

Listas de contenidos disponibles en [CienciaDirecta](#)

# Sistemas Inteligentes con Aplicaciones

revista Página de inicio:[www.elsevier.com/locate/iswa](http://www.elsevier.com/locate/iswa)

## Métodos de aprendizaje automático para el reconocimiento del lenguaje de señas: revisión y análisis críticos

IA Adeyanju, O Bello, MA Adegbeye<sup>a,C,\*</sup><sup>a</sup>Departamento de Ingeniería Informática, Universidad Federal, Oye-Ekiti, Nigeria<sup>b</sup>Departamento de Ingeniería Informática, Universidad Estatal de Ekiti, Ado-Ekiti, Nigeria<sup>c</sup>Grupo de Comunicaciones y Sistemas Autónomos, Escuela de Ingeniería, Universidad Robert Gordon, Aberdeen AB107GJ, Reino Unido

### información del artículo

*Historial del artículo:*

Recibido el 26 de mayo de 2021  
 Revisado el 12 de octubre de 2021  
 Aceptado el 16 de noviembre de 2021  
 Disponible en línea el 8 de diciembre de 2021

*Palabras clave:*

Inteligencia artificial  
 Visión artificial  
 Análisis biométrico  
 Sistemas inteligentes sordomudos para personas con discapacidad auditiva  
 base de datos Scopus  
 Reconocimiento de lenguaje de señas VOSviewer

### abstracto

El lenguaje de señas es una herramienta esencial para cerrar la brecha de comunicación entre las personas normales y las personas con discapacidad auditiva. Sin embargo, la diversidad de más de 7000 lenguajes de señas actuales con variabilidad en la posición del movimiento, la forma de la mano y la posición de las partes del cuerpo hacen que el reconocimiento automático del lenguaje de señas (ASLR) sea un sistema complejo. Para superar tal complejidad, los investigadores están investigando mejores formas de desarrollar sistemas ASLR para buscar soluciones inteligentes y han demostrado un éxito notable. Este artículo tiene como objetivo analizar las investigaciones publicadas sobre sistemas inteligentes en el reconocimiento de lengua de signos durante las últimas dos décadas. De la base de datos Scopus se extraen y analizan un total de 649 publicaciones relacionadas con el apoyo a la decisión y los sistemas inteligentes de reconocimiento de lengua de signos (SLR). Las publicaciones extraídas se analizan utilizando el software biométrico VOSviewer para (1) obtener las distribuciones temporales y regionales de las publicaciones, (2) crear las redes de cooperación entre afiliaciones y autores e identificar instituciones productivas en este contexto. Además, se presentan revisiones de técnicas para el reconocimiento de lenguaje de señas basado en la visión. Se discuten varias técnicas de extracción y clasificación de características utilizadas en SLR para lograr buenos resultados. La revisión de la literatura presentada en este artículo muestra la importancia de incorporar soluciones inteligentes en los sistemas de reconocimiento de lengua de signos y revela que los sistemas inteligentes perfectos para el reconocimiento de lengua de signos siguen siendo un problema abierto. En general, se espera que este estudio facilite la acumulación de conocimientos y la creación de cámaras SLR inteligentes y proporcione a los lectores,

© 2021 El(los) autor(es). Publicado por Elsevier Ltd.

Este es un artículo de acceso abierto bajo la licencia CC BY (<http://creativecommons.org/licenses/by/4.0/>)

### 1. Introducción

La comunicación es una herramienta esencial en la existencia humana. Es una forma fundamental y eficaz de compartir pensamientos, sentimientos y opiniones. Sin embargo, una fracción sustancial de la población mundial carece de esta capacidad (El-Din y El-Ghany, 2020). Muchas personas sufren pérdida de audición, problemas del habla o ambos. La incapacidad parcial o total para oír en uno o ambos oídos se conoce como pérdida auditiva. Por otro lado, el mudo es una incapacidad que dificulta el habla y hace que las personas afectadas no puedan hablar. Si la sordera ocurre durante la infancia, su capacidad de aprendizaje de idiomas puede verse obstaculizada y da como resultado un deterioro del lenguaje, también conocido como mutismo auditivo. Estas dolencias forman parte de las discapacidades más comunes

\*Autor para correspondencia en: Grupo de Sistemas Autónomos y Comunicaciones, Facultad de Ingeniería, Universidad Robert Gordon, Aberdeen AB107GJ, Reino Unido.

Correos electrónicos:[bello.oluwaseyi@eksu.edu.ng](mailto:bello.oluwaseyi@eksu.edu.ng) (OO Bello),  
[m.adegboye@rgu.ac.uk](mailto:m.adegboye@rgu.ac.uk) (MA Adegbeye).

mundial (Hasan et al., 2020). El informe estadístico de niños con discapacidades físicas durante la última década revela un aumento en el número de recién nacidos que nacen con un defecto de discapacidad auditiva y crea una barrera de comunicación entre ellos y el resto del mundo (Krishnaveni et al., 2019).

Según el informe de la Organización Mundial de la Salud (OMS), el número de personas afectadas por discapacidad auditiva en 2005 fue de aproximadamente 278 millones en todo el mundo (Savur & Sahin, 2016). Diez (10) años más tarde, este número saltó a 360 millones, un incremento de aproximadamente el 14% (Savur & Sahin, 2016). Desde entonces, el número ha ido aumentando exponencialmente. El último informe de la OMS reveló que 466 millones de personas padecían pérdida de audición en 2019, lo que representa el 5 % de la población mundial, de los cuales 432 millones (o el 83 %) son adultos, y 34 millones (el 17 %) son adultos. niños (Bin et al., 2019; Hisham y Hamouda, 2019; Saleh e Isa, 2020). La OMS también estimó que el número se duplicaría (es decir, 900 millones de personas) para 2050 (El-Din y El-Ghany, 2020). En estas personas sordomudas de rápido crecimiento, existe la necesidad de romper la comunicación.

barrera de comunicación que afecta negativamente la vida y las relaciones sociales de las personas sordomudas.

Los lenguajes de señas se utilizan como un medio principal de comunicación por personas sordas y con problemas de audición en todo el mundo (Izzah y Suciati, 2014). Es la forma más potente y eficaz de cerrar la brecha de comunicación y la interacción social entre ellos y las personas capaces. Los intérpretes de lenguaje de señas ayudan a resolver la brecha de comunicación con las personas con discapacidad auditiva al traducir el lenguaje de señas a palabras habladas y viceversa. Sin embargo, los desafíos de emplear intérpretes son la estructura flexible de los lenguajes de señas combinada con un número insuficiente de intérpretes expertos en lenguaje de señas en todo el mundo (Kudrinko et al., 2021). Según la Federación Mundial de Sordos, más de 300 lenguajes de señas son utilizados por más de 70 millones en todo el mundo (Rastgoo et al., 2021b). Por lo tanto, la necesidad de un sistema basado en tecnología que pueda complementar a los intérpretes de lenguaje de señas convencionales.

El lenguaje de señas implica el uso de la parte superior del cuerpo, como gestos con las manos (Gupta y Rajan, 2020), expresión facial (Chowdhry et al., 2013), lectura de labios (Cheok et al., 2019), movimientos de cabeza y posturas corporales para difundir información (Butt et al., 2019); (Rastgoo et al., 2021b;Lee et al., 2021a). Las técnicas clave para el reconocimiento del lenguaje de señas son las modalidades de detección portátiles y basadas en la visión, como los guantes sensoriales. Varios investigadores han propuesto sistemas de reconocimiento de lenguaje de señas basados en estos enfoques (Ionescu et al., 2005; Yu et al., 2010; Li et al., 2015; Sonkusare et al., 2015; Bobic et al., 2016; Islam et al., 2017; Islam et al., 2017; Saha et al., 2018; Rastgoo et al., 2021a; Xu et al., 2021). El guante emplea los sensores mecánicos u ópticos adheridos al guante usado por el usuario y convierte los movimientos de los dedos en señales eléctricas para determinar la postura de la mano para el reconocimiento. En un enfoque basado en la visión, se estiman las características correspondientes a las palmas, la posición de los dedos y los ángulos de las articulaciones, que luego se utilizan para realizar el reconocimiento. Este método requiere la adquisición de imágenes o videos de los letreros a través de una cámara y el procesamiento mediante técnicas de procesamiento de imágenes.

El reciente avance de la Inteligencia Artificial (IA) en el reconocimiento del lenguaje de señas ha allanado el camino para que las comunidades de investigación apliquen la IA en las operaciones de interpretación de señas. Hay varias referencias excelentes sobre sistemas inteligentes para el reconocimiento de lengua de signos (Admasu & Raimond, 2010; Deriche et al., 2019; Zapata et al., 2019; Canción et al., 2021; Lee et al., 2021a, 2021b; Gao et al., 2021). Más recientemente, se ha prestado más atención al sistema SLR basado en inteligencia, ya que ahora se aplica en muchas aplicaciones. Esto incluye la robótica (Ryumin et al., 2019), servicios de interpretación, sistemas de reconocimiento de varias personas en tiempo real, juegos, entornos de realidad virtual, comunicaciones en lenguaje natural, seguimiento manual en línea de la comunicación humana en entornos de escritorio e interacción hombre-máquina (Deng et al., 2017; Supančič et al., 2018; Wadhawan y Kumar, 2020; Rastgoo et al., 2021b). A pesar del progreso histórico de la investigación y los notables logros obtenidos en los sistemas inteligentes de reconocimiento de lenguaje de señas, todavía existe un gran potencial en la creación de soluciones inteligentes para los sistemas de reconocimiento de lenguaje de señas. Por lo tanto, este estudio tiene como objetivo proporcionar una revisión sistemática y exhaustiva de los trabajos de investigación publicados en el campo de los sistemas inteligentes de reconocimiento de lenguaje de señas para obtener información sobre la aplicación de los sistemas inteligentes y de apoyo a la decisión en este contexto. Los objetivos principales de este estudio incluyen 1. Análisis de la investigación publicada sobre sistemas inteligentes en el reconocimiento de lenguaje de señas mediante el análisis bibliométrico de 649 publicaciones extraídas de la base de datos Scopus durante el período de dos décadas (2001-2021). 2. Brinda una revisión integral en el contexto de los sistemas de reconocimiento de lenguaje de señas con un enfoque explícito en el apoyo a la toma de decisiones y las tecnologías de sistemas expertos durante las últimas dos décadas. 3. Resaltar temas abiertos y posibles áreas de investigación para futuras consideraciones.

El resto del documento está organizado de la siguiente manera. Sección 2 presenta la metodología para la revisión sistemática y bibli-

análisis métrico de los artículos seleccionados para los estudios. Sección 3 proporciona una revisión de las técnicas utilizadas en los sistemas de reconocimiento de lenguaje de señas. Esto incluye técnicas utilizadas para la adquisición de datos, preprocesamiento, segmentación, extracción de características y clasificación. En Sección 4, se presenta una revisión exhaustiva de los algoritmos de apoyo a la toma de decisiones y de sistemas inteligentes utilizados en el reconocimiento del lenguaje de señas. Las conclusiones y la discusión sobre las necesidades futuras se presentan en Sección 5.

## 2. Análisis bibliométrico

Esta sección emplea una metodología de análisis bibliométrico para analizar las investigaciones publicadas sobre sistemas inteligentes en el reconocimiento de lengua de signos durante las últimas dos décadas. Esto se hace para discernir y comprender las tendencias históricas y los patrones de publicación a lo largo del tiempo por revistas, regiones y cooperación entre instituciones y organizaciones. A lo largo de los años, ha habido un creciente interés de investigación en ASLR. Por lo tanto, investigar la tendencia general de la investigación y la nueva dirección de la investigación será convincente en este campo. Se buscó y analizó un análisis bibliométrico exhaustivo del estado de la tendencia de investigación relacionada con el apoyo a la toma de decisiones y los sistemas inteligentes sobre el reconocimiento del lenguaje de señas desde 2001 hasta 2021. Esto es para identificar la posible brecha de investigación y resaltar los límites del conocimiento durante este período de tiempo. La elección de dos décadas de investigación fue para dar una visión más amplia de cómo ha progresado la investigación durante estos períodos para permitir a los investigadores comprender las actividades del patrón de investigación y sus características. Esta información es útil ya que aclara las actividades científicas con respecto a las tendencias de publicaciones en los países, instituciones, revistas y autores.

*Las palabras clave utilizadas para la recopilación de datos son "Reconocimiento de lenguaje de señas", "Sistemas inteligentes", "Inteligencia artificial", "Sistema de apoyo a la decisión", "Aprendizaje automático", "Red neuronal", "Sistema experto", "Sistema difuso" y "Sistemas basados en el conocimiento". La estrategia de búsqueda para recuperar la publicación durante el período de tiempo seleccionado es la siguiente:*

*(“lenguaje de señas Y reconocimiento”) Y (máquina Y aprendizaje) O (artificial E inteligente) Y TIPO DE DOCUMENTO: (artículo de investigación) Y AÑO DE PUBLICACIÓN:(>2000) ÍNDICE:(base de datos Scopus).* La introducción de comillas (“ ”) en el campo de búsqueda con el objetivo de recuperar la palabra clave o frase de búsqueda exacta. Esto es para identificar los artículos que se centran en el tema del tema de estudio. El análisis se basó en países, instituciones, fuentes de distribución de publicaciones, citas y coincidencia de palabras clave de búsqueda. Estos indicadores se seleccionaron como punto de referencia para destacar los países, las instituciones y los principales expertos a la vanguardia de ASLR. También guía en la exploración de redes de colaboración a través de estos indicadores. Las dos principales herramientas para realizar análisis bibliométricos son el software VOSviewer y CitNetEmployer. El VOSviewer se empleó en este estudio porque se enfoca en el análisis a nivel de la publicación agregada, mientras que CitNetEmployer enfatiza las publicaciones individuales.

### 2.1. Procedimiento de análisis bibliométrico

El procedimiento para el análisis bibliométrico se muestra en Figura 1. De acuerdo con el proceso, los primeros dos pasos involucran la recopilación de literatura extensa de la base de datos de Scopus el 20 de marzo de 2021. La razón para seleccionar la base de datos de Scopus fue que sus revistas indexadas son reconocidas mundialmente como influyentes y como los mejores lugares para resultados de investigación modernos y eminentes. sobre sistemas de reconocimiento de lengua de signos. Por lo tanto, es posible buscar y obtener una proporción significativa de artículos publicados en este contexto.

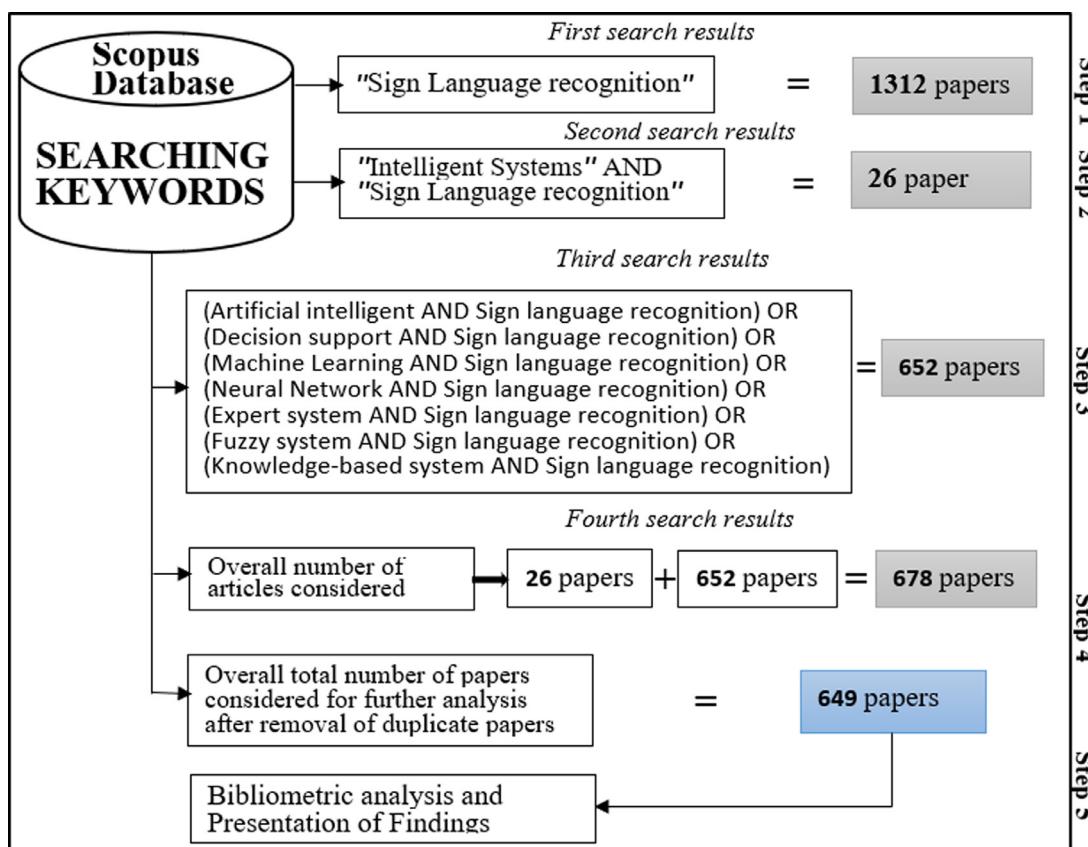


Figura 1. Diagrama de flujo de la recopilación de publicaciones y resultados de investigación de análisis en la base de datos Scopus.

Se utilizaron diferentes palabras clave de forma colectiva para extraer artículos de investigación publicados sobre sistemas inteligentes de reconocimiento de lenguaje de señas a lo largo del proceso de búsqueda. Los autores primero buscaron en la base de datos de Scopus usando la palabra clave "Reconocimiento del lenguaje de señas", y la búsqueda resultó en 1312 artículos de revistas publicados entre 2001 y 2021.

Dado que el enfoque de este estudio es sobre los sistemas inteligentes en el reconocimiento de lenguaje de señas. Despues de realizar el primer paso de búsqueda sobre el reconocimiento general del lenguaje de señas, los autores repitieron este proceso refinando la búsqueda usando palabras clave en el paso 2 ("Sistemas inteligentes" y "reconocimiento de lengua de signos"). Esta búsqueda dio como resultado 26 artículos de revistas que se centran en los sistemas de reconocimiento de signos basados en inteligencia. Sin embargo, el resultado de la búsqueda de 26 artículos probablemente no sea suficiente para determinar las tendencias de investigación en los sistemas de reconocimiento de lenguaje de señas basados en inteligencia durante las últimas dos décadas. Por lo tanto, los autores exploraron más la literatura e identificaron los diferentes tipos de sistemas inteligentes empleados en los sistemas de reconocimiento de lenguaje de señas. Los tipos de sistemas inteligentes identificados incluyen *Inteligencia artificial*, *sistema de apoyo a la decisión*, *aprendizaje automático*, *red neuronal*, *sistema experto*, *sistema difuso* y *sistemas basados en el conocimiento*. Estos términos de búsqueda se combinaron en una cadena de consulta utilizando la conjunción Y y disyunción O operadores. La salida de la cadena de consulta fue 652 páginas y combinado con inicial 26 diarios para un mayor análisis. Para evitar confusiones y duplicaciones en los artículos finales seleccionados para el análisis, los autores filtraron los artículos idénticos en el paso 4. Finalmente, 649 los artículos publicados se consideraron para el análisis bibliométrico a fin de determinar las tendencias de investigación sobre los sistemas de reconocimiento de lenguaje de señas basados en inteligencia. Para el paso 5 se utiliza el VOSviewer, que es una herramienta de software para la construcción y visualización de redes bibliométricas. El resultado obtenido se presenta en Sección 2.2.

## 2.2. Hallazgos y discusión

Los resultados de la evaluación bibliométrica de la investigación publicada sobre sistemas inteligentes en el reconocimiento de lengua de signos durante las últimas dos décadas se presentan en diferentes secciones. Cada una de las tres secciones presenta los hallazgos en relación con una variable específica.

### 2.2.1. Calendario de distribución de publicaciones

Figura 2 muestra el número de publicaciones indexadas en la base de datos Scopus sobre sistemas inteligentes de reconocimiento de lengua de signos entre dos décadas (2001 a 2021). Como se muestra en Figura 2, la mayor cantidad de publicaciones se publicaron en 2020 con una cifra de publicaciones de 114 artículos, seguido del año 2019 registrado 91 documentos. La menor cantidad de artículos se publicó en el año 2003. La tasa anual de producción de investigación en 2019 y 2020 es más del doble de las publicaciones anuales producidas desde 2001. Tenga en cuenta que esta investigación se realizó el 20 de marzo de 2021. Solo 28 los artículos se registraron en la base de datos Scopus para el año 2021 al realizar este estudio. Por lo tanto, el manuscrito presentado en los años anteriores o principios del año 2021 podrá ser publicado en cualquier momento a partir de ahora. Como se puede observar en Figura 2, la tendencia de publicación destaca un aumento en el número de investigaciones en los sistemas inteligentes de reconocimiento de lenguaje de señas a partir de 2011 con 495 documentos; mientras que la década de 2001 se caracterizó 154 artículos fuera de 649 registradas durante las dos décadas completas. Por lo tanto, esta tendencia indica evidentemente que los sistemas inteligentes están ganando atención en el área de investigación del reconocimiento del lenguaje de señas, especialmente a partir de 2011. La tendencia de publicación es consistente con el rápido aumento en el número de personas afectadas por discapacidad auditiva (Bin et al., 2019). Además, la tasa de crecimiento de las publicaciones acumuladas

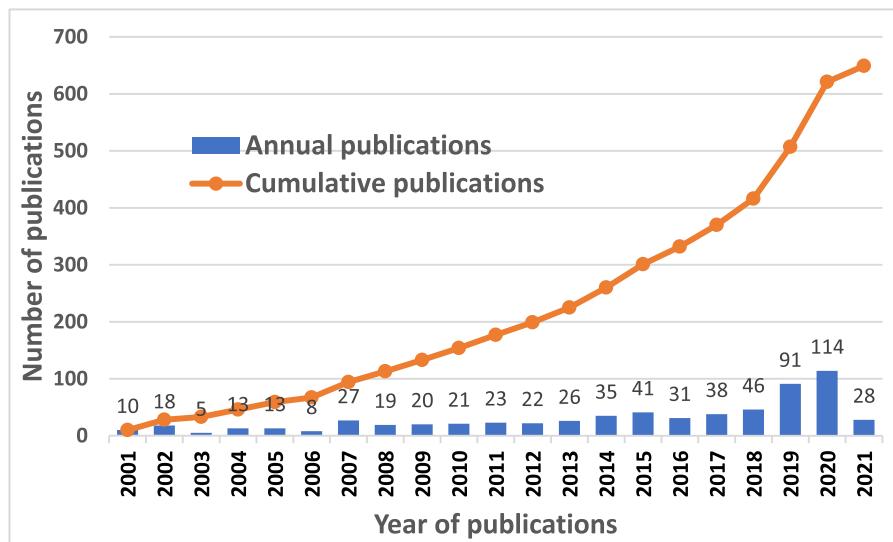


Figura 2. Distribución temporal de las publicaciones anuales y acumulativas de reconocimiento inteligente de lenguaje de señas en la base de datos Scopus desde 2001 hasta 2021.

tabla 1

Las 12 revistas más activas sobre reconocimiento de lenguaje de señas basado en inteligencia.

	Revistas	PT (%)	CT	Cite Score 2021	Editores
1	Sensores Suiza	15	113,885	5.7	MDPI
2	Aplicaciones y herramientas	14	25,844	4.5	naturaleza primaveral
3	multimedia Informática neuronal y	13	19,996	7.1	naturaleza primaveral
4	aplicaciones IEEE Sensors Journal	11	29,411	5.9	IEEE
5	Acceso IEEE	10	191,118	4.9	IEEE
6	Sistemas Expertos Con Aplicaciones	9	33,782	12.5	Elsevier Ltda.
7	Neurocomputación	9	48,529	9.6	Elsevier Ltda.
8	Transacciones IEEE sobre sistemas, hombre y cibernetica, parte B, cibernetica,	8	-	-	IEEE
9	revista de inteligencia ambiental y computación humanizada	7	5153	5.3	naturaleza primaveral
10	Revista de tecnología de la información teórica y aplicada	7	2964	1.3	Pequeño león científico
11	Transacción IEEE en sistemas difusos	6	16,853	18.1	IEEE
12	Revista de informática biomédica	6	5545	7.9	Elsevier Ltda.

significa que la investigación sobre sistemas inteligentes de reconocimiento de lenguaje de señas sigue siendo un área de investigación caliente. Por lo tanto, se espera que la publicación anual siga aumentando.

#### 2.2.2. Fuentes de distribución de publicaciones

Nuestros hallazgos mostraron que se utilizaron 367 revistas diferentes para publicar 649 artículos registrados en la base de datos Scopus desde 2001 hasta 2021 en el área de sistemas de reconocimiento de signos basados en inteligencia. Para evitar una tabla larga, las revistas más activas según el dominio de este estudio se presentan en la tabla 1.

La revista más importante fue publicada por el Instituto Multidisciplinario de Publicaciones Digitales (MDPI), mientras que la revista IEEE encabeza el mayor número de revistas con cuatro en total. El resto de las tres revistas fueron publicadas por Springer Nature, Little Lion Science y Elsevier Ltd. Entre las cuales Springer y Elsevier tienen tres revistas cada una. La editorial Little Lion Scientific tiene una revista que produce 10 revistas. La revista Sensors Switzerland indudablemente juega un papel dominante en el dominio con 15 publicaciones y 113 885 citas de revistas en total. Le sigue Herramientas y Aplicaciones Multimedia (14 publicaciones). Journal of Ambient Intelligence y Journal of Theoretical and Applied Information Technology tienen siete revistas cada una. IEEE Transaction on Fuzzy Systems y Journal of Biomedical Informatics comparten la misma cantidad de artículos (6 publicaciones cada uno). Según el informe CiteScore 2020, ocho revistas tenían CiteScore de 5 o más. El alcance de la revista líder más importante es amplio y cubre diferentes disciplinas de la ciencia y la tecnología sobre el sensor y sus aplicaciones. La frecuencia de pub-

licación y ser una revista de acceso abierto también se puede atribuir a la revista Sensor que publica el mayor número de artículos.

#### 2.2.3. Distribución geográfica de las publicaciones

Analizamos los países donde se realizó la investigación en base a la ocurrencia única de un país en las afiliaciones de cada artículo utilizando la herramienta VOSViewer. A continuación, se suma la presencia de cada país al número total de publicaciones editadas por país. Tenga en cuenta que una publicación puede sumar a más de un país. Eso lleva a 797 resultados de investigación en 78 países que publicaron al menos un artículo de investigación completa sobre sistemas inteligentes en el reconocimiento de lenguaje de señas en las últimas dos décadas. Fig. 3 muestra el mapa geográfico de los 16 países más destacados que contribuyen al crecimiento de la investigación de sistemas SLR en las últimas dos décadas. India, China y EE. UU. aportaron alrededor del 38,14 % de las publicaciones mundiales.

Los tres países juegan un papel clave en el avance de la investigación sobre el reconocimiento del lenguaje de señas, con India a la cabeza a nivel mundial. India lideró con 123 publicaciones en las últimas dos décadas, cubriendo el 15.4% del total de publicaciones globales. Tanto China como EE. UU. contribuyeron con un 13,93 % y un 8,8 %, respectivamente. Además de Fig. 3, también enumeramos las instituciones más activas de los 16 países más importantes en Tabla 2. El Instituto Nacional de Tecnología Maunala Azad, Bhopal, encabeza la lista en India. La Universidad de Ciencia y Tecnología de China, Hefei y la Universidad de Ciencia y Tecnología de Missouri, Rolla, Mo encabezan China y EE. UU., respectivamente.

La colaboración entre instituciones de diferentes países conduce a que un periódico tenga más de una afiliación o país. Un vi-

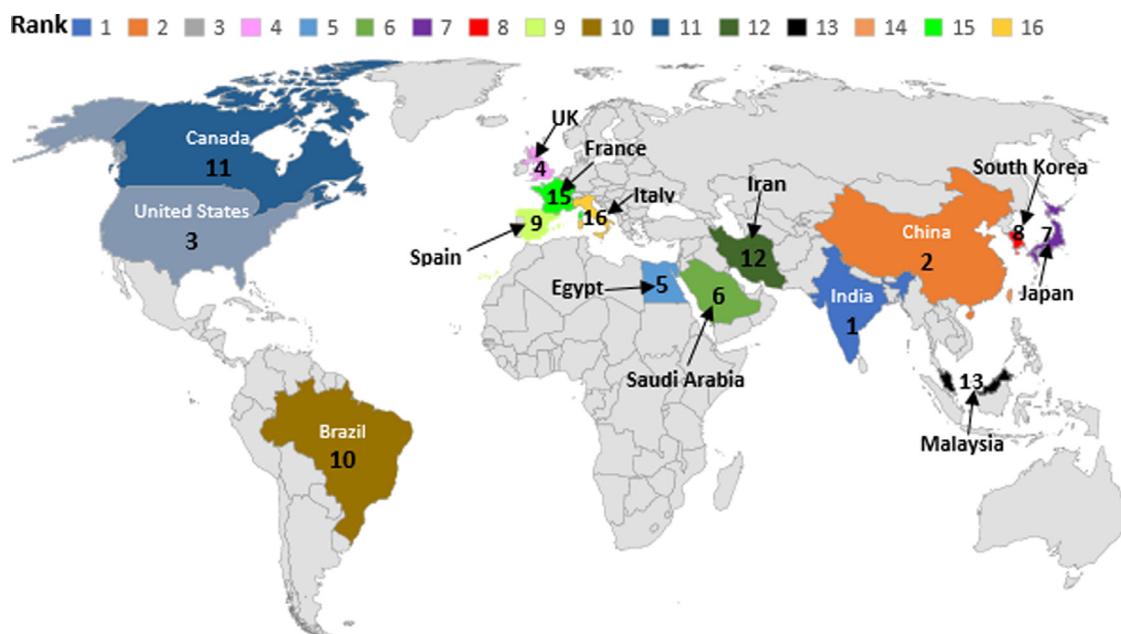


Fig. 3.Ranking regional de los 16 principales países e instituciones académicas con mayor número de publicaciones entre 2001 y 2021.

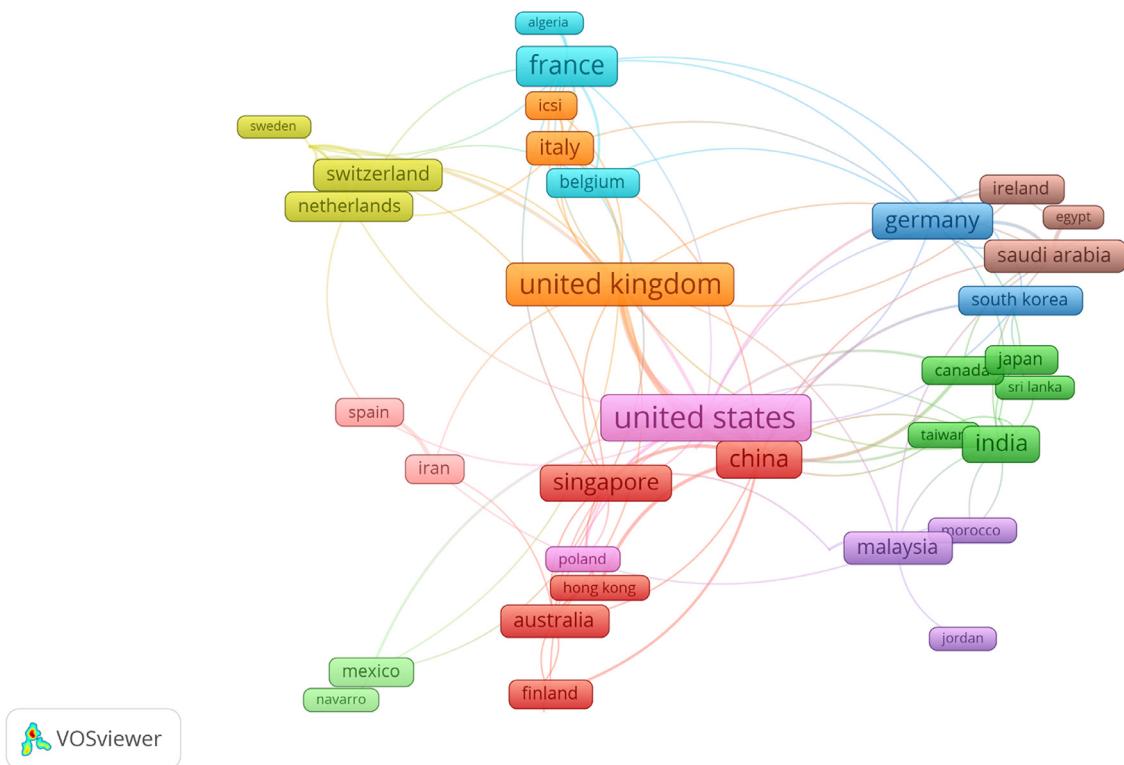
Tabla 2

Los 16 principales países e instituciones académicas más activos en cada ubicación geográfica entre 2001 y 2021. TP: publicación total de un país determinado; TCC: citación total de un determinado país.

Rango	País	TP	TCC	Los más publicados instituto académico	Rango	País	TP	TCC	Los más publicados instituto académico
1	India	123	2118	Nacional Maunala Azad Instituto de Tecnología, Bhopal, India.	9	España	19	360	Centro de Biotecnología y Genómica de Plantas, Upm-inia, España.
2	Porcelana	111	1353	Universidad de Ciencias y Tecnología de China, Hefei, China.	10	Brasil	18	243	AV. joão naves de ávila, 2160, Uberlândia, Mg, Brasil.
3	A NOSOTROS	70	3872	Universidad de Misuri de Ciencia y Tecnología, Rolla, Mo, Estados Unidos.	11	Canadá	17	328	universidad de ottawa, 800 rey Eduardo, Ottawa, Canadá.
4	Reino Unido	29	632	Loughborough Universidad, Loughborough, Estados Unidos Reino.	12	Irán	17	186	Universidad de Semnan, Semnan, 3513119111, Irán.
5	Egipto	26	178	universidad al-azhar, El Cairo, Egipto.	13	Malasia	17	203	universidad multimedios, Jalan Ayer Keroh lama, Melaka, Malasia.
6	Arabia Saudita Arabia	24	169	Universidad Majmaah, Majmaah, Arabia Saudita Arabia.	14	Taiwán	17	588	Universidad china de Tecnología, Taipéi, Taiwán.
7	Japón	23	316	corporación daicel, Osaka, 530-0001, Japón.	15	Francia	desde	288	Chru-nancy, Servicio de Neurologie-Francia.
8	Sur Corea	19	253	universidad de keimyung, Daegu, Corea del Sur.	desde	Italia	desde	610	Universidad de Trento, Italia.

La red de evaluación se desarrolla utilizando el análisis de coautoría de VOSViewer para explorar la cooperación en investigación entre diferentes afiliaciones. La distribución de países o territorios por región se ilustra en Figura 4. Los nodos (cuadrados) en la red simbolizan países, y su tamaño depende de la cooperación de los países. El color en la red de visualización caracteriza la red de colaboración. Cuanto más cerca están los países ubicados entre sí en la red, más sólido es el vínculo, más gruesa es la línea y más fuerte es su relación. Como se muestra en Figura 4, el mayor espesor se observa entre los Estados Unidos y el Reino Unido. Los resultados de la colaboración de coautoría mostraron que Estados Unidos es el

país más afiliado, vinculado a 41 países/territorios. La lista fue seguida por China (35 enlaces), Reino Unido (27 enlaces), Francia (16 enlaces), Arabia Saudita (13 enlaces), Singapur (13 enlaces), India (12 enlaces), Turquía (12 enlaces), Alemania (11 enlaces), Italia (11 enlaces), Suiza (10 enlaces) y otros 55 países conectados desde un rango de nueve a un enlace. Se observó que alrededor del 17% de los países enumerados tenían publicaciones colaborativas internacionales con no menos de diez países. Los posibles factores que contribuyen a la dinámica de la colaboración internacional entre estos países se pueden atribuir a la diversidad de socios de investigación, la financiación sustancial de la investigación y muchos extranjeros que visitan o realizan estudios de posgrado.



**Figura 4.** Una captura de pantalla de la red de cooperación entre diferentes instituciones sobre sistemas de reconocimiento de lenguaje de señas basados en inteligencia en publicaciones de la base de datos Scopus entre 2001 y 2021.

ars. La política de investigación flexible y estable también se puede atribuir a la colaboración internacional en la mayoría de los países altamente vinculados.

### 3. Revisión de las técnicas de reconocimiento de lenguaje de señas basadas en la visión

Las etapas involucradas en el reconocimiento de lenguaje de señas basado en la visión (SLR) se pueden clasificar en cinco etapas: adquisición de imágenes, preprocessamiento de imágenes, segmentación, extracción de características y clasificación, como se muestra en Figura 5. La adquisición de imágenes es la primera etapa en el reconocimiento del lenguaje de señas que se puede adquirir a través de conjuntos de datos públicos creados por uno mismo o disponibles. La segunda etapa es el preprocessamiento para eliminar el ruido no deseado y mejorar la calidad de la imagen. A continuación, después del preprocessamiento, el paso es segmentar y extraer la región de interés de toda la imagen. La cuarta etapa es la extracción de características, que transforma la región de la imagen de entrada en vectores de características para el reconocimiento. La última etapa en SLR basada en visión es la clasificación, que implica hacer coincidir las características de la imagen de la nueva señal con las características almacenadas en la base de datos para el reconocimiento de la señal dada (Raj y Jasuja, 2018; Cheok et al., 2019).

