**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS**



TESIS PARA OBTENER EL TÍTULO DE INGENIERO DE SISTEMAS

**“Sistema traductor de lenguaje de señas peruana usando aprendizaje automático”**

AUTOR

Huaman Lozano, Deivi Justi

ASESOR:

Audante Ramos, Nestor Rafael

Abril 2023

**ÍNDICE**

[**RESUMEN…………………………………………………………………………………3**](#_heading=h.gjdgxs)

1. **INTRODUCCIÓN…………………………………………………………………4**

1.1 [Descripción de la problemática](#_heading=h.1fob9te)…………………………………………………4

1.2 [Formulación del problema](#_heading=h.3znysh7)………………………………………………………5

1.2.1 Problema general…………………………………………………………..5

1.2.2 Problema específico……………………………………………………….6

1.3 [Objetivos de la investigación](#_heading=h.2et92p0)…………………………………………………...6

1.4 [Justificación y delimitación de la investigación](#_heading=h.3rdcrjn)………………………………..7

1.4.1 Justificación………………………………………………………………..7

1.4.2 Delimitación……………………………………………………………….7

1. **ANTECEDENTES ……………………………………………………………….9**
2. **MARCO TEÓRICO…………………………………………………………….14**

3.1 Lenguaje de Señas peruana (LSP)……………………………………………14

3.2 Discapacidad Auditiva………………………………………………………..15

3.3 Aprendizaje Automático..……………………………………………………..17

3.4 Computer Vision………………………………………………………………18

1. **MARCO METODOLÓGICO…………………………………………………..20**

4.1 [Tipo de investigación](#_heading=h.2jxsxqh)………………………………………………………….21

4.2 [Método de investigación](#_heading=h.z337ya) ……………………………………………………...22

4.3 [Población y muestra](#_heading=h.3j2qqm3) …………………………………………………………..24

4.4 [Técnicas e instrumentos para recolección y análisis de datos](#_heading=h.1y810tw) ………………..24

1. **RESULTADOS EXPERIMENTALES ………………………………………..26**
2. **CONCLUSIONES……………………………………………………………….27**
3. **BIBLIOGRAFÍA**………………………………………………………………...28

**RESUMEN**

La investigación que se propone, estudiará el problema de la comunicación entre personas con discapacidad auditiva que se comunican mediante lenguaje de señas peruana con el resto de la sociedad.

Para solucionar el problema descrito se propone desarrollar un sistema usando aprendizaje automático, este sistema podría traducir lo que una persona que usa lenguaje de seña peruana comunica al lenguaje regular usado por la mayoría de la sociedad.

La solución que se propone está basada en la aplicación de los conceptos y técnicas de aprendizaje automático, visión por computador y procesamiento de lenguaje natural.

Los resultados que se pretenden alcanzar es desarrollar un sistema traductor que sea capaz permitir la comunicación de las personas con discapacidad auditiva que usan lenguaje de señas peruana con el resto de sociedad, también lograr una precisión mayor de 90% haciendo uso menos recursos.

DESCRIPTORES TEMÁTICOS

Inteligencia Artificial

Machine Learning

Computer Vision

Lenguaje de señas peruana

1. **INTRODUCCIÓN**

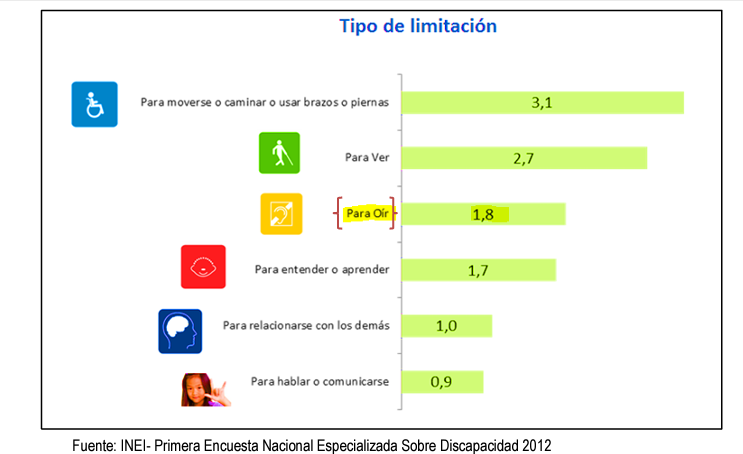
La presente investigación se enfoca en la problemática de la comunicación de las personas sordomudas que utilizan el lenguaje de señas peruana. Esta problemática se desarrolla en un contexto en el que las personas con discapacidad auditiva enfrentan desafíos significativos para expresar sus emociones y comprender las de los demás. La situación actual se caracteriza por una falta de comunicación efectiva en lenguaje de señas peruana, lo que genera múltiples impactos desfavorables en diversos aspectos de la vida de estas personas.

* 1. **Descripción de la problemática**

En Perú, según el último Censo Nacional de Población (INEI-2017) se cuenta 232 176 personas con problemas auditiva, cifra que representa el 7,6 % del total de personas discapacidades en Perú. Según Consejo Nacional para la Integración de la Persona con Discapacidad (CONADIS 2016) el 1.8% de la población peruana presenta una limitación permanente para oír, lo cual el año 2016 ya era más de medio millón de peruanos quienes padecían dificultades permanentes de audición.

