de clasificaciones correctas en las primeras 1000 iteraciones, pero la entropía cruzada con más iteraciones disminuía considerablemente, las cuales se detalla en la siguiente tabla.

**Tabla 16.**  
*Proceso de entrenamiento*

<b>Nº</b>	<b>Iteraciones</b>	<b>Promedio de la entropía cruzada (n=10)</b>
<b>1</b>	1000	0,02912 ± 0,0053
<b>2</b>	2000	0,01313 ± 0,0005
<b>3</b>	3000	0,01028 ± 0,0023
<b>4</b>	4000	0,00754 ± 0,0007
<b>5</b>	5000	0,00538 ± 0,0007
<b>6</b>	6000	0,00507 ± 0,0008

En la Tabla 16., se analiza la entropía cruzada para encontrar el número correcto de iteraciones para el entrenamiento de la red, para esto se realizaron 10 repeticiones por cada iteración, en donde se obtuvo el promedio de la entropía cruzada con su respectiva desviación estándar (dispersión de datos con respecto al promedio). En esta tabla se busca los valores más bajos tanto en el promedio como en la desviación

estándar, con el objetivo de obtener el número de iteraciones necesarios para ajustar el modelo predictivo. Esto se puede ver en las filas 5 y 6, en donde el promedio de la entropía cruzada como la desviación estándar es menor a las demás filas, pero para nuestro caso de estudio se escogió 5000 iteraciones, que pertenece a la fila 5, ya que las iteraciones mayores 5000 presentan datos similares en el promedio y la desviación estándar, cambiando solo en las milésimas, los cuales no mejoran la precisión y predicción del modelo predictivo.

#### D. Pruebas al modelo

A continuación se muestran los resultados de las pruebas realizadas al modelo predictivo. Para estas pruebas se utilizaron nuevas 10 nuevas imágenes de cada una de las 15 palabras del lenguaje de señas.

**Tabla 17**  
*Prueba del modelo predictivo*

Nº	Palabra	Acierto
1	Arriba	97,28% ±0,013
2	Abajo	98,17% ±0,014
3	Color	99,13% ±0,003
4	Rojo	98,28% ±0,005
5	Casa	99,35% ±0,003
6	Delante	99,22% ±0,002
7	Familia	99,13% ±0,003
8	Lunes	97,34% ±0,017



<b>9</b>	Martes	98,50%	$\pm 0,006$
<b>10</b>	Viernes	95,81%	$\pm 0,016$
<b>11</b>	Niño	99,36%	$\pm 0,003$
<b>12</b>	Poco	97,67%	$\pm 0,013$
<b>13</b>	Encima	98,42%	$\pm 0,006$
<b>14</b>	Mamá	99,68%	$\pm 0,003$
<b>15</b>	Pararse	97,45%	$\pm 0,015$
		98,38%	

En la Tabla 17, se ven los resultados al probar el modelo con nuevas imágenes de cada seña, estas pruebas fueron realizadas por los tesisistas, en un PC con las siguientes características: procesador intel core i3 de 2.4 GHz y memoria ram de 8gb.


La tabla está dividida en cinco partes, la segunda fila consta de las palabras con las que fue entrenado en modelo, la tercera fila representa el número de pruebas realizadas por cada seña, en la cuarta fila se muestra el promedio de aciertos que tuvo el modelo al realizar las pruebas con las nuevas imágenes y, mediante la desviación estándar se consiguió los intervalos de acierto que se pueden producir al probar el modelo con otras imágenes. Esta fila refleja la precisión del modelo en un promedio de 98,38% de aceptación.

#### 4.4. Desarrollo de la interfaz de usuario a utilizar en la aplicación móvil.

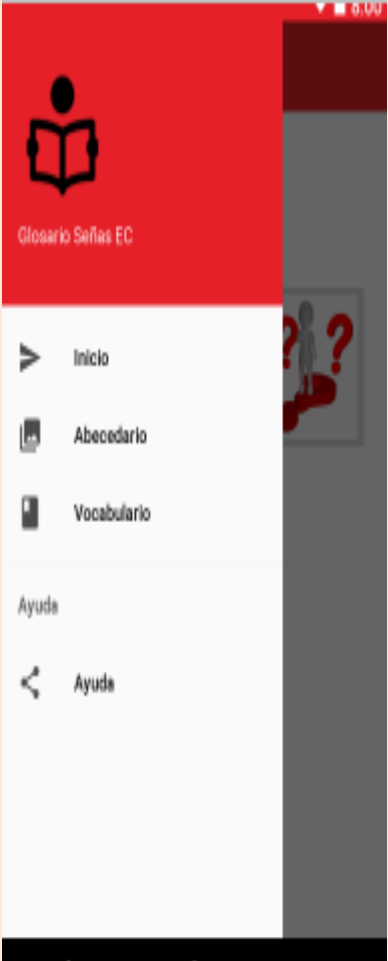
La aplicación móvil cuyo nombre es “En señas” es para el uso tanto de niños para de adultos, por tanto las interfaces de usuario, consta de menús, botones, imágenes multimedia, esto distribuido en las diferentes pantallas de vocabulario, práctica, abecedario, muestra de multimedia y ayuda, que en las tabla 18, tabla 19, tabla 20, tabla 21, tabla 22, tabla 23 y tabla 24, se detalla el funcionamiento de cada una de estas.

**Tabla 18**


*Pantalla de Inicio de la Aplicación*

Interfaz	Descripción
<p><b>Pantalla de Inicio</b></p> 	<p>En esta pantalla de inicio se muestran cuatro iconos, cada uno de ellos al ser seleccionados enviarán a una pantalla distinta:</p> <ol style="list-style-type: none"> <li>1. Play o Vocabulario.- Se muestra en pantalla el listado de palabras que existe en la aplicación, es decir, el vocabulario del lenguaje de señas.</li> <li>2. Abecedario.-Se muestra el abecedario con sus respectivas señas e imágenes.</li> <li>3. Ayuda.- Se muestra en pantalla el funcionamiento detallado tanto de la aplicación como del reconocimiento como ayuda del usuario.</li> <li>4. Salir.-El cuarto botón es para salir de la aplicación.</li> </ol>