#### 3.1. Dispositivos de adquisición de imágenes

Los investigadores han utilizado diferentes dispositivos de adquisición de imágenes para adquirir imágenes de signos para clasificaciones. Estos dispositivos incluyen la cámara o cámara web, el guante de datos, Kinect y el controlador de movimiento de salto (Kamal et al., 2019). Entre estos dispositivos, una cámara o cámara web es el más utilizado por muchos investigadores porque proporciona una interacción hombre-computadora mejor y más natural sin dispositivos adicionales, a diferencia de los datos basados en guantes. El guante de datos ha demostrado ser más preciso en la adquisición de datos pero muy costoso e inconveniente para los usuarios. Kinect es ampliamente utilizado y eficaz. Proporciona video en color y transmisión de video en profundidad simultáneamente. Separa fácilmente el fondo de la imagen del letrero real y extrae trayectorias 3D de movimientos de la mano (Kamal et al., 2019). La desventaja de Kinect, sin embargo, es la implicación del costo, ya que es muy costoso.

El controlador de movimiento de salto puede operar en un rango limitado, pero es un dispositivo de bajo costo con mayor precisión que Kinect (Suharjito et al., 2017). Los usos de una cámara para la adquisición de imágenes de signos se pueden encontrar en la literatura (Wang et al., 2010; Mekala et al., 2011; Kumarage et al., 2011; Pansare & Inglés, 2016; Athira et al., 2019; Sharma et al., 2021). El guante de datos se utilizó para adquirir imágenes de signos en los estudios de reconocimiento de gestos de la mano propuestos por Mehdi Khan (2002), Gao et al. (2004), Phi et al. (2015) y Pan et al. (2020). Kinect se utilizó para adquirir imágenes de signos por Jiang et al. (2015), Wang et al. (2015a, 2015b, 2016), Raheja et al. (2016), Carneiro et al. (2017) y Escobedo y Cámara (2017). Estudios similares emplearon un controlador de movimiento de salto para la adquisición de imágenes de signos (Kiselev et al., 2019; Alnahhas et al., 2020; Enikeev y Mustafina, 2021). Tabla 3 muestra las ventajas y desventajas de los diferentes dispositivos de adquisición de datos para las clasificaciones. Estos estudios revelaron que las imágenes adquiridas eran estáticas o dinámicas capturadas en diferentes posiciones, condiciones de fondo e iluminación como marcos de imágenes.

#### 3.2. Técnicas de preprocessamiento de imágenes

Las técnicas de preprocessamiento se aplican a una imagen de entrada para eliminar el ruido no deseado y también mejorar la calidad. Esto se puede lograr cambiando el tamaño, convirtiendo el color, eliminando el ruido no deseado o una combinación de varias de estas técnicas de la imagen original. El resultado de este proceso puede afectar en gran medida la precisión con una buena selección de técnicas de preprocessamiento. Las técnicas de preprocessamiento de imágenes se pueden clasificar en términos generales en mejora de imágenes y restauración de imágenes. Las técnicas de mejora de imágenes incluyen la ecualización de histogramas (HE), la ecualización de histogramas adaptable (AHE), la ecualización de histogramas adaptable limitada por contraste (CLAHE) y la transformación logarítmica. La restauración de imágenes incluye filtro mediano, filtro medio, filtro gaussiano, filtro adaptativo y filtro Wiener. Figura 6 muestra un esquema detallado de las técnicas de preprocessamiento de imágenes.

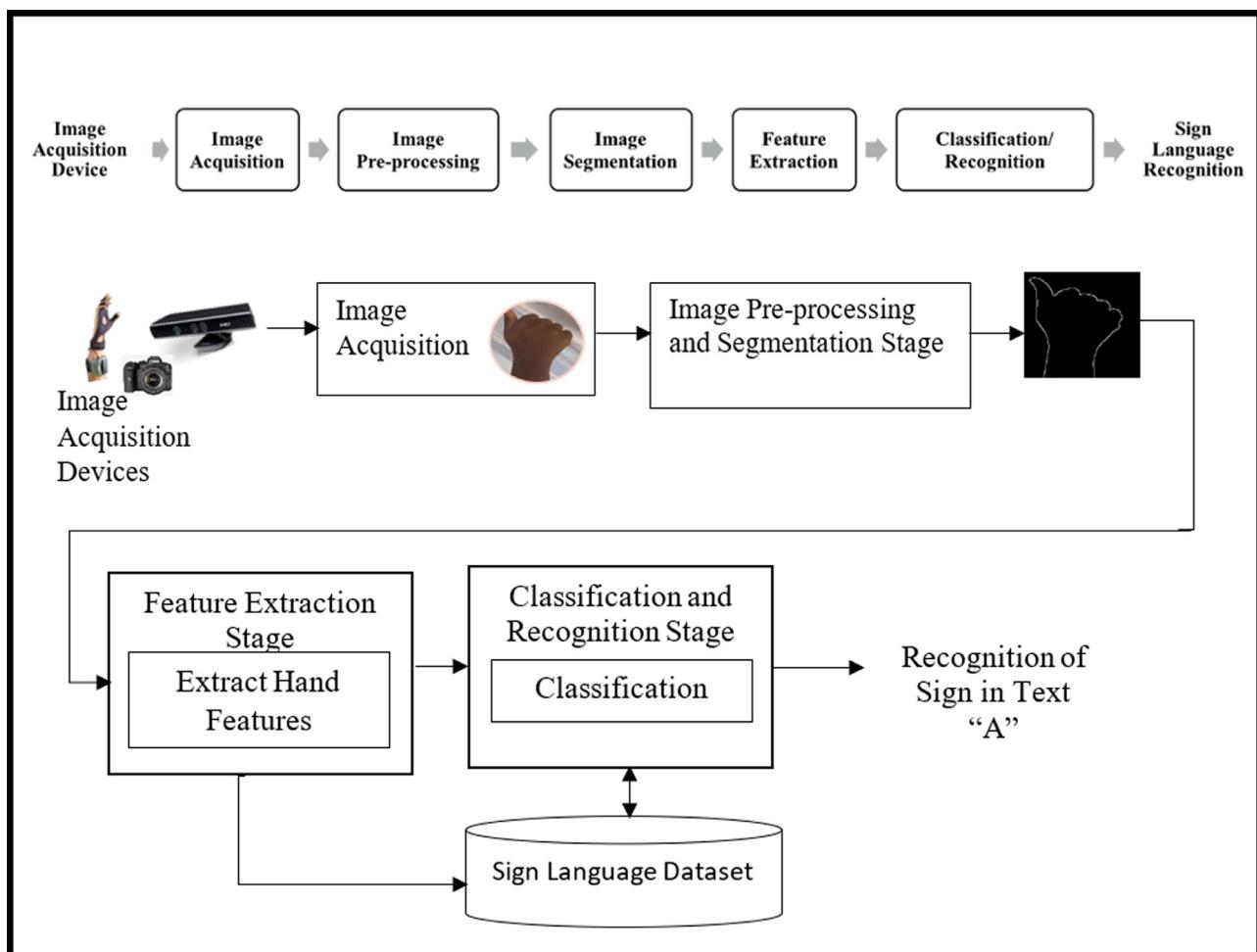


Figura 5. Arquitectura del reconocimiento de lenguaje de señas basado en la visión.

Tabla 3

Ventajas y desventajas de utilizar diferentes controladores de dispositivos de adquisición.

Imagen Adquisición dispositivo	Procedimiento de adquisición	Ventajas	Desventajas
Cámara/cámara web	El dispositivo se utiliza para capturar o grabar la imagen/video de una cámara fija.	El dispositivo no requiere usar otros dispositivos externos, y los usuarios solo necesitan usar sus manos dentro del rango de recolección de la cámara. Bajo costo. Convenient y fácil de usar.	Se ve muy afectado por factores ambientales como la luz, el color de la piel y la oclusión. Requiere varias técnicas de procesamiento de imágenes que podrían afectar el reconocimiento exactitud.
Guante de datos	El guante se adjunta a los sensores para extraer las características necesarias para el reconocimiento de la señal. Las características pueden ser la flexión de los dedos, la orientación de la mano, el movimiento, la rotación y la posición. Estos se transforman en señales eléctricas para determinar la postura de la mano.	Estos dispositivos no se ven afectados por el entorno externo al recopilar datos. Proporciona una mejora precisión de reconocimiento. La extracción de características basadas en sensores es relativamente más fácil.	Reduce la naturalidad de la interacción. Incómodo de usar por el usuario. Muy caro
Kinect	La función Kinect proporciona color y profundidad con información de trayectoria en una imagen simultáneamente.	Es útil para muchas aplicaciones de interacción humano-computadora. La profundidad de detección de distancia es limitada.	Puede verse afectado por las condiciones de iluminación, la segmentación de manos y rostros, fondo complejo y ruido.
Movimiento de salto Controlador	Es un dispositivo basado en sensores que se conecta a una computadora personal que detecta el movimiento de la mano y convierte las señales en comandos de computadora para la adquisición de datos.	Tiene una alta precisión de reconocimiento y una velocidad de procesamiento más rápida de alrededor de 200 cuadros por segundo. Puede detectar y rastrear manos, dedos y objetos similares a dedos.	Kinect no apto para aplicaciones al aire libre y sensibilidad a la luz solar. Debido a su alta sensibilidad, la precisión del reconocimiento puede verse afectada por pequeños movimientos en la posición de la señal.

**Tabla 4**

Comparación de técnicas de mejora de imagen.

Técnicas de mejora	Ventajas	Desventajas
Ecualización de histograma (HE)	Es simple de implementar y altamente efectivo para imágenes en escala de grises.	Puede aumentar el contraste del ruido de fondo. Es difícil distinguir entre el ruido y las características deseadas. Cambia el brillo de una imagen.
Ecualización de histograma adaptable (AHE)	Es adecuado para mejorar el contraste local y los bordes en cada región de una imagen. Supera a la técnica de ecualización de histogramas.	Tiene un efecto adverso en la salida deseada debido a su comportamiento de amplificación de ruido. No logra retener la Brillo en la imagen de entrada. Produce un resultado insatisfactorio cuando la imagen de entrada tiene una relación de contraste desequilibrada y un brillo aumentado.
Ecualización de histograma adaptativo limitado de contraste (CLAHE):	Tiene un ruido reducido en comparación con AHE y HE. Proporciona respuesta de salida local y evita la saturación de brillo.	
Transformación logarítmica	Se utiliza para reducir los valores de píxeles de intensidades más altas a valores de píxeles de intensidades más bajas.	Aplicar la técnica a un valor de píxel más alto mejorará más la imagen y causará la pérdida de información real en la imagen. No se aplica a todo tipo de imágenes.

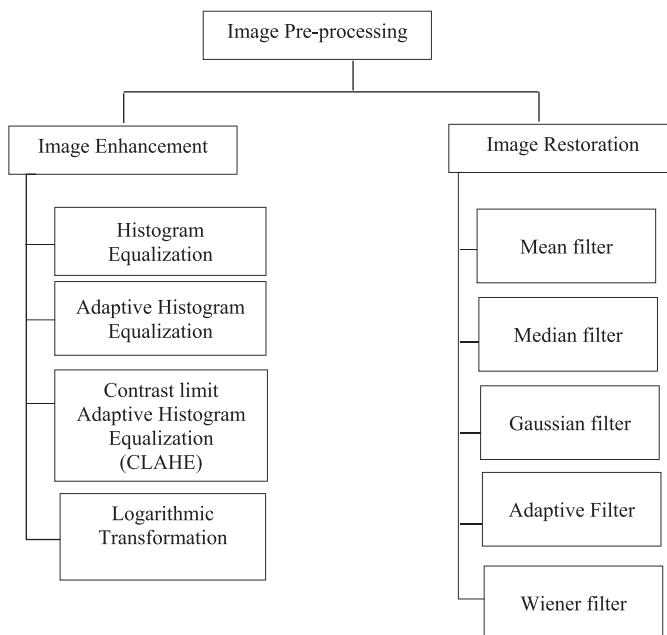


Figura 6.Técnicas de preprocesamiento de imágenes.

### 3.2.1. Técnicas de mejora de imagen

En el aspecto del procesamiento de imágenes, la mejora de imágenes es uno de los problemas más desafiantes. Es un proceso importante para restaurar la apariencia visual de una imagen. El objetivo principal de la mejora de imágenes es mejorar la calidad de la imagen de entrada, que puede ser más adecuada para el análisis humano o automático. La elección de la técnica depende del área de aplicación. Estas técnicas se pueden utilizar para refinar el límite y mejorar la precisión de una imagen de entrada (Majeed e Isa, 2021). Tabla 4 resumen las ventajas y desventajas de varias técnicas de mejora de imágenes.

**Ecualización de histograma (HE):** La ecualización de histogramas se utiliza como técnica de preprocesamiento de imágenes para reforzar el color y aumentar el contraste (González, 2002). Como parte de su operación, la técnica de ecualización de histograma reasigna el nivel de gris de una imagen en particular en función de su distribución de probabilidad.

Este método se aplica distribuyendo el nivel de valor de gris en una imagen. Aumenta el límite inferior de la gama de colores hasta el punto más oscuro y disminuye el límite superior del contenido de colores hasta el punto más claro. El uso de esta técnica mejora la

bordes y límites de diferentes imágenes y reduce los detalles locales de las imágenes (Verma & Dutta, 2017). Para determinar la ecualización del histograma, la función de densidad de probabilidad (pdf) y la función de distribución acumulativa (cdf) se calculan utilizando ecuaciones (1) y (2), respectivamente (Kalhor et al., 2019):

$$\text{pdf} / \text{p}(A_k) = \frac{n_{\text{orte}}}{n_{\text{total}}} \quad (1)$$

$$\text{cdf} / \text{p}(A_k) = \frac{\sum_{j=0}^k n_{\text{orte}}}{n_{\text{total}}} = \frac{k}{n_{\text{total}}} \quad (2)$$

donde  $n$  es el número total de píxeles en una muestra,  $k$  es el rango del valor de gris,  $y n_{\text{orte}}$  es el número de píxeles con el nivel de gris con el nivel de gris de  $A_k$  dentro de la imagen A. La ecualización del histograma se opera en una imagen en tres pasos, como lo indican Bagade y Shandilya (2011):

i Formación del histograma.

ii Se calculan nuevos valores de intensidad para cada nivel de intensidad. iii Reemplace los valores de intensidad anteriores con la nueva intensidad valores.

La ecualización de histogramas es el método predominante para la mejora de imágenes y aumenta el contraste de la imagen, mejorando el contraste y obteniendo un histograma uniforme.

Se utilizó para mejorar el contraste de la imagen de entrada en una ubicación diferente y hace que el brillo y la iluminación de la imagen sean uniformes (Mahmud et al., 2019; Nelson et al., 2019; Sethi et al., 2012; Suharjito et al., 2019).

**Ecualización de histograma adaptable (AHE):** La ecualización de histograma adaptable (AHE) es un algoritmo bien conocido y eficiente para mejorar el contraste de la imagen. Sin embargo, requiere mucho tiempo y es computacionalmente costoso. Se ha aplicado ampliamente a las diversas aplicaciones de procesamiento de imágenes. La ecualización de histograma adaptable varía de la ecualización de histograma ordinaria. El enfoque adaptativo calcula muchos histogramas, cada uno correspondiente a una sección diferente de la imagen, y los usa para redistribuir los valores de luminosidad de la imagen (Sund & Møystad, 2016). Es adecuado para mejorar el contraste local y los bordes en todas las regiones de una imagen, pero tiende a aumentar el ruido en regiones relativamente homogéneas (Sudhakar, 2017). La ecualización adaptativa del histograma es un método de mejora del contraste aplicable a imágenes en escala de grises y en color con muy buena eficiencia. Los pasos en el algoritmo AHE se dan (Longkumer et al., 2014) como:

**Paso 1:** Inicie el programa.

- Paso 2:** Obtenga todas las imágenes de entrada con el número de regiones, rango dinámico y límite de recorte. **Paso 3:** Preprocesar la imagen de entrada. **Etapa 4:** Procese cada región contextual produciendo un nivel de gris cartografía. **Paso 5:** Interpolar el mapeo de niveles de gris para ensamblar la imagen final. **Paso 6:** Detener.

La ecualización de histograma adaptable se utiliza para mejorar el contraste local y el borde en cada región de la imagen (Kaluri y Reddy, 2017; Kamal et al., 2019; Majeed e Isa, 2021; Nelson et al., 2019).

*Ecuación de histograma adaptable de límite de contraste (CLAHE):* Contrast Limited Adaptive Histogram Equalization (CLAHE) proporciona una mejora con respecto a la ecualización de histograma adaptativo y funciona en pequeñas regiones de la imagen en lugar de en la imagen completa. La función de mejora se aplica sobre todos los píxeles de la vecindad y se deriva la función de transformación. Para lograr la técnica CLAHE, la imagen se divide en la región contextual conocida como mosaicos, mientras que la ecualización del histograma se aplica a cada mosaico para obtener una distribución de histograma de salida deseada (Aurangzeb et al., 2021). El algoritmo CLAHE se da como (Rubini y Pavithra, 2019):

#### Algoritmo para CLAHE

Paso 1: Lea la imagen de entrada.

Paso 2: aplique el filtro de media y mediana en la imagen de entrada. Paso 3: encuentre los recuentos de frecuencia para cada valor de píxel. Paso 4: Determine la probabilidad de cada ocurrencia usando el

función de probabilidad

Paso 5: Calcular la probabilidad de distribución acumulada para cada valor de pixel.

Paso 6: Realice el mapeo de ecualización para todos los píxeles.

Paso 7: Muestre la imagen mejorada.

Suharjito et al. (2019) utilizó la ecualización de histograma adaptativo de límite de contraste para mejorar los bordes de la imagen y la detección de la piel de las imágenes. El contraste y el brillo se mejoran de manera que la información original no se pierde y el brillo se conserva (Sykora et al., 2014). Tres técnicas de mejora de imagen, a saber, CLAHE, HE y AHE, fueron comparadas por Nelson et al. (2019).

*Transformación logarítmica:* La transformación logarítmica se utiliza cuando los valores de entrada del nivel de gris son enormes. La técnica se utiliza para reducir el sesgo de las distribuciones muy sesgadas en una imagen al distribuir los píxeles oscuros de la imagen mientras se comprimen los valores más altos (Mahmud et al., 2019). Para lograr la técnica de transformación, una imagen de entrada se convierte primero a escala de grises antes de realizar una transformación logarítmica para que se puedan ver los efectos de la transformación (Chourasiya et al., 2019; Maini y Aggarwal, 2010). La ecuación de transformación logarítmica se da como:

$$s = \text{Registro}(1 + r) \quad (3)$$

dónde  $s$  es la transformación logarítmica,  $c$  es el factor por el cual se mejora la imagen (generalmente se establece en 1) y  $r$  es el píxel actual en la imagen. La comparación de las técnicas de mejora de imágenes en términos de méritos y deméritos se presenta en Tabla 4.

*Otras técnicas de preprocesamiento:* Otras técnicas de preprocesamiento incluyen el escalado de imágenes y la conversión a escala de grises, que también se conocen como cambio de tamaño de imagen. En la técnica de preprocesamiento de imágenes, tienen un papel importante en la ampliación y reducción del tamaño de imagen dado en un formato de píxel (Perumal y Velmurugan, 2018). El cambio de tamaño de la imagen es una de las operaciones básicas para el preprocesamiento de imágenes. Ayuda a ganar mucha importancia en la mejora de imágenes de baja resolución y se puede usar para volver a muestrear la imagen para disminuir o aumentar la resolución (Fadnavis, 2014). Las imágenes de entrada se escalan o se recortan en un tamaño uniforme, ya que las imágenes de entrada pueden ser de diferentes tamaños recopiladas de diferentes fuentes. La salida resultante de este

La técnica produce reducir o ampliar el tamaño de una imagen de entrada. La comunidad de investigación ha utilizado esta técnica para reducir el tiempo de cálculo y el tamaño de almacenamiento (Wang et al., 2010; Jin et al., 2016; Islam et al., 2017; Ramos et al., 2019).

La conversión a escala de grises es una de las técnicas de preprocesamiento. Es una de las técnicas de mejora más simples utilizadas en el procesamiento de imágenes. Se realiza convirtiendo la imagen del espacio de color, como rojo, verde y azul (RGB), en una imagen en escala de grises. Este modelo de color consta de tonos grises de colores, que tienen solo 256 colores grises. Se compone exclusivamente de tonos de gris, que van desde el negro en la intensidad más débil hasta el blanco en la más fuerte (Büyüksahin, 2014). Las ventajas de usar una imagen en escala de grises sobre una imagen en color RGB incluyen su algoritmo simplificado y reduce los requisitos computacionales al tiempo que conserva las características más destacadas de las imágenes en color. Sin embargo, una imagen en escala de grises tiene la ventaja de perder información de color en una imagen (Güneş et al., 2016). Se da la ecuación para la conversión del modelo de color RGB en una imagen en escala de grises ponderada (Biswas et al., 2011) en ecuación (4):

$$GY = 0.56 G + 0.33 R + 0.11B \quad (4)$$

donde  $GY$  denota el nivel de gris resultante (escala de grises) para el píxel calculado,  $R$  es rojo,  $G$  es verde y  $B$  es azul en una imagen dada.

En Mekala et al. (2011), Karami et al. (2011), Oyewumi Jimoh et al. (2018) y Sharma et al. (2020a, 2020b), la imagen en color se convirtió en una imagen en escala de grises de dos dimensiones con dos posibles valores de intensidad de blanco y negro.

#### 3.2.2. Restauración de imagen

La restauración de imágenes es el proceso de restaurar una imagen degradada causada por ruido y desenfoque (Reeves, 2014). El proceso de restauración de la imagen está determinado por el tipo de ruido y corrupción presente en la imagen. Las técnicas de eliminación de ruido deben elegirse en función de la cantidad de ruido que contiene una imagen. Hay varias técnicas de restauración de imágenes disponibles para eliminar el ruido no deseado y el desenfoque de la imagen, entre ellas: filtro mediano, filtro medio, filtro gaussiano, filtro adaptativo y filtro wiener.

*Filtro medio:* El filtro medio es un método de filtrado espacial basado en ventanas deslizantes. Se utiliza para suavizar la imagen y reducir o eliminar el ruido de una imagen. La técnica de filtro medio encuentra el valor del píxel central en la ventana promediando los valores de los píxeles vecinos (Aksoy y Salmán, 2020). Esta técnica de filtrado funciona bien con ruido de sal y pimienta y ruido gaussiano (Singh y Shree, 2016). La ecuación utilizada para calcular el filtro medio se da como:

$$A[y_0, j] = \frac{1}{\text{METRO}} \sum_{i=k-1}^{k+1} \sum_{j=k-1}^{k+1} joder, yo \quad (5)$$

dónde  $\text{METRO}$  es un valor del número de píxeles utilizados en el cálculo,  $k$  y  $y_0$  son valores que representan la ubicación de estos píxeles.  $A[y_0, j]$  es el filtro medio en cualquier punto de la imagen desplazando la ventana a los píxeles vecinos. El filtro medio fue utilizado por Kasmin et al. (2020) para eliminar el ruido de las imágenes de los letreros. Hay un número limitado de investigadores que han usado filtros medios en el reconocimiento del lenguaje de señas. Este trabajo de revisión encuentra que la mayoría de los trabajos de investigación utilizaron filtro mediano y filtro gaussiano en la etapa de preprocesamiento para eliminar el ruido en la imagen, como se informa en el trabajo de revisión (Cheok et al., 2019).

*Filtro mediano:* El filtro mediano es un método no lineal, que se usa más comúnmente como una forma simple de reducir el ruido en una imagen mientras se preservan los bordes (Dhanushree et al., 2019). Logra mejores resultados en la eliminación del ruido de sal y pimienta. El píxel que se considera se reemplaza por el valor de píxel mediano del píxel vecino calculado a partir de los valores de píxel de todos los píxeles vecinos circundantes ordenados en orden numérico y luego se reemplaza el píxel que se considera con el valor de píxel medio

(Ahmad et al., 2019). El promedio de los dos valores de píxel del medio se usa cuando la vecindad en consideración contiene un número par de píxeles. Sus actuaciones son mejores que el filtro medio en la preservación de los bordes y lo hacen más nítido mientras elimina el ruido en la imagen y es fácil de implementar.

La ecuación de filtro de medias aritméticas utilizada se da como (González, 2002):

$$F_{(x, y)} = \frac{1}{\text{Minnesota}} \sum_{(\text{calle}) \in S_{xy}} g(s, t) \quad (6)$$

dónde  $S_{xy}$  representa el conjunto de coordenadas en una ventana de subimagen rectangular (tamaño de kernel) de metro  $x$  norte centrado en cualquier punto  $(X, y)$  en la imagen original,  $s, t$  son las coordenadas de filas y columnas de los píxeles cuyas coordenadas son miembros del conjunto  $S_{xy}$ .

La técnica es muy buena solo para eliminar el ruido de sal y pimienta. Varios investigadores lo utilizaron para reducir el ruido de la imagen y preservar los bordes (Islam et al., 2017; Lahiani et al., 2016; Pansare et al., 2012; Pansare & Inglés, 2016; Zhang et al., 2011).

**filtro gaussiano:** Un filtro gaussiano es un filtro lineal y un filtro de paso bajo no uniforme resultante de desenfocar una imagen por una función gaussiana. Es el más utilizado para reducir el ruido y suavizar los bordes del filtro en una imagen (Basu, 2002). El filtro gaussiano se utiliza generalmente como un filtro más suave ampliamente utilizado en el preprocesamiento de imágenes para reducir el ruido de la imagen al desenfocarla. Actúa como un operador de suavizado y es utilizado principalmente por varios investigadores en el reconocimiento de lenguaje de señas. El desenfoque gaussiano es un filtro utilizado para suavizar o eliminar el desenfoque de una imagen digital (Umamaheswari y Karthikeyan, 2019). Se utilizó un filtro gaussiano para eliminar el ruido y suavizar la imagen en (Oliveira et al., 2017; Pansare et al., 2012; Yusnita et al., 2017). La ecuación del filtro gaussiano de dos dimensiones (2D) viene dada por:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (7)$$

dónde  $G(x, y)$  denota el valor del filtro gaussiano,  $x, y$  denotan valores de fila y columna,  $\sigma$  es la desviación estándar de la distribución gaussiana.

**Filtro adaptativo :** Un filtro adaptativo se aplica a la imagen ruidosa para eliminar el ruido de la imagen mientras se conserva la información acometida en la imagen detallada. Conserva los bordes y otras partes de alta frecuencia de una imagen más que un filtro lineal similar. La media y la varianza son las dos medidas estadísticas utilizadas para determinar los filtros adaptativos. El algoritmo utilizado para lograr el filtro adaptativo se da como (Kaluri y Reddy, 2016a, 2016b):

Paso 1: Leer la imagen de color de entrada.

Paso 2: convierte la imagen en color de entrada en una imagen en escala de grises.

Paso 3: Aumente el ruido gaussiano en la imagen para que la imagen

el tamaño es grande

**Etapa 4: Elimine el ruido de la imagen usando la función wiener2.**  
ción=k=weiner2(j, [XX])

Se usó un filtro adaptativo para eliminar el ruido de la imagen de la señal de entrada en Kaluri y Reddy (2016a), 2016b).

**filtro de salchicha:** El filtro de Wiener se usa principalmente para eliminar el ruido de la imagen y minimiza el error cuadrático medio (MSE) entre el proceso aleatorio estimado y el proceso deseado. Optimiza el compromiso entre suavizar las discontinuidades de la imagen y eliminar el ruido (Tania y Rowaida, 2016). La técnica desenfoca significativamente la imagen debido al uso de un filtro fijo en toda la imagen. La función de restauración del filtro Wiener incluye tanto la función de degradación como la característica estadística del ruido. Utiliza un filtro de paso alto para desconvolución y un filtro de paso bajo para reducción de ruido durante la compresión (Maru & Parij, 2017). Esta técnica se puede utilizar para eliminar varios tipos de ruido, como sal y pimienta, ruido gaussiano y moteado de

la imagen (Kaur, 2015). La ecuación para calcular el filtro Wiener se da como:

$$Gu, V) = \frac{H(tu, V)PAG_s(tu, V)}{| H(tu, V) | _2 PAG_s(tu, V) + PAG_norte(tu, V)} \quad (8)$$

dónde  $t, u, v$  son los valores, representan la ubicación de estos píxeles,  $H(tu, V)$  denota la función de degradación,  $H(tu, V)$  representan el complejo conjugado de la función de degradación,  $PAG_norte(tu, V)$  es la densidad espectral de potencia del ruido,  $PAG_s(tu, V)$  es la densidad espectral de potencia de la imagen original.

Se han utilizado varias técnicas de filtrado en el reconocimiento del lenguaje de señas. Se utilizó un filtro Winer para eliminar el ruido de la imagen del letrero (Kaluri y Reddy, 2017). Algunos investigadores combinaron dos técnicas de filtrado en otra para eliminar el ruido de la imagen. Pansare et al. (2012) combinó filtros medianos y gaussianos para eliminar el ruido y suavizar la imagen de entrada, respectivamente. Tabla 5 resumen diferentes técnicas de filtrado de imágenes con sus ventajas y desventajas.

### 3.3. Técnicas de segmentación de imágenes

La segmentación de imágenes es el proceso de dividir una imagen en regiones significativas llamadas segmentos (Egmont-Petersen et al., 2002). Las imágenes se segmentan para obtener la región de interés. Hay dos enfoques básicos utilizados para la segmentación; segmentación contextual y no contextual (Enikeev y Mustafina, 2020; Jin et al., 2016; Al-Shamayleh et al., 2020). En la segmentación contextual, explota las relaciones entre las características de la imagen, como los bordes, las intensidades similares y la proximidad espacial. Una segmentación no contextual ignora las relaciones espaciales entre las características de la imagen, pero agrupa los píxeles según el valor de los atributos globales (Amigo y amigo, 1993; Sharma et al., 2021). Las técnicas de segmentación de imágenes se clasifican en basadas en detección de bordes, umbrales, basadas en regiones, basadas en agrupaciones y basadas en redes neuronales artificiales. Un esquema detallado de las técnicas de segmentación de imágenes se muestra en Figura 7.

#### 3.3.1. técnica de umbralización es

umbralización es la segmentación más simple y comúnmente utilizada. tón técnica para separar objetos del fondo (Lee et al., 1990; Cheng et al., 2002; Dong et al., 2008; Xu et al., 2013). El Las técnicas separan los píxeles de la imagen según su su grado de fuerza y el rango de valores en el que se encuentra un píxel—ejemplos de las técnicas de umbralización son la umbralización global, la umbralización adaptativa local y la umbralización multinivel.

**Umbra/globat:** La técnica de umbralización global utiliza un único valor de umbral para toda la imagen para separar el primer plano del fondo. Esta técnica asume que la imagen tiene un histograma bimodal. Por lo tanto, la imagen se puede extraer del fondo utilizando una operación simple que compara los valores de la imagen con un umbral  $T$  (Rogowska, 2009).

La técnica del umbral se define por Kaur y Chand (2018) como:

$$g(x, y) = \begin{cases} 1, & \text{sí } (x, y) > T \\ 0, & \text{sí } (x, y) \leq T \end{cases} \quad (9)$$

dónde  $g(x, y)$  es la imagen de salida de las imágenes originales.  $T$  es la constante del valor umbral para imágenes completas.

La técnica de umbralización de Otsu es la técnica de umbralización global más utilizada. La técnica convierte la imagen de varios niveles en una imagen binaria a través de un valor de umbral dado para separar el primer plano del fondo de la imagen. Se basa en el análisis discriminante al maximizar la variación entre clases de los niveles de gris en las partes de la imagen y el fondo. La suma ponderada de la varianza de las dos clases se da como:

$$\Omega_2(t) = w_0(t) \sigma_0^2(t) + w_1(t) \sigma_1^2(t) \quad (10)$$

**Tabla 5**

Ventajas y desventajas de las técnicas de filtrado de imágenes.

Técnicas de filtrado	Ventajas	Desventajas
Filtro medio	Fácil de implementar	Un solo valor de píxel mal representado puede afectar significativamente el valor medio de todos los píxeles en su entorno inmediato. Difumina un borde cuando la vecindad del filtro cruza un límite.
filtro mediano	Conserva los bordes finos y la nitidez de una imagen de entrada. Ambos problemas del filtro medio son abordados por el filtro mediano.	Es relativamente caro y complejo de calcular. Es bueno solo para eliminar el ruido de sal y pimienta. Es menos eficaz para eliminar el tipo de ruido gaussiano de la imagen.
filtro gaussiano	Es eficaz para eliminar el tipo de ruido gaussiano. Los pesos dan mayor importancia a los píxeles cerca del borde.	Tiene un tiempo computacional alto y, a veces, elimina los detalles de los bordes en una imagen.
Filtro adaptativo	Conserva los bordes y otras partes de alta frecuencia mejor que un filtro lineal similar.	Es la complejidad computacional. Todavía hay algunas distorsiones visibles disponibles en la imagen usando un filtro adaptativo.
filtro weiner	Es un filtro popular utilizado para la restauración de imágenes. No es sensible al ruido. Adecuado para explotar las propiedades estadísticas de la imagen. El tamaño de ventana pequeño se puede utilizar para evitar que los bordes se vean borrosos.	El conocimiento previo de la densidad espectral de potencia de la imagen original no está disponible en la práctica. Es comparativamente lento de aplicar porque funciona en el dominio de la frecuencia. La imagen de salida está muy borrosa.

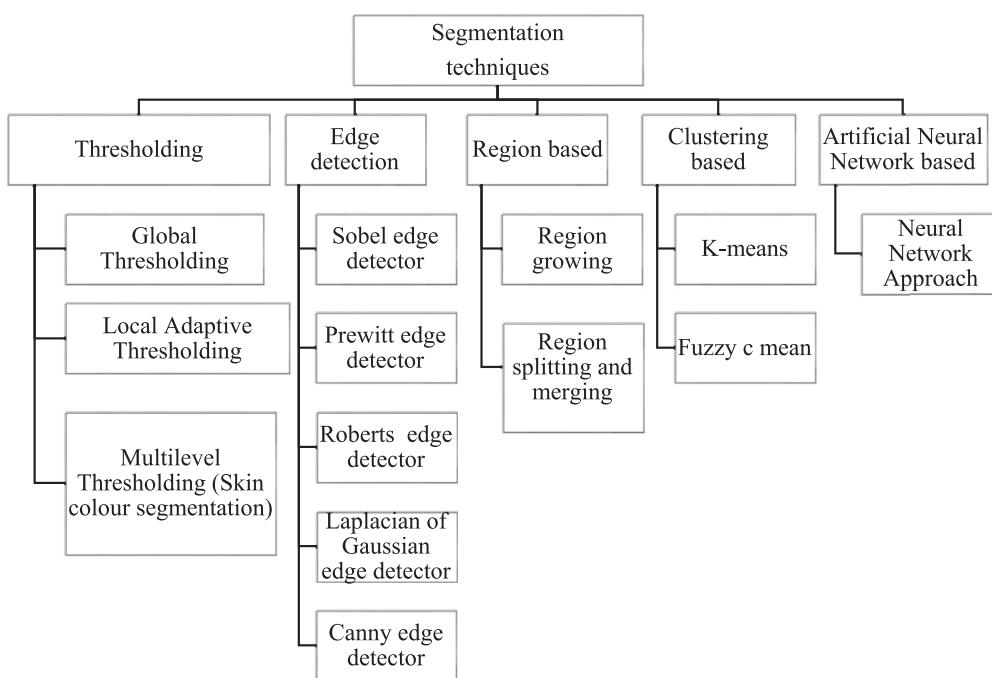


Figura 7. Clasificación de las técnicas de segmentación de imágenes.

dónde;

$w_0$  y  $w_1$  son las probabilidades de las dos clases separadas por un umbral  $t$  (con un rango de valores de 0 a 255),  $\sigma_w$  y  $\sigma_{w_1}$  son variaciones de estas dos clases. Las probabilidades de clase  $w_0$  y  $w_1$  para cada valor de píxel, se calculan a partir de los intervalos del histograma como:

$$w_0(t) = \sum_{j=1}^t P_j \quad (11)$$

$$w_1(t) = \sum_{j=t+1}^I P_j \quad (12)$$

dónde  $I$  es el valor máximo de píxel (255)

La varianza total se da como:

$$\sigma^2 = \sigma_{w_0}^2 + \sigma_{w_1}^2 \quad (13)$$

donde la varianza entre clases está determinada por:

$$\sigma_{\text{between}}^2 = w_0(t) w_1(t) [m_1(t) - m_0(t)]^2 \quad (14)$$

Investigación presentada en Rahim et al. (2020), Pansare et al. (2015), Islam et al. (2017) y Tan et al. (2021) usó el algoritmo de umbralización de Otsu basado en la umbralización global para segmentar la región de la mano de su región de fondo utilizando el valor de umbral calculado. La umbralización de Otsu se fusionó con el canny edge y las transformadas discretas de ondas (DWT) para segmentar la región de interés de la secuencia de video (Kishore y Kumar, 2012). Debido al uso de un solo valor de umbral, esta técnica no produce una segmentación efectiva con respecto a la iluminación de una imagen completa.