**Figura 1:**

*Tipos de limitaciones y el porcentaje de la población peruana que los padece*



Tomado de *Informe Temático N° 5 “Situación de las Personas con Discapacidad Auditivaen el Perú”* por CONADIS, 2016

Las barreras de comunicación que enfrentan las personas con discapacidad auditiva en Perú pueden tener los siguientes impactos desfavorables:

* **Salud y bienestar**: Las barreras de comunicación pueden dificultar el acceso a la atención médica, lo que puede aumentar el riesgo de problemas de salud. Las barreras de comunicación también pueden contribuir a la ansiedad, la depresión y otros problemas de salud mental.
* **Educación:** Las barreras de comunicación pueden dificultar el acceso a la educación, lo que puede limitar las oportunidades de las personas con discapacidad auditiva para obtener un empleo bien remunerado y tener éxito en la vida. Según la Defensoría del Pueblo (2019) el 76% de las I.E. Públicas y 83% de las I.E. Privadas reconocen que no tienen las condiciones necesarias para brindar servicios educativos a estudiantes con discapacidades auditivas.
* **Empleo:** Las barreras de comunicación pueden dificultar el acceso al empleo, lo que puede limitar las oportunidades de las personas con discapacidad auditiva para ganarse la vida y apoyar a sus familias. Se dice que “el principal problema es la dificultad en la inserción laboral para las personas con problemas auditivos y/o vocales, debido a la dificultad en la inserción laboral para que las personas con estos problemas puedan trabajar en una Empresa” (Luyo y Juarez ,2018)
* **Participación social:** Las barreras de comunicación pueden dificultar la participación en actividades sociales y comunitarias, lo que puede limitar las oportunidades de las personas con discapacidad auditiva para formar relaciones y sentirse parte de la sociedad.
  1. **Formulación del problema**

El problema queda claramente formulado partir de plantear la siguiente pregunta de investigación:

**1.2.1. Problema General**

¿Cómo lograr que las personas que no conocen el lenguaje de señas peruana logren entender los mensajes de personas con discapacidad auditiva de manera automática?

**1.2.2. Problemas específicos**

1. ¿Cómo definir los requisitos que usará el modelo para lograr la comunicación entre personas que usan LSP con personas que usan el lenguaje regular?
2. ¿Cómo definir los hiperparametros que usará el modelo para lograr la comunicación entre personas que usan LSP con personas que usan el lenguaje regular?
3. ¿Cómo desarrollar un modelo de aprendizaje automático capaz de traducir el lenguaje de señas peruana?
4. ¿Cómo validar la precisión del modelo de traducción de señas peruanas?

* 1. **Objetivos de la investigación**

**1.3.1. Objetivo general**

Desarrollar un sistema traductor de señas peruana automático que permita el entendimiento de personas que no conocen el lenguaje de señas peruana a personas con discapacidad auditiva que se comunican mediante el lenguaje de señas peruana

**1.3.2. Objetivos específicos**

* Definir los requisitos que usará el modelo para poder lograr la comunicación con personas con discapacidad auditiva.
* Definir hiperparametros que usará el modelo para poder lograr la comunicación con personas con discapacidad auditiva.
* Desarrollar un modelo de aprendizaje automático para la traducción del lenguaje de señas peruana.
* Validar la precisión del modelo de traducción de señas peruanas.
  1. **Justificación y delimitación de la investigación**

**1.4.1. Justificación**

Con el desarrollo de un prototipo de traductor de señas peruanas en este contexto actual presenta un valor significativo, justificado tanto desde perspectivas económicas como sociales. Los beneficios económicos incluyen la inclusión de personas con discapacidad auditiva en el ámbito laboral, lo que aumentaría la fuerza laboral y promovería la diversidad en el lugar de trabajo, contribuyendo a una mejora económica general. A su vez, esto reduce la dependencia de la asistencia pública. Las empresas se beneficiarán al contar con empleados que pueden comunicarse eficazmente y promover una cultura empresarial inclusiva. Además, este prototipo permitiría un acceso más amplio a servicios esenciales, educación y atención médica, reduciendo la carga en los servicios públicos y mejorando la calidad de vida de las personas con discapacidad auditiva. La investigación contribuye a una sociedad más equitativa e inclusiva al eliminar las barreras de comunicación y asegurando que las personas con discapacidad auditiva puedan participar plenamente en la vida cotidiana. Además, se alinea con tendencias que indican un crecimiento continuo de la comunidad sorda y con requisitos legales y normativos que promueven la igualdad de derechos y la accesibilidad.

También con el desarrollo de este sistema se puede lograr la inclusión de las personas con discapacidad auditiva a la sociedad, pues al facilitar una comunicación más accesible y efectiva. Este sistema no sólo ampliará las oportunidades de interacción y participación de las personas con discapacidad auditiva, sino que también promovería un entorno más inclusivo y equitativo para todos los miembros de la sociedad.

No abordar esta problemática podría llevar a una creciente exclusión y dificultades en el futuro, por lo que la investigación es esencial para garantizar una sociedad más inclusiva y equitativa.

**1.4.1. Delimitación**

El trabajo de tesis que se propone realizar se circunscribe al prototipo de traductor de señas peruanas.

La investigación se centrará en los siguientes aspectos:

* Diseño del sistema: El prototipo se diseñará utilizando inteligencia artificial para traducir la lengua de señas peruana (LSP) al español. Se usara específicamente las señas estáticas excluyendo a las señas dinámicas.
* Validación del prototipo: El prototipo se validará con un grupo de personas con discapacidad auditiva en Perú para evaluar su precisión, facilidad de uso y accesibilidad.

El alcance será a nivel prototipo no tendrá una robustez a nivel comercial por lo que en la tesis no se tocará aspectos del comportamiento del sistema frente a situaciones no esperadas.

Se estima que la investigación tomará alrededor de un año.

#POR QUE SE ELIGE EL MARCO TEORICO

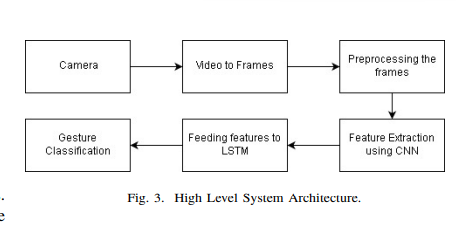
1. **ANTECEDENTES**

En la esta sección, se presenta una breve síntesis de la revisión de la literatura relevante para el desarrollo de la tesis. La inclusión de las personas con discapacidad auditiva ha sido la preocupación, han intentado abordar el problema implementando diversos tipos de soluciones. Para la naturaleza del problema usaremos algoritmos de clasificación donde veremos los más clásicos como Redes Neuronales Convolucionales.