**Tabla 19**  
*Menú de la Aplicación*

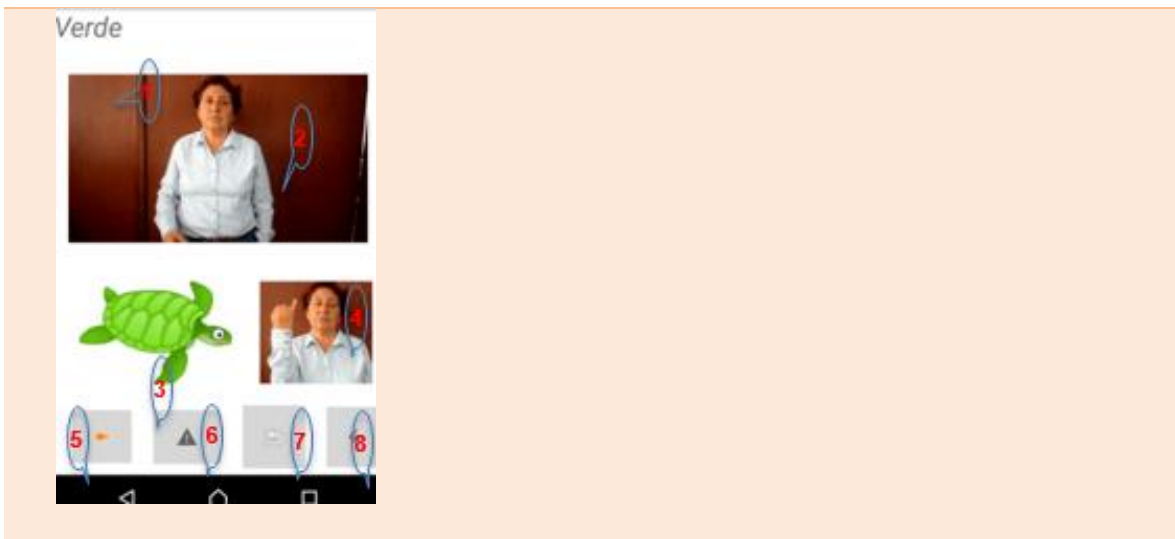
Interfaz	Descripción
<p data-bbox="272 457 363 491"><b>Menú</b></p> 	<p data-bbox="656 457 1367 716">Un menú desplegable de navegación para cuando el usuario esté en una de las pantallas y quiera navegar entre las opciones de abecedario, vocabulario, ayuda fácilmente o regresar al inicio.</p> <ol data-bbox="656 751 1367 1444" style="list-style-type: none"> <li data-bbox="656 751 1367 863">1. Inicio.- volver a la pantalla de inicio de la aplicación.</li> <li data-bbox="656 898 1367 1010">2. Abecedario.-Se muestra el abecedario con sus respectivas señas e imágenes.</li> <li data-bbox="656 1045 1367 1230">3. Vocabulario.- Se muestra en pantalla el listado de palabras que existe en la aplicación, es decir, el vocabulario del lenguaje de señas.</li> <li data-bbox="656 1266 1367 1444">4. Ayuda.- Se muestra en pantalla el funcionamiento detallado tanto de la aplicación como del reconocimiento como ayuda del usuario.</li> </ol>

**Tabla 20***Pantalla de vocabulario*

Interfaz	Descripción
<p><b>Pantalla Vocabulario</b></p> 	<p>La interfaz o pantalla de vocabulario contiene un listado de palabras con sus respectivas imágenes de su significado (2) y la seña de la letra con la que empieza la palabra correspondiente(1) en donde, si se selecciona una de estas se abrirá la pantalla donde muestra cómo se hace en señas la palabra (3)</p>

**Tabla 21***Pantalla Muestra*

Interfaz	Descripción
<p><b>Pantalla Muestra</b></p> <p><b>Multimedia</b></p>	<p>En esta pantalla se simula al docente, se muestra la palabra (1), una imagen (4) en conjunto con un pequeño video (2) de cómo se realiza la seña y una imagen (3) con su significado. En la parte inferior cuenta con cuatro botones, mismo que al seleccionarlos realizarán distintas tareas:</p> <ol style="list-style-type: none"> <li>5. Práctica.-Empieza la práctica para el reconocimiento</li> <li>6. Ayuda.-Botón para ir a la pantalla de ayuda</li> <li>7. Retroceder.-Botón para regresar al vocabulario</li> <li>8. Volver Inicio.-Botón para regresar al inicio</li> </ol>



**Tabla 22**  
*Pantalla de Práctica*  
**Interfaz**

**Descripción**

**Pantalla Práctica**

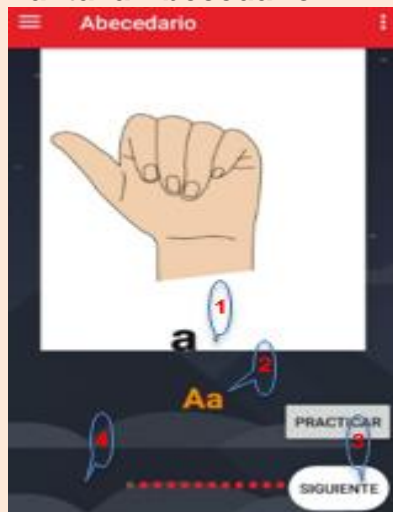


En esta pantalla se activa la cámara delantera para empezar la práctica con el reconocimiento con un tiempo determinado en donde, si este termina y no se ha reconocido la seña correcta saldrá una pantalla con una X de error o, si reconoce saldrá una imagen de bien hecho, la imagen de como se hace la seña y con el botón de volver a la pantalla de muestra.

**Tabla 23**  
*Pantalla Abecedario*

**Interfaz**

**Pantalla Abecedario**



**Descripción**

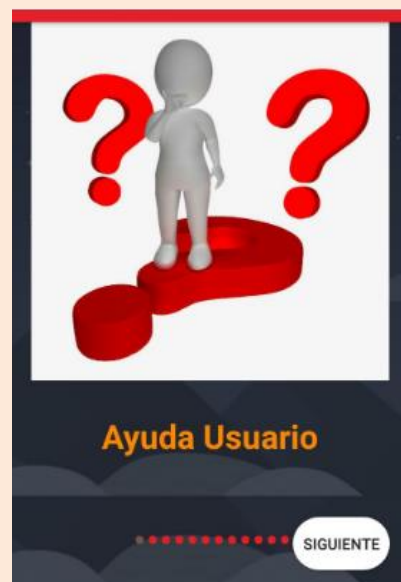
En esta pantalla se muestra imágenes del abecedario en lenguaje de señas (1) y (2) la letra que representa con botones para pasar a la siguiente seña (3) o retroceder de seña (4).

(5) Práctica.-Empieza la práctica para el reconocimiento.

**Tabla 24**  
*Pantalla de Ayuda*

**Interfaz**

**Pantalla Ayuda**



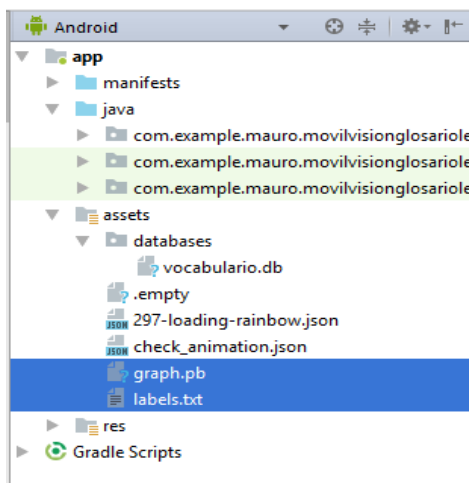
**Descripción**

Esta pantalla muestra imágenes del funcionamiento la aplicación, es decir, lo que necesita saber el usuario en caso de perderse o tener dudas sobre esta.

#### 4.5. Integración de la red neuronal artificial con la aplicación móvil

En la etapa de desarrollo de la red neuronal artificial y entrenamiento, con las imágenes contenidas en el dataset se generó un modelo predictivo, este modelo es integrado a la interfaz de usuario elaborado en la sección 4.4, para el desarrollo de la detección y reconocimiento de la seña en tiempo real.