*Umbral adaptativo local.* La umbralización adaptativa local se emplea para abordar el problema de las técnicas basadas en la umbralización global dividiendo una imagen en subimágenes y calculando umbrales para cada subimagen (Korzynska et al., 2013). Esta técnica de umbral utiliza el valor medio de la distribución de intensidad local u otras métricas estadísticas, como la media más la desviación estándar, para separar el primer plano del fondo en cada subimagen (Senthilkumar y Vaithegi, 2016). El enfoque más básico para

enfoque de umbral adaptativo propuesto por Negro negro (1986), cuyo umbral local se calcula a partir de la media (*metro*) y la desviación estándar (*s*) de la vecindad local del píxel *w*. El valor de umbral adaptativo basado en las técnicas de Niblack se da como:

$$T_{\text{negro}} = \text{metro} + k * s \quad (15)$$

dónde  $T_{\text{negro}}$  es el umbral adaptativo,  $m$  es la media de la ventana de tamaño  $w$ ,  $s$  es la desviación estándar y  $k$  es el valor fijo que se introduce en función del nivel de ruido restante en el fondo de la imagen. La técnica de umbral adaptativo ayuda a separar imágenes de diferentes fondos y extraer regiones diminutas y dispersas. El principal inconveniente de la umbralización adaptativa es que es computacionalmente costosa que la umbralización global. El umbral adaptativo se utilizó para segmentar la región de interés en el reconocimiento del lenguaje de señas para su posterior procesamiento (Dudhal et al., 2019; Rao y Kishore, 2018).

*Umbralización multinivel*: La técnica de umbralización multinivel se emplea para extraer regiones homogéneas en una imagen. Esta técnica determina múltiples umbrales para la imagen dada y divide la imagen en distintas regiones. El método produce resultados adecuados para imágenes con fondos de colores o complejos, donde el umbral de dos niveles ha fallado. La segmentación del color de la piel es uno de los enfoques utilizados para la determinación de umbrales multinivel y se usa ampliamente en diferentes áreas de aplicación, incluida la interacción humano-computadora (HCI), el reconocimiento de imágenes, el sistema de control de tráfico, la videovigilancia y la segmentación de manos (Sallam et al., 2021; Lee et al., 2021a, 2021b; Razmjoo et al., 2021). La técnica usó el modelo de color para separar la región de la piel de la imagen en la segmentación de la mano. Los modelos de color utilizados incluyen el espacio de color rojo, azul y verde (RGB), el valor de saturación de tono (HSV) y el modelo de espacio de color de luminancia (Y) y cromaticidad (CbCr) (YCbCr). Estos modelos de color fueron ampliamente discutidos en García-Lamont et al. (2018) y Zarita et al. (1999). YCbCr y HSV son las técnicas de segmentación del color de la piel más utilizadas en lenguaje de señas y las ecuaciones para transformar la imagen RGB de entrada en YCbCr usando ecuaciones (diecisésis) a (18).

$$\text{H=porque } -\frac{1}{\sqrt{\frac{1}{2}(R-G)-\frac{\text{GRAMO}-B}{(R-\text{GRAMO})^2-(R-B)(G-B)}}} \quad (\text{diecisésis})$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (17)$$

$$V = \max(R, G, B) \quad (18)$$

y la del modelo de color HSV utilizando ecuación (19).

$$\begin{bmatrix} Y \\ R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0 & 0.299 & 0.587 & 0.114 \\ 128 & 128 & 0.500 & \text{GRAMO} \\ 128 & 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (19)$$

El valor de H, S, V, Y, Cb y Cr se determina y utiliza como valores de umbral para obtener el modelo de color requerido.

El espacio de color RGB es el menos preferido para la detección basada en colores y el análisis de colores porque es difícil identificar y establecer la piel humana a partir del tono de color RGB debido a la variación en la piel humana (Shaik et al., 2015). El modelo HSV es un mecanismo eficaz para determinar la piel humana en función del tono y la saturación. Los otros modelos eficientes son los modelos de espacio de color YUV y YIQ (Tabassum et al., 2010).

Investigación en Hartanto et al. (2014) y Huong et al. (2016) espacio de color RGB convertido en espacio de color HSV para identificar regiones de la piel. Tarik et al. (2012) transformó el video RGB de entrada en un modelo de color YCbCr. El valor obtenido contra cada píxel de cuadro se compara con un valor de umbral específico que aísle la región de la mano de toda la imagen. El modelo de color YCbCr se hibridó con el modelo de mezcla gaussiana (GMM) y la operación morfológica para obtener la región de color de la piel en Pan et al. (2016) re-

buscar. Athira et al. (2019) convirtió gestos dinámicos con una sola mano del video de entrada al modelo YCbCr y luego eliminó la región de la cara con solo la región de la mano. La eficiencia del espacio de color CIELAB fue estudiada por Mahmud et al. (2019). Investigaciones similares realizadas por Shaik et al. (2015) y Suharjito et al. (2019) encontró que YCbCr es más robusto y dio resultados más precisos que HSV y otras técnicas de segmentación del color de la piel en diferentes condiciones de iluminación.

### 3.3.2. Técnicas de detección de bordes

El enfoque de detección de bordes es una de las técnicas esenciales de procesamiento de imágenes. La técnica se basa en el cambio rápido del valor de intensidad en una imagen. Los algoritmos de detección de bordes detectan bordes en los que la primera derivada de la intensidad es mayor que un cierto umbral o la segunda derivada contiene cruces por cero. Una buena segmentación basada en bordes requiere tres pasos críticos: detectar bordes, eliminar bordes irrelevantes y unir los bordes. Las técnicas de detección de bordes revisadas en este artículo son el detector de bordes Robert, el detector de bordes Sobel, el detector de bordes Prewitt, el detector de bordes Laplaciano de Gauss y el detector de bordes Canny. Varias técnicas de detección de bordes se discutieron extensamente en Bhardwaj y Mittal (2012); Chinu y Chhabra, 2014; Maini y Aggarwal, 2009; Rashmi y Saxena, 2013; Shrivakshan y Chandrasekar, 2012).

*detector de borde robert*: El detector de bordes de Robert es un operador basado en gradientes que calcula la suma de los cuadrados de la diferencia entre píxeles diagonalmente adyacentes a través de diferenciación discreta y luego calcula el gradiente aproximado de la imagen. La imagen de entrada está convolucionada con  $2 \times 2$  operador de 2 núcleos con magnitud de gradiente y direcciones calculadas (Maini y Aggarwal, 2009). Para obtener la componente de gradiente en cada dirección ( $G_x$  y  $G_y$ ),  $2 \times 2$  el operador de 2 núcleos se aplica de forma independiente a la imagen de entrada. La ecuación utilizada para calcular la magnitud del gradiente  $|GRAMO|$  y direcciones  $\theta$  se dan como:

$$|GRAMO| = \sqrt{GRAMO_x^2 + GRAMO_y^2} \quad (20)$$

$$\theta = \arctan \left( \frac{GRAMO_y}{GRAMO_x} \right) - \frac{3\pi}{4} \quad (21)$$

dónde  $GRAMO_x$  y  $GRAMO_y$  son los gradientes en las direcciones x e y, respectivamente.

*Detector de bordes Sobel*: El detector de bordes de Sobel es un operador de diferenciación discreta de la derivada de primer orden que se utiliza para aproximar el gradiente de la función de intensidad de la imagen para la detección de bordes. La técnica convoluciona la imagen de entrada con un tamaño de núcleo de  $3 \times 3$  para obtener la magnitud del gradiente y enfatiza los píxeles más cercanos al centro del kernel. Esta técnica es sensible al ruido en comparación con la técnica Robert edge pero tiene un tiempo de cálculo rápido debido al tamaño del núcleo (Rashmi y Saxena, 2013; Sujatha y Sudha, 2015). La dirección del gradiente se calcula usando ecuación (22). La dirección del gradiente de  $\theta$  El detector de borde de Sobel se da como:

$$\theta = \arctan \left( \frac{GRAMO_y}{GRAMO_x} \right) \quad (22)$$

*Detector de bordes Prewitt*: El funcionamiento del detector de bordes Prewitt es muy similar al detector de bordes Sobel con un kernel de  $3 \times 3$  y se usa ampliamente para detectar los bordes verticales y horizontales de una imagen (Maini y Aggarwal, 2009). En el detector de bordes de Prewitt, la respuesta máxima de los ocho núcleos para una ubicación de píxel se utiliza para calcular la magnitud del gradiente de borde local. Al mismo tiempo, los énfasis no se colocan en los píxeles más cercanos al centro del núcleo (Shrivakshan y Chandrasekar, 2012). La técnica es menos costosa desde el punto de vista computacional, pero detecta muchos bordes falsos durante la detección, lo que da como resultado una imagen con ruido. Es susceptible al ruido y más efectivo con imágenes sin ruido (Muthukrishnan y Radha, 2011).

La ecuación para calcular la magnitud del gradiente del detector de bordes de Prewitt se da como ([Sujatha y Sudha, 2015](#)):

$$| GRAMO | = \max_i | GRAMO_i |, i=1 \text{ a } n \quad (23)$$

dónde  $| GRAMO |$  es el núcleo/respuesta en la posición de píxel particular, y  $n$  es el número de núcleos de convolución.

*Detector de bordes laplaciano de gaussiano (LoG):* esta técnica utiliza derivadas de segundo orden de las intensidades de los píxeles para localizar los bordes de una imagen. El laplaciano de una imagen se utiliza para resaltar áreas de cambio rápido de intensidad con el fin de detectar bordes. Antes de aplicar la función laplaciana, la imagen se somete a un filtro de suavizado gaussiano para reducir los niveles de ruido. Toma una sola imagen de nivel de gris como entrada y produce otra imagen de nivel de gris como salida ([Maini y Aggarwal, 2009](#)).

El laplaciano  $L(X,y)$  de una imagen con valores de intensidad de píxel  $I(X,y)$  se da como:

$$L(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (24)$$

dónde  $I$  es la intensidad de la imagen,  $X, y$  son las coordenadas del punto.

El filtro de suavizado gaussiano LoG utilizado para reducir el ruido en una imagen antes de aplicar la técnica laplaciana y se da como:

$$LoG(x, y) = -\frac{1}{\pi\sigma^4} \left[ 1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (25)$$

dónde  $\sigma$  denota la desviación estándar gaussiana.

La desviación estándar del filtro de suavizado gaussiano empleado en el filtro LoG influye drásticamente en el comportamiento del detector de bordes LoG. Un aumento en el valor de  $\sigma$  que da como resultado un filtro gaussiano más amplio y más suavizado, lo que puede dificultar la distinción de los bordes en una imagen.

*detector de borde astuto:* La técnica de detección de bordes de Canny ([Astuto, 1986](#)) ha sido sugerido como un detector de bordes óptimo ([Aliyu et al., 2020](#)). Es una de las técnicas estándar de detección de bordes y fue creada por primera vez por John Canny en el MIT en 1983 ([Muthukrishnan y Radha, 2011](#)). El objetivo de esta técnica es detectar bordes en una imagen y al mismo tiempo suprimir el ruido. Encuentra bordes separando el ruido de la imagen antes de encontrar los bordes de la imagen. El algoritmo básico utilizado para detectar Canny edge se da como:

**Paso 1:**Lea la imagen de entrada.

**Paso 2:**Suaviza la imagen con un filtro gaussiano del ker- elegido.

Tamaño del panel para reducir el ruido y los detalles no deseados (a veces se puede usar el filtro mediano porque preserva los bordes más que el filtro gaussiano).

**Paso 3:**El degradado de la imagen se utiliza para determinar el borde.

fortaleza. Para ello, se puede utilizar una máscara de Robert o una máscara de Sobel. La ecuación utilizada para calcular la magnitud del gradiente.  $| GRAMO |$  se da como:

$$\sqrt{| GRAMO_x |^2 + | GRAMO_y |^2} \quad (26)$$

$GRAMO_x$  y  $GRAMO_y$  los gradientes en las direcciones x e y, respectivamente.

**Etapa 4:**Encuentre la dirección del borde usando el gradiente en la x y direcciones y.

La dirección del gradiente  $\theta$  para el detector de borde de Robert se da

como:

$$\theta = \arctan \left( \frac{| GRAMO_y |}{| GRAMO_x |} \right) - \frac{3\pi}{4} \quad (27)$$

**Paso 5:**Las direcciones de borde calculadas se resuelven en horizontal, direcciones positivas, verticales, negativas.

**Paso 6:**Aplicar supresión no máxima para trazar el borde en el dirección del borde y suprimir cualquier valor de píxel por debajo del valor establecido (que no se considera un borde). Esto da una línea delgada en la imagen de salida.

**Paso 7:**Se aplica un umbral de histéresis para eliminar las rayas.

Diferentes investigadores han utilizado varias técnicas de detección de bordes en el reconocimiento de lenguaje de señas. Canny edge funciona mejor que muchas técnicas de detección de bordes que se han desarrollado. Es un método esencial y mejor sin alterar las características de los bordes de la imagen después. Sin embargo, el método enfrenta desafíos de ruido extraído con bordes y, a veces, se obtienen bordes disjuntos al seleccionar manualmente los bordes cuando se trabaja con múltiples imágenes. [Lionie et al. \(2012\)](#) emplearon la detección de bordes de Sobel en su estudio. Compararon el rendimiento de la detección de bordes de Sobel con la segmentación del color de la piel en el espacio de color HSI, el filtrado de paso bajo, la ecualización de histogramas y la desaturación. La técnica de desaturación obtuvo el mejor rendimiento, que convierte la imagen a escala de grises para eliminar el canal cromático y conservar solo el canal de intensidad. [Jayashree et al. \(2012\)](#) utilizó la detección de bordes de Sobel para extraer la región de interés de la imagen. Otras investigaciones de [Thepade et al. \(2013\)](#) realizaron la detección de bordes de Sobel en conjuntos de datos de imágenes. El resultado del sistema propuesto se comparó con otras cuatro técnicas de detección de bordes, a saber, canny, Robert, Prewitt y Laplace. Los resultados muestran que Sobel se desempeñó mejor, pero la detección de bordes Canny tuvo el mejor resultado general en comparación con otros. También, [Prasad et al. \(2016a, 2016b\)](#) usó canny edge con la transformada wavelet discreta para detectar píxeles de límite de la imagen de video de señal. La técnica ayuda a extraer bordes satisfactorios, conserva la mano y hace que la imagen de video sea exacta. En una investigación reciente, se utilizó la detección de bordes astutos para obtener el objeto de interés de la imagen y lograr un rendimiento óptimo en la precisión general ([Jin et al., 2016; Mahmud et al., 2019; Singh et al., 2019](#)).

### 3.3.3. Técnicas basadas en regiones

Las técnicas basadas en regiones agruparon los píxeles que exhibían límites cerrados similares en función de criterios predefinidos ([García-Lamont et al., 2018](#)). Esta técnica también se conoce como segmentación basada en similitudes ([Yogamangalam y Karthikeyan, 2013](#)) y requiere el uso de técnicas apropiadas de umbralización para agrupar sus similitudes. La similitud entre píxeles puede ser en forma de intensidad, color, forma o textura. Las técnicas basadas en regiones se clasifican además en métodos de crecimiento de regiones y división y fusión de regiones ([Divya y Ganesh Babu, 2020; Kaur' & Kaur, 2014; Yogamangalam y Karthikeyan, 2013](#)).

*Técnica de cultivo de regiones:* La técnica de crecimiento de regiones agrupa los píxeles que representan áreas similares en una imagen. Esto se hace agrupando píxeles cuyas propiedades, como la intensidad, el color y la forma, difieren en menos de una cantidad especificada. A cada región cultivada se le asigna una etiqueta entera única en la imagen de salida. El crecimiento de regiones es capaz de segmentar correctamente regiones que tienen las mismas propiedades y están espacialmente separadas. Sin embargo, requiere mucho tiempo y, a veces, produce resultados no deseados. El resultado del crecimiento de la región depende en gran medida de la selección de un criterio similar. Si no se elige correctamente, las regiones se filtran hacia áreas contiguas o se fusionan con áreas que no pertenecen al objeto de interés ([Rogowska, 2009](#)). La técnica de crecimiento de la región fue revisada extensamente por [Ikonomakis et al. \(2000\); Mehnert y Jackway \(1997\)](#) y [Verma et al. \(2011\)](#). Asumiendo  $p(x, y)$  es la imagen original que se va a segmentar,  $s(x, y)$  es la imagen binaria donde se encuentran las semillas y  $T$  es cualquier predicado que se va a probar para cada  $(x, y)$  ubicación. El algoritmo de crecimiento de la región basado en la conectividad 8 está dado por [Kaur' y Kaur \(2014\)](#) como sigue:

**Paso 1:**Todos los componentes conectados de  $s$  se erosionan.

**Paso 2:** Calcular una imagen binaria  $PAG_T$  donde  $PAG_T(X,y) = 1, si T(X,y) = Verdadero$

**Paso 3:** Calcular una imagen binaria  $q$ , donde  $q(X,y) = 1, si PAG_T(X,y) = 1 y (X,y) es 8-conectado a la semilla en S.$

Los componentes conectados en  $q$  son regiones segmentadas que se utilizan como el resultado de la técnica de crecimiento de la región.

**Técnica de división y fusión de regiones:** La técnica de segmentación basada en la división y fusión de regiones utiliza dos enfoques básicos; división y fusión para segmentar una imagen en varias regiones. La división se realiza dividiendo iterativamente una imagen en regiones máximas que tienen características similares, y la fusión contribuye a combinar las regiones adyacentes similares para formar una excelente imagen segmentada de la imagen original (Kaur<sup>1</sup> & Kaur, 2014). Una combinación de estas dos técnicas da un mejor rendimiento. El algoritmo básico para el crecimiento y fusión de regiones es el siguiente:

Sea  $p$  la imagen original,  $R$  representar subdividir la región y  $T$  sea el predicado particular.

**Paso 1:** Todo el  $R$  es igual  $apag$ .

**Paso 2:** Cada región se divide en cuadrantes para los cuales  $T(R) = Falso$ .

**Paso 3:** Si para cada región,  $T(R) = Cierto$ , luego fusionar adyacente regiones  $R_y R_t$  tal que  $T(R_y R_t) = Verdadero$ . **Etapa 4:**

Repita el paso 3 hasta que la fusión sea imposible.

Ikonomakis et al. (2000) resumir el procedimiento para la región de la siguiente manera:

i Dividir en cuatro cuadrantes inconexos cualquier región donde un criterio de geneidad no se sostiene.

ii Combinar cualquier región adyacente que satisfaga un criterio de homogeneidad.  $r_0$

iii Deténgase cuando no sea posible realizar más fusiones o divisiones.

**3.3.4. Técnicas de segmentación basadas en agrupamiento** La segmentación basada en agrupamiento es una técnica de aprendizaje no supervisado que divide un conjunto de elementos en grupos uniformes. Es muy utilizado para la segmentación de imágenes en clusters que tengan píxeles con características similares. Existen varias segmentaciones basadas en agrupaciones. Las dos técnicas principales más utilizadas incluyen K-means y Fuzzy C-means clustering (Cebeci & Yildiz, 2015; Ghosh y Dubey, 2013; Kan, 2013).

**K significa técnica:** K significa que el algoritmo es una técnica basada en agrupamiento, y ha sido ampliamente revisada en Dugirala (2020) y Panwar et al. (2016). Toma la distancia desde el punto de datos hasta el prototipo como la función principal de optimización. Las reglas de ajuste de la operación iterativa se obtienen por el método de encontrar los valores extremos de las funciones. El algoritmo K-means utiliza la distancia euclídea como la medida de similitud que encuentra la clasificación óptima de un vector central de clúster inicial para que el índice de evaluación sea mínimo. La función de criterio de suma de cuadrados de error se utiliza como función de criterio de agrupación. Aunque el algoritmo de K medias es eficiente, el valor de K debe darse por adelantado, y la selección del valor de K es muy difícil de estimar (Zheng et al., 2018).

K significa algoritmo por Ghosh y Dubey (2013) se da de la siguiente manera:

**Paso 1:** Establezca los clústeres deseados,  $K$  e Inicialice para elegir  $k$  inicio-puntos de referencia que se utilizan como estimaciones iniciales de los centroides del conglomerado.

**Paso 2:** Para clasificar todos los datos o imágenes:

(a) Calcule la distancia entre los centroides de los conglomerados y los datos.

(b) Mueva los datos más cerca del grupo que tiene menos distancia en comparación con otros.

**Paso 3:** Cálculo del centroide

- Cuando cada punto del conjunto de datos se asigna a un grupo, es necesario volver a calcular los nuevos k centroides.

**Paso 3:** Criterios de convergencia

- Los pasos (2) a paso (3) se repiten hasta que ningún punto cambie su asignación de clúster o hasta que los centroides ya no se muevan.

**Fuzzy C significa (FCM) Técnica:** La técnica de medios difusos C (FCM) asigna niveles de pertenencia y los utiliza para asignar elementos de datos a uno o más grupos. Proporciona un cálculo más preciso de la membresía del clúster y se ha utilizado con éxito para muchas aplicaciones de agrupación de imágenes (Dhanachandra y Chanu, 2017). FCM es un enfoque iterativo que genera una matriz de partición difusa y requiere un centro de clúster y una función objetivo. Los valores del centro del grupo y de la función objetivo se actualizan para cada iteración, y el proceso finaliza cuando la diferencia entre dos valores sucesivos de la función del objeto es menor que un valor de umbral predeterminado (Kan, 2013). La principal diferencia con el algoritmo de K medias es que, en lugar de tomar una decisión difícil sobre los grupos, a qué grupo pertenece, asigna un valor entre 0 y 1 que describe cada grupo. Requiere mucho tiempo de cálculo, a diferencia del agrupamiento de K-medias que produce casi los mismos resultados con menos tiempo de cálculo. El algoritmo FCM se da de la siguiente manera (Kan, 2013):

**Paso 1:** Asigne los valores para  $c$  (número de conglomerados;  $2 = do < norte$ ),  $q$  (exponente de ponderación de cada miembro borroso) y valor umbral  $\varepsilon$

**Paso 2:** Inicializar la matriz de partición  $tu = [V_{ij}]$  grado de membresía de  $X_i$  en el grupo.

**Paso 3:** Inicializar los centros de clúster y un contador  $pag$ .

**Etapa 4:** Calcule los valores de membresía y guárdelos en una formación.

**Paso 5:** Para cada iteración, calcule los parámetros  $apag$  y  $b_{pag}$  hasta que todos los píxeles se procesan, donde

$$apag = apag + V_{ij} \quad (28)$$

$$b_{pag} = apag + V_{ij} \quad (29)$$

**Paso 6:** Despues de cada iteración, actualice el centro del clúster y comparelo con el valor anterior ( $tub - tub-1$ )

**Paso 7:** Si la diferencia de comparación es menor que la definida valor umbral, detenga la iteración y repita el procedimiento.

**3.3.5. Técnica de segmentación basada en redes neuronales artificiales** Las técnicas de segmentación basadas en redes neuronales artificiales replicaron los mecanismos de aprendizaje del cerebro humano. Está formada por una gran cantidad de nodos conectados, cada uno con su peso, y la neurona corresponde a un píxel en una imagen (Kan, 2014). La imagen se entrena como una red neuronal utilizando muestras de entrenamiento y luego se crean las conexiones entre las neuronas con píxeles. Las imágenes recién creadas se segmentan a partir de la imagen entrenada. Hay dos pasos esenciales para esta técnica; extracción de características relevantes para la imagen y segmentación mediante una red neuronal. La técnica ha funcionado muy bien en un problema difícil de segmentación de imágenes. Red neuronal de retropropagación (BPNN), Red neuronal de alimentación hacia adelante (FFNN), Red neuronal de alimentación hacia adelante multicapa (MLFF), Perceptrón multicapa (MLP), Red neuronal de convolución (CNN) (Kanezaki, 2018) y Mapa de autoorganización (SOM),

**Tabla 6**

Resumen de ventajas y desventajas de varias técnicas de segmentación.

Segmentación Técnica	Ventajas	Desventajas
umbralización Método	Es un enfoque rápido y directo. No requiere información previa para operar. Tiene un bajo costo de cómputo.	Depende en gran medida de los picos, mientras que los detalles espaciales no se consideran. Sensible al ruido. La selección de un valor de umbral óptimo es difícil.
Basado en borde Método	Adecuado para imágenes con mejor contraste entre objetos.	No es adecuado para imágenes con demasiado ruido o demasiados bordes.
basado en la región Método	Es menos susceptible al ruido y más útil cuando es fácil definir criterios de similitud.	Es bastante caro en términos de tiempo de cálculo y consumo de memoria.
Método de agrupamiento	Es más útil para los desafíos del mundo real debido a la membresía parcial difusa empleada.	Determinar las funciones de pertenencia no es fácil.
Neural Artificial Basado en red Método	No requiere un programa complejo para funcionar. Menos propenso al ruido.	El tiempo computacional en entrenamiento es mayor.

se encuentran entre las redes neuronales más utilizadas para la segmentación de imágenes (amza, 2012;Kanezaki, 2018;Moghaddam y Soltanian-Zadeh, 2011). El algoritmo utilizado para separar la región de interés de su fondo se propuso en An y Liu (2019) y Zhao et al. (2010). En una red neuronal, una imagen se puede segmentar según la clasificación de píxeles o la detección de bordes. Secciones 3.3.1 a 3.3.5 explicó varias técnicas de segmentación, que incluyen umbralización, basada en bordes, basada en regiones, basada en agrupaciones y basada en redes neuronales artificiales. Tabla 6 resumen las ventajas y ventajas de las técnicas de segmentación.

Tabla 7 ilustra un resumen de los sistemas SLR basados en visión. En la tabla se presentan los dispositivos de adquisición de datos, las técnicas de preprocesamiento de datos y las técnicas de segmentación utilizadas.

#### 3.4. Técnicas de extracción de características

La extracción de características es una técnica utilizada para obtener las características más relevantes de la imagen de entrada. Su objetivo es encontrar las características más distintivas en la imagen adquirida (Patil y Sinha, 2017). Es una forma de reducción dimensional que representa efectivamente las partes interesantes de una imagen. El vector de características compacto se extrae eliminando una parte irrelevante para aumentar la precisión del aprendizaje y mejorar la visibilidad del resultado (Khalid et al., 2014; Kumar y Bhatia, 2014). La salida de extracción de características admite la etapa de clasificación al verificar las características que se pueden distinguir de manera efectiva entre clases y ayudan a lograr una alta precisión de reconocimiento. Las características extraídas de la región interesada se caracterizan en características de color, textura y forma (Patel y Gamit, 2016). Las técnicas de extracción de características importantes utilizadas en SLR que han logrado un buen rendimiento incluyen el análisis de componentes principales (PCA), el descriptor de Fourier (FD), el histograma de gradiente orientado (HOG), la transformación de características invariantes por desplazamiento (SIFT) y la aceleración de características robustas. (NAVEGAR).

##### 3.4.1. Análisis de componentes principales (PCA)

El análisis de componentes principales (PCA) es una técnica ampliamente utilizada para extraer características en el procesamiento de imágenes (Aliyu et al., 2020). PCA requiere un procedimiento estadístico que utiliza una transformación ortogonal para transformar una serie de observaciones de variables potencialmente correlacionadas en un conjunto de valores para variables no correlacionadas (Kumar y Bhatia, 2014). Opera calculando nuevas variables llamadas componentes principales y usa la nueva variable para crear una combinación lineal de las variables iniciales. La operación PCA es la siguiente: primero, los componentes principales producen la mayor variación potencial, mientras que se extrae la mayor parte. A esto le sigue el cálculo de los vectores propios y los valores propios correspondientes. Los vectores propios calculados se almacenan disminuyendo el orden de

los valores propios, que forman la reducción de la dimensionalidad de los datos en PCA. El algoritmo utilizado para calcular PCA es el siguiente (Cheok et al., 2019; Karamizadeh et al., 2013):

**Paso 1:** La columna o fila del tamaño del vector representa un entrenamiento dado conjunto de imágenes,  $M$  con un  $S$ -vector dimensional.

**Paso 2:** El significado,  $m$  de todas las imágenes en el conjunto de entrenamiento es componer usando ecuación (30):

$$m = \frac{1}{N} \sum_{i=1}^N X_i \quad (30)$$

con  $X_i$  como el  $i^{\text{a}}$  imagen a través de sus columnas concatenadas en un vector.

**Paso 3:** Vectores base PCA que son vectores propios de la dispersión matriz,  $S$  se calculan usando ecuación (31):

$$S = \sum_{i=1}^N (X_i - \mu)(X_i - \mu)^T \quad (31)$$

**Etapa 4:** Los vectores propios y los valores propios correspondientes fueron calculado. Los valores de los vectores propios se almacenan en orden decreciente de valores propios. Los vectores propios con valores propios más bajos contienen menos información sobre la distribución de datos. Estos se filtran para reducir la dimensionalidad de los datos.

Huong et al. (2016) empleó PCA para extraer las características necesarias para reconocer 25 alfabetos de lenguaje de señas vietnamita (VSL) bajo un fondo uniforme. Investigación en Zaki y Shaheen (2011) introdujo PCA con posición de curtosis y código de cadena de movimiento (MCC). La curtosis se fusionó con PCA para calcular los bordes que denotan la articulación, y MCC extrajo los vectores que representan el movimiento de la mano. El hallazgo de la investigación muestra que la combinación de estas tres técnicas de extracción de características mejoró la precisión del reconocimiento en un 89,90 % en comparación con el uso por separado o la combinación de dos métodos. En un estudio similar presentado por Li et al. (2016), PCA se combinó con el algoritmo K-means basado en entropía y se aplicó al modelo oculto de Markov (HMM). La imagen de entrada utilizada para la extracción de características se adquirió a partir de datos basados en guantes utilizando un sensor del Sistema de referencia de rumbo y actitud (AHRS). Simultáneamente, se eliminó el ruido de una imagen poco clara utilizando un filtro de paso bajo (LPF). El modelo desarrollado demostró un mejor rendimiento que el resultado del filtro de Kalman. Prasad et al. (2016a, 2016b) propuso un estudio que combinó PCA y descriptores elípticos de Fourier para la extracción de características en un conjunto de datos de video. Los descriptores elípticos de Fourier se utilizaron para optimizar y preservar los detalles de la forma sin ningún cambio en la rotación. Simultáneamente, usando PCA, la salida del vector de características para un signo dado de múltiples marcos se trata para formar un solo vector.

**Tabla 7**

Resumen de la arquitectura de reconocimiento de lenguaje de señas basada en la visión.

Referencias	Datos Adquisición	Datos adquiridos/ Conjunto de características	conjunto de datos Recopilación	preprocesamiento Técnica	Segmentación Técnica
Kishore y Kumar (2012)	Cámara	Gesto manual y cabeza movimiento.	Adquirido por cuenta propia	escala de grises imagen y gaussiano filtración.	de Otsu umbralización y discreto ondícula transformar con borde astuto detección.
Pansare et al. (2012)	Cámara	Formas de manos.	Adquirido por cuenta propia	escala de grises imagen, mediana filtro, gaussiano filtrado y Morfológico operaciones.	umbralización, gota y cultivo, y borde Sobel detección.
Thepade et al. (2013)	Cámara	Manual (formas de manos).	Adquirido por cuenta propia	-	Borde sobel detección comparado con (astuto, Robert, Prewitt y Laplace).
Jin et al. (2016)	Cámara	formas de manos	Adquirido por cuenta propia	color RGB imagen redimensionada hacia resolución de 320x240 píxeles filtro medio, escala de grises, Contraste estirar.	borde astuto detector y región sembrada creciente.
Carneiro et al. (2017)	Kinect	Gestos con las manos	Adquirido por cuenta propia		Fondo sustracción.
Prasad et al., 2016)	Cámara	formas de manos y cabeza posición	Adquirido por cuenta propia	escala de grises, Morfológico operaciones y filtro medio.	borde astuto con el discreto ondícula transformar (DWT).
Raheja et al. (2016)	Kinect y Cámara	formas de manos	Adquirido por cuenta propia	Filtro de piel	Color de piel (Color VHS espacio).
islam et al. (2017)	Cámara	formas de manos	Adquirido por cuenta propia	-	el método de Otsu, Mediana filtrado y morfológico.
Mahmud et al. (2019)	Cámara	formas de manos	Público	logarítmico Transformación e Histograma igualdad (ÉL).	color CIELAB espacio y borde astuto detección.

### 3.4.2. Descriptor de Fourier (FD)

Los descriptores de Fourier se utilizan para caracterizar la complejidad de la forma y se han utilizado para identificar diferentes formas de signos ([Agrawal et al., 2012;Kishore et al., 2015](#)). Los coeficientes transformados de Fourier forman los descriptores de Fourier de la imagen y representan la forma en el dominio de la frecuencia, descriptores de menor y mayor frecuencia ([Cheok et al., 2019](#)). Los descriptores de frecuencia más baja contienen información sobre las características generales de la forma, mientras que los descriptores de frecuencia más alta contienen información sobre la parte relevante de la forma.

Se da la coordenada compleja de los píxeles del límite ([Shukla et al., 2016](#)) como:

$$yo(k) = [Xk + jyk] \quad (32)$$

El coeficiente complejo de los descriptores de Fourier de las coordenadas de frontera ( $tu$ ) se da como:

$$un(u) = \sum_{k=0}^{K-1} yo(k)^m e^{j2\pi u k} \quad (33)$$

donde K es el número total de píxeles de la imagen,  $k = 0, 1, 2, \dots, K-1$ ,  $u = 0, 1, 2, \dots, K-1$  y  $(X, y)$  son las coordenadas del punto.

El coeficiente del descriptor de Fourier para lograr la invariancia de rotación, escala y traslación se da como:

$$un(u) = \frac{un(u)}{a(0)}, \text{ donde } tu = 0, 1, \dots, K-1 \quad (34)$$

Para el reconocimiento de formas, los descriptores de Fourier son particularmente efectivos debido a su invariancia en términos de escala, rotación y traducción. En la investigación presentada por [Kumar y Bhatia \(2014\)](#), se utilizaron descriptores de Fourier de distancia centroide (CeFD) para la extracción de características. [Izzah y Suciati \(2014\)](#) empleó un descriptor de Fourier genérico (GFD) para extraer vectores de características para el lenguaje de señas en tiempo real. Sus hallazgos mostraron que GFD con el clasificador K-Nearest Neighbor (KNN) funcionó mejor que el resultado de PCA con KNN. En investigaciones similares de [Kishore et al. \(2015\)](#), se utilizaron descriptores elípticos de Fourier de una curva cerrada para obtener vectores de características en la imagen. Simultáneamente, la clasificación se realizó utilizando un algoritmo de red neuronal artificial (ANN). El sistema logró una mejor precisión de reconocimiento del 95,10%. [Shukla et al. \(2016\)](#) empleó el descriptor de Fourier (FD) y la deformación dinámica del tiempo (DTW)

para obtener las funciones necesarias para el reconocimiento de gestos de 26 alfabetos del lenguaje de señas indio (ISL). Su sistema funcionó mejor que otros estudios (Rekha et al., 2011a, 2011b; Agrawal et al., 2012; Kishore et al., 2015). Se empleó una combinación de SIFT, momentos Hu y FD por Pan et al. (2016) para extraer características de imágenes dadas. Las características de la imagen se reducen mediante la aplicación de técnicas PCA y LDA y se realizan tanto para el lenguaje de señas chino (CSL) como para los alfabetos ASL.

#### 3.4.3. Histograma de gradiente orientado

El histograma de gradiente orientado (HOG) es un descriptor de características utilizado para identificar un objeto en el procesamiento de imágenes. Las características obtenidas de HOG ofrecen una representación de imágenes concisa y eficiente para la clasificación de imágenes. Es una de las técnicas de extracción de características más robustas utilizadas en los últimos tiempos para identificar formas o estructuras dentro de una imagen (Torrione et al., 2014). Las funciones HOG se utilizan para medir el grado del gradiente de la imagen de entrada y su ruta de gradiente. El concepto central detrás de las funciones HOG muestra que la entrada y el contorno del objeto se consideran mediante la circulación de las direcciones de los bordes (Mahmud et al., 2019). Tavari y Deorankar (2014) introdujo HOG como una técnica de extracción de características. En su estudio, las características extraídas se calculan contando las ocurrencias de orientación de gradiente en partes localizadas de una imagen. El algoritmo utilizado para el descriptor HOG es el siguiente:

**Paso 1:** El degradado de la imagen se calcula tamizando con máscara derivada distintiva unidimensional horizontal y vertical dada como:

$$\begin{bmatrix} D_x & \\ & D_y \end{bmatrix} = \begin{bmatrix} [-1, 0, 1] & \\ & 0 & \\ & & -1 \end{bmatrix} \quad (35)$$

donde  $D_x$  y  $D_y$  máscaras horizontales y verticales respectivamente y obtenidas de  $\nabla$  y derivados utilizando la siguiente convolución de lución:

$$I_x = I * D_x \quad (36)$$

$$I_y = I * D_y \quad (37)$$

La magnitud del gradiente se obtiene como:

$$|I_{GRAD}| = \sqrt{I_x^2 + I_y^2} \quad (38)$$

La orientación del gradiente se da como:

$$\theta = \arctan \frac{-1 \cdot I_x}{I_y} \quad (39)$$

**Paso 2:** Cree los histogramas de celdas. Cada píxel calcula un binario por un canal de histograma basado en la orientación basado en los valores encontrados en el cálculo del gradiente. Las celdas en sí eran rectangulares, y los canales del histograma se distribuyeron uniformemente entre 0 y 180 o entre 0 y 360, dependiendo de si el gradiente no estaba firmado o no.

**Paso 3:** Para cambiar la iluminación y el contraste, la fuerza del gradiente se normaliza regionalmente, lo que necesita agrupar las celdas en bloques conectados espacialmente más grandes.