Bantupalli (2018) aborda el desafío de la comunicación entre personas que utilizan lenguaje de señas y aquellas que no lo utilizan, proponiendo una solución basada en una aplicación de visión que traduce el lenguaje de señas a texto. Esta solución se fundamenta en el uso de una red neuronal convolucional (CNN) para la extracción de características temporales de los gestos de la mano, seguida de una red neuronal recurrente (RNN) para modelar las dependencias temporales y predecir el significado de dichos gestos. La metodología empleada incluye también la segmentación de color de la piel para aislar la región de la mano en los videos y técnicas de aumento de datos para mejorar el rendimiento del modelo. Se alcanza una precisión del 93% en la clasificación de 100 signos de lenguaje de señas utilizando la capa Softmax de la CNN y un modelo LSTM, aunque la precisión disminuye al 58% al emplear la capa de agrupación global en lugar de la capa Softmax. Es importante destacar que la precisión tiende a disminuir a medida que aumenta la complejidad del conjunto de datos con más signos de lenguaje de señas.

**Figura 2:**

*Arquitectura en alto nivel del sistema*



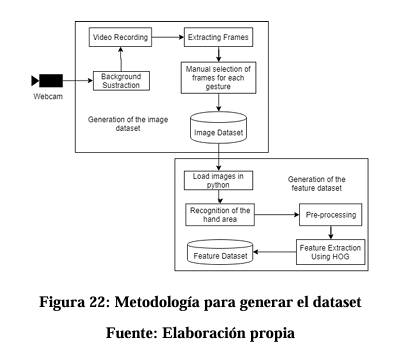
Tomado de *American Sign Language Recognition using Deep Learning and Computer Vision (p. 5) por* Bantupalli, 2018, IEEE

De este artículo puedo usar como base para comprender cómo se aplican las redes neuronales convolucionales y redes neuronales recurrentes en la traducción de lenguaje de señas a texto mediante técnicas de visión por computadora. Además, puedes tomar en cuenta los desafíos mencionados, como la segmentación de la mano y la reducción en la precisión con conjuntos de datos más complejos.

Berru (2019) aborda la carencia de un modelo de clasificación que permita reconocer los gestos de mano estáticos del abecedario de la lengua de señas peruana, con el propósito de utilizar cámaras de baja resolución para reconocer estos gestos y facilitar la enseñanza del lenguaje a las personas sordas. La investigación se justifica por la necesidad de mejorar la calidad de vida de las personas sordas en Perú, dada la lengua de señas como su principal medio de comunicación y la limitación en la enseñanza debido a la falta de recursos y herramientas. La tesis propone un modelo de reconocimiento de gestos de mano que clasifica 24 gestos estáticos del abecedario de la lengua de señas peruana, desarrollado con cámaras de baja resolución y dividido en cinco fases: sustracción de fondo, segmentación de la mano, extracción de características, selección de características y clasificación. Además, se crea un base de datos propio de gestos estáticos del abecedario de la lengua de señas peruana para facilitar la investigación y el desarrollo en esta área. Esta solución contribuye a la enseñanza del español en las escuelas de sordos en Perú y a la investigación en reconocimiento de gestos de mano en el contexto de la lengua de señas peruana.

**Figura 3:**

*Metodología para generar base de datos*



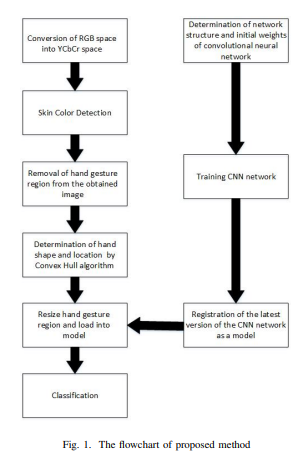
Tomado de *Reconocimiento de gestos estáticos del abecedario de la lengua de señas peruana utilizando cámaras de baja resolución* por Berru (p. 25) UPC

Puedo utilizar esta tesis para entender cómo se aplican técnicas de procesamiento de imágenes y aprendizaje automático en el reconocimiento de gestos de mano en el contexto específico de la lengua de señas peruana y con cámaras de baja resolución. Además, para obtener el dataset y entender cuál fue la metodología que usó para su regeneración.

Taskiran, Killioglu y Kahraman (2018) abordan el desafío de la comunicación entre personas con discapacidad auditiva y aquellas que no conocen el lenguaje de señas, proponiendo un sistema en tiempo real para el reconocimiento del lenguaje de señas americano mediante el aprendizaje profundo. Esta solución tiene como objetivo facilitar la comunicación entre personas con y sin discapacidad auditiva al permitir que aquellos que no conocen el lenguaje de señas puedan comunicarse fácilmente con personas con discapacidad auditiva. El sistema propuesto utiliza una cámara para capturar imágenes en tiempo real y procesarlas para identificar los gestos de la mano en el lenguaje de señas americano.

**Figura 4:**

*Flujograma de la metodología usada*

c

Tomada de *A Real-Time System for Recognition of American Sign Language by using Deep Learning, International Conference on Telecommunications and Signal Processing (TSP)* por Taskiran, Killioglu y Kahraman, 2018, IEEE

Esta referencia puedo utilizarla para entender cómo usar técnicas como el aumento de datos, lo cual genera que mi data sea más rica en cantidad de datos y el entrenamiento pudiese ser más preciso.