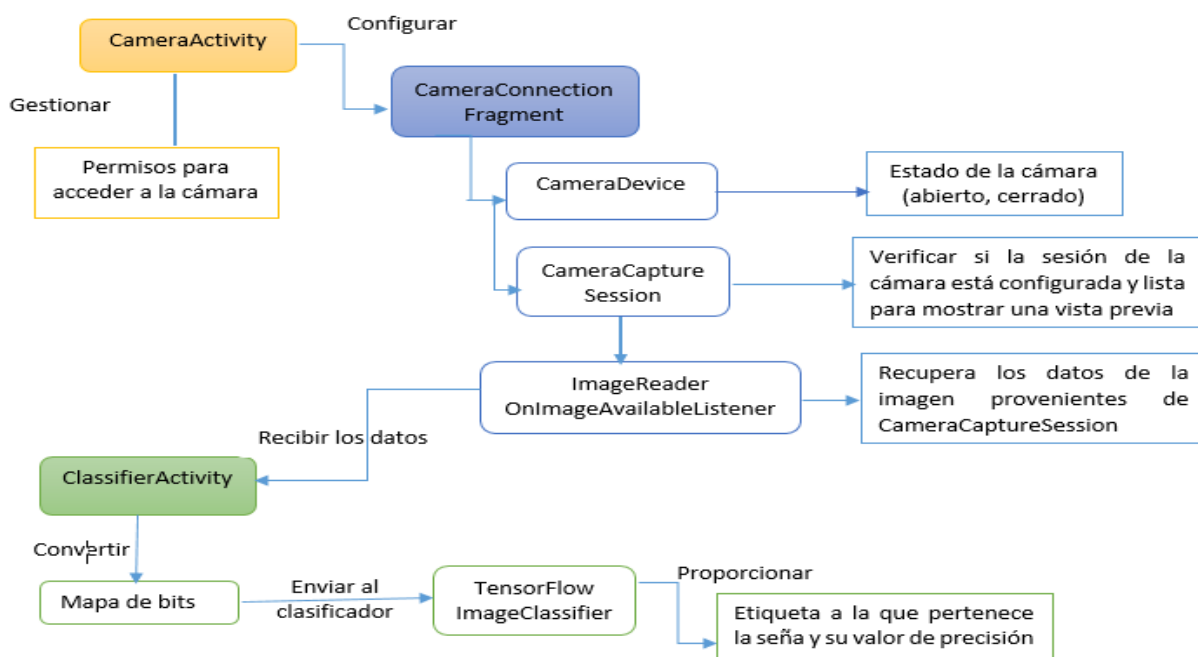
En primera instancia se carga y se llama al modelo con sus etiquetas a un clasificador de TensorFlow Android (TensorFlowImageClassifier), este modelo se encuentra en la carpeta assets, como se muestra en la Figura 28



**Figura 28** Localización del modelo y las etiquetas de las imágenes

La clasificación comienza cuando los datos provenientes de las imágenes son recibidos por la cámara del dispositivo móvil. La aplicación utiliza la API de Android denominada Camera2, el código de esta API es bastante extenso y se encuentra

empaquetado en la clase abstracta `CameraActivity`. Esta clase se encarga de gestionar los permisos necesarios para acceder a la cámara y configurar la clase `Fragment`: `CameraConnectionFragment`. Este `Fragment` maneja toda la configuración relacionada con la cámara, como por ejemplo: abre y cierra la conexión de la cámara, el manejo de la vista previa y la captura de imágenes (ver Figura 29).



**Figura 29.** Proceso de reconocimiento de la contraseña por la cámara del móvil

Toda la clasificación se maneja en la clase `ClassifierActivity`, esta clase hereda todo el contenido de la clase abstracta `CameraActivity`. La clase `ClassifierActivity` recibe los datos (las imágenes) a través del método `onImageAvailable`, el cual está vinculado al función `OnImageAvailableListener`, esta función se encuentra en el la clase `CameraConnectionFragment`, mediante esta función se registra la lectura de



imágenes procedentes de la vista previa de la cámara. Cuando una imagen está disponible, se llama al método `onImageAvailable` de la clase `ClassifierActivity`, se procesa los datos de la imagen y se los almacena en la variable `ImageReader` que es un parámetro del método `onImageAvailable`. Antes de que los datos contenidos en la variable `ImageReader` se puedan pasar a un clasificador, se convierten en un mapa de bits, ya que este clasificador maneja formatos bitmap.

El clasificador que realiza las operaciones de reconocimiento de imágenes es el `TensorFlowImageClassifier` y es llamado por la clase `ClassifierActivity`. Las operaciones de reconocimiento realizadas por este clasificador proporcionan la etiqueta a la que pertenece la seña y su respectivo valor de precisión, con estos datos se realiza la validación de la seña.

#### **4.6. Pruebas e Implantación**

Para validar el modelo predictivo del lenguaje de señas se han efectuado varias pruebas con los niños de la Unidad Educativa de segundo y tercer grado con la docente a cargo quien controlaba las pruebas, con las siguientes palabras: lunes, rojo, color, delante y pararse.

#### **4.6.1. Pruebas para validar el modelo y el dispositivo en ambientes controlados**

Estas pruebas fueron realizadas en ambientes controlados es decir, disminuyendo los efectos de la luz, distancia, la velocidad con la que se efectúa el movimiento de la seña frente a la cámara del dispositivo móvil.

Para descubrir una mejor edificación del dataset se experimentó en primera instancia con las manos descubiertas y con 90 imágenes de cada seña, teniendo como resultado un modelo predictivo, que produce valores de predicción menores al 60%, al momento de realizar el reconocimiento de la configuración manual de una palabra del lenguaje de señas. Por este motivo se optó por incrementar la cantidad de imágenes a un mínimo de 250 por seña y la integración de guantes de un color distintivo (tomate), que se diferencie del resto de objetos que se encuentran en el entorno.

Las pruebas fueron realizadas en un aula de la unidad educativa, en donde se intentó controlar la iluminación, que afecta el color de los guantes y la forma de la seña. Para evitar esto, se buscó en el aula un lugar en donde no influya la exposición de luz.

Para determinar los parámetros de precisión del modelo se realizaron pruebas con las palabras lunes, pararse y delante, en donde solo interviene la configuración manual. También se determinó la precisión de señas en las cuales intervenga la cara y las manos como son color y rojo.

En la Tabla 25 se muestra las palabras: lunes, rojo, color, delante y pararse, el rango de precisión y la distancia que debe tener las señas con respecto al dispositivo móvil, para un buen reconocimiento. En donde la distancia mínima es de 15cm y máxima es de 20cm de reconocimiento de la configuración manual del lenguaje de señas

**Tabla 25**

*Reconocimiento de imágenes. Pruebas al dispositivo móvil con el modelo predictivo en un ambiente controlado*

<b>Palabras</b>	<b>Lunes</b>	<b>Color</b>	<b>Rojo</b>	<b>Pararse</b>	<b>Delante</b>
<b>Rango de precisión</b>	0.6 - 0.86	0.6 - 0.9	0.6- 0.79	0.6 - 0.93	0.6 - 0.83
<b>Distancia</b>	15-20 cm	15-20 cm	15-20 cm	15-20 cm	15-20 cm

Para validar los datos de la tabla 25 en un ambiente controlado, se utilizó un umbral de aprobación, para evitar la confusión entre señas, ya que al momento de realizar los procesos de detección y reconocimiento los valores que el modelo proporcionaba eran inferiores a 0.1, esto producía respuestas incorrectas al evaluar la seña por parte de la aplicación. Para evitar esto se utilizó un umbral.

Para conseguir este umbral de aprobación se hicieron pruebas con los siguientes valores de 0.4, 0.5, 0.6, 0.7 y 0.8. Cada valor era un reflejo de los datos proporcionados por el modelo al realizar el reconocimiento de señas. Para los valores 0.4, 0.5 se observó que la seña que se detectaba y reconocía era diferente a la que se enseñaba por parte de la aplicación, por otro lado con los

valores 0.7 y 0.8 no se presentaba este problema, pero el tiempo que se debía brindar para la detección y el reconocimiento era alto, de unos 4 a 5 segundos, esto genero inconformidad por parte de los niños, por eso se optó por utilizar un valor intermedio de 0.6, ya que minimiza el tiempo que se necesita para reconocer la seña y evita la confusión entre la seña que se enseña con la que es validada por la aplicación.

Si el modelo predictivo al momento de realizar el reconocimiento se obtenía valores inferiores a este umbral, la aplicación no validaba la seña, pero al ser igual o mayor al umbral de aprobación, se entendía que la seña realizada era la que se había enseñado por parte de la aplicación “En señas” y corrobora por la docente como se observa en la *Figura 30* y *Figura 31*, en donde se muestra el proceso de enseñanza, detección, reconocimiento y validación de la seña.