**Etapa 4:** Se obtienen los vectores de características finales.

Mahmud et al. (2019) usó HOG para segmentar la imagen en 64 bloques, mientras que cada bloque constituye 2x2 celdas. La orientación del borde del histograma se manipula para cada píxel en una celda creada para obtener la dirección del gradiente, el agrupamiento de orientación y los bloques descriptores que son la característica extraída y almacenada en la matriz de características. Su sistema funcionó mejor con el clasificador K-NN que el resultado obtenido usando Bolsa de características con una máquina de vectores de soporte en el mismo experimento. Raj y Jasuja (2018) usó

HOG como técnica de extracción de características para los alfabetos del lenguaje de señas británico. Las características se obtienen constituyendo un 8xVentana de 8 píxeles de tamaño, que se desliza sobre la imagen. Investigaciones similares (Butt et al., 2019) utilizaron HOG con Local Binary Pattern (LBP) y características estadísticas para extraer imágenes esenciales. LBP se utiliza para mejorar la salida de las características extraídas. A pesar de la eficiencia de HOG en muchos estudios sobre el reconocimiento del lenguaje de señas, se observó que la elección de los parámetros de HOG afecta el tamaño del vector de características, lo que genera problemas de precisión y tiempo de cálculo. Joshi et al. (2020) propuso un vector de características HOG multinivel para abordar los desafíos de selección de parámetros HOG. El estudio adopta un método combinado de toma de decisiones basado en Taguchi y Técnica para el orden de preferencia por similitud con la solución ideal (TOPSIS) para determinar el conjunto óptimo de parámetros de vectores de características HOG multinivel. El sistema propuesto demuestra un mejor rendimiento que el método de última generación en conjuntos de datos ISL de fondo complejos.

#### 3.4.4. Transformación de características invariante al desplazamiento (SIFT)

SIFT extrae características de una imagen con una variante diferente, como traslación, escala y rotación. Ofrece un marco sólido para detectar características únicas e invariantes de la imagen que coinciden de manera sólida con diferentes vistas de una imagen, como la traducción, la rotación y el escalado, y algunos cambios invariantes en la iluminación y la perspectiva de la cámara. Los cuatro pasos computacionales SIFT (bajo, 2004) incluir:

**Paso 1:** Determine la ubicación aproximada y la escala de las características sobresalientes. puntos de tura

La función de diferencia de Gaussian (DoG) se utiliza para identificar puntos de interés potenciales y mejorar la velocidad de cálculo, que son invariantes a ubicaciones y escalas (Mistry y Banerjee, 2017).

El espacio de escala de una imagen  $L(x, y, \sigma)$  se calcula realizando la convolución entre la función gaussiana  $G(x, y, \sigma)$  con una imagen de entrada  $y(x, y)$ .

La función escala-espacio se da como:

$$L(x, y, \sigma) = G(x, y, \sigma) * y(x, y) \quad (40)$$

La función gaussiana se da como:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (41)$$

$D(x, y, \sigma)$  se puede calcular a partir de la diferencia de dos escalas cercanas separadas por un factor multiplicativo constante  $k$ .

La diferencia de la ecuación de Gauss se da como:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * y(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (42)$$

#### Paso 2: Localización de puntos clave

La localización del punto clave se determina eliminando el bajo contraste (respuesta de borde) y se selecciona en función de sus medidas de estabilidad. La ubicación del punto clave,  $\theta(x, y)$ , escala,  $m(x, y)$  se calcula como se indica en ecuaciones (43) y (44), respectivamente:

$$m(x, y) = \sqrt{\frac{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}{4}} \quad (43)$$

$$\theta(x, y) = \arctan \frac{(L(x, y+1) - L(x, y-1))}{(L(x+1, y) - L(x-1, y))} \quad (44)$$

**Paso 3:** Determine la(s) orientación(es) para cada punto clave.

En el tercer paso (cálculo de orientación), las orientaciones se asignan en función del gradiente de la imagen en cada ubicación de punto clave

**Etapa 4:** Obtenga un descriptor para cada punto clave.

Los puntos clave se transforman en una representación que permite niveles significativos de distorsión de la forma local y cambios en la iluminación.

En la literatura se pueden encontrar varios estudios sobre sistemas de reconocimiento de signos en los que se emplea SIFT para la extracción futura (Agrawal et al., 2012; Tharwat et al., 2015; Yasir et al., 2016; Patil y Sinha, 2017; Shanta et al., 2018). Patil y Sinha (2017) desarrolló un marco que se centró en el tiempo necesario para implementar varias fases del algoritmo SIFT. Propuso un sistema que hibridaba SIFT con umbralización adaptativa y suavizado de imágenes borrosas gaussianas para la extracción de características. Duddal et al. (2019), Agrawal et al. (2012) descriptor de forma combinado con métodos HOG y SIFT. Se utilizó un descriptor de forma para analizar la forma general de la imagen segmentada, mientras que los descriptores HOG se emplearon para la invariancia al cambio de iluminación, la orientación y los gestos articulados u ocluidos asociados con los gestos con las dos manos. Los puntos clave para cada imagen de signo se determinaron mediante la técnica SIFT y se almacenaron como un vector de características para su clasificación. Sus hallazgos muestran un mejor desempeño con la combinación de tres técnicas.

### 3.4.5. Acelerar función robusta (SURF)

Speed Up Robust Feature (SURF) es una técnica eficaz para la extracción de puntos característicos (Raj y José, 2016). Es un marco recientemente desarrollado para mejorar el rendimiento de un sistema de reconocimiento de objetos. Está diseñado como una alternativa eficiente a SIFT, que es mucho más rápido y robusto que SIFT (Sykora et al., 2014). Los descriptores se derivan del píxel que rodea un punto interesante. El SURF puede detectar objetos en imágenes tomadas bajo diferentes entornos intrínsecos y extrínsecos (Bhosale et al., 2014). SURF usa imágenes integrales en lugar de la Diferencia de Gaussianas (DoG) empleada en SIFT. La imagen integral es la suma del valor de intensidad de los puntos de la imagen con una ubicación menor o igual que la imagen original (Cheok et al., 2019).

la imagen integral  $y_o(x, y)$  ecuación presentada por Mistry y Banerjee (2017) se da como:

$$y_o(x, y) = \sum_{j=0}^x \sum_{i=0}^y y_o(y, j) \quad (45)$$

Para obtener puntos de interés, SURF utiliza un detector de manchas de arpillería. El determinante de la matriz hessiana define la magnitud de la respuesta. La imagen integral está complicada con el filtro de caja, mientras que el filtro de caja es el filtro aproximado del filtro gaussiano (Mistry y Banerjee, 2017).

La matriz hessiana se da como:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{yx}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (46)$$

dónde  $H(X, \sigma)$  es arpillería matriz,  $X = (x, y)$  de una imagen  $I$  a escala  $\sigma$ ,  $L_{xx}(X, \sigma)$  (Laplaciano de Gauss) es la convolución de la derivada de segundo orden gaussiana;  $\partial_x^2 g(\sigma)$  con la imagen  $I$  en el punto  $X$  y de manera similar para  $L_{xy}(X, \sigma)$  y  $L_{yy}(X, \sigma)$ .

SURF es una técnica de extracción de características muy adecuada para la imagen debido a sus atributos eficientes que incluyen invariancia de escala, invariancia de traducción e invariancia de rotación. SURF ha atraído una gran cantidad de investigación sobre los sistemas de reconocimiento de lenguaje de señas (Rekha et al., 2011a; Yang y Peng, 2014; Fang Hu et al., 2014; Wang et al., 2015a, 2015b; Jin et al., 2016; Ravi et al., 2018; Kaluri y Reddy, 2017; Narayanan et al., 2018).

Hartanto et al. (2014) llevó a cabo un estudio sobre el reconocimiento de alfabetos utilizando imágenes de gestos con las manos. El algoritmo SURF se emplea para detectar y extraer características de puntos clave. Sykora et al. (2014) compararon los métodos SIFT y SURF de diez gestos. Los hallazgos muestran que la precisión general de SURF es mayor que la del método SIFT. En el estudio de Merin Joy y Rajesh (2017), el detector de características y

los descriptores se extrajeron del video de vigilancia utilizando el método SURF. Los descriptores obtenidos de diferentes imágenes para el comparador SURF se utilizaron posteriormente encontrando los pares coincidentes para el reconocimiento. Rekha et al. (2011a, 2011b) hibridaron los métodos SURF y Hu Moment Invariant para lograr una buena tasa de reconocimiento junto con una baja complejidad de tiempo. Elouariachi et al. (2021) propuso nuevas técnicas de extracción de características llamadas Quaternion Krawtchouk Moments (QKM) para el reconocimiento del lenguaje de señas. Los autores compararon la técnica QKMs con el clasificador KNN con varias técnicas tradicionales de extracción de características (HOG, LBP, SIFT, SURF y Gabor). Su hallazgo muestra un mejor rendimiento para el método propuesto en comparación con las técnicas convencionales.

El tipo de características extraídas determina el rendimiento del reconocimiento. Por lo tanto, es necesario elegir una buena técnica de extracción de características para lograr un buen rendimiento. Tabla 8 resume las ventajas y desventajas de varias técnicas de extracción de características utilizadas en los sistemas de reconocimiento de lenguaje de señas basados en visión.

### 3.5. Categorías de lenguaje de señas

Los lenguajes de señas se utilizan para lograr las mismas funciones que los lenguajes hablados y, al mismo tiempo, interpretar las palabras habladas en lenguaje de señas. La estructura del lenguaje de señas es diferente de las palabras habladas y tiene su propia fonología, morfología sintáctica, vocabulario y gramática distintas de las lenguas habladas (Agrawal et al., 2016; Sahoo et al., 2014). Para lograr esto, las Señales se hacen en o cerca del cuerpo del firmante con una o dos manos en una configuración particular de manos (Schembri et al., 2013; Wadhawan y Kumar, 2021). Los cinco parámetros básicos utilizados en la comunicación del lenguaje de señas incluyen la forma de la mano, la ubicación de la mano, la orientación de la palma, el movimiento de la mano y la expresión facial; éstos influyen en el significado de un signo. El significado del signo cambia cuando cambia cualquiera de estos parámetros (Wilcox y Occhino, 2016).

Los lenguajes de señas no son universales como los lenguajes hablados, y varían entre países debido a diferentes ubicaciones geográficas, nacionalidades, fronteras sociales y vocabularios (Lemaster & Mon-Aghan, 2007). Hay varios lenguajes de señas en todo el mundo; estos incluyen lenguaje de señas americano (ASL), lenguaje de señas árabe (ArSL), lenguaje de señas británico (BSL), lenguaje de señas australiano (Auslan), lenguaje de señas chino (CSL), lenguaje de señas francés (FSL), lenguaje de señas indio (ISL), lenguaje de señas danesa (DSL), lenguaje de señas alemana (DGS), lenguaje de señas de Taiwán (TSL), lenguaje de señas tailandesa (TLS), lenguaje de señas finlandesa (FinSL), lenguaje de señas brasileña (LSB) y muchas otras que han evolucionado en comunicación para personas con discapacidad auditiva corbatas (Agrawal et al., 2016; Cheok et al., 2019; Kadhim y Jamees, 2020; Sahoo et al., 2014; Wadhawan y Kumar, 2021). Es importante saber que la mayoría de los países que comparten un idioma hablado no comparten el mismo lenguaje de señas. Esto sucede porque las personas sordas y con discapacidad auditiva de los dos o más países que hablan el mismo idioma no estaban en contacto. Por lo tanto, el lenguaje de señas de estos países se desarrolló de forma independiente. Los desafíos en el desarrollo de un sistema de reconocimiento de lenguaje de señas incluyen la velocidad de señas, que varía según el firmante. La segmentación de la región de interés de imágenes completas también es problemática debido a los diferentes entornos en los que se toman las imágenes, las iluminaciones, el tiempo de cómputo, el seguimiento de gestos, la naturaleza de los parámetros utilizados para la comunicación, como la forma de la mano, la expresión facial, la orientación, etc. (Agrawal et al., 2016).

### 3.6. Conjuntos de datos de referencia

Hay varios conjuntos de datos de referencia de acceso público para evaluar el rendimiento de los sistemas de reconocimiento de lenguaje de señas estáticos, aislados y continuos. Para el lenguaje de señas americano

**Tabla 8**

Resumen de varias ventajas y desventajas de las técnicas de extracción de características.

Técnica de extracción de características	Ventajas	Desventajas
Análisis de componentes principales (PCA)	Baja sensibilidad al ruido. Reduzca las características redundantes y la gran dimensionalidad de los datos. Por lo tanto, mejora el rendimiento del algoritmo. Ayuda a superar los problemas de sobreajuste de datos al disminuir la cantidad de funciones.	La interpretación de los datos de pérdida PCA solo tiene en cuenta las relaciones por pares entre los véxoles de las imágenes cerebrales. Es difícil evaluar adecuadamente la matriz de covarianza. Es la invariancia a la escala.
Descriptor de Fourier (FD)	Invariantes de traslación, escala y rotación. La complejidad computacional es más diminuta. Una técnica muy robusta	La transformación afín y la resistencia al ruido son malas. Las características locales no se pueden ubicar porque solo se conocen las magnitudes de las frecuencias en lugar de las ubicaciones.
Transformación de características invariantes de escala (SIFT)	Es más preciso que cualquier otro descriptor. Es rotación y escala invariante. Está cerca del rendimiento en tiempo real.	SIFT es relativamente lento en comparación con SURF. No es bueno en los cambios de iluminación.
Acelerar funciones robustas (SURF)	Es más robusto a los cambios de escala y rotación. SURF es más rápido que SIFT en aplicaciones en tiempo real. Tiene baja dimensionalidad en comparación con SIFT. Reduce el tiempo de cálculo.	Alta dimensión del descriptor de características. Alto tiempo computacional. No es estable a la rotación y no funciona correctamente con iluminación. No es efectivo para dispositivos de baja potencia.
Histograma de Gradiente Orientado (HOG)	Captura la estructura de borde o degradado que es muy característica de la imagen local.  Es invariable a las transformaciones de escala e iluminación.	Es susceptible a la rotación de imágenes. Es computacionalmente costoso debido a su enfoque de escaneo en profundidad sobre toda la región de interés.

(ASL), el conjunto de datos Purdue RVL-SLLL ([Martínez et al., 2002](#)) consta de gestos, movimientos, palabras y frases firmadas por catorce (14) firmantes. Comprende 2576 videos, que componen 184 videos por firmante. Treinta y nueve (39) de los videos son primitivas de movimiento aisladas, 62 formas de manos y oraciones. Conjunto de datos de video del léxico del lenguaje de señas estadounidense (ASLLVD) ([Athitsos et al., 2008](#)) abarca secuencias de video de alta calidad de alrededor de 3800 signos ASL correspondientes a alrededor de 3000 señales firmadas por cuatro hablantes nativos. RWTHBOSTON-104 ([Zahedi et al., 2006](#)) y RWTHBOSTON-400 ([Dreuw et al., 2008](#)) los conjuntos de datos consisten en ASL aislado y continuo. El conjunto de datos RWTHBOSTON-104 contiene lenguaje de señas aislado con un vocabulario de 104 señas y 201 oraciones firmadas por tres señas. El conjunto de datos RWTHBOSTON-400 se crea para desarrollar el reconocimiento continuo de ASL. Comprende 843 oraciones con un tamaño de vocabulario de 406 palabras firmadas por cuatro (4) firmantes. Además, el conjunto de datos de la Universidad de Massey ([Barczak et al., 2011](#)) consta de 36 clases de alfabetos (AZ) y números (0-9). El número total de imágenes es 2160. Para el conjunto de datos del lenguaje de señas árabe (ArSL), la base de datos del lenguaje de señas ([Assaleh et al., 2012](#)) contiene 40 oraciones, y cada oración se repite 19 veces. Fue adquirido de ochenta firmantes. El Atlas Mundial de Signos ([Shohieb et al., 2015](#)) tiene alrededor de 500 gestos estáticos (deletreo con los dedos y movimientos de las manos) con gestos dinámicos (signos no manuales) que involucran lenguaje corporal, lectura de labios y expresiones faciales. El conjunto de datos del lenguaje de señas brasileño (LIBRAS-HCRGBDS) consta de sesenta y una configuraciones manuales de Libras ([Porfirio et al., 2013](#)). El sensor Kinect se utilizó para recopilar el conjunto de datos de aproximadamente 610 videoclips obtenidos de cinco firmantes diferentes. El conjunto de datos del lenguaje de señas británico fue creado por British Sign Language (BSL) Corpus ([Schembri et al., 2013](#)) y compuesto por videos que contenían 249 personas conversando en BSL con anotaciones de 6330 gestos de la conversación. RWTH-PHOENIX-Tiempo 2014 ([Forster et al., 2014](#)) y base de datos SIGNUM ([Agris y Kraiss, 2007](#)) fueron creados para el reconocimiento del lenguaje de señas alemán. RWTH-

PHOENIX-Weather 2014 contiene lenguaje de señas continuo de 6861 oraciones y 1558 vocabularios. La base de datos SIGNUM contiene un tamaño de vocabulario de 450 gestos básicos y 780 oraciones firmadas por 25 firmantes. [Tabla 9](#) conjuntos de datos de referencia resumidos de diferentes países utilizados por varios investigadores. Estos conjuntos de datos se utilizaron como base para la comparación o evaluación del rendimiento del modelo desarrollado.

#### 4. Revisión de arquitecturas de clasificación inteligente empleadas en reconocimientos de lengua de signos

Una vez que se ha completado el preprocesamiento, la segmentación y la extracción de características de las imágenes, es necesario utilizar un algoritmo de predicción para ayudar a dar un significado valioso a las características extraídas. Al igual que los humanos aprenden haciendo tareas repetidamente, las máquinas también están entrenadas para aprender y el aprendizaje automático mejora su rendimiento. El aprendizaje automático es un subcampo de la informática y también se clasifica como un método de inteligencia artificial ([Voyant et al., 2017](#)). Las técnicas de inteligencia artificial utilizadas para el reconocimiento del lenguaje de señas incluyen supervisadas o no supervisadas. El aprendizaje automático supervisado tomó un conjunto de datos de entrenamiento conocidos y los usó para inferir una función a partir de datos de entrenamiento etiquetados, mientras que; El aprendizaje automático no supervisado se utiliza para extraer inferencias de conjuntos de datos con datos de entrada sin respuesta etiquetada. Después de una revisión exhaustiva de la literatura, los predictores inteligentes comúnmente utilizados para el reconocimiento del lenguaje de señas son el vecino más cercano (KNN), la red neuronal artificial (ANN), la máquina de vectores de soporte (SVM), el modelo oculto de Markov (HMM), la red neuronal convolucional (CNN), lógica difusa y aprendizaje conjunto. Esta sección informa sobre las técnicas de aprendizaje automático utilizadas para el reconocimiento del lenguaje de señas.

**Tabla 9**

Conjuntos de datos de referencia para el lenguaje de señas en diferentes países.

Idioma	Nombre del conjunto de datos	tipo de señal	Fuente
Lenguaje de señas americano (ASL)	La base de datos RVL-SLLL de Purdue.	Estático, aislado y continuo.	(Martínez et al., 2002) (
	Conjunto de datos de video del léxico del lenguaje de señas estadounidense (ASLVD), RWTHBOSTON-104	Continuo	Athitsos et al., 2008) (
	RWTH BOSTON-400	Aislado y Continuo. Estática	Zahedi et al., 2006) (
	Conjunto de datos de Massey	Aislada y Continua.	Dreuw et al., 2008) (
Lengua de señas árabe (ArSL)	Base de datos de lenguaje de señas árabe	Continuo	Barczak et al., 2011) (
Lenguaje de señas brasileño (Libras)	Signs World Atlas, una lengua de signos árabe de referencia	Estático y continuo	Assaleh et al., 2012) (
Lenguaje de señas británico (BSL)	LIBRAS-HCRGBDS		Shohieb et al., 2015) (
Lenguaje de señas alemán (DGS)	Corpus de lenguaje de señas británico RWTH-PHOENIX-Tiempo	Continuo	Porfirio et al., 2013) (
	2014 La base de datos SIGNUM	Continuo	(Schembri et al., 2013) (
			Forster et al., 2014) (
			Agris y Kraiss, 2007)

#### 4.1. Algoritmo K-vecino más cercano (K-NN)

El algoritmo K-Nearest Neighbor también se conoce como aprendizaje perezoso. Se basa en el principio de que las instancias dentro de un conjunto de datos generalmente existirán muy cerca de otras instancias con propiedades similares (Selim et al., 2019). Predice la clase de un nuevo objeto en función de las clases de sus k vecinos más cercanos realizando una votación por mayoría simple para decidir la clase de la instancia de prueba (Alamelu et al., 2013). El parámetro 'k' en K-NN se refiere al número de vecinos más cercanos de un punto de datos de prueba para incluir en la mayoría de los procesos de votación (Kotsiantis et al., 2006). El procedimiento para diseñar el algoritmo de clasificación K-NN presentado en Mahmud et al. (2019) se da de la siguiente manera:

**Paso 1:** conjunto de datos de entrenamiento y nuevo conjunto de datos de imagen de entrada para prueba  
ing fueron cargados.

**Paso 2:** Se eligió el valor K, que es el punto de datos más cercano.

**Paso 3:** Para cada punto en los datos de prueba, haga lo siguiente:

- i Distancia entre cada fila de datos de prueba y cada fila de entrenamiento  
Los datos de medición se calcularon utilizando métricas de distancia. La ecuación de distancia euclíadiana se da como:

$$\text{distancia euclíadiana} = \sqrt{\sum_{i=1}^{norte} (a_i - b_i)^2} \quad (47)$$

dónde  $a$  son los datos de prueba,  $b$  son los datos de entrenamiento y  $n$  es el numero de vecino

i Los valores de distancia obtenidos fueron ordenados en orden ascendente. ii Se eligieron los valores de las K filas superiores de la matriz ordenada.

iii La clase fue asignada para probar el voto mayoritario basado en puntos de datos y una clase más frecuente de estas filas.

**Etapa 4:** Repita los pasos 1 a 3 para todas las imágenes de prueba en el

conjunto de datos

K-NN es uno de los algoritmos de aprendizaje automático que utiliza medidas de distancia como sus características principales. Algunos de los métodos de clasificación que se basan en la medición de la distancia utilizados para el reconocimiento del lenguaje de señas incluyen la distancia de Mahalanobis (Huong et al., 2016), Manhattan (Yanutara et al., 2012) y la distancia euclíadiana (Pansare et al., 2012; Tara et al., 2012; Hartanto et al., 2014; Huong et al., 2016). Tara et al. (2012) utilizó el método de distancia de Manhattan para lograr una precisión de reconocimiento del 95 % con una latencia computacional pequeña. Un estudio similar de Huong et al. (2016) investigó y comparó los métodos de distancia euclíadiana y distancia de Mahalanobis. El rendimiento del sistema reveló una precisión de reconocimiento del 90,4 % y el 91,5 % para la distancia euclíadiana y la distancia de Mahalanobis, respectivamente.

Izzah y Suciati (2014) aplicó KNN para reconocer 120 imágenes almacenadas de ASL y 120 imágenes capturadas en tiempo real usando una cámara web. El sistema logró una precisión del 86% con imágenes almacenadas en la base de datos y una precisión de reconocimiento del 69% con la imagen capturada en tiempo real. Su sistema tiene un mejor rendimiento en comparación con PCA con KNN. K-NN se utiliza como clasificador para obtener una precisión de la imagen de prueba para el reconocimiento estático del lenguaje de señas (Jasim y Hasanuzzaman, 2015; Tharwat et al., 2015; Gupta et al., 2016; Anand et al. (2016) desarrolló un sistema para el reconocimiento de 13 alfabetos del lenguaje de señas indio. Las características extraídas mediante la Transformada de onda discreta (DWT) se alimentaron al clasificador KNN y lograron un reconocimiento con una precisión del 99,23 %. También, Mahmud et al. (2019) empleó K-NN para el reconocimiento de alfabetos ASL. El sistema utilizado L-a-b espacio de color para segmentar la región de interés y canny edge con HOG para extraer características. El sistema logró una precisión de reconocimiento del 94,23%. En la misma investigación, se probó el rendimiento del sistema utilizando una combinación de una bolsa de características (BoF) con k-means para la extracción de características y SVM como clasificación con una precisión de reconocimiento del 86%, que es inferior al resultado con K-NN .Butt et al. (2019) usó K-NN para el reconocimiento del lenguaje de señas. Compararon el rendimiento de su modelo con el modelo lineal generalizado (GLM) y los algoritmos de aprendizaje profundo. La precisión obtenida para GLM, K-NN y aprendizaje profundo es del 100 %, 98,03 % y 99,9 %,

respectivamente. Los autores informaron que la actuación de Naive Bayes, árbol de decisión y otros algoritmos considerados fueron mucho menores que GLM, K-NN y aprendizaje profundo para el mismo experimento. Tabla 10 resume algunos de los estudios de SLR basados en visión utilizando varias técnicas de extracción de características con KNN.

#### 4.2. Red neuronal artificial (RNA)

La Red Neural Artificial (ANN) es una herramienta analítica computacional inspirada en el sistema nervioso biológico del cerebro en un intento por imitar el razonamiento humano. Consiste en redes altamente interconectadas que pueden calcular valores de entrada y realizar cálculos paralelos para el procesamiento de datos y la representación del conocimiento. Es una rama de la inteligencia artificial (IA) que ayuda a construir modelos predictivos a partir de grandes bases de datos. Debido a su naturaleza robusta y adaptativa, ANN se ha utilizado para realizar cálculos como reconocimiento de patrones, coincidencia, clasificación (Jielai et al., 2015; Adegbeye et al., 2020). Generalmente se define por tres parámetros: el patrón de interconexión entre diferentes capas de neuronas, el peso de las interconexiones y la función de activación. Una neurona tiene entradas  $x_1, x_2, x_3, \dots, x_{pag}$ , cada uno está etiquetado con un peso  $w_1, w_2, w_3, \dots, w_{pag}$ , y mide la permeabilidad,  $K$  es la función de activación. Figura 8 muestra la estructura de las capas de la red de neuronas.

Función de salida ( $y$ ) se da como:

$$y = k \cdot \sum_{j=1}^{pag} X_j \cdot w_j = X_1 w_1 + X_2 w_2 + X_3 w_3 + \dots + X_{norte} w_{norte} \quad (48)$$

**Tabla 10**

Resumen de KNN usando varias técnicas de extracción de características.

Referencias	Extracción de características	Clasificación	Exactitud	Observación
Pansare et al. (2012)	Centroide y área del borde	distancia euclídea	90,19%	Reconocimiento en tiempo real de 26 alfabetos ASL. Logró una precisión de reconocimiento del 69 % para probar datos capturado en tiempo real.
Izzah y Suciati (2014)	Fourier genérico Descriptor (GFD) y PCA	K-NN	86 %	Veinticuatro alfabetos de signo indonesio lenguaje predicho. j y Z excluidos.
Hartanto et al. 2014)	NAVEGAR	distancia euclídea	63,0%	Señales dinámicas de ISL obtenidas de video aporte
Shukla et al. (2016)	Detección de bordes, Descriptor de Fourier y Tiempo dinámico Pandeo.	K-NN	96,15%	Gestos ISL de 13 alfabetos ingleses.
Anand et al. (2016)	Tiempo dinámico Deformación (DWT)	K-NN	99,23%	26 alfabetos ASL. Obtuvo mejores resultados en comparación con el uso de Bag of Features (BoF) con k-means y SVM.
Mahmud et al. (2019)	Histograma de Orientado Gradiente (HOG)	K-NN	94,23%	Weka-3-9-3 y rápido Miner 9.1.0 se utilizó para entrenar y probar modelos de datos usando varias clasificaciones algoritmos
Butt et al. (2019)	HOG, binario local Patrón (LBP) y características estadísticas	KNN	98,03%	MLP funciona mejor que otro métodos de clasificación.
Sharma et al. (2020)	ORB K-medias Bolsa de agrupamiento de palabras (BoW)	K-NN	95,81% con KNN 96,96% para MLP	Gesto ASL estático utilizando el conjunto de datos de Keggle. MLP función mejor

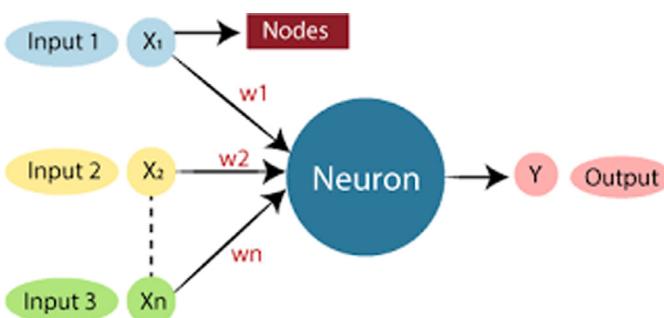


Figura 8.Estructura de una neurona y capas de red (Vandeginste et al., 1998).

Los algoritmos de redes neuronales utilizados para el reconocimiento de gestos incluyen; red neuronal de alimentación hacia adelante, algoritmos de red neuronal de retropropagación y perceptrón multicapa (MLP). Kishore et al. (2015) empleó ANN con un algoritmo de retropropagación para reconocer los gestos de las manos del lenguaje de señas indio. El sistema obtuvo un vector de características utilizando descriptores elípticos de Fourier y se utilizaron cuatro cámaras para mejorar el resultado con una precisión de reconocimiento del 95,10 %. Prasad et al. (2016a, 2016b) empleó el algoritmo NN de retropropagación para reconocer gestos estáticos y dinámicos de alfabetos y números ISL. Las palabras seleccionadas consisten en 59 gestos de signos con una precisión de reconocimiento del 92,34%. El conjunto de datos utilizado son datos adquiridos en video y extraídos mediante una combinación de PCA y descriptores elípticos de Fourier. Un enfoque similar usando el algoritmo NN de retropropagación hacia adelante (Islam et al., 2017) fue propuesto para alfabetos y números de ASL y alcanzó una precisión de reconocimiento del 94,32 %. En reconocimiento del lenguaje de señas bengalí de 20 alfabetos bengalíes, el algoritmo de propagación hacia atrás NN fue empleado por Hassan et al. (2017). Las características extraídas de una combinación de canny edge y FCC se introducen en el sistema neuronal

red. El sistema logró una precisión de reconocimiento del 96,5%. También, Raj y Jasuja (2018) propuso un sistema basado en ANN para reconocer los alfabetos del lenguaje de señas británico. El conjunto de datos de prueba contenía 780 imágenes de signos y el rendimiento del sistema alcanzó una precisión del 99,01 %, lo que es mejor que una investigación similar (Liwicki y Everingham, 2009) para los alfabetos de la lengua de signos británica con una precisión de reconocimiento del 98,9 %. Shaik et al. (2015) Red neuronal de retropropagación híbrida con un modelo de algoritmo genético. Informaron que su sistema funcionaba mejor que el reconocimiento de gestos de mano basado en retropropagación existente. La extensa revisión que realizamos reveló varios investigadores (Tariq et al., 2012; Adithya et al., 2013; Yasir & Khan, 2014) había utilizado ANN para el reconocimiento del lenguaje de señas. Los sistemas propuestos han demostrado un buen desempeño. Investigaciones recientes de Sharma et al. (2020a, 2020b) utilizaron varios algoritmos de aprendizaje automático para reconocer características extraídas de la combinación de ORB con K-means y Bag of words (BoW). Su hallazgo muestra una mejor precisión de reconocimiento del 96,96% utilizando el modelo de perceptrón multicapa. El resumen de los estudios SLR que utilizan varias técnicas de extracción de características con ANN se presenta en Tabla 11.

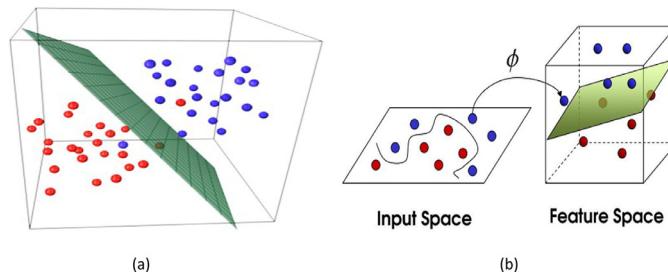
#### 4.3. Máquina de vectores de soporte (SVM)

Support Vector Machine (SVM) es un modelo de aprendizaje supervisado con algoritmos de aprendizaje asociados que no son probabilísticos. Es una técnica popular de aprendizaje de reconocimiento de patrones para el análisis de regresión de clasificación (Xanthopoulos y Razzaghi, 2014). SVM se puede utilizar para resolver problemas de clasificación de patrones y de regresión no lineal (Adeyanju et al., 2015), pero es más útil para resolver problemas difíciles de clasificación de patrones (Martiskainen et al., 2009). La clasificación se realiza en SVM diferenciando entre dos o más clases de datos. Esto se hace definiendo un hiperplano óptimo que divide todas las categorías, como se muestra en Figura 9(a). Sin embargo, suponga que es imposible encontrar una sola línea para separar los dos

**Tabla 11**

Resumen del método ANN utilizando varias técnicas de extracción de características.

Referencias	Extracción de características	Clasificación	Exactitud	Observación
Tarik et al. (2012)	defectos de convexidad	ANN (alimentación Red neuronal con retropropagación algoritmo)	62% en el conjunto de datos de prueba	Bajo reconocimiento tasa de "M" y "N" caracteres vencidos a las similitudes en forma.
Kishore et al. (2015)	Fourier elíptico Descriptor (EFD)	ana con retropropagación algoritmo	95,10%	Las imágenes fueron adquirido usando cuatro cámaras
Prasad et al. (2016)	Fourier elíptico Descriptor (EFD) y PCA	ANA	92,34%	se logró mejor resultados comparados con un combinación de Morfológico operación sobel, DCT y ANN.
islam et al. (2017)	Curvatura K y casco convexo algoritmo	Neural Artificial Red (ANN)- retroalimentación retropropagación Red neuronal	94,32%	El sistema logrado mejor reconocimiento con una combinación de todos Las características extraídas.
Hassan et al. (2017)	borde astuto detector. Hombre libre Código de cadena (FCC)	Neural Artificial Red (RNA)	96,50%.	Reconocimiento la precisión es baja bajo una oscuridad medio ambiente y mala iluminación.
Raj y Jasuja (2018)	Histograma de Gradiente orientado (CERDO)	ANN (alimentación retropropagación Algoritmo)	99,0%	26 signo británico idioma (BSL) alfabetos firmados con dos manos Gesto ASL estético usando Kaggle
Sharma et al. (2020)	ORB K-medias Bolsa de agrupamiento de palabras (BoW)	multicapa Perceptrón (MLP) bosques aleatorios, SVM, bayesiano ingenuo, Regresión logística, K-NN	96,96% para MLP	conjunto de datos MLP realizar mejor que otros clasificación métodos.

**Figura 9.**(a) Hiperplano que separa dos clases de datos, (b) truco kernel ([Merenda et al., 2020](#)).

clases en el espacio de entrada. En ese caso, se puede realizar una clasificación usando el truco del núcleo, como se muestra en [Figura 9\(b\)](#).

SVM opera de acuerdo con el concepto de cálculo de margen y se trazan límites entre clases. Los márgenes se dibujan de modo que la distancia entre el margen y las clases sea máxima. Por lo tanto, el error de clasificación se reduce. Las técnicas de optimización se emplean para encontrar el hiperplano óptimo ([Cheok et al., 2019](#)). La ecuación del hiperplano que hace la separación se da como:

$$\vec{w} \cdot \vec{x} + b = 0 \quad (49)$$

donde esta  $\vec{w}$  es el vector de peso para  $w$ , para datos de entrenamiento  $(\vec{x}_1, y_1), \dots, (\vec{x}_{norte}, y_{norte})$ ,  $y$  son 1 o -1, lo que indica a qué clasificar los datos  $\vec{x}_i$  pertenecen (salida de destino). El vector de peso decide la orientación del límite de decisión, mientras que el punto de sesgo  $b$  decide su ubicación.

Distancia,  $D_1$  entre el vector de soporte 1 y el plano se calcula de acuerdo con:

$$D_1 = \frac{1}{\|\vec{w}\|} \quad (50)$$

Distancia,  $D_2$  entre el vector de soporte 2 y el plano se calcula de acuerdo con:

$$D_2 = \frac{1}{\|\vec{w}\|} \quad (51)$$

El Margen,  $M$  es la suma de la distancia entre los vectores de soporte  $D_1$  y  $D_2$  dado por:

$$METRO = \frac{2}{\|\vec{w}\|} \quad (52)$$

SVM ha sido propuesto para el reconocimiento de la lengua de signos por muchos investigadores ([Rashid et al., 2009; Rekha et al., 2011b; Mohandes, 2013; Moreira Almeida et al., 2014; Kong y Ranganath, 2014; Dahmani y Larabi, 2014; Sol et al., 2015; Jayshree R. Pansare & Inglés, 2016; Raheja et al., 2016; Lee et al., 2016; Chong y Lee, 2018; Rahim et al., 2019; Abiev et al., 2020](#)). [Agrawal et al. \(2012\)](#) desarrolló un modelo para el reconocimiento del lenguaje de señas indio en tiempo real usando SVM. El sistema combinó descriptorios de forma con HOG y SIFT para extraer características. El modelo logró un precisión de reconocimiento del 93%. El rendimiento de SVM se ha comparado con K-NN en ([Tharwat et al., 2015; Yasir et al., 2016](#)). Informó que SVM tiene un mejor rendimiento en comparación con el clasificador KNN. También, [Tharwat et al. \(2015\)](#) usó SVM para reconocer el lenguaje de señas árabe (ArSL). El modelo logró una precisión de reconocimiento del 99%. Una investigación similar también usó Kinect con el clasificador SVM para desarrollar el sistema de reconocimiento de lenguaje de señas indio (SLR).

**Tabla 12**

Resumen del reconocimiento de lenguaje de señas basado en SVM.