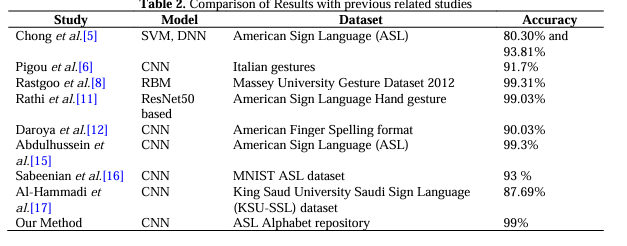
Jaime (2019) investiga el reconocimiento del lenguaje de señas mediante técnicas de aprendizaje profundo, centrándose en el uso de modelos de redes neuronales convolucionales. El documento presenta una revisión de la literatura en el campo, describiendo modelos como Inception y Xception, y comparando su precisión y velocidad de respuesta en la clasificación de imágenes de lenguaje de señas. La solución propuesta consiste en el uso de estos modelos para clasificar imágenes de lenguaje de señas con alta precisión y velocidad de respuesta, con el objetivo de mejorar la comunicación entre personas con discapacidad auditiva y aquellas que no conocen el lenguaje de señas.

Esta referencia puede ser utilizada para la compresión de los modelos basados en redes neuronales convolucionales, Inception y Xception comparando sus métricas y su comportamiento en diversos escenarios.

Krismono, Ahmat y Yuda (2023) proponen un enfoque basado en aprendizaje profundo para el reconocimiento de lenguaje de señas en tiempo real, utilizando un modelo de red neuronal convolucional para clasificar imágenes capturadas por una cámara web. Los autores emplearon un conjunto de datos de lenguaje de señas americano compuesto por 24 categorías, el cual fue preprocesado para eliminar el ruido y estandarizar el tamaño de las imágenes. Los resultados experimentales demostraron que el modelo alcanzó una precisión de prueba del 97.2% y fue capaz de clasificar las señas en tiempo real con una velocidad de respuesta de 25 cuadros por segundo. Los autores concluyen que su enfoque representa una herramienta prometedora para mejorar la comunicación entre personas con discapacidad auditiva y aquellas que no conocen el lenguaje de señas.

**Figura 5:**

*Comparación del modelo propuesto en la tesis con otros modelos*



Tomada de Sign *Language Recognition for Arabic Alphabets Using Transfer Learning Technique* por Krismono, Ahmat y Yuda, 2023, IEEE

Esta referencia puede ser utilizada como una base sólida para el desarrollo de sistemas de reconocimiento de lenguaje de señas en tiempo real, lo que puede ser beneficioso para mejorar la comunicación entre personas sordas y personas que no conocen el lenguaje de señas. Además, los resultados experimentales proporcionados por los autores pueden servir como punto de referencia para evaluar y comparar el rendimiento de otros enfoques en el reconocimiento de gestos de mano en el contexto del lenguaje de señas.

1. **MARCO TEÓRICO**

**3.1 Lenguaje de Señas peruana (LSP)**

El Lenguaje de Señas peruana (LSP) es un sistema lingüístico que va más allá de la comunicación gestual básica. Sus raíces se sumergen en la expresión cultural y en la identidad de la comunidad sorda en Perú. Según el Ministerio de Educación (2014), el lenguaje de señas es un sistema de comunicación producido por el cuerpo y que se percibe a través de la vista.

Esta forma de comunicación no solo se limita a la transmisión de palabras o conceptos, sino que es un vehículo para compartir la riqueza de una cultura visualmente vibrante. Cada gesto, cada movimiento, no solo comunica ideas, sino que también encierra una historia, un contexto cultural y emocional.

Oficialmente solo existe el lenguaje de señas peruana como único lenguaje de señas en el Perú, sin embargo, según Clark (2017) afirma que existen dos nuevas lenguas de señas en el país. Una de ellas la ha designado como la lengua de señas de Sivia (LSSiv), en referencia al pueblo en Arequipa donde se originó, mientras que la otra es la lengua de señas de la Inmaculada (LSSin), supuestamente utilizada por personas sordas que asistieron a la institución educativa Centro de Educación Especial-CEBE La Inmaculada antes de 1960. Según sus descubrimientos, estas lenguas no son comprensibles para los usuarios de la LSP.

La estructura gramatical del LSP es intrincada, donde la posición de las manos, la expresión facial y la fluidez del movimiento contribuyen a una comprensión precisa y completa. Esta lengua no solo se enseña, se vive; se arraiga en la historia y la identidad de las personas sordas, actuando como un elemento esencial de conexión y pertenencia a una comunidad que ha desarrollado su propia narrativa visual.

El conocimiento del LSP no solo implica la capacidad de comunicarse con personas sordas, sino también el reconocimiento y respeto hacia una cultura que ha encontrado en la visualidad una forma única y expresiva de existir en un mundo predominantemente auditivo. Su dominio no solo es un acto lingüístico, es un reconocimiento de la diversidad humana y un paso vital hacia la inclusión y la igualdad de oportunidades para todos en la sociedad peruana.

**Figura 6:**

*Lenguaje de señas peruana*



Tomada de Discapacidad Auditiva por Salma Jimenez, 2022, ISSUU

Existen 2 tipos gestos en el lenguaje de señas peruana, gestos estáticos y gestos dinámicos.

Los gestos estáticos son aquellos que implican una forma de mano que se mantiene relativamente fija sin movimiento. En el contexto del LSP, los gestos estáticos pueden representar letras del alfabeto, números, palabras y conceptos específicos. Algunos ejemplos de gestos estáticos en el LSP incluyen:

* Letras del alfabeto: Cada letra del alfabeto se representa mediante un gesto de mano específico. Por ejemplo, la letra "A" puede representarse extendiendo los dedos índice y pulgar y formando un círculo con el resto de los dedos.
* Números: Los números del 0 al 9 pueden representarse utilizando gestos específicos de mano que indican la cantidad correspondiente.
* Palabras o conceptos específicos: Algunas palabras o conceptos específicos pueden representarse mediante gestos de mano únicos en el LSP. Por ejemplo, el gesto para "casa" puede implicar formar un techo con las manos.