**Figura 30:** Detección y reconocimiento de la palabra “Color”



**Figura 31.** Detección y reconocimiento de la palabra “Lunes”

En cuanto al parámetro de distancia existen dos valores: 15 y 20 cm, en la figura 31 se observa que para realizar el reconocimiento de la seña color, se necesita que la mano con respecto a la cámara del dispositivo este a una distancia de 15 cm, en cambio para otras señas como lunes, pararse y delante esta distancia es de 20cm, esto se debe a las imágenes que fueron utilizados en el entrenamiento de la red.

Para el entrenamiento de la red con las palabras color y rojo se utilizaron dos tipos de imágenes, los cuales son: imágenes con el contorno de la configuración manual e imágenes que además de tener el contorno de la seña tienen la parte del rostro en donde está ubicada la mano, como se observa en la *Tabla 14*. Por esta razón es necesario estar a una distancia de 15cm entre la cámara y las manos con la parte de la cara en donde esta ubica la seña, para realizar el reconocimiento.

Por otro lado las imágenes que solo tienen el contorno de la configuración manual con diferentes fondos, como por ejemplo: lunes, delante y pararse. Mejora la precisión del modelo y permite que la distancia del reconocimiento de la configuración manual sea de 20cm.

#### **4.6.2. Pruebas del refuerzo de enseñanza-aprendizaje con niños de primer y segundo grado de la Unidad Educativa Camilo Gallegos con la aplicación “En Señas”**

Para el reconocimiento de señas a través de la aplicación móvil realizada a niños/niñas, se efectuó varias pruebas a 6 niños estando en la mitad del año lectivo, ya que en este periodo de tiempo los niños tienen un conocimiento inicial de las palabras que constan en la aplicación. Lo cual permitió iniciar con la validación del funcionamiento de la aplicación y a su vez reforzar el proceso de la enseñanza-aprendizaje del lenguaje de señas. Para llevar a cabo el proceso de inserción de la aplicación indicado se realizó pruebas: sin la aplicación, con la aplicación “En señas” y con la experiencia en la aplicación “En Señas”, se tomó una muestra de 10 palabras (ver *Tabla 26: Palabras*) del dataset. Los niños/as realizaron 5 repeticiones de cada palabra, en un intervalo de tiempo de 30 min a la semana, en sesiones de una hora treinta minutos, quince minutos por participante, durante un mes. Detallado a continuación.

- **Pruebas sin la aplicación:** Para este tipo de pruebas se lo realizó en conjunto entre el docente con los niños en el aula, en donde los niños observaban los movimientos de las manos de la docente, para posteriormente imitar, repetir y aprender los movimientos de configuración manual. La precisión de estas pruebas se las realizan evaluando directamente a los niños con la aplicación sin ningún aprendizaje previo con esta, resultados que se obtuvieron fueron de un 68.3% como puede observar en la *Tabla 26: Precisión antes de la aplicación.*
- **Pruebas con la aplicación “En Señas”:** Para llevar a cabo el proceso de inserción de la aplicación en el entorno tanto a docentes como a niños, se siguieron los siguientes pasos:

Primero, la docente encargada de manipular el sistema seleccionaba una de las palabras del vocabulario de la app, el sistema inicia con un contenido multimedia de cómo se realiza la seña y una vez entendido el movimiento de esta, pasan a la práctica del gesto manual, en donde, el niño se coloca frente a la cámara del dispositivo móvil y empieza a imitar dicho movimiento, finalmente la aplicación se encargaba de detectar, reconocer y validar esta seña. Cabe notar que algunos niños tenían un conocimiento previo de las palabras contenidas en la aplicación móvil por lo que, solo veían el contenido multimedia por pocos segundos y pasaban a la práctica.

- **Pruebas con experiencia en la aplicación:** Una vez que los niños conocen el entorno de la aplicación, se realizaron pruebas para evaluar la interacción del niño con la app y el refuerzo de la enseñanza-aprendizaje. Cuyos resultados obtenidos existieron tanto aciertos(A) como fallos (F) como se muestra en la *Tabla 26*: Precisión después de la aplicación y, para validar estos resultados se utilizó las métricas de precisión<sup>24</sup> para evaluar el aprendizaje, ecuación (1): (Abu-Nimeh, Nappa, Wang, & Nair, 2008)

$$\text{Precisión} = P = \frac{A}{A+B}(1)$$

**Tabla 26**

*Métricas de precisión para evaluar el refuerzo del aprendizaje con la aplicación "En Señas"*

Palabras	Número de Aciertos Número de Fallos	Precisión antes de la aplicación	Número de Aciertos Números de Fallos	Precisión después de la aplicación
Arriba	20 10	0.666	21 9	0.733
Color	25 5	0.833	27 3	0.9
Lunes	24 6	0.8	26 4	0.866
Rojo	22 8	0.733	24 6	0.80
Viernes	22 8	0.733	23 7	0.766

<sup>24</sup> Métrica **Precisión**.- Es el porcentaje de clasificaciones correcta o incorrectamente del clasificador (modelo predictivo).(Abu-Nimeh et al., 2008)



<b>Delante</b>	22 8	0.733	25 5	0.833
<b>Familia</b>	21 9	0.7	21 9	0.7
<b>Encima</b>	23 7	0.766	25 5	0.833
<b>Pararse</b>	25 5	0.833	28 2	0.933
<b>Sentarse</b>	23 7	0.766	24 6	0.8
		0.756	<----->	0.816
				<b>6% MEJORA</b>

Con la experiencia previa de las señas, contenidas en la aplicación por parte de los niños, sobre todo en las pruebas con la aplicación y con experiencia en la aplicación, se efectuó el refuerzo de este conocimiento del lenguaje de señas mediante pruebas realizadas por la aplicación “En Señas”, donde los resultados obtenidos en porcentaje promedio de reconocimiento de señas es del 81,6%, resultados que se muestran en la tabla 26, donde el porcentaje de precisión se encuentra en el rango de 0.7 a 0.9333, mientras que antes de la utilización de la aplicación el rango varía entre 0.666 a 0.833 con un porcentaje promedio de reconocimiento de señas del 75,6%, lo que muestra que hubo un incremento del 6% en el grado de precisión en la forma de la configuración manual que debe tener la mano al realizar la seña en el aprendizaje del lenguaje de señas y en un periodo de tiempo corto, ya que con el paso del tiempo las sesiones de refuerzo del aprendizaje iban disminuyendo hasta llegar a 5 minutos por estudiante en

donde no era necesario pasar por la sección de visualización y enseñanza de contenido multimedia de la aplicación porque los niños ya tenían el conocimiento reforzado de la seña a realizar, pasando directamente al modo de práctica, detección y reconocimiento de la seña. Cabe resaltar que no siempre va haber un aumento de aprendizaje en algunas palabras ya que no observó que varió el valor de precisión.

#### 4.6.3. Pruebas de Integración

Estas pruebas se verificaron el cumplimiento de los requisitos, en cuanto al funcionamiento de la aplicación, a través del checklist de la Tabla 27.

**Tabla 27**

*Pruebas de integración*

Nº	Identificador	CRITERIOS	Cumple
1	RF01	Consulta del abecedario del lenguaje de señas	✓
2	RF02	Consulta del vocabulario del lenguaje de señas	✓
3	RF03	Selecciona y consulta datos del vocabulario.	✓
4	RF04	Reconoce la configuración manual de las palabras del lenguaje de señas que están en la aplicación con esta opción	✓
5	RF05	Valida el reconocimiento de la configuración manual	✓
6	RIU01	Se visualiza pantalla principal	✓
7	RIU02	Visualiza menú deslizante	✓
8	RF06	Visualización de la pantalla de ayuda	✓
9	RNF01	Rendimiento de la	✓ X*

**CONTINÚA** 

		aplicación móvil	
<b>10</b>		Facilidad para la comprensión de la aplicación	✓
<b>11</b>		Facilidad para el uso de la aplicación	✓
<b>12</b>	RF05	Intervalo de tiempo para la detección, reconocimiento y validación de la seña	✓

\* Depende de la tecnología del teléfono móvil por ejemplo la velocidad de procesamiento del teléfono, resolución de cámara, memoria del teléfono.