Referencias	Extracción de características	Clasificación	Exactitud	Observación
(Agrawal et al., 2012)	descriptores de forma, HOG y TAMIZAR	MVS	93%	Bajo reconocimiento tarifa con similar signos como "M" y "N"
(Tharwat et al., 2015)	TAMIZAR LDA	SVM K-NN (con k=1 y k=5)	99%	SVM logrado Mejores resultados comparado con K-NN
(Yasir et al., 2016)	SIFT, K-medias agrupamiento y bolsa de palabras (BoW)	SVM K-NN	-	SVM logró un mejor resultado con un gran conjunto de datos que K-NN.
(Pan et al., 2016)	SIFT, Hu-momentos, Descriptor de Fourier (FD), PCA y LDA	MVS	99,8 % (CSL) 94 % (ASL)	Logró el reconocimiento precisión bajo ninguno complejo fondo
(Raheja et al., 2016)	Hu Momentos y rejeción de movimiento	MVS	97,5%	Cuatro signos gestos ("A", "B", "C" y "Hola") de indio lenguaje de señas.
(Rahim et al., 2019)	Fusión de características de CNN	MVS	97,28%	Bajo reconocimiento bajo complejo fondo, pobre iluminación y ocultación parcial.
(Athira et al., 2019)	Momentos Zernike	MVS	Precisión para el estático y dinámico signo es 91% y 89%, respectivamente.	Bajo reconocimiento con conjunto de datos capturado bajo desordenado fondo y diferente iluminación.
(Joshi et al., 2020)	HOG multivelvel utilizando Taguchi y Técnica para Orden de Preferencia por Similitud con Ideal Solución (TOPSIS).	MVS	92% para ISL y 99,2% para ASL	Firmar imágenes obtenido de Conjuntos de datos de Kaggle. Buen rendimiento bajo complejo fondo.
Barbhuiya et al. (2021)	Neural de convolución Red (CNN) basado en AlexNet y modelos VGG16	MVS	99,82%	Requiere alta computacional complejidad cuando tratar con un gran conjunto de datos.

El sistema demostró una precisión de reconocimiento del 97,5 % (Raheja et al., 2016). En Yasir et al. (2016), SVM se empleó para predecir el lenguaje de señas de Bangla. Se usó una bolsa de palabras con agrupamiento k-means para reducir los vectores de características obtenidos de la secuencia de video de signos. SVM se usó como clasificador y los momentos de Zernike se emplearon para encontrar el fotograma clave en una secuencia de video en Athira et al. (2019). El sistema alcanzó una precisión del 91% y 89% para estático y dinámico, respectivamente. Sin embargo, el rendimiento de ambos sistemas es bajo en entornos complejos y con poca iluminación. SVM se implementó en una aplicación móvil para reconocer 16 alfabetos ASL y logró una precisión de reconocimiento del 97,13 % en Jin et al. (2016). Otros estudios que emplearon SVM como un clasificador con mejor rendimiento incluyen (Joshi et al., 2020; Barbhuiya et al., 2021; Sharma et al., 2020a, 2020b). El resumen de SLR utilizando varias técnicas de extracción de características junto con SVM se presenta en Tabla 12.

#### 4.4. Modelo oculto de Markov (HMM)

Una de las técnicas de detección y reconocimiento de secuencias más efectivas es el modelo oculto de Markov (HMM). Es un modelo estadístico que se supone que es un proceso de Markov con un conjunto desconocido de parámetros ocultos. Los parámetros ocultos se pueden obtener de los parámetros de observación relacionados (Lan et al., 2017). Varios re-

Los buscadores han utilizado esta técnica en el reconocimiento del lenguaje de señas, donde la imagen a procesar puede estar en una secuencia de video y ha logrado una mayor precisión de reconocimiento. En HMM, se genera un nuevo estado cuando se aplica la entrada. Lo "oculto" en el modelo de Markov oculto implica que los cambios del estado anterior al estado nuevo no se pueden medir claramente, y la probabilidad de transición depende de cómo se entrena el modelo con los conjuntos de entrenamiento. La técnica funciona entrenando al modelo usando conjuntos de entrenamiento. El entrenamiento tiene que ser válido y clasificar todas las clases porque el modelo solo aprenderá de lo que se entrena. Algunos de los HMM utilizados como clasificador para el reconocimiento del lenguaje de señas incluyen el Modelo de Markov Oculto continuo (usando Gaussian) y el Modelo Markov Oculto discreto (usando Multinomial) (Suharjito et al., 2019).

Aloysius y Geetha (2020a), 2020b) dio pasos de procedimiento para el reconocimiento continuo de lenguaje de señas (CSLR) usando HMM. El clasificador HMM está estructurado para ser modelado por cada etiqueta de signo, con un número predefinido de estados ocultos como se muestra en Figura 10. Dado un videoclip como una secuencia de imágenes  $X_{norte} = X_1, X_2, \dots, X_n$ , un sistema CSLR encuentra una secuencia de palabras de signos Brillometro, para la cual  $X_{norte}$  ajusta mejor a los modelos aprendidos. Para fines de decodificación, se consideró un enfoque de ventana deslizante para cada segmento de video del enunciado continuo. Al mismo tiempo, el algoritmo de Viterbi (Forney, 1973) se llevó a cabo para cada HMM, para encontrar el gesto más probable en función de sus puntuaciones. El

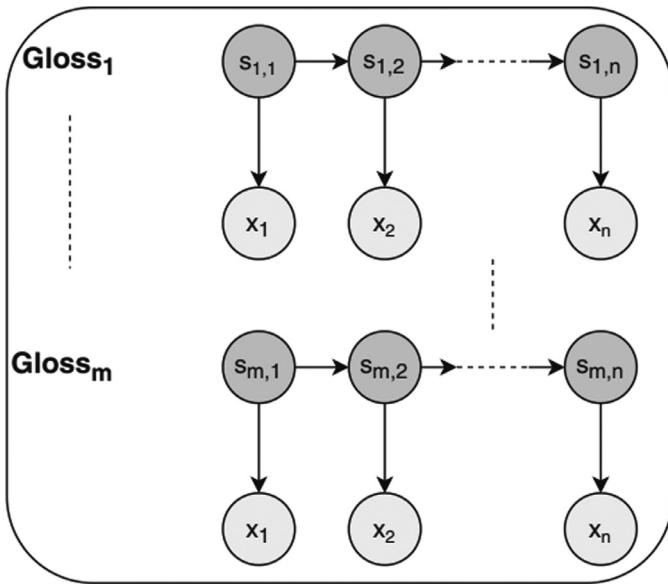


Figura 10. Modelo HMM para reconocimiento de gestos (Aloysius y Gita, 2020b).

modelo con mejores puntajes fue considerado como el signo (glosa o etiqueta).

La ecuación HMM se empleó para maximizar el signo óptimo y es dado como:

$$\{ \prod_i^{\text{Brillo}_{\text{meteo}} = \text{argumento m\'aximo}} \text{pag}(s_{norte}, \text{Brillo}_{\text{meteo}}) \cdot \text{pag}_{\frac{s_{\text{norte}}}{s_{\text{norte}-1}}, \text{Brillo}_{\text{meteo}}} \} \} \quad (53)$$

Parametroetiquetas y observaci\'on. \$x\$ de longitud \$norte\$, \$x\$ es una observaci\'on de la secuencia de entrada \$x\$ en la posici\'on \$j\$ y \$s\_0\$ representa el estado de HMM para la etiqueta de la señal (\$Brillo\$) y \$x\_j\$.

Zaki y Shaheen (2011) han propuesto el modelo de lenguaje de se\'nas basado en HMM utilizando una t\'ecnica de extracci\'on de caracter\'isticas basada en la apariencia. Se utiliz\'o la hibridaci\'on de tres t\'ecnicas de extracci\'on de caracter\'isticas para obtener caracter\'isticas que definen la disposici\'on del punto de articulaci\'on, la orientaci\'on de la mano y el movimiento. El sistema utiliz\'o umbrales de color de piel e identificaci\'on de componentes conectados para extraer solo la mano dominante, la cara y rastrearla. Su sistema logr\'o una tasa de error general del 10,9 % en la base de datos RWTH-BOSTON-50. Tambi\'en, Li et al. (2016) us\'o HMM como clasificador para reconocer 11 palabras del lenguaje de se\'nas de Taiw\'an relacionadas con el servicio a domicilio. El sistema combin\'o PCA con el algoritmo K-means basado en entrop\'ia para la transformaci\'on de datos, eliminando dimensiones redundantes y evaluando la cantidad de grupos en el conjunto de datos sin perder la informaci\'on principal. El algoritmo ABC y el algoritmo Baum-Welch est\'an integrados para resolver problemas de optimizaci\'on y lograron una precisi\'on de reconocimiento del 91,30 %. Algunos investigadores hibridan HMM con otras t\'ecnicas para mejorar la precisi\'on de SLR, como CNN-HMM (Koller et al., 2018), HMM-SVM (Lee et al., 2016) con Kinect para crear un modelo 3D de la imagen adquirida y logr\'o una precisi\'on de reconocimiento del 85,14 % y Coupled-HMM (Kumar et al., 2017) utiliz\'o Kinect y Leap para adquirir datos con una precisi\'on de reconocimiento del 90,80 %.

Kaluri y Reddy (2017) introdujo un algoritmo gen\'etico para mejorar HMM para el reconocimiento de gestos y compar\'o el rendimiento del modelo propuesto con SVM y Neural Network. La precisi\'on de reconocimiento para el m\'etodo propuesto, la red neuronal y SVM son 83%, 72% y 79%, respectivamente. Suharjito et al. (2019) modelos propuestos para la predicci\'on de diez se\'nas de la lengua de se\'nas argentina utilizando se\'nas Gaussianas HMM y Multinomial HMM. los modelos son

en comparaci\'on con las t\'ecnicas de detecci\'on de bordes y detecci\'on de piel. Su resultado mostr\'o que Gaussian HMM funcion\'o mejor cuando se utiliz\'o la detecci\'on de bordes con una precisi\'on de reconocimiento del 83 %. Otros investigadores han desarrollado algoritmos basados en HMM para el reconocimiento del lenguaje de se\'nas, incluido el HMM acoplado (CHMM) (Marca et al., 1997; Kumar et al., 2017), HMM de densidad de mezcla ligada (Zhang et al., 2004), HMM paralelo (PaHMM) (Vogler y Metaxas, 1999) y HMM param\'etrico (PHMM) (Wilson y Bobick, 1999).

#### 4.5. Red neuronal convolucional (CNN)

La red neuronal de convoluci\'on ofrece una amplia gama de aplicaciones, que incluyen reconocimiento facial, etiquetado de escenas, clasificaci\'on de im\'agenes, reconocimiento de voz y procesamiento de lenguaje natural (Gopika et al., 2020). Es un tipo de algoritmo de aprendizaje profundo que toma una imagen de entrada, asigna valores a varias caracter\'isticas en la imagen y usa el valor para diferenciar varias caracter\'isticas (Sah\'a, 2018). Normalmente tiene capas de entrada, convoluci\'on, agrupaci\'on y totalmente conectadas con salida (Nisha y Meral, 2021). Figura 11 muestra el funcionamiento de una red neuronal de convoluci\'on. La capa de convoluci\'on extrae los datos de entrada mediante una operaci\'on de convoluci\'on. La capa de agrupaci\'on se da cuenta de la reducci\'on de la dimensionalidad de las caracter\'isticas y controla la carga de c\'alculo mediante la selecci\'on de las caracter\'isticas m\'as importantes para evitar el sobreajuste. Las caracter\'isticas extra\'idas se alimentan a la capa totalmente conectada, que consta de una funci\'on de activaci\'on (Jiang et al., 2020). En lugar de crear funciones complejas hechas a mano, las CNN pueden automatizar el proceso de extracci\'on de funciones y funcionar mejor en comparaci\'on con otras t\'ecnicas tradicionales de procesamiento de im\'agenes (Am\'en & Vadera, 2017).

Un sistema de reconocimiento de ASL en tiempo real usando CNN fue propuesto por Taskiran et al. (2018). La imagen de entrada se segmenta mediante un algoritmo de casco convexo para obtener la regi\'on de la mano mediante la segmentaci\'on del color de la piel YCbCr. El modelo CNN propuesto consta de la capa de entrada, dos capas de convoluci\'on 2D, agrupaci\'on, aplanamiento y dos capas densas. Se logr\'o una precisi\'on del 98,05 % cuando se prob\'o con los datos en tiempo real. Tolentino et al. (2019) desarroll\'o un sistema en tiempo real para aprender lenguaje de se\'nas para principiantes que requieren reconocimiento de manos. Este sistema se basa en una t\'ecnica de modelado del color de la piel para extraer la regi\'on de la mano del fondo e integrarla en la Red Neural Convolucional (CNN) para clasificar las im\'agenes. El sistema logr\'o una precisi\'on del 90,04 % para los alfabetos ASL, 93,44 % para el reconocimiento de n\'umeros y 97,52 % para el reconocimiento de palabras est\'aticas.

Dudhal et al. (2019) propuso un enfoque basado en CNN para el reconocimiento de ISL utilizando SIFT para extraer las caracter\'isticas de una imagen con una variante diferente. El resultado del algoritmo SIFT se ingres\'o a Convolution Neural Network (CNN) para su clasificaci\'on con una precisi\'on de reconocimiento del 92,78 %. La investigaci\'on tambi\'en utiliz\'o un umbral adaptativo en la imagen del letrero con CNN y logr\'o una precisi\'on del 91,84 %. El estudio muestra que se logra un mejor rendimiento con la hibridaci\'on de SIFT, mientras que la hibridaci\'on de la umbralizaci\'on adaptativa y el desenfoque gaussiano es mejor que la umbralizaci\'on adaptativa solamente. Investigaciones similares (Barbhuiya et al., 2021) propusieron un enfoque similar para el modelado robusto del reconocimiento est\'atico del lenguaje de se\'nas utilizando redes neuronales convolucionales (CNN) basadas en el aprendizaje profundo. El sistema de reconocimiento de lenguaje de se\'nas propuesto incluye cuatro fases principales: adquisici\'on de im\'agenes, preprocesamiento de im\'agenes, entrenamiento y prueba del clasificador CNN. El sistema desarrollado logr\'o la mayor precisi\'on de entrenamiento del 99,72 % y del 99,90 % en im\'agenes en color y en escala de grises. M\'as recientemente, Barbhuiya et al. (2021) dise\'n\'o un marco de redes neuronales convolucionales (CNN) basado en un AlexNet preentrenado modificado y un modelo VGG16 preentrenado modificado para la extracci\'on de caracter\'isticas con una m\'quina de vectores de soporte multiclas (SVM) para el clasificador. El sistema logr\'o una precisi\'on de reconocimiento del 99,82%.

Otro enfoque de aprendizaje profundo que utiliza un VGG16 ajustado con precisi\'on basado en el modelo CNN Nelson et al. (2019) logr\'o una precisi\'on de reconocimiento del 97 % en comparaci\'on con un VGG19 ajustado con precisi\'on (Khari et al., 2019) con una precisi\'on de reconocimiento del 94,8%. Adithya y Rajesh (2020) en-

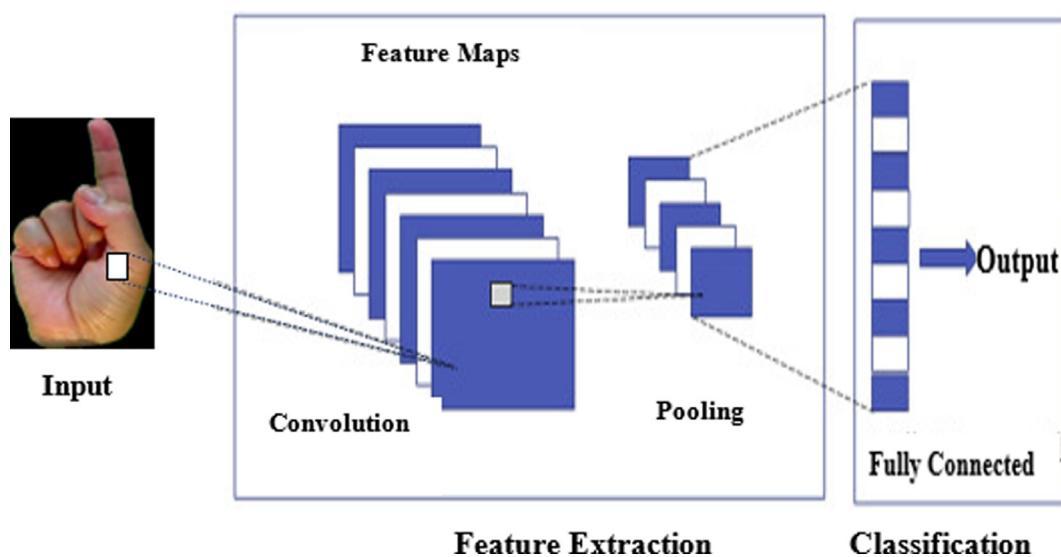


Figura 11. Arquitectura de red neuronal convolucional.

introdujo un modelo de aprendizaje profundo basado en CNN para reconocer alfabetos ASL estáticos. El modelo se probó en el conjunto de datos de postura de la mano de la Universidad Nacional de Singapur (NUS) y en el conjunto de datos A de deletreo manual estadounidense que contiene 24 letras del alfabeto ASL. El modelo desarrollado se desempeñó bien en los dos conjuntos de datos con una precisión de reconocimiento del 94,7 % y del 99,96 % en los conjuntos de datos NUS y American Fingerpelling A.

#### 4.6. Sistemas de lógica difusa

La lógica difusa es una forma de lógica multivaluada que utiliza la teoría matemática de los conjuntos borrosos para el razonamiento en lugar de la precisión (Ngai et al., 2014). Zadeh presentó un algoritmo de lógica difusa en 1965 para resolver los modelos que se ocupan del razonamiento lógico, como impreciso y vago (Canción y Chissom, 1993). El enfoque imita cómo los humanos toman decisiones que involucran todas las posibilidades intermedias entre los valores digitales 0 y 1 (Scott et al., 2003; Nuhu et al., 2021). Esta técnica de clasificación se basa en la teoría de conjuntos borrosos y se utiliza en un área diferente para resolver problemas. Opera en base a tres etapas que son fuzzificación, inferencia y defuzzificación. En la etapa de fuzzificación, los vectores de características de entrada se convierten en valores difusos con la ayuda de una función de pertenencia difusa, que se correlaciona con la puntuación de cada valor difuso. La segunda etapa es el motor de inferencia difuso, donde el mapeo entre la función de pertenencia de entrada y salida se realiza en función de las reglas difusas. Al mismo tiempo, la defuzzificación es la fase final de la lógica difusa que combina la salida en un único valor numérico para la predicción (Kaluri y Reddy, 2016a, 2016b).

El sistema de lógica difusa incluye el Sistema de Inferencia Fuzzy (FIS) y el Sistema de Inferencia Neuro-Fuzzy Adaptativo (ANFIS) (Nedeljkovic et al., 2004) se han aplicado en el sistema de reconocimiento de lengua de signos.

##### 4.6.1. Sistema de inferencia borrosa (FIS)

La inferencia difusa utiliza la lógica difusa para crear una asignación de una entrada determinada a una salida. Implica asignar a cada regla un peso entre 0 y 1, luego multiplicar el valor de membresía asignado al resultado de las reglas (Lenard et al., 1999). El proceso de inferencia difusa implica funciones de pertenencia, operadores de lógica difusa y reglas si-entonces. La implementación del Sistema de Inferencia Difusa (FIS) se puede clasificar en Mamdani y Sugeno (Wang, 2001).

*Sistema de inferencia borrosa Mamdani:* El sistema de inferencia borrosa Mamdani fue introducido por primera vez por Mamdani y Assilian (1975). El

La técnica de Mamdani es uno de los sistemas de inferencia borrosa más utilizados. Se emplea porque el diseño combina el conocimiento experto con más intuición en forma de reglas SI-ENTONCES expresadas en lenguaje natural (Izquierdo & Izquierdo, 2018).

En Mamdani, las reglas de los sistemas borrosos se dan como:

$$q_i: \text{SI } x_i \text{ es } A_{i,j} \text{ y } x_j \text{ es } B_{j,k} \text{ ENTONCES } y \text{ es } B_k \quad (54)$$

dónde  $q_i$  denota el  $i$ -a regla,  $j=1, \dots, n$ ,  $k=1, \dots, m$ ,  $A_{i,j}$  y  $B_{j,k}$  es el número de reglas,  $x_n$  es  $n$ -ésima entrada al sistema difuso,  $A_{i,j}$  y  $B_{j,k}$  son borrosos conjuntos descritos por  $FMM_{A_{i,j}}(X_i) \rightarrow [0,1]$  y  $M_{B_{j,k}}(Y) \rightarrow [0,1]$ .

Las proposiciones en la parte IF de la regla se combinan aplicando operadores mínimos. A veces se calcula el producto, pero depende principalmente de la situación. El número de proposiciones en la parte de consecuencia de la regla depende del número de salidas del sistema difuso (García-Díaz et al., 2013).

##### 4.7. Sistema de inferencia borrosa Sugeno

La técnica Sugeno de inferencia difusa es similar a la técnica Mamdani. Los pocos pasos del proceso de inferencia difusa, la confusión de las entradas y la aplicación del operador difuso, son exactamente los mismos. En comparación, la inferencia difusa de Sugeno emplea funciones de pertenencia de salida singleton, que son funciones constantes o lineales de los valores de entrada. En comparación con un sistema Mamdani, el procedimiento de defuzzificación de Sugeno es computacionalmente más eficiente.

Las reglas en el sistema de inferencia difuso funcional de Sugeno son dado por:

$$q_i: \text{SI } x_i \text{ es } A_{i,j} \text{ y } x_j \text{ es } B_{j,k} \text{ ENTONCES } y \text{ es } f_i(X) \quad (55)$$

dónde  $f_i(X)$  es una función nítida de las variables de entrada, en lugar de una proposición difusa.

La principal diferencia entre el FIS de tipo Mamdani y el FIS de Sugeno es que el FIS de tipo Mamdani utiliza la técnica de desfuzzificación de una salida difusa. Por el contrario, el FIS de tipo Sugeno utiliza un promedio ponderado para calcular la salida nítida, por lo que se pasa por alto la etapa de desfuzzificación. Mamdani FIS se adapta bien a la entrada humana con una base de reglas mejor interpretable (Kaur y Kaur, 2012; Shleeg y Elabib, 2013).

##### 4.7.1. Sistema de inferencia neurodifuso adaptativo (ANFIS)

Adaptive Neuro-Fuzzy Inference System (ANFIS) propuesto por (Jan, 1993) ) es un sistema de integración basado en redes neuronales que optimiza el sistema de inferencia borrosa. Crea un conjunto difuso de si-entonces

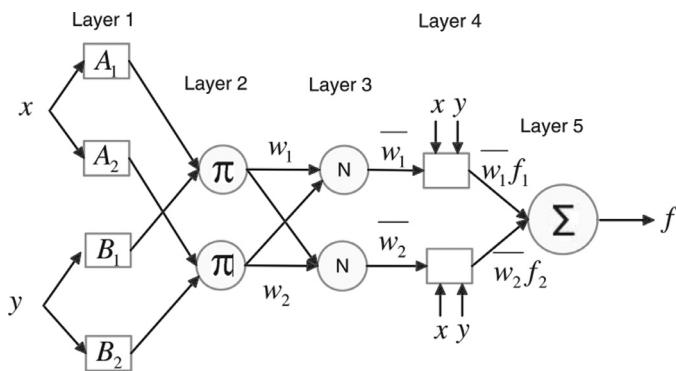


Figura 12. Arquitectura ANFIS (Al-Hmouz et al., 2012).

reglas con funciones de membresía apropiadas. La experiencia humana con respecto a los resultados que se modelarán se utiliza para establecer las reglas difusas iniciales y las funciones de pertenencia. ANFIS puede modificar estas reglas difusas ifthen y funciones de membresía basadas en una red neuronal interconectada para mapear las entradas numéricas en un error de salida reducido (Ahmed y Shah, 2017; Al-Hmouz et al., 2012) como se muestra en la arquitectura ANFIS de dos entradas en Figura 12. Esta técnica refina las reglas difusas SI-ENTONCES para describir el comportamiento de un sistema complejo y permite un mecanismo de aprendizaje rápido y preciso. Ha combinado con éxito las ventajas de la lógica difusa y las técnicas de redes neuronales en un solo enfoque (Jan, 1993; Kassem et al., 2017).

Dadas las reglas si-entonces de Sugeno como:

Regla 1:  $y_1 = f_1 \text{ si } x_1 \in A_1 \text{ y } x_2 \in B_1, \text{ entonces } f_1 = p_{11}x_1 + q_{11}x_2 + r_1$

Regla 2:  $y_2 = f_2 \text{ si } x_1 \in A_2 \text{ y } x_2 \in B_2, \text{ entonces } f_2 = p_{22}x_1 + q_{22}x_2 + r_2$  El resultado global del modelo se calcula mediante:

$$\sum_i \frac{\overline{w}_i f_i}{\sum_i \overline{w}_i} \quad (56)$$

dónde  $\overline{w}_i f_i$  es la salida del nodo  $i$  en la Capa 4 como se ilustra en Figura 12,  $x$  y  $y$  son las entradas,  $A_i$  y  $B_j$  son los conjuntos borrosos,  $f_i$  son las salidas dentro de la región difusa especificada por la regla difusa, y  $p_{ij}$ ,  $q_{ij}$  y  $r_i$  son los parámetros de diseño que se determinan durante el proceso de entrenamiento.

Al-Jarrah y Halawani (2001) desarrolló un sistema de inferencia neurofuzzy adaptativo para reconocer el lenguaje de señas árabe. El sistema se aplicó en imágenes de señales con la mano desnuda y logró una precisión de reconocimiento del 93,55 %. Al-Jarrah y Al-Omari (2007) también presentó un modelo similar utilizando un algoritmo mejorado. El sistema logró una precisión de reconocimiento del 97,5 % y del 100 % para las reglas 10 y 19, respectivamente. Kausar et al. (2008) utilizaron guantes marcados con colores para extraer las características de la yema de los dedos y las articulaciones de los dedos de los alfabetos del lenguaje de señas pakistaní. El ángulo entre la yema del dedo y la articulación se distinguió utilizando el Sistema de Inferencia Fuzzy (FIS). Lech y Kostek (2012) presentó un sistema de inferencia basado en reglas difusas para el reconocimiento dinámico de gestos. Su sistema ha mostrado un mejor rendimiento que el método de umbrales de velocidad fijos utilizando la inferencia difusa de Sugeno como clasificador para el conjunto de datos de video.

en el estudio de Kishore et al. (2016), se propuso el Fuzzy Inference Engine para el reconocimiento del lenguaje de señas indio. Los gestos continuos de 50 palabras se segmentaron y extrajeron utilizando el algoritmo Horn Schunck Optical Flow (HSOF) y los contornos activos, respectivamente. El sistema combinó las características más dominantes de los datos para formar las características utilizadas para el reconocimiento. Mufarroha y Utaminingrum (2017) utilizó un sistema de inferencia difuso basado en una red adaptativa para agrupar características y un clasificador KNN para acelerar el proceso de reconocimiento. El sistema logró una precisión de reconocimiento del 80,77% con las diez épocas. Elatawy et al. (2020) propuso un modelo predictivo para reconocer los alfabetos del lenguaje de señas árabe. El

El conjunto de datos utilizado se convirtió primero en el dominio neutrosófico para agregar más información. El sistema utilizó la técnica fuzzy c-means y logró una precisión de reconocimiento del 91%. Tabla 13 resume los artículos revisados sobre los sistemas de lógica difusa utilizados para el reconocimiento del lenguaje de señas.

#### 4.8. aprendizaje conjunto

Ensemble learning (EL) es un modelo de aprendizaje híbrido que combina varios algoritmos clasificadores para mejorar el rendimiento de predicción del modelo (Simské, 2019). Puede hacer un clasificador fuerte con errores de entrenamiento bajos a partir de la combinación de múltiples clasificadores débiles. Un método de aprendizaje conjunto es un meta-algoritmo que combina varias técnicas de aprendizaje automático para resolver muchos problemas de la vida real y crear un modelo predictivo con precisión mejorada (Policar, 2012; Huang y Chen, 2020; Zhao et al., 2021). El algoritmo se ha utilizado ampliamente en varias aplicaciones de aprendizaje automático, incluido el reconocimiento de rostros y objetos, la corrección de errores, el seguimiento de objetos y la selección de funciones. Ejemplos de técnicas de aprendizaje en conjunto son Adaptive Boosting (AdaBoost) y XGBoost (Extreme Gradient Boosting).

Los enfoques de aprendizaje por conjuntos comúnmente adoptados son embolsado y potenciado. El embolsado combina las clasificaciones predichas (predicción) de múltiples modelos o el mismo tipo de modelo para diferentes datos de aprendizaje. El embolsado también aborda la inestabilidad inherente de los resultados cuando se aplican modelos complejos a conjuntos de datos relativamente pequeños. En Boosting, las decisiones de diferentes modelos, como el embolsado, fusionan los diversos resultados en una sola predicción, pero derivan los modelos individuales de diferentes maneras. Sin embargo, en el enfoque de embolsado, los modelos reciben el mismo peso, mientras que la ponderación se utiliza para influir en los que tienen más éxito en el impulso.

Investigaciones anteriores discutieron ampliamente las técnicas de aprendizaje en conjunto (Dong et al., 2008; Policar, 2012; Sagi y Rokach, 2018; Talia et al., 2016; Zhou, 2009).

Aloysius y Geetha (2020a), 2020b presentó un conjunto promedio ponderado de CNN, que incluye una red de baja resolución (LRN), una red de resolución intermedia (IRN) y una red de alta resolución (HRN). Su sistema demostró un rendimiento mejorado con una alta precisión del 91,76 % para gestos manuales estáticos que el modelo entrenado por CNN. Zhang et al. (2005a) propuso un clasificador de conjunto basado en HMM continuo (CHMM) y AdaBoost para mejorar el rendimiento de las subpalabras del lenguaje de señas chino (CSL). Zhang et al. (2005b) desarrolló una clasificación de votación jerárquica (HVC) basada en una combinación de modelos de Markov ocultos continuos (CHMM) con una precisión del 95,4 %. Simón et al. (2007) propuso un sistema conjunto para reconocer el lenguaje de señas y el reconocimiento del comportamiento humano. También, Huang et al. (2012) desarrolló un sistema de conjunto basado en la hibridación de HMM y DWT para el reconocimiento de gestos con las manos. Su método supera el enfoque de vanguardia que utiliza HMM o DWT. En la investigación (Savur & Sahin, 2017; Mustafá y Lhasasna, 2016), el sistema de reconocimiento de lenguaje de señas estadounidense se desarrolló utilizando el aprendizaje conjunto y logró una buena precisión de reconocimiento. Yang et al. (2016) presentó un sistema basado en una combinación de construcción de nivel basada en deformación dinámica del tiempo (LB-DTW) y Fast HMM para mejorar la precisión de reconocimiento de la señal continua. Se utilizó Fast HMM para manejar la complejidad del cálculo. El sistema logró una precisión de reconocimiento del 87,80% en 100 oraciones en lenguaje de señas chino (CSL) con solo 21 vocabularios de señas. Ji et al. (2017) propusieron comandos de robots interactivos en lenguaje de señas utilizando un clasificador híbrido de CNN-SVM. Su sistema alcanzó una precisión de reconocimiento del 97,72 % con el conjunto de datos II de postura de la mano de NUS. Kim et al. (2018) presentó una red neuronal artificial de conjunto. El sistema utilizó un sensor de brazalete de electromiografía con 8 canales y ocho clasificadores ANN. La precisión de reconocimiento óptima del 97,40 % se logró con el método propuesto. Nguen et al. (2019) ensambló un sistema utilizando CNN basado en ResNet

**Tabla 13**

Resumen del reconocimiento de lengua de signos basado en sistemas de lógica difusa.

Autor	Extracción Técnicas	Clasificación	Exactitud	Observaciones
Al-Jarrah y Halawani (2001)	-	Adaptado Neuro-boroso Sistema de inferencia (ANFIS).	93,55%	Reconocimiento de imágenes de signos árabes con la mano desnuda.
Al Jarrah y Al Omari (2007)	Límite y características de la región.	Adaptado Neuro-boroso Sistema de Inferencia.	100% con 19 reglas y 97,5% con diez reglas	El sistema funcionó mejor que (Al-Jarrah y Al-Omari, 2007) con gestos de límites similares.
Prasad et al. (2016)	Contorno activo	sugeno boroso sistema de inferencia	92,5%	El conjunto de datos de video de signos indios contiene 80 palabras y pruebas de oraciones.
Kishore et al. (2016)	Contornos activos	Inferencia difusa Motor	96%	El sistema logró mejores resultados en comparación con otros modelos en las mismas categorías.
Lech y Kostek (2012)		basado en reglas difusas sistema de inferencia con filtros Kalman		Se observó un defecto de efecto pequeño al usar los filtros de Kalman. Se implementó con unos pocos gestos simples.
Elatawy et al. (2020)	nivel de gris co-ocurrencia matriz (GLCM)	c borrosa significa	91%	Reconoció 28 alfabetos de signos del lenguaje de señas árabe.

**Tabla 14**

Resumen de los métodos de aprendizaje de conjunto revisados en el reconocimiento del lenguaje de señas.

Autor	Clasificación	Exactitud	Observaciones
Luis y Gita (2020a), 2020b)	conjunto promedio ponderado de CNN (LRN, IRN y HRN)	91,76%	El modelo de conjunto se desempeñó mejor que varios modelos individuales seleccionados para las comparaciones.
Zhang et al. (2005)	HMM continuo (CHMM) y AdaBoost	92,70%	Logró una precisión de reconocimiento mejorada del 3 % en comparación con el uso exclusivo del clasificador CHMM.
Zhang y Gao (2005)	Combinación de CHMM	95,4%	Es eficaz y mejora el rendimiento con un conjunto de datos de entrenamiento limitado.
Huang et al. (2012)	HMM y DWT	-	El modelo demostró ser mejor que DWT o HMM solo con un conjunto de datos más grande pero con un alto costo de cálculo debido a las técnicas de hibridación.
Savur y Sahin (2016)	SVM, aprendizaje en conjunto (Árbol embolsado) clasificador	60,85% 80%	El clasificador de aprendizaje de conjuntos (Bagged Tree) superó al clasificador SVM para los conjuntos de datos de un solo tema y de varios sujetos.
Yang et al. (2016)	LB-DTW y HMM rápido	87,8%	El sistema propuesto resuelve el problema de segmentación y reconocimiento de secuencias. Además, se utilizaron restricciones gramaticales y de longitud de signos para mejorar la precisión del reconocimiento.
Ji et al. (2017)	CNN-SVM	97,72 %	Un reconocimiento de postura de la mano basado en la visión egocéntrica es adecuado para el control de robots de reconocimiento en el campo de batalla para comandos.
Gupta Jha (2020)	Aprendizaje en conjunto usando Múltiple SVM	98,7%	Funcionó bien en tiempo real y mejor que un clasificador único y otros clasificadores múltiples.
Nguen et al. (2019)	CNN basada en ResNet y ResNet cuaternión CNN	94,1%	Logró un mejor resultado usando la red neuronal convolucional de cuaterniones ResNet que el resultado obtenido usando ResNet CNN.
Sharma et al. (2020)	RF, TrAdaboost, TrBagg y TrRemuestreo	97,04%	Trbaggboost ha mejorado el rendimiento cuando se proporciona el conjunto de datos etiquetado en comparación con RF, TrAdaboost, TrBagg y TrResampling.
Kim et al. (2018)	E-ANN	97,4%	Tuvo una mayor precisión y un rendimiento más estable que una ANN general al comparar las precisiones promedio de los clasificadores y sus desviaciones estándar
Yuan et al. 2020)	Bosque aleatorio (RF) ANA MVS	95,48%	RF logró una mejor precisión de reconocimiento en comparación con ANN y SVM

y ResNet quaternion CNN para reconocer el lenguaje de señas japonés. Gupta y Jha (2020) propuso un sistema de reconocimiento en tiempo real para oraciones continuamente firmadas del lenguaje de señas indio. El sistema propuesto utilizó una técnica de conjunto basada en múltiples clasificadores SVM en características extraídas de duraciones de ventanas fijas y logró una precisión del 98,7% para 11 oraciones. Raghuveera et al. (2020) usó un conjunto de tres características (histograma de gradiente orientado (HOG), características robustas aceleradas (SURF) y patrón binario local (LBP)) para entrenar SVM. Su sistema logró una precisión de reconocimiento del 71,85 % con un tiempo de respuesta de 35 s. Sharma et al. (2020a, 2020b) introdujo un novedoso aprendizaje de transferencia basado en conjuntos denominado algoritmo Trbaggboost. El rendimiento de la

El método de conjunto propuesto se compara con algoritmos de aprendizaje por transferencia que incluyen TrAdaboost (Yao y Doretto, 2010), TrRemuestreo (Liu et al., 2017), Tr Bagg (Kamishima et al., 2009) y bosque aleatorio (Yuan et al., 2020). Un algoritmo Trbaggboost mejorado supera al Trbaggboost de última generación con una precisión de reconocimiento del 97,04 %. Yuan et al. (2020) propuso una técnica de conjunto basada en bosque aleatorio (RF) para el reconocimiento de alfabetos de lenguaje de señas chino fuera de línea. Se extrajeron cuatro características del antebrazo de los sujetos y se logró una precisión de reconocimiento del 95,48 %, mejor que SVM y ANN. El resumen de los métodos de aprendizaje en conjunto revisados en el reconocimiento del lenguaje de señas se presenta en Tabla 14.