Gestos Dinámicos:

Los gestos dinámicos son aquellos que implican movimientos de las manos, los brazos, la cara u otras partes del cuerpo. Estos gestos pueden transmitir información adicional sobre el significado, la intensidad o el contexto de un mensaje en el LSP. Algunos ejemplos de gestos dinámicos en el LSP incluyen:

* Movimientos de mano: Los movimientos de mano pueden incluir gestos como mover la mano de un lugar a otro para indicar dirección o movimiento, o agitar la mano para expresar saludo o despedida.
* Expresiones faciales: Las expresiones faciales son una parte importante del LSP y pueden cambiar el significado de un gesto estático o dinámico. Por ejemplo, una ceja levantada puede indicar interrogación, mientras que una sonrisa puede indicar satisfacción o felicidad.
* Expresiones corporales: Los gestos dinámicos también pueden implicar movimientos del cuerpo, como inclinarse hacia adelante para indicar atención o retroceder para indicar rechazo.

**3.2 Discapacidad Auditiva**

La discapacidad auditiva, en sus diversas formas y grados, no solo limita la capacidad de percibir sonidos, sino que también puede afectar profundamente la interacción con el entorno y las relaciones sociales. Aquellos que experimentan esta condición pueden encontrarse en un mundo donde el lenguaje oral es predominante, lo que puede generar barreras significativas para la comunicación y la participación plena en la sociedad. Según Forner (2021), la discapacidad auditiva se define como la pérdida o anormalidad de una función anatómica y/o fisiológica del sistema auditivo, lo que resulta en una discapacidad para oír y un déficit en el acceso al lenguaje oral. Esta pérdida auditiva afecta tanto a la comunicación en el lenguaje oral como a la percepción de los sonidos del entorno. Cuanto mayor sea el grado de discapacidad auditiva, mayor será la desconexión con el entorno, lo que impacta negativamente en la calidad de vida.

La discapacidad auditiva puede presentarse desde el nacimiento o desarrollarse a lo largo de la vida debido a diversos factores, como infecciones, lesiones, exposición prolongada a ruidos fuertes o el proceso natural de envejecimiento. La adaptación a esta condición es única para cada individuo y puede implicar el uso de diversas estrategias para comunicarse. La lengua de señas, un medio visual y expresivo de comunicación, no solo es un idioma en sí mismo, sino también una ventana hacia la cultura y la comunidad sorda. La lectura de labios, los implantes cocleares y otros dispositivos de asistencia auditiva también desempeñan un papel vital al proporcionar herramientas que permiten a las personas con discapacidad auditiva acceder al mundo sonoro que los rodea.

Superar las barreras comunicativas no solo implica el acceso a estas herramientas, sino también la sensibilización y la creación de entornos inclusivos que fomenten la comprensión y el respeto hacia la diversidad auditiva. La promoción de la accesibilidad y el entendimiento de las necesidades individuales de las personas con discapacidad auditiva son pasos cruciales hacia una sociedad más inclusiva, donde todos puedan participar activamente y ser valorados por sus contribuciones, independientemente de su capacidad auditiva.

**Figura 7:**

*Elaboración propia*



**3.3 Machine Learning**

El Machine Learning, como parte integral de la inteligencia artificial, representa una revolución en la capacidad de las computadoras para analizar y procesar datos de manera inteligente y autónoma. Este campo se enfoca en desarrollar algoritmos y modelos capaces de aprender patrones complejos a partir de datos y mejorar continuamente su desempeño sin una programación específica para cada tarea.

Según AWS (2018), el machine learning se define como la ciencia de desarrollo de algoritmos y modelos estadísticos que utilizan los sistemas de computación para llevar a cabo tareas sin instrucciones explícitas, basándose en patrones e inferencias. Estos sistemas de computación emplean algoritmos de machine learning para procesar grandes cantidades de datos históricos e identificar patrones, lo que les permite generar resultados con mayor precisión a partir de un conjunto de datos de entrada. Por ejemplo, los científicos de datos pueden entrenar una aplicación médica para diagnosticar el cáncer utilizando imágenes de rayos X y millones de diagnósticos correspondientes almacenados.

Según SAP (2019), el machine learning se define como un subconjunto de la inteligencia artificial (IA) que se centra en enseñar a las computadoras a aprender de los datos y mejorar con la experiencia, en lugar de ser explícitamente programadas para hacerlo. En el ámbito del machine learning, los algoritmos se entrenan para descubrir patrones y correlaciones en grandes conjuntos de datos, con el objetivo de tomar las mejores decisiones y proyecciones en base a dicho análisis. Además, las aplicaciones de machine learning tienden a mejorar con el uso y a volverse más precisas a medida que tienen acceso a más datos.

En el contexto de un traductor de señas, el Machine Learning se convierte en un aliado invaluable. La capacidad inherente de estos algoritmos para reconocer patrones en conjuntos de datos extensos les permite captar los matices de los gestos específicos del Lenguaje de Señas peruana. La adaptabilidad del Machine Learning es clave, ya que puede aprender de variaciones en los movimientos, expresiones y estilos individuales de señas, lo que posibilita una traducción más precisa y contextualizada.

La aplicación del Machine Learning en el reconocimiento y traducción de gestos del lenguaje de señas no solo implica la identificación de movimientos individuales, sino también la comprensión de la gramática gestual y la interpretación de la intención detrás de cada signo. Al emplear conjuntos de datos amplios y diversos, estos sistemas pueden mejorar su precisión y eficiencia, acercándose cada vez más a una comunicación fluida y natural entre personas sordas y oyentes, eliminando las barreras lingüísticas y fomentando la inclusión.

**3.4 Computer Vision**

La Computer Vision, como rama fundamental de la inteligencia artificial, va más allá de la simple percepción visual; se trata de dotar a las máquinas con la capacidad de entender y procesar imágenes o vídeos de manera similar a la percepción humana. Este campo de estudio se adentra en el reconocimiento de patrones, formas, colores y movimientos, permitiendo a los sistemas informáticos no solo identificar elementos visuales, sino también comprender su contexto y significado.