## **CAPÍTULO V**

### **VALIDACIÓN Y ANÁLISIS**

#### **5.1. Validación de la aplicación móvil para la enseñanza y aprendizaje del lenguaje de señas**

En este apartado se analiza y valida los resultados obtenidos en el capítulo IV, en la sección pruebas e implantación, para lo cual se utiliza métodos y técnicas estadísticas. El capítulo consta de dos partes: en la primera se analiza y valida la hipótesis presentada en el perfil de investigación, utilizando el método de inferencia estadística para dos muestras en base a los datos obtenidos por medio de las pruebas del refuerzo de la enseñanza-aprendizaje de lenguaje de señas con los niños de primer y segundo grado antes y después de utilizar la aplicación móvil y, en la segunda se validó la aplicación en base a encuestas realizadas a los docentes para afianzar los resultados obtenidos en la primera parte de esta sección.

##### **5.1.1. Análisis y validación de la hipótesis**

La hipótesis propuesta en el perfil de investigación es: Si se desarrolla una aplicación móvil con visión artificial entonces se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada

Camilo Gallegos para personas con discapacidad auditiva, donde las variables de la investigación que estructuran esta hipótesis son:

Variable independiente: si se desarrolla una aplicación móvil con visión artificial.

Variable dependiente: Se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

Los indicadores considerados para la validación son:

- a) Incremento del aprendizaje
- b) Reducción del tiempo de enseñanza (refuerzo)
- c) Validación de la aproximación de la imitación/simulación de los rasgos manuales del lenguaje de señas
- d) Porcentaje de precisión del algoritmo de aprendizaje superior al 70%

En la Tabla 28, sección 4.6.2 Pruebas del refuerzo de enseñanza-aprendizaje con los niños de primer y segundo grado, se demuestra que existe un incremento del aprendizaje en un 6% en el término de un mes, lo que permite validar el indicador a) y c), ya que la aplicación al tener el modo de práctica de la configuración manual, necesita evaluar la imitación correcta de la seña. Para el caso del indicador b) de acuerdo a las pruebas realizadas en sección 4.6.2 Pruebas del refuerzo de enseñanza-aprendizaje con los niños de primer y segundo grado, se muestra que en cada sesión existo una disminución en

tiempo que se dedicaba a esta actividad, validando así este indicador. Lo cual también fue rectificado por los docentes de la unidad educativa.

Para el último indicador d), se validó de acuerdo a la Tabla 25 de la sección 4.7. Pruebas para validar el modelo y el dispositivo en ambientes controlados cuya precisión tienen un pico 90%, por lo que podemos decir que el cuarto indicador está admitido.

Para realizar la validación de la hipótesis en conjunto con los indicadores, se utilizó inferencia estadística para dos muestras con un grado de confianza de 0,05. En nuestro caso el número de muestras es pequeño y se adapta al método estadístico planteado. En las pruebas de Levene (prueba de normalidad, para evaluar la igualdad de varianzas) se verifica si existen datos anormales que afecte el método de inferencia estadística, los resultados de la prueba de Levene se obtiene ingresando los datos de la Tabla 28 a un programa estadístico especializado en nuestro caso SPSS (Statistical Package for the Social Sciences) en donde la probabilidad del estadístico arrojado es de 0.418, por ende es mayor al grado de confianza de 0.05, y se puede decir que la distribución es normal y simétrica, por lo que se procede a ocupar el método de inferencia estadística de dos muestras, el cual ayuda a comprobar el fortalecimiento del proceso de enseñanza-aprendizaje mediante una aplicación móvil con visión artificial.

La ecuación estadística que se implementó para este método es (2):

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} \quad (2)$$

Donde

$t$ : Inferencias estadísticas de muestras,

$\bar{d}$ : es la media de las diferencias, entre el después y antes del aprendizaje

$\mu_d$ : es el planteamiento del promedio real de las diferencias individuales es cero,

$s_d$ : Desviación estándar (dispersión de los datos en base al promedio)

$n$ : es el número total de la muestra (Palabras)

Para ello se integra la hipótesis nula y la hipótesis alternativa:

**Hipótesis nula ( $H_0$ ):** Si se desarrolla una aplicación móvil con visión artificial entonces no se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

**Hipótesis Alternativa ( $H_1$ ):** Si se desarrolla una aplicación móvil con visión artificial entonces se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo

grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

En la *Tabla 28*

Calculo de la media de las diferencias, se muestra los resultados de la evaluación en el refuerzo del aprendizaje con la aplicación "En Señas", en la sección 4.6.2 Pruebas del refuerzo de enseñanza-aprendizaje con los niños de primer y segundo grado, se calcula la diferencia entre los valores: del antes y después de utilizar la aplicación, donde el promedio de la diferencia de medias es 0.0601 y la desviación estándar es 0.03061

**Tabla 28**

*Calculo de la media de las diferencias*

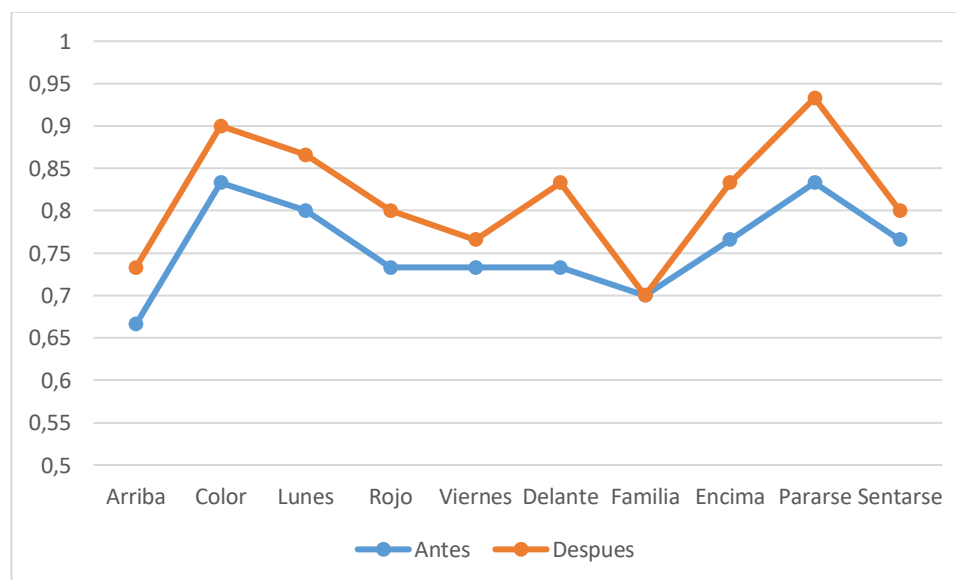
<b>Palabras</b>	<b>Precisión antes de la aplicación</b>	<b>Precisión después de la aplicación</b>	<b>Diferencia de medias <math>\bar{d} = \text{despues} - \text{antes}</math></b>
<b>Arriba</b>	0.666	0.733	0.067
<b>Color</b>	0.833	0.9	0.067
<b>Lunes</b>	0.8	0.866	0.066
<b>Rojo</b>	0.733	0.80	0.067
<b>Viernes</b>	0.733	0.766	0.033
<b>Delante</b>	0.733	0.833	0.1
<b>Familia</b>	0.7	0.7	0
<b>Encima</b>	0.766	0.833	0.067

**CONTINÚA** 



<b>Pararse</b>	0.833	0.933	0.1
<b>Sentarse</b>	0.766	0.8	0.034
$\bar{x}$ (Promedio)	0.756	0.816	0.0601
<b>Desviación Estándar</b>			0.03061

En la *Figura 32* se observa de manera gráfica como los valores contenidos en la línea azul representan el nivel de conocimiento previo de cada palabra en función a su configuración manual, la cual está por debajo de la roja. Esta última representa el refuerzo de la configuración manual mediante la aplicación móvil.