**Tabla 15**

Ventajas y desventajas de los clasificadores.

Lenguaje de señas métodos de clasificador	Ventajas	Desventajas
K-más cercano vecinos	Es fácil de implementar. Es un algoritmo simple de interpretar.	Muy sensible a las características irrelevantes. No funciona bien con un gran conjunto de datos. No funciona bien con dimensiones altas.
Redes Artificiales (RNN)	Es útil cuando se requiere una evaluación rápida de la función objetivo aprendida. Es bastante resistente al ruido en el conjunto de datos de entrenamiento. Tiene tolerancia a fallas.	Es computacionalmente costoso y tiene dificultades para encontrar una estructura de red adecuada.
Vector de soporte Máquina (SVM)	Funciona mejor cuando se trata de multidimensiones y características continuas. Es aplicable en numerosos dominios.	Requiere una gran muestra del conjunto de datos para lograr su máxima precisión de predicción. Los hiperparámetros suelen ser un desafío al interpretar su impacto.
Markov oculto Modelo (HMM)	Tolerancia a los atributos irrelevantes. Se desempeña relativamente bien en el reconocimiento. Es más fácil de implementar y analizar. Elimina el problema del sesgo de la etiqueta.	Los HMM a menudo tienen una gran cantidad de parámetros no estructurados. Requiere una gran cantidad de entrenamiento para obtener mejores resultados. Requiere un gran conjunto de datos para el entrenamiento.
Neural de convolución Red (CNN)	Detecta automáticamente funciones importantes sin supervisión humana. Maneja la clasificación de imágenes con éxito y alta precisión.	Alto costo computacional. Requiere una gran cantidad de datos de entrenamiento para lograr una buena precisión. Falta de capacidad para ser espacialmente invariante a los datos de entrada. No codifica la posición y orientación del objeto.
Lógica difusa	Es un sistema robusto donde no se requieren entradas precisas. Es flexible y también puede permitir modificaciones. Es una técnica de base experta que aporta soluciones a soluciones complejas. Se ocupa de problemas complicados de una manera sencilla.	Depende completamente de la inteligencia y la experiencia humanas. Tiene poca precisión y sus predictores no siempre son correctos.
Aprendizaje en conjunto	Mejora el rendimiento medio de la predicción. Proporciona alta precisión y un modelo más estable. Reduce la varianza de los errores predictivos.	Puede ser más difícil de interpretar. A veces, el modelo se puede sobreajustar o no ajustar utilizando el método de aprendizaje por conjuntos.

**Tabla 16**

Resumen de los algoritmos de aprendizaje automático menos populares revisados empleados en el reconocimiento del lenguaje de señas.

Autor	Técnicas de extracción	Clasificación	Exactitud	Observaciones
Wong y Cipolla (2005)	Orientación de gradiente de movimiento (MGO)	Bayesiano disperso Clasificador RVM		El sistema propuesto funciona consistentemente en tiempo real y brinda una salida probabilística útil en el análisis de movimiento complejo.
Karami et al. (2011)	DWT	MLP-NN	94,06%	El modelo se probó con un conjunto de datos de 640 muestras de imágenes de signos persas que contenían 20 imágenes para cada signo.
Verma y Dev (2009)	Sobel edge detector, harris corner y fuzzy c media	MEV	-	La técnica propuesta se aplicó con éxito en gestos como agitar la mano izquierda, agitar la mano derecha, señalar para detener, avanzar y retroceder.
Maraqá y Abu Zaiter (2008)	medios de agrupamiento difusos (FCM)	RNN	95,11%	El sistema se probó con 900 imágenes de 30 gestos cada una, firmadas por dos personas. RNN supera a Feedforward, las redes neuronales de Elman y Jordan.
Bhat et al. (2013)	borde astuto	MOS	92%	Reconocimiento en tiempo real de 18 gestos del lenguaje de señas americano.
Lim et al. (2016)	-	Filtro de partículas en serie con función Matriz de covarianza	87,33%	Usó un filtro de partículas en serie para rastrear con precisión la posición de la mano y funcionó bien en un fondo despejado.
Camgoz et al. (2017)	-	CNN y BLSTM con CTC		Supera al modelo existente seleccionado para la comparación.
Vicente et al. (2019)	-	DeepConvLSTM™ con datos aumento	91,1%	La hibridación de los tres modelos (DeepConvLSTM) demostró ser mejor que otros modelos en comparación con ConvNet, LSTM, RF, SVM y KNN.
meng y li (2021)		Grafico convolucional Red (GCN)	98,08%	El sistema logró un mejor rendimiento y redujo el desenfoque de movimiento, la variación de signos y la ocultación de los dedos.

En este documento se han revisado varios clasificadores de lenguaje de señas. Las técnicas tienen méritos y deméritos sobre los demás.

**Tabla 15** muestra varias ventajas y desventajas de los clasificadores revisados en este estudio.

#### 4.9. Otros métodos de clasificación

Esta sección presenta brevemente la revisión de los algoritmos de aprendizaje automático menos populares empleados en el reconocimiento del lenguaje de señas. Estos algoritmos incluyen memoria a largo plazo modificada,

Clasificador bayesiano, máquina de estados finitos y basado en lógica difusa, perceptrón multicapa y mapas autoorganizados. Wong y Cipolla (2005) utilizó una máquina de vectores de relevancia (RVM) con la naturaleza probabilística del clasificador bayesiano para mejorar la precisión de reconocimiento de diez gestos en análisis de movimiento complejo. El sistema logró una precisión del 91,8%.

Verma y Dev (2009) propuso una máquina de estados finitos y un método basado en lógica difusa para el reconocimiento de gestos con las manos. Su sistema extrae características de las imágenes que comprenden posiciones de manos en 2D usando el detector de esquina de Harris y las agrupó junto con el algoritmo fuzzy c-means (FCM). El

grupos de postura de la mano determinan los estados de FSM y dan el reconocimiento de gestos apropiado.

Karami et al. (2011) desarrolló un sistema basado en MultiLayered Perceptron (MLPNN) para el reconocimiento de 32 clases de lenguaje de señas persa (PSL). Las características para la clasificación se extrajeron mediante transformada wavelet discreta (DWT) y lograron una precisión de reconocimiento del 94,06 %. Marqa y Abu-Zaiter (2008) introdujo redes neuronales recurrentes (RNN) para reconocer signos árabes estáticos. El modelo propuesto alcanzó una precisión del 95,11%. Bhat et al. (2013) propuso Self Organizing Maps (SOM) para el reconocimiento de gestos con las manos. Su sistema transformó la imagen en bordes cerrados radiales utilizados para entrenar SOM y logró una precisión de reconocimiento del 92 %. Lim et al. (2016) propuso una matriz de covarianza de características con un filtro de partículas en serie para el reconocimiento del lenguaje de señas aislado. El sistema de reconocimiento de lenguaje de señas estadounidense propuesto logró una precisión de reconocimiento del 87,33 %. Camgoz et al. (2017) introdujo una nueva arquitectura de aprendizaje profundo llamada SubUNets. Este algoritmo se basa en redes neuronales convolucionales (CNN) y memoria bidireccional a largo y corto plazo (BLSTM) con clasificación temporal conexiónista (CTC) para el reconocimiento continuo del lenguaje de señas. El modelo se entrenó en un conjunto de datos de un millón de gestos con las manos con una precisión mejorada de aproximadamente un 30 % en comparación con el sistema de vanguardia anterior.

Abrahán et al. (2019) propuso un sistema de reconocimiento de lenguaje de señas indio en tiempo real utilizando redes neuronales basadas en LSTM. El modelo propuesto alcanzó una precisión de reconocimiento del 98% para 26 gestos. Mittal et al. (2019) diseñó una red modificada de memoria a corto plazo (LSTM) para reconocer gestos aislados y continuos del lenguaje de señas indio. El sistema propuesto logró una precisión de reconocimiento promedio de 72,3% y 89,5% tanto para signos continuos como para palabras de signos aislados. Vicente et al. (2019) desarrolló un sistema de reconocimiento de lenguaje de señas estadounidense para actividades humanas. El sistema incorpora una Red Neuronal Convolutional (ConvNet) y una Red Neuronal Recurrente con Memoria a Largo-Corto Plazo (LSTM) (Bantupalli y Xie, 2019) llamado DeepConvLSTM. Logró una precisión de reconocimiento del 91,1 % con aumento de datos en conjuntos de datos de prueba. Tornay et al. (2020) presentó el HMM de divergencia Kullback Leibler (KL-HMM) para un sistema de reconocimiento de lenguaje de señas multilingüe. El desempeño del sistema fue validado utilizando tres lenguajes de señas diferentes. Meng y Li (2021) propuso una nueva Red de reconocimiento de lenguaje de señas dual y multiescala (SLR-Net) basada en una red convolucional de gráficos (GCN) para superar el desafío de la información redundante, la oclusión de los dedos, el desenfoque de movimiento y la variación en las formas en que las personas firman. Su modelo consta de tres submódulos: red de atención multiescala (MSA), red de atención espaciotemporal multiescala (MSSTA) y red de convolución temporal mejorada de atención (ATCN). Tabla 16 presenta un resumen de los algoritmos de aprendizaje automático menos populares revisados empleados en el reconocimiento del lenguaje de señas.

## 5. Conclusión y trabajo futuro

Con el reciente avance en el aprendizaje automático y los métodos de inteligencia computacional, los sistemas inteligentes en el reconocimiento del lenguaje de señas continúan atrayendo la atención de los investigadores académicos y los profesionales de la industria. Este estudio presenta un análisis sistemático de los sistemas inteligentes empleados en estudios relacionados con el reconocimiento del lenguaje de señas entre 2001 y 2021. Se proporciona una descripción general de las tendencias de investigación del reconocimiento del lenguaje de señas basado en inteligencia basada en 649 artículos de investigación completos recuperados de la base de datos Scopus. Usando las tendencias de publicación del artículo recuperado de la base de datos Scopus, este estudio muestra que el aprendizaje automático y las tecnologías inteligentes en el reconocimiento del lenguaje de señas están proliferando en los últimos 12 años. Los países e instituciones académicas con una gran cantidad de artículos publicados y sólidas colaboraciones internacionales han sido identificados y presentados en este documento. Es ex-

Se espera que este estudio sea una oportunidad para que el investigador de los países con menos colaboraciones amplíe sus colaboraciones en investigación.

Como parte de este trabajo, esta revisión brinda un análisis perspicaz de las técnicas anteriores utilizadas por varios investigadores durante dos décadas en diferentes etapas involucradas en SLR basado en visión, incluye la adquisición de imágenes, la segmentación de imágenes, la extracción de características y los algoritmos de clasificación empleados por varios investigadores para alcanzar precisión de reconocimiento. El estudio también estableció numerosas deficiencias y desafíos que enfrenta un enfoque basado en la visión para el reconocimiento del lenguaje de señas, a saber, el costo de implementación, las técnicas, la precisión del sistema, la naturaleza de la señal, el fondo complejo de la imagen, la variación en la iluminación de la imagen y tiempo computacional. Se han utilizado varios dispositivos para adquirir datos de señales, como Dataglove, Kinect, el controlador de movimiento Leap Camera en la adquisición de datos. A pesar de que estos dispositivos han contribuido al rendimiento y la precisión del sistema ASLR. Estos dispositivos tienen algunas deficiencias, como el alto costo y la inconveniencia de usar asociados con el guante de datos. La imagen adquirida de una cámara de baja resolución también afecta la precisión de reconocimiento del sistema. Por lo tanto, es necesario realizar más investigaciones que fusionen imágenes de múltiples dispositivos, como una cámara, un guante de datos y Kinect, con imágenes adquiridas para producir mejores resultados sin extracción de características. Las técnicas de segmentación del color de la piel y detección de bordes han demostrado un sólido rendimiento de segmentación mejorado. También se ha demostrado que la hibridación de dos o más técnicas de extracción de características produce características de reconocimiento más sólidas, como el alto costo y la inconveniencia del uso asociado con dataglove. La imagen adquirida de una cámara de baja resolución también afecta la precisión de reconocimiento del sistema. Por lo tanto, es necesario realizar más investigaciones que fusionen imágenes de múltiples dispositivos, como una cámara, un guante de datos y Kinect, con imágenes adquiridas para producir mejores resultados sin extracción de características. Las técnicas de segmentación del color de la piel y detección de bordes han demostrado un sólido rendimiento de segmentación mejorado. También se ha demostrado que la hibridación de dos o más técnicas de extracción de características produce características de reconocimiento más sólidas, como el alto costo y la inconveniencia del uso asociado con dataglove. La imagen adquirida de una cámara de baja resolución también afecta la precisión de reconocimiento del sistema. Por lo tanto, es necesario realizar más investigaciones que fusionen imágenes de múltiples dispositivos, como una cámara, un guante de datos y Kinect, con imágenes adquiridas para producir mejores resultados sin extracción de características. Las técnicas de segmentación del color de la piel y detección de bordes han demostrado un sólido rendimiento de segmentación mejorado. También se ha demostrado que la hibridación de dos o más técnicas de extracción de características produce características de reconocimiento más sólidas, y Kinect a imágenes adquiridas para producir mejores resultados sin extracción de características. Las técnicas de segmentación del color de la piel y detección de bordes han demostrado un sólido rendimiento de segmentación mejorado. También se ha demostrado que la hibridación de dos o más técnicas de extracción de características produce características de reconocimiento más sólidas.

Se han propuesto numerosos enfoques sobre las imágenes de signos adquiridas manualmente, lo que implica formas de manos, y se ha logrado un éxito notable. Para lograr un rendimiento de referencia en este contexto, los siguientes puntos merecen más atención para futuras investigaciones:

1. El estudio adicional sobre un signo no manual involucra la región de la cara, incluido el movimiento de la cabeza, el parpadeo de los ojos, el movimiento de las cejas y la forma de la boca.
2. Necesidad de abordar el reconocimiento de signos con expresión facial, gestos con las manos y movimiento del cuerpo simultáneamente con una mejor precisión de reconocimiento en tiempo real con un rendimiento mejorado. Los investigadores prevén que estos desafíos se pueden lograr utilizando un enfoque de aprendizaje profundo con un sistema de alta configuración para procesar los datos de entrada con poco tiempo de cómputo.
3. Se han realizado diferentes investigaciones sobre palabras, alfabetos y números. Sin embargo, existe la necesidad de más investigación en el futuro para el reconocimiento de oraciones en lenguaje de señas.
4. La mayor parte del trabajo sobre sistemas de reconocimiento de señales basados en inteligencia se encuentra en la etapa de investigación y prototipo. La implementación del modelo propuesto encontrará aplicación práctica a la lengua de signos automática.

## Declaración de interés en competencia

Los autores declaran que no tienen intereses financieros en competencia ni relaciones personales conocidas que pudieran haber influido en el trabajo informado en este documento.

## Referencias

- Pal, NR y Pal, SK Una revisión de las técnicas de segmentación de imágenes, 26 patrones reconocimiento 1277 (1993).[https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J). Abiyev, RH, Arslan, M, e Idoko, JB (2020). Traducción de lenguaje de señas usando deep Redes neuronales convolucionales. *Transacciones KSII en Internet y Sistemas de Información*, 14(2), 631–653.<https://doi.org/10.3837/tis.2020.02.009>. Abraham, E., Nayak, A. e Iqbal, A. (2019). Traducción en tiempo real del lenguaje de señas indio calibre utilizando LSTM. En *Actas de la Conferencia Global para el Avance de la Tecnología, GCAT 2019*.<https://doi.org/10.1109/GCAT47503.2019.8978343>. Adegbeye, MA, Aibinu, AM, Kolo, JG, Aliyu, I., Folorunso, TA y Lee, SH (2020). Incorporación de inteligencia en el sistema de alimentación de peces para dispensadores

- alimentación basada en la intensidad de alimentación de los peces. *Acceso IEEE*, 8, 91948–91960. <https://doi.org/10.1109/ACCESO.2020.2994442>.
- Adeyanju, IA, Omidiola, EO y Oyedokun, OF (2015). Evaluación del desempeño de diferentes núcleos de máquinas de vectores de soporte para el reconocimiento de emociones faciales. En *Actas de la conferencia de sistemas inteligentes SAI, intellisys 2015* (págs. 804–806). <https://doi.org/10.1109/IntelliSys.2015.7361233>.
- Adithya, V. y Rajesh, R. (2020). Un enfoque de red neuronal convolucional profunda para el reconocimiento estático de gestos con las manos. *Procedia Informática*, 171(2019), 2353–2361. <https://doi.org/10.1016/j.procs.2020.04.255>.
- Adithya, V., Vinod, PR y Gopalakrishnan, U. (2013). Red neuronal artificial basada método para el reconocimiento del lenguaje de señas indio. En *Actas de la conferencia IEEE sobre tecnologías de la información y la comunicación, ICT 2013*, Ict(págs. 1080–1085). <https://doi.org/10.1109/ICT.2013.6558259>.
- Admasuf, YF y Raimond, K. (2010). Reconocimiento del lenguaje de señas étnico usando arred neuronal artificial. En *Actas de la 10º conferencia internacional sobre diseño y aplicaciones de sistemas inteligentes, ISDA'10*(págs. 995–1000). <https://doi.org/10.1109/ISDA.2010.5687057>.
- Agrawal, SC, Jalal, AS y Bhatnagar, C. (2012). Reconocimiento del lenguaje de señas indio calibre utilizando la fusión de características. En *Actas de la 4.ª conferencia internacional sobre interacción inteligente entre humanos y ordenadores: avance de la tecnología para la humanidad, IHCI 2012*. <https://doi.org/10.1109/IHCI.2012.6481841>.
- Agrawal, SC, Jalal, AS y Tripathi, RK (2016). Una encuesta sobre manuales y no Reconocimiento manual de lenguaje de señas para señas aisladas y continuas. *Revista internacional de reconocimiento de patrones aplicados*, 3(2), 99. <https://doi.org/10.1504/ijapr.2016.709048>.
- Agris, UV y Kraiss, K. F (2007). Hacia un corpus de video para firmantes independientes reconocimiento continuo del lenguaje de señas. En *Actas del 7º taller internacional sobre gestos en la interacción y simulación humano-computadora: 7*(págs. 10 y 11).
- Ahmad, K., Khan, J. e Iqbal, MSUD (2019). Un estudio comparativo de diferentes técnicas de eliminación de ruido en el procesamiento de imágenes digitales. En *Actas de la 8.ª conferencia internacional sobre simulación de modelos y optimización aplicada, ICMSAO 2019*(págs. 1–6). <https://doi.org/10.1109/ICMSAO.2019.8880389>.
- Ahmed, AAM y Shah, SMA (2017). Aplicación de neuro-fuzzy adaptativo sistema de inferencia (ANFIS) para estimar la demanda bioquímica de oxígeno (DBO) del río Surma. *Revista de la Universidad King Saud - Ciencias de la Ingeniería*. <https://doi.org/10.1016/j.jksues.2015.02.001>.
- Aksay, B. y Salman, OK M (2020). Un nuevo enfoque de filtro híbrido para el procesamiento de imágenes cesando. *Revista de Ciencias de la Información y la Computación de la Universidad de Sakarya*, 3(3), 334–342. <https://doi.org/10.35377/saucis.03.03.785749>.
- Alamelu, J., Satej, W. y Santhosh, V. (2013). Un k vecino más cercano mejorado clasificador bor usando medidas de interés para la minería de imágenes médicas. *Academia Mundial de Ciencias, Ingeniería y Tecnología*, 7(9), 550–554. <http://waset.org/publications/16638/an-improved-k-nearest-neighborclassifier-using-interestingness-measures-for-medical-image-mining>.
- Al-Hmouz, A., Shen, J., Al-Hmouz, R. y Yan, J (2012). Modelado y simulación de un sistema de inferencia neurodifuso adaptativo (ANFIS) para el aprendizaje móvil. *Transacciones IEEE sobre tecnologías de aprendizaje*. <https://doi.org/10.1109/TLT.2011.36>.
- Aliyu, I., Gana, KJ, Musa, AA, Adegbeye, MA y Lim, CG (2020). Incorporando reconocimiento en el algoritmo de conteo de bagres usando red neuronal artificial y geometría. *Transacciones KSII en Internet y Sistemas de Información*, 14(12), 4866–4888. <https://doi.org/10.3837/tis.2020.12.014>.
- Al-Jarrah, O. y Al-Omari, F. A (2007). Mejorar el reconocimiento de gestos en árabe lenguaje de señas mediante análisis de texturas. *Inteligencia Artificial Aplicada*, 21(1), 11–33. <https://doi.org/10.1080/08839510600938524>.
- Al-Jarrah, O. y Halawani, A. (2001). Reconocimiento de gestos en lengua de signos árabe mediante sistemas neurodifusos. *Inteligencia Artificial*, 13(1–2), 117–138. [https://doi.org/10.1016/S0016-4370\(01\)00141-2](https://doi.org/10.1016/S0016-4370(01)00141-2).
- Alnahhas, A., Alkhatib, B., Al-Boukaee, N., Alhakim, N., Alzabibi, O. y Ajalya-agudo, N (2020). Mejorar el reconocimiento del lenguaje de señas árabe mediante el aprendizaje profundo y el controlador de movimiento de salto. *Revista Internacional de Investigación Científica y Tecnológica*, 9(4), 1865–1870.
- Aloysius, N. y Geetha, M. (2020a). Un modelo de espacio a escala de CNN promedio ponderado conjunto para el reconocimiento de deletreo manual ASL. *Revista Internacional de Ciencias Computacionales e Ingeniería*, 22(1), 154–161. <https://doi.org/10.1504/IJCE.2020.107268>.
- Aloysius, N. y Geetha, M. (2020b). Entender el lenguaje de señas continuo basado en la visión reconocimiento de calibre. *Herramientas y aplicaciones multimedia*, 79(31–32), 22177–22209. <https://doi.org/10.1007/s11042-020-08961-z>.
- Al-Shamayleh, AS, Ahmad, R., Jomhari, N. y Abushariah, MAM (2020). Auto-Reconocimiento del lenguaje de señas árabe mítico: una revisión, taxonomía, desafíos abiertos, hoja de ruta de investigación y direcciones futuras. *Revista de Ciencias de la Computación de Malasia*, 33(4), 306–343. <https://doi.org/10.22452/mjcs.vol33no4.5>.
- Ameneen, S. y Vadera, S. (2017). Una red neuronal convolucional para clasificar América puede deletrear con los dedos el lenguaje de señas a partir de imágenes profundas y en color. *Sistemas Expertos*, 3(3). <https://doi.org/10.1111/exsy.12197>.
- Anza, C. (2012). Una revisión sobre la técnica de segmentación de imágenes basada en redes neuronales. *Delaware Universidad de Montfort, Ingeniería Mecánica y de Manufactura*.
- An, FP y Liu, ZW (2019). Algoritmo de segmentación de imágenes médicas basado en feed-mecanismo trasero CNN. *Medios de contraste e imagen molecular*. <https://doi.org/10.1155/2019/6134942>.
- Anand, MS, Kumar, NM y Kumaresan, A. (2016). Un marco eficiente para Reconocimiento del lenguaje de señas indio mediante transformada wavelet. *Circuitos y Sistemas*, 07(08), 1874–1883. <https://doi.org/10.4236/cs.2016.78162>.
- Assaleh, K., Shanableh, T. y Zourob, M. (2012). Sistema de clasificación de baja complejidad. tem para el reconocimiento de lenguaje de señas árabe basado en guantes. *Apuntes de clase en computadora ciencia*(pp. 262–268) (incluidas las notas de clase de la subserie sobre inteligencia artificial y las notas de clase sobre bioinformática), 7665 LNCS (PARTE 3). [https://doi.org/10.1007/978-3-642-34487-9\\_32](https://doi.org/10.1007/978-3-642-34487-9_32).
- Athira, PK, Sruthi, CJ y Lijiyaa, A. (2019). Un lenguaje de señas independiente de firmantes reconocimiento con eliminación de coarticulación de videos en vivo: un escenario indio. *Revista de la Universidad King Saud - Informática y Ciencias de la Información*, 0–10 xxxx. <https://doi.org/10.1016/j.jksuci.2019.05.002>.
- Athitsos, V., Neidle, C., Sclaroff, S., Nash, J., Stefan, A., Yuan, Q. y Thangali, A. (2008). El conjunto de datos de video del léxico del lenguaje de señas estadounidense. En *Actas de la conferencia de la sociedad informática del IEEE sobre visión artificial y talleres de reconocimiento de patrones*, talleres CVPR. <https://doi.org/10.1109/CVPRW.2008.4563181>.
- Aurangzeb, K., Aslam, S., Alhussein, M., Naqvi, RA, Arsalan, M. y Haider, SI (2021). Mejora del contraste de las imágenes del fondo de ojo mediante el empleo de PSO modificado para mejorar el rendimiento de los modelos de aprendizaje profundo. *Acceso IEEE*. <https://doi.org/10.1109/ACCESO.2021.3068477>.
- Bantupalli, K. y Xie, Y. (2019). Reconocimiento del lenguaje de señas americano usando deep aprendizaje y visión artificial. En *Actas de la conferencia internacional IEEE sobre big data, big data 2018*(págs. 4896–4899). Instituto de Ingenieros Eléctricos y Electrónicos Inc. <https://doi.org/10.1109/BigData.2018.8622141>.
- Barbhuiya, AA, Karsh, RK y Jain, R. (2021). Extracción de características basada en CNN y Clasificación de la lengua de signos. *Herramientas y aplicaciones multimedia*, 80(2), 3051–3069. <https://doi.org/10.1007/s11042-020-09829-y>.
- Barczak, ALC, Reyes, NH, Abastillas, M., Piccio, A., y Susnjak, T. (2011). un nuevo Conjunto de datos de imágenes en color de gestos manuales estáticos 2D para gestos ASL. *Res. Letón. información Matemáticas. ciencia*, 15, 12–20. <http://liims.massey.ac.nz/research/letters/>.
- Basu, M. (2002). Métodos de detección de bordes basados en Gauss: una encuesta. *transacción IEEE ciones sobre Sistemas, Hombre y Cibernética Parte C: Aplicaciones y Revisiones*. <https://doi.org/10.1109/TSIMC.2002.804448>.
- Bessa Carneiro, S., De Santos, EDFDM, De Barbosa, TMDA, Ferreira, JO, Alcalá, SGSS y Da Rocha, AF (2017). Reconocimiento de gestos estáticos para lengua de señas brasileña con sensor kinect. En *Procedimientos de sensores IEEE*. <https://doi.org/10.1109/ICSENS.2016.7808522>.
- Bhardwaj, S. y Mittal, A. (2012). Una encuesta sobre varias técnicas de detección de bordes. *Procedia tecnologia*. <https://doi.org/10.1016/j.protcy.2012.05.033>.
- Bhat, NN, Venkatesh, Y., Varni, U. y Vig, D (2013). Reconocimiento de gestos de mano usando un mapa autoorganizado para la interacción humano-computadora. En *Actas de la conferencia internacional de 2013 sobre avances en computación, comunicaciones e informática, ICACCI 2013*(págs. 734–738). <https://doi.org/10.1109/ICACCI.2013.6637265>.
- Bin, LY, Huann, GY y Yun, L. K (2019). Estudio de red neuronal convolucional en el reconocimiento del lenguaje de señas americano estático. En *Actas de la conferencia internacional IEEE sobre aplicaciones de procesamiento de señales e imágenes, ICSIPA 2019*(págs. 41–45). <https://doi.org/10.1109/ICSIPIA5851.2019.8977767>.
- Biswas, D., Nag, AN, Ghosh, S., Pal, A., Biswas, SB y Banerjee, S. (2011). Novedoso técnicas de conversión de escala de grises basadas en la profundidad de pixel. *Revista de Investigación Global en Ciencias de la Computación*, 2(6), 118–121.
- Bobic, V., Tadic, P. y Kvasev, G. (2016). Reconocimiento de gestos de mano mediante red neuronal. técnicas basadas en el trabajo. En *Actas del 13º simposio sobre redes neuronales y aplicaciones, NEUREL 2016*(págs. 14–17). <https://doi.org/10.1109/NEUREL.2016.7800104>.
- Marca, M., Oliver, N. y Pentland, A. (1997). Modelos de Markov ocultos acoplados para reconocimiento de acciones complejas. En *Actas de la conferencia de la sociedad informática IEEE sobre visión artificial y reconocimiento de patrones*. <https://doi.org/10.1109/cvpr.1997.609450>.
- Butt, UM, Husnain, B., Ahmed, U., Tariq, A., Tariq, I., Butt, MA y Zia, M. S (2019). Análisis algorítmico basado en características en el conjunto de datos del lenguaje de señas estadounidense. *Revista internacional de informática avanzada y aplicaciones*, 10(5), 583–589. <https://doi.org/10.14569/jiacsa.2019.0100575>.
- Buyukahin, U. (2014). Tecnologías de sensores basadas en la visión: cámara web: una sensor de función. *Procesamiento integral de materiales*. <https://doi.org/10.1016/B978-0-08-096532-1.01314-5>.
- Camgoz, NC, Hadfield, S., Koller, O. y Bowden, R. (2017). SubUNets: de extremo a extremo forma de la mano y reconocimiento continuo del lenguaje de señas. En *Actas de la conferencia internacional IEEE sobre visión artificial*(págs. 3075–3084). 2017-octubre. <https://doi.org/10.1109/ICCV.2017.332>.
- Canny, J. (1986). Un enfoque computacional para la detección de bordes. *Transacciones IEEE sobre análisis de patrones e inteligencia artificial*. <https://doi.org/10.1109/TPAMI.1986.4767851>.
- Cebeci, Z. y Yıldız, F (2015). Comparación de algoritmos K-means y fuzzy C-means en diferentes estructuras de clúster. *Revista de Informática Agrícola*. <https://doi.org/10.17700/jai.2015.6.3.196>.
- Cheng, HD, Jiang, XH y Wang, J (2002). Segmentación de imágenes en color basada en hombralización de mogramas y fusión de regiones. *Reconocimiento de patrones*, 35(2), 373–393. [https://doi.org/10.1016/S0031-3203\(01\)00054-1](https://doi.org/10.1016/S0031-3203(01)00054-1).
- Cheok, MJ, Omar, Z. y Jaward, MH (2019). Una revisión del gesto de la mano y el signo. Técnicas de reconocimiento de lenguaje. *Revista internacional de aprendizaje automático y cibernetica*, 10(1), 131–153. <https://doi.org/10.1007/s13042-017-0705-5>.
- Chhabra, ACHINU. (2014). Descripción general y análisis comparativo de la tecnología de detección de bordes. niques en el procesamiento de imágenes digitales. *Revista internacional de tecnología de la información y la computación*.
- Chong, TW y Lee, BG (2018). Reconocimiento del lenguaje de señas americano usando salto controlador de movimiento con enfoque de aprendizaje automático. *Sensores (Suiza)*, 10, 18. <https://doi.org/10.3390/s18103554>.
- Chourasiya, A., Khare, N., Maini, R., Aggarwal, H., Chourasiya, A. y Khare, N. (2019). Una revisión completa de las técnicas de mejora de imágenes. *Revista Internacional-*