Según SAS (2018), Computer Vision se define como un campo de la inteligencia artificial que capacita a las computadoras para interpretar y comprender el mundo visual. Mediante el uso de imágenes digitales de cámaras y vídeos, así como modelos de aprendizaje profundo, las máquinas pueden identificar y clasificar objetos con precisión, y posteriormente reaccionar ante lo que perciben visualmente.

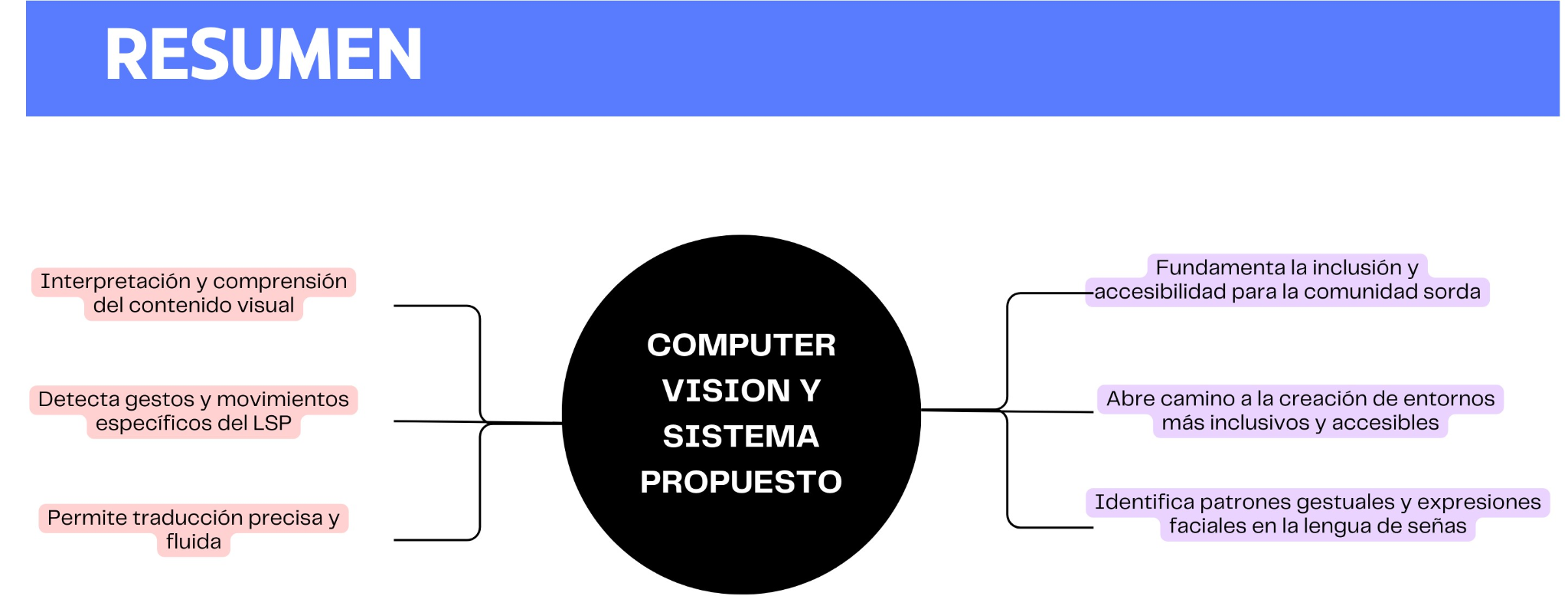
Según IBM (2019), la visión artificial se define como un campo de la inteligencia artificial (IA) que capacita a los ordenadores y sistemas para extraer información significativa a partir de imágenes digitales, videos y otras entradas visuales, y tomar medidas o realizar recomendaciones en función de esa información. Mientras que la IA permite a los ordenadores pensar, la visión artificial les permite ver, observar y comprender.

En el ámbito del desarrollo de un traductor de señas, la Computer Vision juega un rol esencial. Sus algoritmos y técnicas permiten a las máquinas detectar con precisión los gestos específicos, expresiones faciales y movimientos característicos utilizados en el Lenguaje de Señas peruana (LSP). Esta capacidad de análisis visual no se limita a la identificación de formas o movimientos aislados, sino que abarca la comprensión de la estructura gramatical y la sintaxis visual inherente a la lengua de señas.

Al analizar y entender la secuencia y el contexto de los gestos, la Computer Vision facilita la interpretación exacta de las expresiones visuales utilizadas en el LSP. Esto permite una traducción más precisa y contextualizada, contribuyendo directamente a la eliminación de barreras comunicativas entre personas sordas y oyentes. Asimismo, al promover una comunicación fluida y comprensible, la Computer Vision desempeña un papel crucial en la construcción de entornos inclusivos y accesibles para la comunidad sorda, respaldando su plena participación en la sociedad.

**Figura 8:**

*Elaboración propia*



1. **MARCO METODOLÓGICO**
   1. **Tipo de investigación**

El tipo de investigación que se realizará en esta tesis será una investigación aplicada debido a que se busca solucionar el problema de entendimiento de las personas con discapacidad auditiva que usan lenguaje de señas peruana por parte de personas que no conocen este lenguaje usando un traductor automático de lenguaje de seña peruana a voz, lo cual facilita el entendimiento de mensaje realizados por personas con discapacidad auditiva a personas que desconocen el este lenguaje.

* 1. **Método de investigación**

**Para el objetivo 1:**

1. Definir los requisitos que usará el modelo para lograr la comunicación con personas con discapacidad auditiva.

* Revisar la literatura: Explorar papers y tesis relacionadas para entender los requisitos clave identificados previamente en soluciones propuestas.
* Analizar necesidades específicas: Entender las necesidades de la comunidad sorda en cuanto a comunicación y comprensión.
* Establecer requisitos técnicos: Definir requisitos de reconocimiento de gestos, comprensión gramatical y contexto de la lengua de señas.