**Figura 32** Análisis del refuerzo del conocimiento del lenguaje de señas

Una vez encontrado todos los valores de la ecuación (2) se calcula los grados de libertad (gl) que es igual a  $n - 1$ , entonces nuestro  $gl = 9$  y con el nivel de confianza de 0.05 considerado para este proyecto, se busca en la tabla de distribución t student, donde se consigue la regla de decisión:

Si  $t \leq 1.833$  no se rechaza  $H_0$

Si  $t > 1.833$  se rechaza  $H_0$

Y se procede a emplear la ecuación de la inferencia estadística de muestras (2) para determinar la veracidad de una de las reglas de decisión:

$$t = \frac{0.0601 - 0}{\frac{0.03061}{\sqrt{10}}}$$

$$t = 6.208$$

De acuerdo a los datos obtenidos en el análisis estadístico, nos revela que  $t = 6.208$ , el cual indica que es mayor a 1.833. Por lo tanto se rechaza a la hipótesis nula y se acepta la hipótesis alternativa, donde se afirma que se fortalece el proceso de enseñanza-aprendizaje de configuración manual del lenguaje de señas en los alumnos de primer y segundo grado de la Unidad Educativa Especializada Camilo Gallegos para personas con discapacidad auditiva.

### **5.1.2. Validación de la aplicación**

Para la validación de la aplicación, también se basó en los resultados obtenido por la encuesta realizada a los docentes (7) de Unidad Educativa Especializada

Dr. Camilo Gallegos Domínguez (ver **¡Error! No se encuentra el origen de la referencia.**), lo que permite conocer más afondo el grado de aceptación y satisfacción, referentes al refuerzo de la enseñanza-aprendizaje del lenguaje de señas.

Al analizar los resultados obtenidos de las encuestas realizadas a los docentes de la Unidad Educativa Especializada Camilo Gallegos se destaca los siguientes aspectos:

- i. Según las preguntas 1,2 y 3, se puede observar que el conocimiento del lenguaje de señas por parte de los alumnos de primer y segundo grado de la unidad educativa, se encuentra por debajo del nivel deseado por los docentes, ya que el tiempo que se dedica a esta actividad no es suficiente para reforzar los conocimientos en la unión del significado de la palabra con su respectivo movimiento de la seña adquiridos en clases. Cabe destacar que solo una minoría de niños tiene el nivel requerido de conocimiento del lenguaje de señas.
- ii. La utilización de la tecnología como medio didáctico con videos, imágenes, y/o animaciones en el fortalecimiento del conocimiento del lenguaje de señas, ayuda a los estudiantes con un mejor desempeño en las actividades escolares, de acuerdo a los resultados obtenidos en las preguntas 4 y 5.

- iii. Se destaca en el análisis de las preguntas 10 y 11, que los alumnos al interactuar con la aplicación móvil fortalecieron el nivel de atención, disminuyó el tiempo de aprendizaje y mejoró la retención de información del conocimiento del lenguaje de señas.
- iv. La facilidad de manejo que tiene la aplicación es factible para su utilización ya que la información proporcionada en las diferentes opciones de esta es competente para los usuarios que utilizaron la aplicación, de acuerdo a las preguntas 6,7 ,8 y 9 del análisis de datos.

De la pregunta 11 se puede indicar que los docentes recomiendan aumentar el vocabulario de la aplicación y disminuir los factores externos a la aplicación (que afectan a la cámara) para realizar el reconocimiento de la configuración manual de las palabras del lenguaje de señas.

Algunos docentes de niveles avanzados encuestados indicaron que la podría ayudar a personas que recién estén empezando o quieran aprender este lenguaje.

## CAPÍTULO VI

### CONCLUSIONES Y RECOMENDACIONES

#### 6.1. Conclusiones

- La investigación realizada en nuestro proyecto de investigación nos permitió determinar que el lenguaje de señas varían de un país a otro, por ejemplo en países como México, las palabras color, rojo, encima, martes, miércoles, familia, entre otros, tienen distinciones en el contexto y la configuración manual, con respecto a Ecuador. Estas pueden llegar a limitar la comunicación del lenguaje de señas entre países.
- Cabe notar que la construcción del dataset se realizó de forma manual, debido a la carencia de imágenes vinculadas a la configuración manual del lenguaje de señas ecuatoriano. Por tal motivo, se hizo un tratamiento y pre-procesamiento de las imágenes obtenidas, descartando y verificando los tipos de colores que ayuden a mejorar la configuración manual para la exactitud y precisión del reconocimiento de la seña por parte de la aplicación móvil “En Señas”, se considera a esta etapa como la que consume la mayor parte del tiempo para el desarrollo de este proyecto de investigación.
- El dataset para entrenar la red neuronal, comprendió varias etapas de retroalimentación: adquisición de imágenes, procesamiento, segmentación y clasificación, las cuales permitieron determinar que las imágenes con fondos

variados mejoran el desempeño de la red neuronal convolucional. Lo que se vio reflejado al momento de evaluar el modelo predictivo, observándose mejores resultados por cada refinamiento de este.

- De acuerdo con los resultados obtenidos en las pruebas de la aplicación para el reconocimiento del lenguaje de señas ecuatoriano realizado a niños de la Unidad Educativa Especializada Camilo Gallegos aplicados en un ambiente controlado se obtuvo un porcentaje promedio de reconocimiento de 81,6% es decir, un 6% de mejora en el refuerzo de la configuración manual que debe tener la mano al realizar la seña, lo que indica que la aplicación refuerza el aprendizaje del lenguaje.
- Una de las principales limitaciones en el reconocimiento de las manos es la influencia de factores externos como el hardware del dispositivo móvil (cámara, procesador, memoria), luminosidad, la velocidad del movimiento con el que se realiza la seña y la confusión que se da por el color del entorno y el color de piel. Por ello se realizó pruebas con distintos tipos de guantes dando mejores resultados con guantes tomates que se distinguen tanto del entorno como del color de ropa y de la piel, los cuales mejoran la predicción de señas.
- De acuerdo a las encuestas realizadas a los docentes, se concluye que la aplicación “En Señas” brinda apoyo a los niños de la unidad educativa, o a personas que requieran aprender-reforzar el lenguaje de señas, fortaleciendo

y disminuyendo el tiempo en el proceso de enseñanza-aprendizaje de este lenguaje.

## **6.2. Recomendaciones**

- De acuerdo a las conclusiones obtenidas en nuestra tesis de grado podemos recomendar: La aplicación “En Señas” puede ser usada por personas, niños o adultos, que deseen aprender y aplicar el lenguaje de señas ecuatoriano, ya sean estos hablantes o tengan poco vocabulario o estén iniciando con el aprendizaje de este lenguaje y reforzar la difusión de las tecnologías del lenguaje de señas o aplicaciones relacionadas así a la enseñanza-aprendizaje de este lenguaje, con la finalidad fomentar la inclusión de personas sordas a la sociedad.
- Se recomienda que se siga realizando este tipo de investigaciones con el objetivo de mejorar la calidad de vida de las personas sordas, teniendo como base la aplicación “En Señas”, a la cual se podría agregar más funcionalidades por ejemplo reconocimiento de gestos faciales o localización de objetos.
- Se recomienda un incremento sustancial en el desarrollo de la materia de Inteligencia Artificial ya que en la actualidad está en auge y sus conocimientos son demandados por la mayoría de empresas que prefieren la automatización

de tareas de mayor complejidad como toma de decisiones, percepción visual y reconocimiento del habla.