- nal de Investigación Innovadora y Crecimiento, 8(6), 60-71.<https://doi.org/10.26671/ijirg.2019.6.8.101>.
- Chowdhry, DA, Hussain, A., Ur Rehman, MZ, Ahmad, F., Ahmad, A. y Per-Váiz, M. (2013). Sistema de seguridad inteligente para áreas sensibles mediante reconocimiento facial. En *Actas de la conferencia IEEE sobre utilización y desarrollo sostenibles en ingeniería y tecnología, IEEE CSUDET*(págs. 11-14).<https://doi.org/10.1109/CSUDET.2013.6670976>.
- Dahmani, D. y Larabi, S. (2014). Sistema independiente del usuario para el dedo en lengua de signos reconocimiento ortográfico. *Revista de Comunicación Visual y Representación de Imágenes*, 25(5), 1240-1250.<https://doi.org/10.1016/j.jvcir.2013.12.019>.
- Deng, X., Yang, S., Zhang, Y., Tan, P., Chang, L. y Wang, H. (2017). Mano3D: Mano estimación de poses utilizando una red neuronal 3D. ArXiv.
- Deriche, M., Aliyu, S. y Mohandes, M. (2019). Un lenguaje de señas árabe inteligente sistema de reconocimiento utilizando un par de LMC con clasificación basada en GMM. *Revista de sensores IEEE*, 19(18), 1-12.<https://doi.org/10.1109/JSEN.2019.2917525>. Dhanachandra, N. y Chanu, YJ (2017). Una encuesta sobre los métodos de segmentación de imágenes utilizando técnicas de agrupamiento. *Revista europea de investigación y ciencia en ingeniería*. <https://doi.org/10.24014/ejers.2017.2.1.237>.
- Dhanushree, M., Priyadarsini, R. y Sree Sharmila, T. (2019). Eliminación de ruido de imagen acústica utilizando diversas técnicas de filtrado espacial. *Revista internacional de tecnología de la información*(Singapur).<https://doi.org/10.1007/s41870-018-0272-3>.
- Divya, D. y Ganesh Babu, TR (2020). Una encuesta sobre técnicas de segmentación de imágenes. *Apuntes de clase sobre ingeniería de datos y tecnologías de la comunicación*.[https://doi.org/10.1007/978-3-030-32150-5\\_112](https://doi.org/10.1007/978-3-030-32150-5_112).
- Dong, L., Yu, G., Ogunbona, P. y Li, W. (2008). Un algoritmo iterativo eficiente para umbralización de imágenes. *Letras de reconocimiento de patrones*, 29(9), 1311-1316.<https://doi.org/10.1016/j.patrec.2008.02.001>.
- Drewu, P., Neidle, C., Athitsos, V., Sclaroff, S. y Ney, H. (2008). Bases de datos comparativas para el reconocimiento automático de lenguaje de señas basado en video. En *Actas de la 6ª conferencia internacional sobre recursos lingüísticos y evaluación, LREC 2008*(págs. 1115-1120).
- Dudhal, A., Mathkar, H., Jain, A., Kadam, O. y Shirole, M. (2019). Función SIFT híbrida enfoque de extracción para el sistema de reconocimiento de lenguaje de señas indio basado en CNN. *Apuntes de clase en visión computacional y biomecánica*(30, págs. 727-738).[https://doi.org/10.1007/978-3-030-00665-5\\_72](https://doi.org/10.1007/978-3-030-00665-5_72).
- Duggirala, RK (2020). Segmentación de imágenes usando hibridación de K-means y algoritmos fuzzy C-means. *Introducción a la ciencia de datos y el aprendizaje automático*.<https://doi.org/10.5772/intechopen.86374>.
- Egmont-Petersen, M., De Ridder, D. y Handels, H. (2002). Procesamiento de imágenes con Redes neuronales: una revisión. *Reconocimiento de patrones*, 35(10), 2279-2301.[https://doi.org/10.1016/S0031-3203\(01\)00178-9](https://doi.org/10.1016/S0031-3203(01)00178-9).
- Elatawy, SM, Hawa, DM, Ewees, AA y Saad, AM (2020). Sistema de reconocimiento para el lenguaje de señas del alfabeto árabe usando c-means neutrosóficos y difusos. *Educación y Tecnologías de la Información*, 25(6), 5601-5616.<https://doi.org/10.1007/s10639-020-10184-6>.
- El-Din, SAE y El-Ghany, MAA (2020). Sistema de interpretación de lenguaje de señas: un sistema alternativo para el aprendizaje automático. En *Actas de la segunda conferencia de ciencias emergentes inteligentes y líderes en novelas, NILEs 2020*, M(págs. 332-337).<https://doi.org/10.1109/NILESS0944.2020.9257958>.
- Elouariachi, I., Benouini, R., Zenkoura, K., Zarghili, A. y El Fadili, H. (2021). Explícito quaternion krawtchouk momento invariantes para reconocimiento de lenguaje de señas de deletreo de dedos. En *Actas de la conferencia europea de procesamiento de señales*(págs. 620-624). 2021-enero.<https://doi.org/10.23919/Eusipco47968.2020.9287845>.
- Enikeev, DG y Mustafina, SA (2021). Reconocimiento de lenguaje de signos a través de Leap Controlador de movimiento y algoritmo de predicción de entrada. En KS V Mikhailev GA Mikhailev GA (Ed.) *Journal of Physics: Serie de conferencias*, 1715(1) IOP Publishing Ltd.<https://doi.org/10.1088/1742-6596/1715/1/012008>.
- Enikeev, D. y Mustafina, S. (2020). Reconocimiento de la lengua de signos mediante salto modos del controlador de cón. En *Actas de la 2ª conferencia internacional sobre sistemas de control, modelado matemático, automatización y eficiencia energética, SUMMA 2020*(págs. 393-397).<https://doi.org/10.1109/SUMMA50634.2020.9280795>. Escobedo, E., & Camara, G. (2017). Un nuevo enfoque para el reconocimiento dinámico de gestos nición usando la representación de la trayectoria del esqueleto e histogramas de magnitudes acumuladas. En *Actas de las 29 jornadas de gráfica, patrones e imágenes, SIBGRAPI 2016*(págs. 209-216).<https://doi.org/10.1109/SIBGRAPI.2016.037>.
- Fadnavis, S. (2014). Técnicas de interpolación de imágenes en el procesamiento de imágenes digitales: una descripción general. *Revista de investigación y aplicaciones de ingeniería*, 4(10), 70-73. Fang Hu, Z., Yang, L., Luo, L., Zhang, Y. y Chuan Zhou, X. (2014). La investigación y aplicación del algoritmo SURF basado en el algoritmo de selección de puntos característicos. *Sensores y transductores*, 16(94), 67-72.
- Forney, GD (1973). El algoritmo de Viterbi. *Actas del IEEE*<https://doi.org/10.1109/PROC.1973.9030>.
- Forster, J., Schmidt, C., Koller, O., Bellgardt, M. y Ney, H. (2014). Extensiones de la corpus de traducción y reconocimiento de lenguaje de señas RWTH-PHOENIX-Weather. En *Actas de la 9ª conferencia internacional sobre recursos lingüísticos y evaluación, LREC 2014*(págs. 1911-1916).
- Gao, L., Li, H., Liu, Z., Liu, Z., Wan, L. y Feng, W. (2021). Basado en transductor RNN Reconocimiento de lenguaje de señas chino. *Neurocomputación*, 434, 45-54.<https://doi.org/10.1016/j.neucom.2020.12.006>.
- Gao, W., Fang, G., Zhao, D. y Chen, Y (2004). Un lenguaje de señas chino que reconoce sistema de gestión basado en SOFM/SRN/HMM. *Reconocimiento de patrones*, 37(12), 2389-2402.[https://doi.org/10.1016/S0031-3203\(04\)00165-7](https://doi.org/10.1016/S0031-3203(04)00165-7).
- García-Díaz, N., Lopez-Martin, C., & Chavoya, A. (2013). Un estudio comparativo de dos modelos de lógica difusa para la estimación del esfuerzo de desarrollo de software. *Tecnología Procedia*, 7, 305-314.<https://doi.org/10.1016/j.protcy.2013.04.038>.
- García-Lamont, F., Cervantes, J., López, A., & Rodríguez, L. (2018). Segmentación de imágenes por características de color: una encuesta. *neurocomputación*.<https://doi.org/10.1016/j.neucom.2018.01.091>.
- Ghosh, S. y Dubey, SK (2013). Análisis comparativo de K-means y fuzzy C-significa algoritmos. *Revista internacional de informática avanzada y aplicaciones*, 4 (4), 35-39.<https://doi.org/10.14569/rijaca.2013.040406>. González, RC (2002). *Digital\_image\_processing\_2ndEd.pdf* (págs. 1-793).[www.prehn.com/gonzalezwoods](http://www.prehn.com/gonzalezwoods)
- Gopika, P., Krishnendu, CS, Hari Chandana, M., Ananthakrishnan, S., Sowmya, V., Gopalakrishnan, EA y Soman, KP (2020). Red neuronal de convolución de una sola capa para la clasificación de enfermedades cardíacas utilizando señales de electrocardiograma. *Aprendizaje profundo para análisis de datos*.<https://doi.org/10.1016/b978-0-12-819764-6.00003-x>. Güneş, A., Kalkan, H. y Durmuş, E. (2016). Optimización de la conversión de color a escala de grises versión para clasificación de imágenes. *Procesamiento de Señal, Imagen y Video*.<https://doi.org/10.1007/s11760-015-0828-7>.
- Gupta, B., Shukla, P. y Mittal, A. (2016). K-clasificación de vecino correlacionado más cercano para el reconocimiento de gestos del lenguaje de señas indio mediante la fusión de funciones. En *Actas de la conferencia internacional sobre comunicación por computadora e informática, ICCCI* (págs. 7 a 11).<https://doi.org/10.1109/ICCCI.2016.7479951>.
- Gupta, R. y Jha, N. (2020). Clasificación continua en tiempo real del lenguaje de señas us-conjunto de ventanas. En *Actas de la sexta conferencia internacional sobre sistemas avanzados de computación y comunicación, ICACCS 2020*(págs. 73-78).<https://doi.org/10.1109/ICACCS48705.2020.9074319>.
- Gupta, R. y Rajan, S. (2020). Análisis comparativo de redes neuronales de convolución modelos para la clasificación continua del lenguaje de señas indio. *Procedia Informática*, 171(2019), 1542-1550.<https://doi.org/10.1016/j.procs.2020.04.165>.
- Hartanto, R., Susanto, A. y Santosa, PI (2014). Gestor manual estático en tiempo real prototipo de sistema de reconocimiento para el lenguaje de señas de Indonesia. En *Actas de la sexta conferencia internacional sobre tecnología de la información e ingeniería eléctrica: aprovechamiento de la investigación y la tecnología a través de la colaboración entre la universidad y la industria, ICI-TEE*.<https://doi.org/10.1109/ICTEED.2014.7007911>.
- Hasan, Md Mehedi, Srizon, AY, Sayeed, A. y Hasan, MA M (2020). Clasificación de caracteres del lenguaje de señas mediante la aplicación de una red neuronal convolucional profunda. En *Actas de la 2da conferencia internacional sobre tecnologías avanzadas de información y comunicación, ICACIT*(págs. 434-438). 2020, noviembre.<https://doi.org/10.1109/ICACIT51780.2020.9333456>.
- Hasan, Mohammad Mahadi, Khaliluzzaman, M., Himel, SA y Chowdhury, R. T. (2017). Reconocimiento de lenguaje de señas manual para el alfabeto bengalí basado en el código de cadena Freeman y ANN. En *Actas de la 4ª conferencia internacional sobre avances en ingeniería eléctrica, ICAEE*(págs. 749-753). 2017, 2018-enero.<https://doi.org/10.1109/ICAEE.2017.8255454>.
- Hisham, B. y Hamouda, A. (2019). Clasificadores de aprendizaje supervisado para gestos árabes reconocimiento mediante Kinect V2. *SN Ciencias Aplicadas*, 1(7), 1-21.<https://doi.org/10.1007/s42452-019-0771-2>.
- Huang, Y., Monekosso, D., Wang, H. y Augusto, J. C (2012). Un método híbrido para reconocimiento de gestos de mano. En *Actas de la 8ª conferencia internacional sobre entornos inteligentes, IE 2012*(págs. 297-300).<https://doi.org/10.1109/IE.2012.30>
- Huang, YF y Chen, PH (2020). Detección de noticias falsas usando un aprendizaje conjunto modelo basado en algoritmos de búsqueda de armonía autoadaptativos. *Sistemas expertos con aplicaciones*, 159.<https://doi.org/10.1016/j.eswa.2020.113584>.
- Huong, TNT, Huu, TV, Xuan, T, Le y Van, S. V (2016). Reconocimiento estático del gesto de la mano nición para el lenguaje de señas vietnamita (VSL) utilizando el análisis de componentes principales. En *Actas de la conferencia internacional sobre informática, gestión y telecomunicaciones, ComManTel 2015*(págs. 138-141).<https://doi.org/10.1109/ComManTel.2015.7394275>.
- Ikonomakis, N., Plataniotis, KN y Vénetsanopoulos, AN (2000). Segmento de imagen en color mentación para aplicaciones multimedia. *Diario de Sistemas Inteligentes y Robóticos: Teoría y Aplicaciones*.<https://doi.org/10.1023/a:100816391397>.
- Ionescu, B., Coquin, D., Lambert, P. y Buzuloiu, V. (2005). Gesto dinámico de la mano reconocimiento utilizando el esqueleto de la mano. *Revista Eurasip sobre procesamiento de señales aplicadas*, 2005(13), 2101-2109.<https://doi.org/10.1155/ASP.2005.2101>.
- Islam, MM, Siddiqua, S. y Afnan, J (2017). Reconocimiento de gestos de mano en tiempo real usando diferentes algoritmos basados en el lenguaje de señas americano. En *Actas de la conferencia internacional IEEE sobre imágenes, visión y reconocimiento de patrones, ICIPVR 2017*. 13(págs. 1031-1036).<https://doi.org/10.1109/ICIPVR.2017.7890854>. Izquierdo, SS e Izquierdo, LR (2018). Sistemas difusos Mamdani para modelado y simulación: una evaluación crítica. *Revista de Sociedades Artificiales y Simulación Social*.<https://doi.org/10.18564/jasss.3660>.
- Izzah, A. y Suciati, N. (2014). Traducción de lenguaje de señas usando fourier genérico descriptor y vecino más cercano. *Revista Internacional de Cibernética e Informática*, 3 (1), 31-41.<https://doi.org/10.5121/ijci.2014.3104>.
- Jang, JSR (1993). ANFIS: Sistema de inferencia difusa basado en redes adaptativas. *IEEE Transacciones sobre Sistemas, Hombre y Cibernética*.<https://doi.org/10.1109/21.256541>.
- Jasim, M. y Hasanuzzaman, M. (2015). Interpretación de lenguaje de signos mediante lineal análisis discriminante y patrones binarios locales. 1-5. 10.1109/iciev.2014.7136001 Ji, P., Song, A., Xiong, P., Yi, P., Xu, X. y Li, H. (2017). Basado en una visión egocéntrica sistema de control de postura de la mano para robots de reconocimiento. *Revista de Sistemas Inteligentes y Robóticos: Teoría y Aplicaciones*, 87(3-4), 583-599.<https://doi.org/10.1007/s10846-016-0440-2>.
- Jiang, X., Lu, M. y Wang, S.-HSH (2020). Una neural convolucional de ocho capas red con agrupación estocástica, normalización por lotes y abandono para el reconocimiento de deletreo manual del lenguaje de señas chino. *Herramientas y aplicaciones multimedia*, 79 (21-22), 15697-15715.<https://doi.org/10.1007/s11042-019-08345-y>. Jiang, Y., Tao, J., Ye, W., Wang, W. y Ye, Z (2015). Un reconocimiento aislado del lenguaje de señas sistema de nición usando sensor RGB-D con codificación dispersa. En *Actas de la*

- 17.<sup>a</sup> conferencia internacional IEEE sobre ciencia e ingeniería computacional, CSE 2014, junto con la 13.<sup>a</sup> conferencia internacional IEEE sobre informática y comunicaciones ubicuas, ICCC 2014, 13.<sup>o</sup> simposio internacional sobre sistemas ubicuos (págs. 21 a 26).<https://doi.org/10.1109/CSE.2014.38>.
- Jielai, X., Hongwei, J. y Qiyi, T. (2015). Introducción a las redes neuronales artificiales. *Estadísticas médicas avanzadas*, 1431–1449.[https://doi.org/10.1142/9789814583312\\_0037](https://doi.org/10.1142/9789814583312_0037).
- Jin, CM, Omar, Z. y Jaward, MH (2016). Una aplicación móvil de signo americano. traducción de idiomas a través de algoritmos de procesamiento de imágenes. En*Actas del simposio de la región 10 de IEEE de 2016, TENSYMP 2016*(págs. 104-109).<https://doi.org/10.1109/TENCONSpring.2016.7519386>.
- Joshi, G., Singh, S. y Vig, R. (2020). Selección de parámetros HOG basada en Taguchi-TOPSIS para el reconocimiento de lenguaje de señas de fondo complejo. *Revista de Comunicación Visual y Representación de Imágenes*, 71, Artículo 102834.<https://doi.org/10.1016/j.jvcir.2020.102834>.
- Kadhim, RA y Khamees, M. (2020). Un reconocimiento de lenguaje de señas americano en tiempo real sistema de ción utilizando red neuronal convolucional para conjuntos de datos reales. *Revista TEM*. <https://doi.org/10.18421/TEM93-14>.
- Kalhor, M., Kajouei, A., Hamidi, F. y Asem, MM (2019). Evaluación de Técnicas de mejora de contraste de imágenes médicas basadas en histogramas: una implementación. En*Actas del noveno taller y conferencia anual de computación y comunicación de IEEE, CCWC 2019*.<https://doi.org/10.1109/CCWC.2019.8666468>.
- Kaluri, R. y Reddy, PC (2016a). Un marco para el reconocimiento de gestos de signos usando algoritmo genético mejorado y filtro adaptativo. *Inginería convincente*, 3(1), Artículo 1251730.<https://doi.org/10.1080/23311916.2016.1251730>.
- Kaluri, R. y Reddy, PC (2016b). Firmar el reconocimiento de gestos usando la región modificada algoritmo de crecimiento y clasificador difuso genético adaptativo. *Revista Internacional de Ingeniería y Sistemas Inteligentes*, 9(4), 225–233.<https://doi.org/10.22266/iies2016.1231.24>.
- Kaluri, R. y Reddy, PC (2017). Un marco mejorado para el reconocimiento de gestos de signos nición utilizando el modelo oculto de Markov y la técnica del histograma adaptativo. *Revista Internacional de Ingeniería y Sistemas Inteligentes*, 10(3), 11–19.<https://doi.org/10.22266/iies2017.0630.02>.
- Kamal, SM, Chen, Y., Li, S., Shi, X. y Zheng, J. (2019). Enfoques técnicos para Procesamiento del lenguaje de señas chino: una revisión. *Acceso IEEE*, 7, 96926–96935. <https://doi.org/10.1109/ACCESO.2019.2929174>.
- Kamishima, T., Hamasaki, M. y Akaho, S. (2009). TrBagg: un aprendizaje de transferencia simple y su aplicación a la personalización en el etiquetado colaborativo. En*Actas de la conferencia internacional IEEE sobre minería de datos, ICDM*(págs. 219–228).<https://doi.org/10.1109/ICDM.2009.9>.
- Kanezaki, A. (2018). Segmentación de imágenes no supervisada por backpropagation. En*Pro- actas de la conferencia internacional ICASSP, IEEE sobre acústica, voz y procesamiento de señales*.<https://doi.org/10.1109/ICASSP.2018.8462533>.
- Karami, A., Zanj, B. y Sarkaleh, AK (2011). Reconocimiento del lenguaje de señas persa (PSL) mediante transformada wavelet y redes neuronales. *Sistemas expertos con aplicaciones*, 38 (3), 2661–2667.<https://doi.org/10.1016/j.eswa.2010.08.056>.
- Karamizadeh, S., Abdullah, SM, Manaf, AA, Zamani, M. y Hooman, A. (2013). Una visión general del análisis de componentes principales. *Revista de Procesamiento de Señales e Información*.<https://doi.org/10.4236/jspip.2013.43031>.
- Kasmiri, F., Rahim, NSA, Othman, Z., Ahmad, SSS y Sahri, Z. (2020). un com- análisis parativo de filtros de brillo en el reconocimiento de lenguaje de signos. *Revista internacional de tendencias avanzadas en informática e ingeniería*.<https://doi.org/10.3053/ijatcse/2020/84942020>.
- Kassem, Y., Çamur, H. y Esenel, E. (2017). Sistema de inferencia adaptativo neuro-boroso tem (ANFIS) y metodología de superficie de respuesta (RSM) predicción de la viscosidad dinámica del biodiesel a 313K. *Procedia Informática*.<https://doi.org/10.1016/j.procs.2017.11.274>.
- Kaur, A. y Kaur, A. (2012). Comparación de tipo mamdani y tipo sugeno fuzzy sistemas de inferencia para sistema de aire acondicionado. *Revista internacional de computación blanda e ingeniería*.
- Kaur, M. y Chand, L. (2018). Revisión de la segmentación de imágenes y sus técnicas. *Revista de Tecnologías Emergentes e Investigación Innovadora (JETIR)*, 5(7), 974–981.
- Kaur, S. (2015). Tipos de ruido y diversas técnicas de eliminación. *Revista Internacional de Investigación Avanzada en Ingeniería Electrónica y de las Comunicaciones (IJAREC)*, 4(2), 226–230.
- Kaur', D. y Kaur, Y (2014). Diversas técnicas de segmentación de imágenes: una revisión. *Enterr- revista nacional de informática y computación móvil*.
- Kausar, S., Javed, M. y Sohail, S. (2008). Reconocimiento de gestos en pakistání lenguaje de señas usando clasificador difuso. En*Actas de la 8<sup>a</sup> conferencia internacional sobre procesamiento de señales, geometría computacional y visión artificial (Iscgav'08)*(págs. 101–105).<http://www.wseas.us/e-library/conferences/2008/rhodes/iscgav/iscgav16.pdf>.
- Khalid, S., Khalil, T. y Nasreen, S. (2014). Una encuesta de selección de características y características Técnicas de extracción en aprendizaje automático. En*Actas de la conferencia de ciencia e información de 2014, SAI 2014*(págs. 372–378).<https://doi.org/10.1109/SAI.2014.6918213>.
- Khan, AM (2013). Métodos de segmentación de imágenes: un estudio comparativo. *Internacional Revista de Ingeniería y Computación Suave (IJSCe)*.
- Kan, W. (2014). Técnicas de segmentación de imágenes: una encuesta. *Revista de Imagen y Gráficos*.<https://doi.org/10.12720/roig.1.4.166-170>.
- Khatri, M., Garg, AK, Gonzalez-Crespo, R., y Verdú, E. (2019). Reconocimiento de gestos de Imágenes estáticas RGB y RGB-D utilizando redes neuronales convolucionales. *Revista Internacional de Multimedia Interactiva e Inteligencia Artificial*, 5(7), 22.<https://doi.org/10.9781/ijmai.2019.09.002>.
- Kim, S., Kim, J., Ahn, S. y Kim, Y. (2018). Reconocimiento del lenguaje de los dedos basado en ensamblar el aprendizaje de redes neuronales artificiales utilizando sensores EMG de brazalete. *Tecnología y Salud*, 26(51), S249–S258.<https://doi.org/10.3233/THC-174602>. Kiselev, V., Khlamov, M. y Chuvilin, K. (2019). Reconocimiento de gestos con la mano con mu- Dispositivos de movimiento de salto triple. En*Actas de la conferencia de asociación de innovación abierta, FRUCT*(págs. 163–169). 2019-abril.<https://doi.org/10.23919/FRUCT.2019.8711887>.
- Kishore, PV, Kumar, DA, Goutham, END y Manikanta, M (2016). Continuo reconocimiento de lenguaje de señas a partir de características de seguimiento y forma usando Fuzzy Inference Engine. En*Actas de la conferencia internacional IEEE 2016 sobre comunicaciones inalámbricas, procesamiento de señales y redes, WISPNET 2016*(págs. 2165–2170).<https://doi.org/10.1109/WISPNET.2016.7566526>.
- Kishore, PVV, Prasad, MV, Prasad, CR y Rahul, R (2015). Modelo de 4 cámaras para el reconocimiento de lenguaje de señas usando descriptores elípticos de Fourier y ANN. En*Actas de la conferencia internacional sobre procesamiento de señales y sistemas de ingeniería de comunicaciones, SPACES 2015*, en asociación con IEEE(págs. 34–38).<https://doi.org/10.1109/ESPACIOS.2015.7058288>.
- Kishore, PVV y Kumar, Rajesh (2012). Un reconocimiento de lenguaje de señas indio basado en video sistema de nición (INSLR) usando transformada wavelet y lógica difusa. *Revista Internacional de Ingeniería y Tecnología*, 4(5), 537–542.<https://doi.org/10.7763/ijet.2012.v4.i4.427>.
- Koller, O., Zargaran, S., Ney, H. y Bowden, R. (2018). Señal profunda: Habilidades robusta Reconocimiento estadístico continuo del lenguaje de señas a través de CNN-HMM híbridos. *Revista Internacional de Visión por Computador*, 12(12), 1311–1325.<https://doi.org/10.1007/s11263-018-1121-3>.
- Kong, WW y Ranjanath, S. (2014). Hacia sujeto continuo independiente reconocimiento de lenguaje de señas: un enfoque de segmento y fusión. *Reconocimiento de patrones*, 47(3), 1294–1308.<https://doi.org/10.1016/j.patcog.2013.09.014>.
- Korzynska, A., Roszkowiak, L., Lopez, C., Bosch, R., Witkowski, L. y Leje- une, M. (2013). Validación de varios métodos de umbral adaptativo de segmentación aplicados a imágenes digitales de linfoma folicular teñidas con 3,3'-diaminobencidina y hematoxilina. *Patología diagnóstica*.<https://doi.org/10.1186/1746-1596-8-48>.
- Kotsiantis, SB, Zaharakis, ID y Pintelas, PE (2006). Aprendizaje automático: una revisión de técnicas de clasificación y combinación. *Revista de inteligencia artificial*, 26(3), 159–190.<https://doi.org/10.1007/s10462-007-9052-3>.
- Krishnaveni, M., Subashini, P. y Dhivyaprabha, TT (2019). Un marco asertivo para el sistema de reconocimiento automático de lenguaje de señas tamil utilizando inteligencia computacional. *Biblioteca de referencia de sistemas inteligentes*: 150. Publicaciones internacionales de Springer.[https://doi.org/10.1007/978-3-319-96002-9\\_3](https://doi.org/10.1007/978-3-319-96002-9_3).
- Kudrinko, K., Flavin, E., Zhu, X. y Li, Q. (2021). Lenguaje de señas portátil basado en sensores reconocimiento: una revisión exhaustiva. *Revisiões de IEEE en ingeniería biomédica*, 14, 82–97.<https://doi.org/10.1109/RBME.2020.3019769>.
- Kumar, G. y Bhata, PK (2014). Una revisión detallada de la extracción de características en la imagen sistemas de procesamiento En*Actas de la conferencia internacional sobre tecnologías avanzadas de computación y comunicación, ACCT, febrero*(págs. 5–12).<https://doi.org/10.1109/ACCT.2014.74>.
- Kumar, P., Gauba, H., Roy, PP y Dogra, DP (2017). acoplado basado en HMM multi- Fusión de datos de sensores para reconocimiento de lenguaje de señas. *Letras de reconocimiento de patrones*, 86, 1–8.<https://doi.org/10.1016/j.patrec.2016.12.004>.
- Kumarage, D., Fernando, S., Fernando, P., Madushanka, D. y Samarasinghe, R. (2011). Reconocimiento de gestos en lenguaje de señas en tiempo real mediante comparación de imágenes fijas y reconocimiento de movimiento. En*Actas de la sexta conferencia internacional sobre sistemas industriales y de información, ICII 2011*(págs. 169–174).<https://doi.org/10.1109/ICIINS.2011.6038061>.
- Lahiani, H., Elleuch, M. y Kherallah, M. (2016). Reconocimiento de gestos de mano en tiempo real sistema para dispositivos android. En*Actas de la conferencia internacional sobre diseño y aplicaciones de sistemas inteligentes, ISDA*.<https://doi.org/10.1109/ISDA.2015.7489184>.
- Lan, Y., Zhou, D., Zhang, H. y Lai, S. (2017). Desarrollo de modelos de alerta temprana. *Alerta temprana para brotes de enfermedades infecciosas: teoría y práctica*.<https://doi.org/10.1016/B978-0-12-812343-0-00003-5>.
- Lech, M. y Kostek, B. (2012). Reconocimiento de gestos con las manos respaldado por reglas difusas y filtros de Kalman. *Revista internacional de información inteligente y sistemas de bases de datos*, 6(5), 407–420.<https://doi.org/10.1504/IJIDS.2012.049304>.
- Lee, CKM, Ng, KH, Chen, CH, Lau, HCW, Chung, SY y Tsui, T. (2021). Método de entrenamiento y reconocimiento de lenguaje de señas americano con red neuronal recurrente. *Sistemas expertos con aplicaciones*, 16(Agosto de 2020), artículo 114403. <https://doi.org/10.1016/j.eswa.2020.114403>.
- Lee, GC, Yeh, FH y Hsiao, YH (2016). Lenguaje de señas taiwanés basado en Kinect sistema de reconocimiento *Herramientas y aplicaciones multimedia*, 79(1), 261–279.<https://doi.org/10.1007/s11042-014-2290-x>.
- Lee, K., Jin, K., Kim, Y., Lee, JH y Lee, EC (2021). Un análisis comparativo sobre el impacto del rastreador facial y la segmentación de la piel en la mejora del rendimiento de la fotopletilsmografía remota en tiempo real. *Notas de clase en informática*(págs. 27–37) (incluidas las notas de clase de la subserie sobre inteligencia artificial y las notas de clase sobre bioinformática), 12616 LNCS.[https://doi.org/10.1007/978-3-030-68452-5\\_3](https://doi.org/10.1007/978-3-030-68452-5_3).
- Lee, SU, Yoon Chung, S. y Park, RH (1990). Un desempeño comparativo estudio de varias técnicas globales de umbralización para la segmentación. *Visión artificial, gráficos y procesamiento de imágenes*, 52(2), 171–190.[https://doi.org/10.1016/0734-189X\(90\)90053-X](https://doi.org/10.1016/0734-189X(90)90053-X).
- Lemaster, B. y Monaghan, L. (2007). Variación en lenguajes de señas. *un compañero de antropología lingüística*.<https://doi.org/10.1002/9780470996522.ch7>.
- Lenard, JG, Pietrzyk, M. y Cser, L. (1999). Capítulo 9 - Modelo basado en el conocimiento- En g. *Simulación matemática y física de las propiedades de productos laminados en caliente*

- (págs. 279–318) En JG Lenard, M. Pietrzyk y L. Cser (Eds.). Elsevier Science Ltd.<https://doi.org/10.1016/B978-008042701-0/50009-7>.
- Li, SZ, Yu, B., Wu, W., Su, SZ y Ji, R. R (2015). Aprendizaje de características basado en SAE-Red PCA para reconocimiento de gestos humanos en imágenes RGBD. *Neurocomputación*, 151(P2), 565–573.<https://doi.org/10.1016/j.neucom.2014.06.086>.
- Li, TSH, Kao, MC y Kuo, P. H (2016). Sistema de reconocimiento para servicio a domicilio-lenguaje de señas relacionado utilizando el algoritmo K-means basado en entropía y HMM basado en ABC. *Transacciones IEEE en sistemas, hombre y cibernetica: sistemas*, 46(1), 150–162.<https://doi.org/10.1109/TSMC.2015.2435702>.
- Lim, KM, Tan, AWYC y Tan, SC (2016). Una matriz de covarianza de características con filtro de partículas en serie para reconocimiento de lenguaje de señas aislado. *Sistemas expertos con aplicaciones*, 54, 208–218.<https://doi.org/10.1016/j.eswa.2016.01.047>. Lionnie, R., Timotius, IK y Setyawan, I. (2012). Comparación de rendimiento de varios métodos de preprocesamiento en un sistema de reconocimiento de gestos de la mano basado en el vecino más cercano para diferentes condiciones de fondo. *Revista ITB de Tecnologías de la Información y la Comunicación*, 6(3), 183–194.<https://doi.org/10.5614/itbj.ict.2012.6.3.1>.
- Liu, X., Liu, Z., Wang, G., Cai, Z. y Zhang, H. (2017). Algoritmo de aprendizaje de transferencia en conjunto ritmo. *Acceso IEEE*, 6, 2389–2396.<https://doi.org/10.1109/ACCESO.2017.2782884>. Liwicki, S. y Everingham, M. (2009). Reconocimiento automático de palabras deletreadas con los dedos en lenguaje de señas británico. En *Actas de la conferencia IEEE de 2009 sobre visión artificial y reconocimiento de patrones, CVPR 2009* (págs. 50–57).<https://doi.org/10.1109/CVPR.2009.5204291>.
- Longkumer, N., Kumar, M. y Saxena, S. (2014). Técnicas de realce de contraste utilizando la ecualización de histogramas: una encuesta. *Revista Internacional de Ingeniería Tecnología Actual*, 4(3), 1561–1565.
- Lowe, G. (2004). SIFT: la transformación de características invariantes de escala. *revista internacional*. Mahmud, I., Tabassum, T., Uddin, MP, Ali, E., Nitu, AM y Afjal, MI (2019). Reducción de ruido eficiente y extracción de funciones HOG para el reconocimiento de lenguaje de señas. En *Actas de la conferencia internacional sobre avances en ingeniería eléctrica y electrónica, ICAEEE 2018* (págs. 1–4).<https://doi.org/10.1109/ICAEEE.2018.8642983>.
- Maini, R. y Aggarwal, H. (2009). Estudio y comparación de varios bordes de imagen. técnicas de protección. *Revista internacional del procesamiento de imágenes*.
- Maini, R. y Aggarwal, H. (2010). Una revisión completa de la mejora de la imagen técnicas. *Revista de Computación*, 2(3), 8–13.
- Majeed, SH e Isa, NAM (2021). Ecualización del histograma del índice de entropía adaptativa para imágenes de bajo contraste. *Acceso IEEE*, 9, 6402–6437.<https://doi.org/10.1109/ACCESO.2020.3048148>.
- Mamdani, EH y Assilian, S (1975). Un experimento de síntesis lingüística con un controlador de lógica difusa. *Revista internacional de estudios hombre-máquina*.[https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2).
- Maraqa, M. y Abu-Zaiter, R. (2008). Reconocimiento del uso de la lengua de señas árabe (ArSL) en redes neuronales recurrentes. En *Actas de la 1ra conferencia internacional sobre las aplicaciones de la información digital y las tecnologías web, ICADIWT 2008* (págs. 478–481).<https://doi.org/10.1109/ICADIWT.2008.4664396>.
- Martínez, AM, Wilbur, RB, Shay, R. y Kak, AC (2002). Purdue RVL-SLL. Base de datos ASL para el reconocimiento automático del lenguaje de señas estadounidense. En *Actas de la cuarta conferencia internacional IEEE sobre interfaces multimodales, ICMI 2002* (págs. 167–172).<https://doi.org/10.1109/ICMI.2002.1166987>.
- Martiskainen, P., Järvinen, M., Sköni, J.-P., Tiirikainen, J., Kolehmainen, M. y Mononen, J. (2009). Reconocimiento de patrones de comportamiento de vacas utilizando un acelerómetro tridimensional y máquinas de vectores de soporte. *Ciencias aplicadas del comportamiento animal*, 119(1), 32–38.<https://doi.org/10.1016/j.applanim.2009.03.005>.
- Maru, M. y Parikh, CM (2017). Técnicas de restauración de imágenes: una encuesta. *Internacional Revista Nacional de Aplicaciones Informáticas*, 16(06), 15–19.<https://doi.org/10.5120/ijca2017913060>.
- Mehdi, SA y Khan, YN (2002). Reconocimiento de lengua de signos mediante guantes sensores. En *Actas de ICONIP, 9.ª conferencia internacional sobre procesamiento de información neuronal: Inteligencia computacional para la E-Age: 5* (págs. 2204–2206).<https://doi.org/10.1109/ICONIP.2002.1201884>.
- Mehnert, A. y Jackway, P. (1997). Un algoritmo mejorado de crecimiento de la región sembrada. *Letras de reconocimiento de patrones*.[https://doi.org/10.1016/S0016-8655\(97\)00131-1](https://doi.org/10.1016/S0016-8655(97)00131-1). Mekala, P., Gao, Y., Fan, J. y Davari, A. (2011). Reconocimiento de lenguaje de señas en tiempo real basado en la arquitectura de redes neuronales. En *Actas del simposio anual del sudeste sobre teoría de sistemas* (págs. 195–199).<https://doi.org/10.1109/SSST.2011.5753805>.
- Meng, L. y Li, R. (2021). Un lenguaje de señas dual y de múltiples escalas que mejora la atención red de reconocimiento basada en una red de convolución de grafos. *Sensores (Suiza)*, 27(4), 1–22.<https://doi.org/10.3390/s21041120>.
- Merenda, M., Porcaro, C. e Iero, D (2020). Aprendizaje automático perimetral para IA habilitado Dispositivos IoT: una revisión. *Sensores (Suiza)*, 19, 20.<https://doi.org/10.3390/s20092533>.
- Merin Joy, D. y Rajesh, DM (2017). Reconocimiento de gestos de mano de vigilancia video usando la función de navegación. *Revista internacional de investigación y ciencia en ingeniería*, 3 (5), 82–89.<https://doi.org/10.25125/engineering-journal-joer-may-2017-31>. Mistry, D. y Banerjee, A. (2017). Comparación de detección de características y aplicación de coincidencia enfoques: SIFT y SURF. *Revistas GRD - Revista de investigación y desarrollo global para ingeniería*.
- Mittal, A., Kumar, P., Roy, PP, Balasubramanian, R. y Chaudhuri, BB (2019). A modelo LSTM modificado para el reconocimiento continuo de lenguaje de señas usando el movimiento de salto. *Revista de sensores IEEE*, 19(16), 7056–7063.<https://doi.org/10.1109/JSEN.2019.2909837>.
- Moghaddam, MJ y Soltanian-Zadeh, H. (2011). Segmentación de imágenes médicas usando redes neuronales artificiales. *Redes neuronales artificiales - Avances metodológicos y aplicaciones biomédicas*.<https://doi.org/10.5772/16103>.
- Mohandes, MA (2013). Reconocimiento de signos árabes a dos manos utilizando el ciber-guante. *Revista árabe de ciencia e ingeniería*, 36(3), 669–677.<https://doi.org/10.1007/s13369-012-0378-z>.
- Moreira Almeida, SG, Guimarães, FG, & Arturo Ramírez, J. (2014). Característica extractiva en el reconocimiento de la lengua de signos brasileña a partir de la estructura fonológica y utilizando sensores RGB-D. *Sistemas expertos con aplicaciones*, 41(16), 7259–7271.<https://doi.org/10.1016/j.eswa.2014.05.024>.
- Mufarroha, FA y Utaminingrum, F. (2017). Reconocimiento de gestos de mano mediante sistema de inferencia difuso basado en red adaptativa y K-vecino más cercano. *Revista Internacional de Tecnología*, 8(3), 559–567.<https://doi.org/10.14716/ijtech.v8i3.3146>.
- Mustafa, MB y Lahsasna, A. (2016). Un reconocedor de lenguaje de señas preciso que usa oráculo de esfera aleatoria. En *Actas de la 1.ª conferencia internacional sobre multimedia y procesamiento de imágenes de 2016, ICIMP 2016* (págs. 36–40).<https://doi.org/10.1109/ICIMP.2016.15>.
- Muthukrishnan, R. y Radha, M. (2011). Técnicas de detección de bordes para la segmentación de imágenes. *mentación Revista internacional de informática y tecnología de la información*.<https://doi.org/10.5121/ijcsit.2011.3620>.
- Narayanan, A., Sharmila, TS y Gajalakshmi, P. (2018). Reconocimiento de signos de palabra de imágenes invariantes basadas en SURF con mapas propios laplacianos. En *Actas de la 2.ª conferencia internacional sobre informática, comunicación y procesamiento de señales: enfoque especial en tecnología e innovación para entornos inteligentes, ICCSP 2018*.<https://doi.org/10.1109/ICCCSP.2018.8452825>.
- Nedeljkovic, I., Vi, C. y Vi, WG (2004). Clasificación de imágenes basada en lógica difusa. *Los Archivos Internacionales de Fotogrametría, Teledetección y Ciencias de la Información Espacial*, 34(30), 1–6.
- Nelson, A., Price, KJ y Multari, R. (2019). Diccionario inverso ASL - ASL traducción usando aprendizaje profundo. *Revisión de ciencia de datos de SMU*, 2(1).<https://scholar.smu.edu/datascience/reviewhttp://digitalrepository.smu.edu>. Disponible en: <https://scholar.smu.edu/datascience/review/vol/iss1/21>.
- Ngai, EWT, Peng, S., Alexander, P. y Moon, KKL (2014). Apoyo a la decisión y sistemas inteligentes en la cadena de suministro de textiles y prendas de vestir: una revisión académica de artículos de investigación. *Sistemas expertos con aplicaciones*, 41(1), 81–91.<https://doi.org/10.1016/j.eswa.2013.07.013>.
- Nguen, NT, Sako, S. y Kwolke, B. (2019). Profundo reconocimiento basado en CNN del dedo JSL ortografía. En *CLMQPHCRE Pérez García H. Sánchez González L. (Ed.) Notas de clase en informática.. Springer International Publishing (incluidas las subseries Lecture Notes in Artificial Intelligence y Lecture Notes in Bioinformatics): vol. 11734 LNIA*.[https://doi.org/10.1007/978-3-03-29859-3\\_51](https://doi.org/10.1007/978-3-03-29859-3_51). Niblack, W. (1986). Una introducción al procesamiento digital de imágenes. *una introducción a procesando imagen digital*.<https://doi.org/10.1002/0470035528.ch15>.
- Nisha, SS y Meerab, NM (2021). Aplicaciones del aprendizaje profundo en biomedicina ingeniería. *manual de aprendizaje profundo en ingeniería biomédica*.<https://doi.org/10.1016/B978-0-12-823014-5.00008-9>.
- Nuhu, BK, Aliyu, I., Adegbeye, MA, Ryu, JK, Olaniyi, OM y Lim, CG (2021). Sistema experto de monitoreo de salud estructural basado en red distribuida. *Investigación e información sobre edificios*, 49(1), 144–159.<https://doi.org/10.1080/09613218.2020.1854083>.
- Oliveira, M., Sutherland, A. y Farouk, M. (2017). PCA de dos etapas con interpolado datos para el reconocimiento de la forma de la mano en lengua de signos. En *Actas del taller de reconocimiento de patrones de imágenes aplicadas*.<https://doi.org/10.1109/APIR.2016.8010587>.
- Oyewumi Jimoh, K., Morufat Adepoju, T., Sobowale, AA y Ayilara, O. A (2018). Sistema de reconocimiento de gestos sin conexión para el conteo de números yorubá. *Revista asiática de investigación en informática*, del 1 al 11 de septiembre de 2018.<https://doi.org/10.9734/ajrcos/2018/v1424753>.
- Pan, J., Luo, Y., Li, Y., Tham, C.-K., Heng, C.-H. y Thean, AVY (2020). un inalámbrico Sistema de sensor capacitivo multicanal para un reconocimiento de gestos eficiente basado en guantes con IA en el borde. *Transacciones IEEE en circuitos y sistemas II: Resúmenes expresos*, 67(9), 1624–1628.<https://doi.org/10.1109/TCII.2020.3010318>.
- Pan, T. Y., Lo, L., Yeh, C. W., Li, J., Liu, HT y Hu, MC (2016). señal en tiempo real reconocimiento de idioma en una escena de fondo compleja basado en un método de clasificación de agrupamiento jerárquico. En *Actas de la segunda conferencia internacional de IEEE sobre big data multimedia, BigMM 2016* (págs. 64–67).<https://doi.org/10.1109/BiMM.2016.44>.
- Pansare, JR, Dhurmal, H., Babar, S. e International, KS (2015). estática en tiempo real sistema de reconocimiento de gestos con la mano en un fondo complejo que usa el sistema numérico del lenguaje de señas indio reconocimiento estático de gestos con la mano para alfabetos de lenguaje de señas usando histograma orientado al borde y SVM multiclasificación con la mano en tiempo real R. 2(3), 2014–2015.
- Pansare, JR, Gawande, SH e Ingle, M. (2012). Gesto manual estático en tiempo real reconocimiento del lenguaje de señas americano (ASL) en contextos complejos. *Revista de Procesamiento de Señales e Información*, 03(03), 364–367.<https://doi.org/10.4236/sip.2012.33047>.
- Pansare, Jayshree R. e Ingle, M. (2016). Enfoque basado en la visión para el signo americano reconocimiento de idioma utilizando histograma de orientación de borde. En *Actas de la conferencia internacional sobre imagen, visión y computación, ICIVC 2016* (págs. 86–90).<https://doi.org/10.1109/ICIVC.2016.7571278>.
- Panwar, P., Gopal, G. y Kumar, R. (2016). Segmentación de imágenes usando K-means agrupamiento y umbralización segmentación de imágenes usando agrupamiento y umbralización de K-means. *Revista internacional de investigación de ingeniería y tecnología*.
- Patel, JM y Gamit, NC (2016). Una revisión de las técnicas de extracción de características en recuperación de imágenes basadas en tienda. En *Actas de la conferencia internacional IEEE 2016 sobre comunicaciones inalámbricas, procesamiento de señales y redes, WiSPNET 2016* (págs. 2259–2263).<https://doi.org/10.1109/WiSPNET.2016.7566544>.
- Patil, SB y Sinha, GR (2017). Extracción de características distintivas para el lenguaje de señas indio. gesto de calibre (ISL) utilizando la transformación de características invariantes de escala (SIFT). *diario de la*