**Para el objetivo 2:**

1. Definir hiperparámetros para lograr la comunicación con personas con discapacidad auditiva.

* Estudiar modelos existentes: Analizar modelos de aprendizaje automatico aplicados a traducción de señas para comprender los hiperparámetros más efectivos.
* Experimentar y ajustar: Definir y ajustar hiperparámetros como tasa de aprendizaje, número de capas, funciones de activación, entre otros.
* Optimizar modelos: Evaluar el impacto de los hiperparámetros en el rendimiento del modelo y ajustarlos para mejorar la precisión.

**Para el objetivo 3:**

1. Desarrollar un modelo de aprendizaje automático para la traducción de señas peruanas.

* Implementar la arquitectura del modelo: Escribir el código para la estructura del modelo de traducción de señas peruana.
* Entrenar el modelo: Utilizar conjuntos de datos adecuados para entrenar el modelo en la traducción de gestos del LSP.
* Desarrollar el prototipo: Integrar el modelo en un prototipo funcional de traductor de señas.

**Para el objetivo 4:**

1. Validar la precisión del modelo de traducción de señas peruanas.

* Definir métricas de evaluación: Establecer métricas como precisión, recall y F1-score para evaluar el desempeño del modelo.
* Evaluar inicialmente: Evaluar el modelo con un conjunto de datos de prueba para obtener métricas iniciales.
* Ajustar y mejorar: Reajustar el modelo basado en los resultados de las métricas para mejorar su precisión.
* Validar finalmente: Realizar pruebas exhaustivas con datos de validación para garantizar la fiabilidad del modelo y su precisión en la traducción de señas peruanas.
  1. **Población y muestra**

La población identificada es el conjunto total de personas que usan el lenguaje de señas peruana, sin embargo, la muestra vendría ser la cantidad de personas que ayudara a entrenar y validar la eficiencia del sistema. También se requerirá un experto que sirva como nexo entre las personas que se comunican mediante lenguaje de señas y las personas que no conocen este lenguaje. Con respecto al lenguaje como tal se considera el uso solo de las señas estáticas excluyendo el uso de las señas dinámicas, esto por simplicidad del modelo. El estudio será en base a una data ya existente.

El dataset que se usará es el obtenido de la tesis “Reconocimiento de gestos estáticos del abecedario de la lengua de señas peruana utilizando cámaras de baja resolución” que consta de 150 imágenes por cada señal estática.

* 1. **Técnicas e instrumentos para recolección y análisis de datos**

Se empleará algoritmos y técnicas de visión por computador para filtrar o mejorar las imágenes antes de ingresar al modelo de forma que esta pueda tener un mejor desempeño tales como ecualización o contraste de manera que los objetos sean más claros para la computadora.

Se usará modelos de redes neuronales convolucionales para lograr traducción de las imágenes a texto y posteriormente a voz.

Las herramientas específicas que se han pensado para el sistema son:

Para el procesamiento de imágenes:

* **OpenCV:** OpenCV es una de las bibliotecas más utilizadas para el procesamiento de imágenes y visión por computadora en Python, C++, Java, y otras plataformas.

Ofrece una amplia gama de funciones para el procesamiento de imágenes, incluyendo filtrado, transformaciones geométricas, detección de características, seguimiento de objetos, y mucho más.

Es de código abierto y cuenta con una gran comunidad de usuarios y desarrolladores que contribuyen con nuevas características y mejoras.

* **Pillow:** Pillow es una biblioteca de procesamiento de imágenes en Python que proporciona una interfaz fácil de usar para abrir, manipular y guardar imágenes en varios formatos. Es una bifurcación de la biblioteca PIL y ofrece una API más moderna y mejorada. Es ideal para tareas básicas de procesamiento de imágenes, como redimensionamiento, rotación, recorte, conversión de formato, entre otros.

Para el desarrollo del modelo de redes neuronales convolucionales:

* **TensorFlow / Keras:** TensorFlow es una de las bibliotecas más populares para el aprendizaje profundo, y Keras proporciona una interfaz de alto nivel fácil de usar para construir y entrenar modelos de redes neuronales. Puedes utilizar TensorFlow junto con Keras para diseñar y entrenar tu modelo de traducción de lenguaje de señas. TensorFlow ofrece una amplia variedad de herramientas y funciones para el procesamiento de imágenes y el aprendizaje profundo, lo que lo hace ideal para tu aplicación.

Para el convertir el texto a voz:

* **Pyttsx3:** permite a los desarrolladores integrar fácilmente la síntesis de voz en sus aplicaciones de Python, lo que les permite agregar funcionalidades de texto a voz de manera rápida y sencilla. Ofrece opciones de configuración para ajustar el tono, la velocidad y otros aspectos de la voz generada, y es compatible con múltiples motores de síntesis de voz disponibles en diferentes sistemas operativos.

**Matriz de consistencia**

|  |  |  |
| --- | --- | --- |
|  | Problema | Objetivo |
| General | No se conoce cómo lograr la comunicación entre personas que usan lenguaje de señas peruana y personas que no usan lenguaje de señas peruana | Desarrollar un sistema traductor de señas peruana automático que permita el entendimiento de personas que no conocen el lenguaje de señas peruana a personas con discapacidad auditiva que se comunican mediante el lenguaje de señas peruana |
| Específico | No se conocen los requisitos específicos que permitan lograr desarrollar un modelo que permita la traducción de lenguaje de señas peruana | Definir los requisitos que usará el modelo para poder lograr la comunicación con personas con discapacidad auditiva |
|  | No se conoce los hiperparametros que se usarán en el modelo | Definir hiperparametros que usará el modelo para poder lograr la comunicación con personas con discapacidad auditiva. |
|  | No se conoce cómo traducir lenguaje de señas peruana a voz | Desarrollar un modelo de aprendizaje automático para la traducción de señas peruanas. |
|  | No se conoce cómo asegurar la confiabilidad de que se detecten correctamente los gestos del lenguaje de señas peruanas en las imágenes para validar la precisión del modelo de traducción de señas peruanas. | Validar la precisión del modelo de traducción de señas peruanas. |