## REFERENCIAS BIBLIOGRÁFICAS

- AICE. (2015). *Federación AICE*. Recuperado el 18 de noviembre de 2018, de [http://implantecoclear.org/index.php?option=com\\_content&view=article&id=76&Itemid=82](http://implantecoclear.org/index.php?option=com_content&view=article&id=76&Itemid=82)
- Alberto, A. V. (31 de 3 de 2009). *mailxmail*. Recuperado el 12 de diciembre de 2018, de Lenguaje de señas: <http://www.mailxmail.com/curso-lenguaje-senas-signos/senas-clasificacion>
- Andrew G. Howard, M. Z. (17 de Abril de 2017). *arXiv.org*. Recuperado el 12 de enero de 2019, de arXiv.org: <https://arxiv.org/abs/1704.04861>
- Anguas Pérez, J. M. (Enero de 2005). *Análisis y diseño de dispositivo de radiocomunicación mediante redes neuronales*. Sevilla: e-Reding. Obtenido de bibing.
- Arndt P, S. A. (1999). Within Subject Comparison of Advanced Coding Strategies in the Nucleus 24 Cochlear Implant. *Cochlear Corporation*.
- ASPAS. (2018). Recuperado el 16 de noviembre de 2018, de <http://aspasvalencia.com/que-tipos-de-sordera-existen/>
- Bendemra, H. (30 de Abril de 2018). *Ciencia y Datos*. Recuperado el 23 de marzo de 2019, de Ciencia y Datos: <https://medium.com/datos-y-ciencia/construye-tu-primero-clasificador-de-deep-learning-con-tensorflow-ejemplo-de-razas-de-perros-ed218bb4df89>

- Biblioteca virtual Miguel de Cervantes*. (2016). Recuperado el 5 de diciembre de 2018, de Biblioteca virtual Miguel de Cervantes: [http://www.cervantesvirtual.com/obra-visor/lenguaje-de-signos--0/html/ffbeaf86-82b1-11df-acc7-002185ce6064\\_5.html](http://www.cervantesvirtual.com/obra-visor/lenguaje-de-signos--0/html/ffbeaf86-82b1-11df-acc7-002185ce6064_5.html)
- Calvo, D. (20 de Julio de 2017). *Red Neuronal Convolutacional CNN*. Recuperado el 25 de noviembre de 2018, de Red Neuronal Convolutacional CNN: <http://www.diegocalvo.es/red-neuronal-convolutacional/>
- Calvo, D. (8 de Diciembre de 2018). *Perceptrón Multicapa – Red Neuronal*. Recuperado el 21 de diciembre de 2018, de Perceptrón Multicapa – Red Neuronal: <http://www.diegocalvo.es/perceptron-multicapa/>
- CAMELO, F. A. (Enero de 2009). *researchgate*. Recuperado el 10 de febrero de 2019, de researchgate: <https://www.researchgate.net/>
- Castro Heredia, C. E. (diciembre de 2012). ANÁLISIS CLÚSTER COMO TÉCNICA DE ANÁLISIS EXPLORATORIO DE REGISTROS. *eidemar*, 11-20.
- Cesar Hernández, J. L. (2015). Las tecnologías de la información en el aprendizaje de la lengua de señas. *Salud pública*, 61-72.
- Chacoa, W. (2015). *ESTUDIO LÉXICO-GRAMATICAL DE LA LENGUA DE SEÑAS*. UNIVERSIDAD POLITÉCNICA TERRITORIAL DEL ESTADO MÉRIDA, Barinas. Obtenido de Cultura sorda.
- Claudio Javier Tablada, G. A. (2009). REDES NEURONALES ARTIFICIALES. *REVISTA DE EDUCACIÓN MATEMÁTICA*, 22-28. Obtenido de REVISTA DE EDUCACIÓN MATEMÁTICA.

- contributors, E. (8 de junio de 2013). *Cofosis o anacusia*. Recuperado el 12 de noviembre de 2018, de [https://www.ecured.cu/Cofosis\\_o\\_anacusia](https://www.ecured.cu/Cofosis_o_anacusia)
- Developers Android Studio*. (25 de 4 de 2018). Recuperado el 4 de febrero de 2019, de <https://developer.android.com/studio/intro/?hl=es-419>
- Developers, A. S. (24 de octubre de 2018). *Android Studio*. Recuperado el 12 de febrero de 2019, de <https://developer.android.com/studio/intro/?hl=ES>
- E.H.U, U. (20 de 12 de 2018). Recuperado el 20 de diciembre de 2018, de <http://www.ehu.eus/ccwintco/uploads/d/d4/PresentacionMundoVirtual.pdf>
- eltelegrafo*. (2017). Recuperado el 15 de noviembre de 2018, de *eltelegrafo*.: <https://www.eltelegrafo.com.ec/noticias/sociedad/4/la-discapacidad-auditiva-afecta-a-360-millones-de-personas-en-el-mundo>
- FENASEC. (2018). *Federación Nacional de Personas Sordas del Ecuador*. Recuperado el 12 de noviembre de 2018, de <http://fenasec.ec/index.html>
- Foto Nostra*. (2018). Recuperado el 26 de enero de 2019, de <https://www.fotonostra.com/glosario/puntofocal.htm>
- Fuzi, C. (2013). Metodología ágil para diseño de aplicaciones móviles.
- García, J. F. (Noviembre de 2006). *biblioteca.usac*. Recuperado el 17 de enero de 2019, de [http://biblioteca.usac.edu.gt/tesis/08/08\\_0155\\_EO.pdf](http://biblioteca.usac.edu.gt/tesis/08/08_0155_EO.pdf)
- Glosario. (16 de 10 de 2018). *Glosarios*. Recuperado el 24 de noviembre de 2018, de <https://glosarios.servidor-alicante.com/terminos-estadistica/grados-de-libertad>
- Gonzales, A. (30 de Julio de 2014). *cleverdata*. Recuperado el 10 de noviembre de 2018, de <https://cleverdata.io/conceptos-basicos-machine-learning/>

hear-it. (diciembre de 2017). *Definición de pérdida de audición*. Recuperado el 12 de noviembre de 2018, de Definición de pérdida de audición: <https://www.hear-it.org/es/definicion-de-perdida-de-audicion>

Howard, A. G. (14 de Julio de 2017). *Google AI blog*. Recuperado el 20 de noviembre de 2018, de Google AI blog: <https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html>

HUESO, A. D. (2018). *Visión Por Computador*. Madrid.

Ignacio, M. (Enero de 2009). *researchgate*. Recuperado el 21 de diciembre de 2018, de researchgate: [https://www.researchgate.net/publication/319411221\\_Historia\\_y\\_gramatica\\_de\\_la\\_lengua\\_de\\_senas](https://www.researchgate.net/publication/319411221_Historia_y_gramatica_de_la_lengua_de_senas)

INFAIMON. (18 de enero de 2018). *Visión por computador*. Recuperado el 16 de diciembre de 2018, de <https://blog.infaimon.com/vision-computador-soluciones-permite/>

Jain, A. (9 de Julio de 2018). *Data Science*. Recuperado el 23 de enero de 2019, de Data Science: <https://medium.com/data-science-101/transfer-learning-57ce3b98650>

Jiménez, J. J. (Septiembre de 2006). *Repositorio Digital de la Universidad Politécnica de Cartagena*. Recuperado el 4 de enero de 2019, de Repositorio Digital de la Universidad Politécnica de Cartagena: <http://repositorio.upct.es/bitstream/handle/10317/241/pfc1958.pdf?sequence=1>