- Institución de Ingenieros (India): Serie B, 98(1), 19–26.<https://doi.org/10.1007/s40031-016-0250-8>.
- Perumal, S. y Velmurugan, T. (2018). Preprocesamiento por tecnología de mejora de contraste níquies para imágenes médicas. *Revista Internacional de Matemática Pura y Aplicada*, 118(18), 3681–3688.
- Phi, LT, Nguyen, HD, Bui, TT y Vu, TT (2015). Un gesto a base de guantes sistema de reconocimiento de la lengua de señas vietnamita. En *Actas de la ICCAS, 15<sup>a</sup> conferencia internacional sobre control, automatización y sistemas*(págs. 1555–1559). <https://doi.org/10.1109/ICCAS.2015.7364604>.
- Polikar, R. (2012). Aprendizaje automático de conjuntos. *Aprendizaje automático de conjuntos*. <https://doi.org/10.1007/978-1-4419-9326-7>.
- Porfirio, AJ, Wiggers, KL, Oliveira, LES y Weingaertner, D (2013). signo LIBRAS Reconocimiento de la configuración de la mano del lenguaje basado en mallas 3D. En *Actas de la conferencia internacional IEEE sobre sistemas, hombre y cibernetica, SMC 2013*(págs. 1588–1593). <https://doi.org/10.1109/SMC.2013.274>.
- Prasad, MVD, Kishore, PVV, Kiran Kumar, E. y Anil Kumar, D. (2016). indio Sistema de reconocimiento de lenguaje de señas que utiliza un nuevo operador de borde basado en fusión. *Revista de Tecnología de la Información Teórica y Aplicada*, 88(3), 574–584. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84976576556&partnerID=40andmd5=6057d7251cae34213b054e885cc97967>.
- Prasad, MVD, Kishore, PVV, Anil Kumar, D. y Prasad, C. R (2016). Clase difusa sifier para el reconocimiento continuo del lenguaje de señas a partir de características de seguimiento y forma. *Revista India de ciencia y tecnología*, 30, 9.<https://doi.org/10.17485/ijst/2016/v9i30/98726>.
- Raghuveera, T., Deepthi, R., Mangalashri, R. y Akshaya, R. (2020). una profundidad- basado en el reconocimiento del lenguaje de señas indio usando Microsoft Kinect. *Sadhana - Actas de la Academia en Ciencias de la Ingeniería*, 45(1), 1–13.<https://doi.org/10.1007/s12046-019-1250-6>.
- Raheja, JL, Mishra, A. y Chaudhary, A. (2016). Reconocimiento del lenguaje de señas indio utilizando SVM. *Reconocimiento de patrones y análisis de imágenes*, 26(2), 434–441.<https://doi.org/10.1134/S1054661816020164>.
- Rahim, MA, Shin, J. y Yun, K. S (2020). Reconocimiento del alfabeto de signos basado en gestos manuales e interpretación de oraciones utilizando una red neuronal convolucional. *Anales de tecnologías emergentes en informática*, 4(4), 20–27.<https://doi.org/10.33166/AETIC.2020.04.003>.
- Rahim, MA, Islam, MR y Shin, J (2019). Basado en el reconocimiento de palabras de signos sin contacto sobre gestos dinámicos con la mano usando segmentación híbrida y fusión de características de CNN. *Ciencias Aplicadas (Suiza)*, (18), 9.<https://doi.org/10.3390/app9183790>. Raj, RD y Jasuja, A. (2018). Reconocimiento de lenguaje de signos británica mediante HOG. En *Actas de la conferencia internacional de estudiantes del IEEE sobre electricidad, electrónica e informática, SCEECS 2018*(págs. 1–4).<https://doi.org/10.1109/SCEECS.2018.8546967>.
- Raj, R. y Joseph, N (2016). Extracción de puntos clave utilizando el algoritmo SURF para CMFD, *procedia informática*93, págs. 375–381 En JJ Krishna DD Mathew J. (Ed.). Elsevier BV <https://doi.org/10.1016/j.procs.2016.07.223>.
- Ramos, ALA, Dalhag, GDM, Daygon, MLD, Omar, J., Cruz, KDL, Macaranas, AA y Almodóvar, KLJ V (2019). Reconocimiento de gestos alfabéticos con la mano usando histograma de gradientes orientados, máquina de vectores de soporte y K: algoritmo vecino más cercano. *Revista Internacional de Investigación de Ciencias de la Computación*, 6 (04), 200–205.
- Rao, GA y Kishore, PVV (2018). Selfie video basado en lenguaje de señas indio continuo sistema de reconocimiento de calibre. *Revista de ingeniería Ain Shams*.<https://doi.org/10.1016/j.jasej.2016.10.013>.
- Rashid, O., Al-Hamadi, A. y Michaelis, B. (2009). Un marco para la integración de reconocimiento de gestos y posturas mediante HMM y SVM. En *Actas de la conferencia internacional IEEE sobre computación inteligente y sistemas inteligentes, ICIS 2009*: 4(págs. 572–577).<https://doi.org/10.1109/ICICISYS.2009.5357615>. Rashmi, Kumar, M. y Saxena, R (2013). Algoritmo y técnica en varios bordes. detección: una encuesta. *Procesamiento de señales e imágenes: una revista internacional*. <https://doi.org/10.5121/sipij.2013.4306>.
- Rastgoo, R., Kiani, K. y Escalera, S. (2021a). Pose de mano consciente multimodal aislado reconocimiento de lenguaje de señas. *Herramientas y aplicaciones multimedia*, 80(1), 127–163. <https://doi.org/10.1007/s11042-020-09700-0>.
- Rastgoo, R., Kiani, K. y Escalera, S. (2021b). Reconocimiento del lenguaje de señas: una profunda encuesta. *Sistemas expertos con aplicaciones*, 164(febrero 2020), artículo 113794. <https://doi.org/10.1016/j.eswa.2020.113794>.
- Ravi, S., Malaji, S., Poluri, VVK y Eepuri, KK (2018). Reconocimiento de lenguaje de señas ción con fusión de características múltiples y clasificador ANN. *Revista turca de ingeniería eléctrica y ciencias de la computación*, 26(6), 2871–2885.<https://doi.org/10.3906/elk-1711-139>.
- Razmjoo, N., Razmjoo, S., Vahedi, Z., Estrela, V. y de Oliveira, GG (2021). Piel segmentación de color basada en una red neuronal artificial mejorada por un algoritmo de optimización de saltamontes modificado. *Apuntes de clase en ingeniería eléctrica*, 696, 169–185.[https://doi.org/10.1007/978-3-030-56689-0\\_9](https://doi.org/10.1007/978-3-030-56689-0_9).
- Reeves, SJ (2014). Restauración de imágenes: Fundamentos de la restauración de imágenes. 10.1016/b978-0-12-396501-1.00006-6
- Rekha, J., Bhattacharya, J. y Majumder, S. (2011a). Reconocimiento de gestos de mano para lenguaje de señas: un nuevo enfoque híbrido. En *Actas de la conferencia internacional sobre procesamiento de imágenes, visión artificial y reconocimiento de patrones, IPCV 2011*: 1(págs. 80–86).
- Rekha, J., Bhattacharya, J. y Majumder, S. (2011b). Forma, textura y movimiento local. Características de los gestos de la mano para el reconocimiento del lenguaje de señas indio. En *Actas de TISC, 3ra conferencia internacional sobre trendz en ciencias de la información y computación*(págs. 30–35).<https://doi.org/10.1109/TISC.2011.6169079>. Rogowska, J. (2009). Descripción general y fundamentos de la segmentación de imágenes médicas.
- manual de procesamiento y análisis de imágenes médicas(2<sup>a</sup> edición). Elsevier inc. <https://doi.org/10.1016/B978-012373904-9.50013-1>.
- Rubini, C. y Pavithra, N. (2019). Mejora de contraste de imágenes de resonancia magnética usando AHE y técnicas CLAHE. *Revista internacional de tecnología innovadora e ingeniería de exploración*, 9(2), 2442–2445.<https://doi.org/10.35940/ijite.b7017.129219>. Ryumin, D., Ivanko, D., Axyonov, A., Kagirov, I., Karpov, A. y Zelezny, M. (2019). Interacción humano-robot con carro de la compra inteligente mediante lenguaje de señas: Recogida de datos. En *Actas de la conferencia internacional IEEE sobre talleres de computación y comunicaciones omniresidentes, Talleres PerCom*(págs. 949–954).<https://doi.org/10.1109/PERCOMW.2019.8730886>.
- Saha, HN, Tapadar, S., Ray, S., Chatterjee, SK y Saha, S. (2018). Una maquina enfoque basado en el aprendizaje para el reconocimiento de gestos manuales mediante la extracción de características distintivas. En *Actas del octavo taller y conferencia anual de computación y comunicación de IEEE, CCWC*(págs. 91–98). 2018, 2018-enero.<https://doi.org/10.1109/CCWC.2018.8301631>.
- Saha, S. (2018). Una guía completa de redes neuronales convolucionales: el ELISI camino | de Sumit Saha | Hacia la ciencia de datos.
- Sahoo, AK, Mishra, GS y Ravulakollu, KK (2014). Reconocimiento de lenguaje de señas: Lo último. *ARPV Revista de Ingeniería y Ciencias Aplicadas*, 9(2), 116–134. Saleh, Y. e Issa, GF (2020). Reconocimiento del lenguaje de señas árabe a través de redes neuronales profundas puesta a punto de redes. *Revista internacional de ingeniería biomédica y en línea*, 16 (5), 71–83.<https://doi.org/10.3991/IJOE.V16I05.13087>.
- Sallam, AA, Kabir, MN, Shamhan, ANM, Nasser, HK y Wang, J (2021). A método de reconocimiento racial basado en el color y la textura facial para mejorar la clasificación demográfica, *Notas de clase en ingeniería eléctrica*,(666, págs. 843–852). [https://doi.org/10.1007/978-981-15-5281-6\\_61](https://doi.org/10.1007/978-981-15-5281-6_61).
- Savur, C. y Sahn, F. (2016). Sistema de reconocimiento de lenguaje de señas estadounidense en tiempo real utilizando la señal EMG de superficie. En *Actas de la 14.<sup>a</sup> conferencia internacional de IEEE sobre aplicaciones y aprendizaje automático, ICMLA 2015*(págs. 497–502).<https://doi.org/10.1109/ICMLA.2015.212>.
- Savur, C. y Sahn, F. (2017). Sistema de reconocimiento de lenguaje de señas estadounidense mediante el uso de señal EMG de superficie. En *Actas de la conferencia internacional IEEE sobre sistemas, hombre y cibernetica, SMC 2016*(págs. 2872–2877).<https://doi.org/10.1109/SMC.2016.7844675>.
- Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S. y Cormier, K. (2013). Edificio el corpus de lenguaje de señas británico. *Documentación y conservación de lenguas*, 7, 136–154.
- Scott, SM, O'Hare, WT y Ali, Z. (2003). Lógica difusa y tecnología de clasificación difusa. niques 95–134. 10.1007/978-3-540-36213-5\_5
- Selim, S., Elhinnamy, E., Othman, H., Abouelsaadat, W. y Salem, MAM (2019). Una revisión de los enfoques de aprendizaje automático para la predicción de ataques epilépticos. En *Actas de la ICICES, 14<sup>a</sup> conferencia internacional sobre ingeniería y sistemas informáticos*(págs. 239–244). Instituto de Ingenieros Eléctricos y Electrónicos Inc. <https://doi.org/10.1109/ICCES4960.2019.9068190>.
- Senthilkumaran, N. y Vaithegi, S. (2016). Segmentación de imágenes mediante umbral-técnicas de captura de imágenes médicas. *Informática e ingeniería: una revista internacional*.<https://doi.org/10.5121/cseij.2016.6101>.
- Sethi, A., Hemant, S., Kumar, K., N, BR y Krishnan, R (2012). SignPro-una aplicación-suite para sordos y mudos. *Revista Internacional de Ingeniería y Tecnología Informática*, 2(5), 1203–1206.
- Shaikh, KB, Ganeshan, P., Kalist, V., Sathish, BS y Jenitha, JMM (2015). Compar-estudio innovador de detección y segmentación del color de la piel en el espacio de color HSV y YCbCr. En SAK Lobiyal DK (Ed.). *Procedia Informática*, 57, 41–48 Elsevier. <https://doi.org/10.1016/j.procs.2015.07.362>.
- Shanta, SS, Anwar, ST y Kabir, MR (2018). Detección de lenguaje de señas bengalí us-ing SIFT y CNN. En *Actas de la 9.<sup>a</sup> conferencia internacional sobre tecnologías informáticas, de comunicación y de redes, ICCNT 2018*.<https://doi.org/10.1109/ICCNT.2018.8493915>.
- Sharma, A., Badal, T., Gupta, A., Gupta, A. y Anand, A (2021). Una señal inteligente máquina de comunicación para personas con discapacidades auditivas y del habla. *Comunicaciones en Informática y Ciencias de la Información*, 1367, 75–86.[https://doi.org/10.1007/978-981-16-0401-0\\_6](https://doi.org/10.1007/978-981-16-0401-0_6).
- Sharma, A., Mittal, A., Singh, S. y Awatramani, V (2020). Reconocimiento de gestos con las manos utilizando técnicas de procesamiento de imágenes y extracción de características. *Procedia Informática*, 173(2019), 181–190.<https://doi.org/10.1016/j.procs.2020.06.022>.
- Sharma, S., Gupta, R. y Kumar, A. (2020). Trbagboost: una trans-fer método de aprendizaje aplicado al reconocimiento del lenguaje de señas indio. *Revista de inteligencia ambiental y computación humanizada*, 2018.<https://doi.org/10.1007/s12652-020-01979-z>.
- Shleeg, AA y Ellabib, IM (2013). Comparación de mamdani y sugeno fuzzy sistemas de interferencia para el riesgo de cáncer de mama. *Revista Internacional de Ingeniería Informática, Eléctrica, Automatización, Control e Información*.
- Shohieb, SM, Elminir, HK y Riad, AM (2015). Atlas mundial de signos; Un punto de referencia Base de datos de lenguaje de señas árabe. *Revista de la Universidad King Saud - Informática y Ciencias de la Información*, 27(1), 68–76.<https://doi.org/10.1016/j.jksuci.2014.03.011>.
- Shrivakshan, GT y Chandrasekar, C. (2012). Una comparación de varios detectores de bordes ciones técnicas utilizadas en el procesamiento de imágenes. *Revista Internacional de Problemas de Ciencias de la Computación*.
- Shukla, P., Garg, A., Sharma, K. y Mittal, A. (2016). Un descriptor de DTW y Fourier enfoque basado en el reconocimiento del lenguaje de señas indio. En *Actas de la 3ra conferencia internacional sobre procesamiento de información de imágenes*, ICIP 2015(págs. 113–118). <https://doi.org/10.1109/ICIP.2015.7414750>.
- Simon, C., Meessen, J., Tzovaras, D. y De Vleeschouwer, C (2007). decisión de uso árboles para el análisis de datos estructurados topológicamente asistidos por el conocimiento. En *Actas del 8<sup>o</sup> taller internacional sobre análisis de imágenes para servicios interactivos multimedia*, WIAMIS 2007(págs. 6–9).<https://doi.org/10.1109/WIAMIS.2007.84>.

- Simske, S. (2019). Introducción, descripción general y aplicaciones. *Meta-análisis*. <https://doi.org/10.1016/b978-0-12-814623-1.00001-0>.
- Singh, P. y Shree, R. (2016). Un estudio comparativo de modelos de ruido e imagen técnicas de restauración. *Revista Internacional de Aplicaciones Informáticas*, 14(1), 18-27. <https://doi.org/10.5120/ijca2016911336>.
- Singh, S., Gupta, AK y Singh, T. (2019). Reconocimiento de lengua de signos mediante híbrido Redes neuronales. *Revista internacional de tecnología innovadora e ingeniería de exploración*, 9(2), 1092-1098. <https://doi.org/10.35940/ijitee.I3349.129219>. Song, Q. y Chissom, BS (1993). Series temporales difusas y sus modelos. *Conjuntos borrosos y Sistemas*, 5(4), 269-277. [https://doi.org/10.1016/0165-0114\(93\)90372-O](https://doi.org/10.1016/0165-0114(93)90372-O). Song, T., Zhao, H., Liu, Z., Liu, H., Hu, Y. y Sun, D (2021). mano humana inteligente reconocimiento de gestos mediante funciones de reconocimiento de calidad de fusión local-global. *Sistemas informáticos de generación futura*, 115, 298-303. <https://doi.org/10.1016/j.future.2020.09.013>.
- Sonkusare, JS, Chopade, NB, Sor, R. y Tade, SL (2015). Una reseña a la mano Sistema de reconocimiento de gestos. En *Actas de la primera conferencia internacional sobre computación, comunicación, control y automatización, ICCUBE: 2015*(págs. 790-794). <https://doi.org/10.1109/ICCUBEA.2015.158>.
- Suharjito, Thiracitta, N., Gunawan, H. y Witjaksono, G (2019). la comparacion de algunos modelos ocultos de markov para el reconocimiento de lenguaje de señas. En *Actas de la primera conferencia internacional de la asociación de Indonesia para el reconocimiento de patrones, INAPR 2018*(págs. 6-10). <https://doi.org/10.1109/INAPR.2018.8627031>.
- Suharjito, Anderson, R., Wiryana, F., Arestia, MC y Kusuma, G. P (2017). lenguaje de señas Sistemas de aplicación de reconocimiento de lenguaje para personas sordomudas: una revisión basada en entrada-proceso-salida. *Procedia Informática*, 116, 441-448. <https://doi.org/10.1016/j.procs.2017.10.028>.
- Sujatha, P. y Sudha, KR (2015). Análisis de rendimiento de diferentes detección de bordes. Técnicas de segmentación de imágenes. *Revista india de ciencia y tecnología*. <https://doi.org/10.17485/ijst/2015/v8i14/72946>.
- Sur, C., Zhang, T. y Xu, C. (2015). Modelado de máquina de vector de soporte latente para signo Reconocimiento de idiomas con kinect. *Transacciones ACM sobre Sistemas Inteligentes y Tecnología*, 6(2). <https://doi.org/10.1145/2629481>.
- Supančič, JS, Rogez, G., Yang, Y., Shotton, J. y Ramanan, D (2018). Profundidad-estimación basada en la postura de la mano: métodos, datos y desafíos. *Revista Internacional de Visión por Computador*, 12(11), 1180-1198. <https://doi.org/10.1007/s11263-018-1081-7>.
- Sykora, P., Kamencay, P. y Hudec, R. (2014). Comparación de los métodos SIFT y SURF para usar en el reconocimiento de gestos con la mano basado en el mapa de profundidad. *Procedia AASRI*, 9(Csp), 19-24. <https://doi.org/10.1016/j.aasri.2014.09.005>.
- Tabassum, MR, Gias, AU, Kamal, MM, Islam, S., Muctadir, HM, Ibrahim, M., Shakir, AK, Imran, A., Islam, S., Rabbani, MG, Khaled, SM, Islam, MS y Begum, Z. (2010). Estudio comparativo de algoritmos estadísticos de detección de piel para imágenes humanas subcontinentales. *Revista de tecnología de la información*, 9(4), 811-817. <https://doi.org/10.3923/rtj.2010.811.817>.
- Tan, ZY, Basah, SN, Yazid, H. y Safar, MJA (2021). Análisis de rendimiento de Umbral de Otsu para la segmentación del lenguaje de señas. *Herramientas y aplicaciones multimedia*. <https://doi.org/10.1007/s11042-021-10688-4>.
- Tania, S. y Rowaida, R. (2016). Un estudio comparativo de varias tecnologías de filtrado de imágenes. niques para eliminar varios píxeles ruidosos en una imagen aérea. *Revista internacional de procesamiento de señales, procesamiento de imágenes y reconocimiento de patrones*, 9(3), 113-124. <https://doi.org/10.14257/ijsp.2016.9.3.10>.
- Tariq, M., Iqbal, A., Zahid, A., Iqbal, Z. y Akhtar, J. (2012). Localización del lenguaje de señas ción: Aprender a eliminar los dialectos del idioma. En *Actas de la 15<sup>a</sup> conferencia internacional multitemática, INMIC 2012*(págs. 17-22). <https://doi.org/10.1109/INMIC.2012.6511463>.
- Taskiran, M., Kilioglu, M. y Kahraman, N. (2018). Un sistema en tiempo real para reconocer el lenguaje de señas americano mediante el uso de aprendizaje profundo. En *Actas de la 41<sup>a</sup> conferencia internacional sobre telecomunicaciones y procesamiento de señales, TSP 2018*(págs. 1 a 5). <https://doi.org/10.1109/TSP.2018.8441304>.
- Tavari, N. V y Deorankar, PAV (2014). Reconocimiento del lenguaje de señas indio basado en histogramas de gradiente orientado. *Revista Internacional de Ciencias de la Computación y Tecnologías de la Información*, 5(3), 3657-3660.
- Tharwat, A., Gaber, T., Hassaniene, AE, Shahin, MK y Refaat, B. (2015). Tamizar-sistema de reconocimiento de lenguaje de señas árabe basado. *Avances en Sistemas Inteligentes y Computación*, 334, 359-370. <https://doi.org/10.1007/978-3-319-13572-430>.
- Thepade, SD, Kulkarni, G., Narkhede, A., Kelvekar, P. y Tathe, S. (2013). lenguaje de señas reconocimiento de calibre utilizando medios de color de imágenes de borde de magnitud de pendiente de gradiente. En *Actas de la conferencia internacional sobre sistemas inteligentes y procesamiento de señales, ISSP 2013*(págs. 216-220). <https://doi.org/10.1109/ISSP.2013.6526905>.
- Toletino, LKS, Serfa Juan, RO, Thio-ac, AC, Pamahoy, MAB, Forteza, JRR, & García, XJO (2019). Reconocimiento estático de lenguaje de señas mediante aprendizaje profundo. *Revista internacional de aprendizaje automático y computación*, 9(6), 821-827. <https://doi.org/10.18178/jimc.2019.9.6.879>.
- Tornay, S., Razavi, M. y Magimai-Doss, M. (2020). Hacia un lenguaje de señas multilingüe reconocimiento de calibre. En *Actas de la conferencia internacional ICASSP, IEEE sobre acústica, voz y procesamiento de señales*(págs. 6309-6313). <https://doi.org/10.1109/ICASSP40776.2020.9054631>.
- Torrione, PA, Morton, KD, Sakaguchi, R. y Collins, LM (2014). Histogramas de gradientes orientados para la detección de minas terrestres en datos de radar de penetración terrestre. *Transacciones IEEE sobre geociencia y teledetección*, 52(3), 1539-1550. <https://doi.org/10.1109/TGRS.2013.2252016>.
- Ummaheswari, D. y Karthikeyan, E. (2019). Análisis comparativo de varios filtros. técnicas de procesamiento de imágenes. *Revista Internacional de Investigación Científica y Tecnológica*.
- Vandeginste, BGM, Massart, DL, Buydens, LMC, De Jong, S., Lewi, PJ y Smeyers-Verbeke, J. (1998). Capítulo 44 - Redes neuronales artificiales, *manual de quimiometría y cualimetría: Parte B*(20, págs. 649-699) En BGM Vandeginste, DL Massart, LMC Buydens, S. De Jong, PJ Lewi y J. Smeyers-Verbeke (Eds.). Elsevier. [https://doi.org/10.1016/S0922-3487\(98\)80054-3](https://doi.org/10.1016/S0922-3487(98)80054-3).
- Verma, N. y Dutta, M. (2017). Técnicas de realce de contraste: una breve y revisión concisa. *Revista Internacional de Investigación de Ingeniería y Tecnología (IRIET)*.
- Verma, OP, Hanmandlu, M., Susan, S., Kulkarni, M. y Jain, PK (2011). un sim- Algoritmo de crecimiento de una sola región sembrada para la segmentación de imágenes en color mediante el umbral adaptativo. En *Actas de la conferencia internacional sobre sistemas de comunicación y tecnologías de red, CSNT 2011*. <https://doi.org/10.1109/CSNT.2011.107>.
- Verma, R. y Dev, A. (2009). Reconocimiento de gestos de la mano basado en visión usando estado finito máquinas y lógica difusa. En *Actas de la conferencia internacional sobre telecomunicaciones y talleres ultramodernos*. <https://doi.org/10.1109/ICUMT.2009.5345425>.
- Vincent, H., Tomoya, S. y Gentiane, V. (2019). Neural convolucionar y recurrente red para el reconocimiento de la acción humana: Aplicación en lenguaje de señas americano. *BioRxiv*. 10.1101/535492
- Vogler, C. y Metaxas, D (1999). Modelos paralelos ocultos de Markov para el signo americano reconocimiento de lenguaje. En *Actas de la Conferencia Internacional IEEE sobre Visión por Computador*. <https://doi.org/10.1109/iccv.1999.791206>.
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F. y Fouill-loy, A. (2017). Métodos de aprendizaje automática para la predicción de la radiación solar: una revisión. *Energía renovable*, 105, 569-582. <https://doi.org/10.1016/j.renene.2016.12.095>.
- Wadhawan, A. y Kumar, P. (2020). Reconocimiento de lenguaje de señas basado en aprendizaje profundo sistema para señales estáticas. *Informática neuronal y aplicaciones*, 32(12), 7957-7968. <https://doi.org/10.1007/s00521-019-04691-y>.
- Wadhawan, A. y Kumar, P. (2021). Sistemas de reconocimiento de lenguaje de signos: una década revisión sistemática de la literatura. *Archivos de métodos computacionales en ingeniería*. <https://doi.org/10.1007/s11831-019-09384-2>.
- Wang, H., Chai, X., Hong, X., Zhao, G. y Chen, X (2016). Lenguaje de señas aislado reconocimiento con matrices de covarianza de Grassmann. *Transacciones de ACM en informática accesible*, 8(4). <https://doi.org/10.1145/2897735>.
- Wang, H., Chai, X., Zhou, Y. y Chen, X (2015). Beneficios del reconocimiento rápido del lenguaje de señas ajustando a partir de una aproximación de bajo rango. En *Actas de la 11.<sup>a</sup> conferencia internacional IEEE y talleres sobre reconocimiento automático de rostros y gestos, FG 2015*. <https://doi.org/10.1109/FG.2015.7163092>.
- Wang, K. (2001). Inteligencia computacional en la ingeniería de fabricación ágil. *Ag-Manufactura ile: La estrategia competitiva del siglo XXI*. <https://doi.org/10.1016/b978-008043567-1/50016-4>.
- Wang, SJ, Zhang, LB, Zhang, DC, Jia, CC, Zhang, N. y Zhou, CG (2010). Un reconocimiento de lenguaje de signos basado en tensor. En *Actas de la conferencia internacional sobre multimedia y tecnología de la información, MMIT 2010*(págs. 192-195). <https://doi.org/10.1109/MMIT.2010.21>.
- Wang, Y., Yang, MZ y Luo, Y. (2015). Algoritmo SURF mejorado basado en distancia y ángulo optionales para el reconocimiento estático del lenguaje de señas. En *Actas de ingeniería de comunicación y seguridad de red, conferencia internacional sobre ingeniería de comunicación y seguridad de red, NSCE 2014*(págs. 265-270). Prensa CRC/ Balkema. <https://doi.org/10.1201/b18660-59>.
- Wilcox, S. y Occhino, C. (2016). Cambio histórico en las Lenguas de Signos. *Oxford mano - libroEn línea*.
- Wilson, AD y Bobick, AF (1999). Modelos paramétricos ocultos de Markov para gestos reconocimiento. *Transacciones IEEE sobre análisis de patrones e inteligencia artificial*. <https://doi.org/10.1109/34.790429>.
- Wong, SF y Cipolla, R. (2005). Reconocimiento adaptativo del movimiento de la mano en tiempo real ing un clasificador bayesiano disperso. *Apuntes de clase en informática*(pp. 170-179) (incluidas las notas de lectura de la subserie sobre inteligencia artificial y las notas de lectura sobre bioinformática), 3766 LNCS. [https://doi.org/10.1007/11573425\\_17](https://doi.org/10.1007/11573425_17). Xanthopoulos, P. y Razzaghi, T. (2014). Un método de máquina de vector de soporte ponderado para el reconocimiento de patrones de gráficos de control. *Informática e Ingeniería Industrial*, 70, 134-149. <https://doi.org/10.1016/j.cie.2014.01.014>.
- Xu, B., Huang, S. y Ye, Z (2021). Aplicación de la descomposición del tren tensorial en Modelo S2VT para reconocimiento de lenguaje de signos. *Acceso IEEE*, 9. <https://doi.org/10.1109/ACCESS.2021.3059660>.
- Xu, W., Li, Q., Feng, H., Xu, Z. y Chen, Y (2013). Umbralización de una nueva imagen estelar método para segmentación efectiva y estadísticas centroides. *óptica*, 124(20), 4673-4677. <https://doi.org/10.1016/j.jileo.2013.01.067>.
- Yang, Q. y Peng, JY (2014). Método de reconocimiento de lenguaje de señas chino basado en información de imagen de profundidad y SURF-BoW. *Moshi Shibie Yu Rengong Zhineng/ Reconocimiento de patrones e inteligencia artificial*, 27(8), 741-749.
- Yang, W., Tao, J. y Ye, Z (2016). Reconocimiento continuo de lenguaje de señas usando nivel edificio basado en el modelo oculto rápido de Markov. *Letras de reconocimiento de patrones*, 78, 28-35. <https://doi.org/10.1016/j.patrec.2016.03.030>.
- YanuTara, R., Insap Santosa, P. y Bharata Adji, T. (2012). Reconocimiento de lenguaje de señas en la teleoperación de robots utilizando descriptores de Fourier de distancia centroide. *Revista Internacional de Aplicaciones Informáticas*, 48(2), 8-12. <https://doi.org/10.5120/7318-0100>. Yao, Y. y Doretto, G. (2010). Impulso al aprendizaje por transferencia con múltiples fuentes. *En Actas de la conferencia de la sociedad informática IEEE sobre visión artificial y reconocimiento de patrones*(págs. 1855-1862). <https://doi.org/10.1109/CVPR.2010.5539857>.
- Yasir, F., Prasad, PWCC, Alsadoon, A. y Elchouemi, A. (2016). aplicación basada en SIFT Proach sobre el reconocimiento de la lenguaje de signos bengali. *En Actas del octavo taller internacional de IEEE sobre inteligencia computacional y aplicaciones*, IWCIA 2015 (págs. 35-39). <https://doi.org/10.1109/IWCIA.2015.7449458>.
- Yasir, R. y Khan, RA (2014). Reconocimiento de gestos con las dos manos para bengali lenguaje de señas usando LDA y ANN. *En Actas de la SKIMA, 8th international*

- congreso nacional sobre software, conocimiento, gestión de la información y aplicaciones.*<https://doi.org/10.1109/SKIMA.2014.7083527>.
- Yogamangalam, R. y Karthikeyan, B. (2013). Comparación de técnicas de segmentación en el procesamiento de imágenes.***Revista Internacional de Ingeniería y Tecnología*. Yu, C., Wang, X., Huang, H., Shen, J. y Wu, K. (2010). Gestos manuales basados en la visión Reconocimiento de turas usando características combinacionales. En*Actas de la sexta conferencia internacional sobre ocultación inteligente de información y procesamiento de señales multimedia, IIHMSP 2010*(págs. 543–546).<https://doi.org/10.1109/IIHMSP.2010.138>.
- Yuan, S., Wang, Y., Wang, X., Deng, H., Sun, S., Wang, H., Huang, P. y Li, G. (2020). Reconocimiento del alfabeto del lenguaje de señas chino basado en un algoritmo de bosque aleatorio. En*Actas del taller internacional IEEE sobre metrología para la industria 4.0 e IoT, MetroInd 4.0 e IoT 2020*(págs. 340–343).<https://doi.org/10.1109/MetroInd4.0IoT48571.2020.9138285>.
- Yusnita, L., Rosalina, R., Roestam, R. y Wahyu, RB (2017). Implementación de real-Reconocimiento estático de gestos de la mano en el tiempo mediante una red neuronal artificial. *Revista Commit (tecnologías de la información y la comunicación)*.<https://doi.org/10.21512/commit.v1i2.2282>.
- Zahedi, M., Dreuw, P., Rybach, D., Deselaers, T. y Ney, H. (2006). característica geométrica turas para mejorar el reconocimiento continuo del lenguaje de señas basado en la apariencia. En*Actas de la BMVC, conferencia británica sobre visión artificial*(págs. 1019–1028).<https://doi.org/10.5244/c.20.104>.
- Zaki, MM y Shaheen, SI (2011). Reconocimiento de lenguaje de señas usando una combinación de nuevas funciones basadas en la visión. *Letras de reconocimiento de patrones*, 3(4), 572–577. <https://doi.org/10.1016/j.patrec.2010.11.013>.
- Zapata, MM, Madrenas, J., Zapata, MM y Álvarez, J. (2019). Avances en arti-inteligencia oficial, ingeniería de software y sistemas (Vol. 787).<https://doi.org/10.1007/978-3-319-94229-2>.
- Zarit, BD, Super, BJ y Quek, FKH (1999). Comparación de cinco modelos de color en Clasificación de los pixeles de la piel. En*Actas del taller internacional sobre reconocimiento, análisis y seguimiento de rostros y gestos en sistemas en tiempo real, RATFG-RTS*.<https://doi.org/10.1109/RATFG.1999.799224>.
- Zhang, H., Wang, Y. y Deng, C. (2011). Aplicación de reconocimiento de gestos basada en la red neuronal BP de recodido simulado. En*Actas de la conferencia internacional sobre ingeniería electrónica y mecánica y tecnología de la información, EMEIT 2011*.<https://doi.org/10.1109/EMEIT.2011.6022891>.
- Zhang, LG, Chen, Y., Fang, G., Chen, X. y Gao, W. (2004). Un lenguaje de señas basado en la visión sistema de reconocimiento de calibre utilizando HMM de densidad de mezcla ligada. En*Actas de la ICMI'04, sexta conferencia internacional sobre interfaces multimodales*.<https://doi.org/10.1145/1027933.1027967>.
- Zhang, L. G, Chen, X., Wang, C., Chen, Y. y Gao, W. (2005). reconocimiento de signo subpalabras del idioma basadas en modelos de Markov ocultos potenciados. En*Actas de la séptima conferencia internacional sobre interfaces multimodales, ICMI'05*(págs. 282–287).<https://doi.org/10.1145/1088463.1088511>.
- Zhang, LG, Chen, X., Wang, C. y Gao, W. (2005). Clasificación de votación jerárquica. esquema de ción para mejorar el reconocimiento visual del lenguaje de señas. En*Actas de la 13.ª conferencia internacional ACM sobre multimedia, MM 2005*(págs. 339–342).<https://doi.org/10.1145/1101149.1101220>.
- Zhao, S., Zhang, S., Liu, J., Wang, H., Zhu, J., Li, D. y Zhao, R. (2021). Solicitud del aprendizaje automático en la acuicultura inteligente de peces: una revisión. *acuicultura*, 540, Artículo 736724.<https://doi.org/10.1016/j.aquaculture.2021.736724>.
- Zhao, W., Zhang, J., Li, P. y Li, Y. (2010). Estudio del algoritmo de segmentación de imágenes basado en características texturales y redes neuronales. En*Actas de la conferencia internacional sobre computación inteligente e informática cognitiva, ICICCI*.<https://doi.org/10.1109/ICICCI.2010.53>.
- Zheng, X., Lei, Q., Yao, R., Gong, Y. y Yin, Q (2018). Segmentación de imágenes basada en el algoritmo adaptativo de K-medias. *Revista Eurasip sobre Procesamiento de Imágenes y Video*.<https://doi.org/10.1186/s13640-018-0309-3>.