**5. RESULTADOS EXPERIMENTALES**

**6. CONCLUSIONES**

**7. BIBLIOGRAFÍA**

Aural Centro Auditivo (17 de mayo de 2021). La discapacidad auditiva: concepto y soluciones.<https://www.aural.es/blog/que-es-y-en-que-consiste-la-discapacidad-auditiva>

AWS (s.f.) ¿Qué es el aprendizaje automático?<https://aws.amazon.com/es/what-is/machine-learning/>

Bantupalli, K. and Xie, Y. (2018) *American Sign Language Recognition using Deep Learning and Computer Vision*, 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4896-4899. <https://ieeexplore.ieee.org/document/8622141>

Berrú, B. (2019) Reconocimiento de gestos estáticos del abecedario de la lengua de señas peruana utilizando cámaras de baja resolución. Tesis de pregrado, Departamento de sistemas, UPC. <https://repositorioacademico.upc.edu.pe/handle/10757/628221>

Clark, Brenda (2017). Sign Language Varieties in Lima, Peru.Sign Language Studies, 17(2): 222-264. <https://eric.ed.gov/?id=EJ1270234>

CONADIS (2016) Informe Temático N° 5 “Situación de las Personas con Discapacidad Auditivaen el Perú”. <https://es.scribd.com/document/347044586/Informe-Tematico-N-5-Situacion-de-las-Personas-con-Discapacidad-Auditiva-en-el-Peru>

Defensoría del Pueblo Perú (24 de octubre de 2020). Defensoría del Pueblo: debe facilitarse el aprendizaje de la lengua de señas peruana y promover la identidad lingüística y cultural de las personas sordas. <https://www.defensoria.gob.pe/defensoria-del-pueblo-debe-facilitarse-el-aprendizaje-de-la-lengua-de-senas-peruana-y-promover-la-identidad-linguistica-y-cultural-de-las-personas-sordas/>

Del Valle, J. (2021). Lengua de señas peruana: Un recorrido por su historia, estructura y desafíos. Revista Peruana de Psicología, 28(3), 547-562. <https://lucidez.pe/la-lengua-de-senas-es-un-desafio1/>

Diccionario LSP al español (s. f.) Pontificia Universidad Católica del Perú. <https://diccionariolsp.pucp.edu.pe/search-by-text>

Encuesta Nacional de Hogares (2022). Estadísticas modulo vivienda <https://www.inei.gob.pe/media/encuestas/documentos/enaho/2022/02_Enaho01A_2022_Educacion_Salud_y_Empleo.pdf>

IBM. (s. f.). Qué es Visión por Computador?. <https://www.ibm.com/es-es/topics/computer-vision>

IBM. (s. f.). ¿Qué es Machine Learning? <https://www.ibm.com/es-es/topics/machine-learning>

International Journal of Computer Vision. (2023). Scimago Journal & Country Rank. Recuperado de [https://www.scimagojr.com](https://www.scimagojr.com/)

Jaime, A. (2022). Classification of Sign-Language Using Deep Learning - A Comparison between Inception and Xception models – 2022. International Journal of Academic Engineering Research (IJAER), Gaza, Palestina. <https://philarchive.org/archive/ABUCOS-4>

Jimenez, S. (2022) Discapacidad Auditiva. ISSUU. https://issuu.com/salmajl/docs/revista\_discapacidades\_/s/28454758

Krismono, Ahmat y Yuda (2023) Sign Language Recognition for Arabic Alphabets Using Transfer Learning Technique. 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9054420/>

LeeCun, Y., Bengio, Y. y Hinton, G. (2017). Deep Learning for Computer Vision. Pyimagesearch. <https://pyimagesearch.com/2017/01/13/table-of-contents-deep-learning-for-computer-vision-with-python/>

Leek, J. (2018) Probability and Statistics for Computer Science. Springer. <https://link.springer.com/book/10.1007/978-3-319-64410-3>

Lorenzon, E. (2005) Teoría General de Sistemas Aplicada. Editorial de la Universidad de la Plata. <https://libros.unlp.edu.ar/index.php/unlp/catalog/book/1505>

Luyo Tuesta, W y Juarez Poémape, C (2018). System to facilitate the labor insertion of people with hearing and / or vocal problems through a Peruvian Sign Language Translator System using Microsoft Kinect technology. <https://www.urp.edu.pe/pdf/id/7415/n/system-to-facilitate-the-labor-insertion-of-people-with-hearing-and-or-vocal-problems-through-a-peruvian-sign-language-translator-system-using-microsoft-kinect-technology>

Ministerio de Educación (2014) Lengua de señas peruana guía para el aprendizaje de la lengua de señas peruana. Ministerio de educación.<https://hdl.handle.net/20.500.12799/5545>

Rodríguez, S. (2018). Discapacidad auditiva. Universidad de La Laguna.<https://riull.ull.es/xmlui/bitstream/handle/915/11441/Discapacidad%20auditiva.pdf>

Sampieri, H., Fernández, C. y Baptista, P. (2014). Metodología de la Investigación. 6ta Edición. Editorial McGRAW-HILL. https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y %20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf

SAP (s.f.) ¿Qué es machine learning?<https://www.sap.com/latinamerica/products/artificial-intelligence/what-is-machine-learning.html>

SmartPanel. (2019, 9 octubre). Algoritmos de machine learning. SmartPanel. <https://www.smartpanel.com/algoritmos-de-machine-learning>

Smith, K. (2021). Real-time sign language recognition using a deep learning approach - 2022. International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece. <https://www.researchgate.net/publication/377631034_Real_Time_Static_and_Dynamic_Sign_Language_Recognition_using_Deep_Learning>

Taskiran, N., Killioglu, M. y Kahraman, N. (2018). A Real-Time System for Recognition of American Sign Language by using Deep Learning, International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece. <https://ieeexplore.ieee.org/document/8441304>