- John, D. (9 de Abril de 2018). *Diferencia Entre*. Recuperado el 24 de noviembre de 2018, de Diferencia Entre: <http://www.diferenciaentre.net/diferencias-entre-el-aprendizaje-supervisado-y-el-aprendizaje-no-supervisado/>
- Lenoir, R. P. (19 de octubre de 2016). *Viaje al Mundo de la Audición*. Recuperado el 20 de noviembre de 2018, de <http://www.cochlea.eu/es/celulas-ciliadas>
- Matich, D. J. (Marzo de 2001). *Redes neuronales conceptos básicos y aplicación* . Universidad Tecnológica Nacional, Rosario. Recuperado el 2019, de Universidad Tecnológica Nacional – Facultad Regional Rosario: [https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5\\_anio/orientadora1/monograias/matich-redesneuronales.pdf](https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monograias/matich-redesneuronales.pdf)
- NIOSH. (2008). Recuperado el 11 de noviembre de 2018, de [https://www.cdc.gov/spanish/niosh/docs/2008-102\\_sp/pdfs/2008-102.pdf](https://www.cdc.gov/spanish/niosh/docs/2008-102_sp/pdfs/2008-102.pdf)
- NOISSES. (12 de Febrero de 2015). *NOISSES*. Recuperado el 17 de noviembre de 2018, de <http://www.noisess.com/el-decibelio-no-es-una-unidad-de-medida-de-sonido/>
- Nolla, F. C. (19 de Enero de 2014). Recuperado el 20 de enero de 2019, de <http://users.dsic.upv.es/~fcn/Students/ml/t7aa2p.pdf>
- OMS. (15 de marzo de 2018). *OMS*. Recuperado el 16 de noviembre de 2018, de Sordera y pérdida de la audición: <https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss>
- OpenCV. (21 de 01 de 2018). *OpenCV*. Recuperado el 13 de febrero de 2019, de <https://opencv.org/about.html>

- Óptica*. (16 de Octubre de 2017). Recuperado el 18 de enero de 2019, de <https://www.significados.com/optica/>
- Orellana, L. (2008). *dm - Regresión Lineal Simple*. Buenos Aires.
- Oviedo, A. (2007). *cultura-sorda*. Recuperado el 5 de diciembre de 2018, de cultura-sorda: <http://www.cultura-sorda.org/abad-de-lepee/>
- Oviedo, A. (2007). *cultura-sorda*. Recuperado el 5 de diciembre de 2018, de cultura-sorda: <https://cultura-sorda.org/abad-de-lepee/>
- Perez, G. (2 de octubre de 2018). *Quora*. Recuperado el 18 de diciembre de 2018, de Quora: <https://es.quora.com/C%C3%B3mo-funcionan-las-redes-neuronales-convolucionales>
- Pérez, J. A. (2018). *Diseño de dispositivos de radiocomunicación*. Universidad de Sevilla, Sevilla.
- Quiza, J. (11 de mayo de 2018). *Ciencia&Datos*. Recuperado el 14 de febrero de 2019, de Ciencia&Datos: <https://medium.com/datos-y-ciencia/deep-learning-inception-v3-con-pyspark-b97678c9616b>
- Ramírez Aldama, E. (2015). *Reconocimiento del hablante mediante extracción de características empleando la transformada Wavelet*. Universidad de las Américas Puebla, Puebla.
- RENKEMA, J. (2011). *THE TEXTURE OF DISCOURSE*.
- RobertoGonzález, A. B. (Diciembre de 2017). Aplicación de las Máquinas de Soporte Vectorial (SVM) al diagnóstico clínico de la Enfermedad de Párkinson y el

- Temblor Esencial. *Revista Iberoamericana de Automática e Informática Industrial RIAI*, 394-405. Obtenido de sciencedirect.
- Rodriguez, A. (21 de octubre de 2016). *l álgebra lineal y el procesamiento digital de imágenes*. Recuperado el 9 de diciembre de 2018, de <https://www.nibcode.com/es/blog/1137/algebra-lineal-y-el-procesamiento-digital-de-imagenes-parte-iii-transformaciones-afines>
- Rodriguez, M. I. (2000). Historia y Gramatica de la Lengua de Señas. *Pedagogia y Saberes*, 91-94.
- Salicci, P. M. (2007). La inteligencia y los Sordos: derribemos mitos.
- Santiago Prado, R. B. (2003). *Didactica lengua signos*. Barina. Obtenido de Cultura-sorda.eu.
- Sarrió, G. A. (2015). *Aplicación de técnicas de aprendizaje automático sobre juegos*. Universitat Politècnica de València, Valencia.
- Software Narciso*. (30 de 07 de 2017). Recuperado el 14 de febrero de 2019, de <https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0>
- Sordos Ecuador*. (Marzo de 2017). Recuperado el 12 de noviembre de 2018, de Sordos Ecuador: <http://www.sordosecuador.com/p/diccionarios-de-lengua-de-senas.html>
- sordosecuador*. (Julio de 2017). Recuperado el 15 de noviembre de 2018, de sordosecuador: <http://www.sordosecuador.com/p/lenguaje-de-senas-guia-basica.html>

- Sosa, S. (25 de abril de 2016). *Discapacidad Auditiva*. Recuperado el 15 de noviembre de 2018, de <https://www.lifeder.com/discapacidad-auditiva/#Tipos>
- Sykes, F. (9 de 11 de 2009). *La Lengua de Señas*. Universidad de Palermo, Buenos Aires. Obtenido de palermo.edu.
- tensorflow. (2019). *tensorflow.org*. Recuperado el 15 de marzo de 2019, de <https://github.com/tensorflow/hub>
- Tensorflow TM*. (30 de 12 de 2018). Recuperado el 16 de enero de 2019, de <https://www.tensorflow.org/?hl=es>
- Torres, J. (8 de Junio de 2018). *torres.ai*. Recuperado el 10 de noviembre de 2018, de [torres.ai: https://torres.ai/redes-neuronales-convolucionales/](https://torres.ai/redes-neuronales-convolucionales/)
- uniovied. (2012). Recuperado el 17 de diciembre de 2018, de [https://www.unioviedo.es/compnum/laboratorios\\_web/laborat03\\_intro\\_imagen/laborat03.htmlxdp0](https://www.unioviedo.es/compnum/laboratorios_web/laborat03_intro_imagen/laborat03.htmlxdp0)
- UNIVERSAL., S. (1986). SALVAT UNIVERSAL. Barcelona, España: Editores S.A. p. Tomo 15; 351.
- UOC. (19 de Abril de 2012). *Informatica*. Recuperado el 12 de diciembre de 2018, de <http://informatica.blogs.uoc.edu/2012/04/19/la-vision-por-computador-una-disciplina-en-auge/>
- USEFEDORA. (2017). *Lengua de Signos Española (Nivel Inicial)*. Recuperado el 17 de febrero de 2019, de <https://sefacil.usefedora.com/courses/3640/lectures/71282>
- Vargas Jose, C. B. (2012). MÁQUINAS DE SOPORTE VECTORIA. *Conferencias de la Universidad Nacional de Córdoba*, 231-265.



Venegas, J. V. (2016). *docplayer.es*. Recuperado el 26 de noviembre de 2018, de docplayer.es: <https://docplayer.es/15680367-Capitulo-ii-la-linguistica-en-el-lenguaje-de-senas-en-el-presente-capitulo-se-expondran-temas-como-lenguaje-lengua-y-habla.html>

# ANEXOS



**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA**

**CARRERA DE INGENIERÍA EN SOFTWARE  
CERTIFICACIÓN**

Se certifica que el presente trabajo fue desarrollado por los señores: Chicaiza Rocha, Mauro Lenin y Rengifo Tulcanaz, Paola Belén como requerimiento parcial a la obtención del título de ingeniería en Software

En la ciudad de Latacunga, miércoles 5 de junio del 2019

**Aprobado por:**

