

DetECCIÓN DEL ABECEDARIO DE LENGUA DE SEÑAS MEXICANAS (LSM) USANDO MEDIAPIPE, SVM Y RANDOM FOREST

Detection of the Mexican Sign Language (LSM) alphabet using MediaPipe, SVM and Random Forest

Edgard Jose Maciel Cataño^a, Juan Carlos Estrada Gutiérrez^a,
Ariadna Berenice Flores Jiménez^a, Miroslava Cano Lara^b,
Maricela Jiménez Rodríguez^c

Resumen / Abstract

En la actualidad la mayoría de las personas cuenta con dispositivos electrónicos que les permiten acceder a Internet, investigar información, comunicarse con otros usuarios en el mundo, realizar compras, mantener la colaboración entre proveedores y clientes, etc.; esos dispositivos también se pueden aprovechar como herramienta para desarrollar o implementar software que ayude a eliminar la brecha de comunicación existente entre personas con discapacidad auditiva y las oyentes. Por tal razón, en esta investigación se propone un software que permite detectar los gestos del abecedario de la Lengua de Señas Mexicanas, utilizando la tecnología MediaPipe para analizar y tomar información de fotografías donde aparecen las letras, con la cual se obtiene un vector con las características de las imágenes y para la clasificación se realizaron pruebas con SVM y Random Forest donde se obtuvieron resultados entre 97.3 % y 99.4%.

Palabras clave: Lengua de Señas Mexicanas, MediaPipe, SVM, Random Forest.

a. Departamento de Ciencias Tecnológicas, Centro Universitario de la Ciénega, Universidad de Guadalajara.

b. Departamento de Ingeniería Mecatrónica, Tecnológico Nacional de México.

c. Departamento de Ciencias Básicas, Centro Universitario de la Ciénega, Universidad de Guadalajara.

Currently, most people have electronic devices that allow them to access the Internet, research information, communicate with other users in the world, make purchases, maintain collaboration between suppliers and customers, etc.; These devices can also be used as a tool to develop or implement software that helps eliminate the communication gap between people with hearing disabilities and hearing people. For this reason, in this research a software is proposed that allows detecting the gestures of the alphabet of the Mexican Sign Language, using MediaPipe technology to analyze and take information from photographs where the letters appear, with which a vector is obtained with the characteristics of the images and for classification tests were carried out with SVM and Random Forest where results between 97.3% and 99.4% were obtained.

Keywords: *Mexican Sign Language, MediaPipe, SVM, Random Forest.*

1. INTRODUCCIÓN

Hoy en día existen personas que tienen problemas auditivos, por alguna disminución en la capacidad de escuchar, la cual es generada muchas de las veces de forma congénita, heredada por alguno de sus familiares, en ocasiones también se puede adquirir durante el parto o por alguna enfermedad como: otitis media o meningitis [1]. Actualmente es indispensable que personas con diferentes capacidades puedan comunicarse sin problemas, por tal razón, es de suma importancia desarrollar herramientas que puedan eliminar la brecha de comunicación existente entre dichas personas y la sociedad en general. La lengua de señas es un lenguaje con gramática que permite la comunicación a través de movimientos con la mano, en México se reconoce la Lengua de Señas Mexicana (LSM) [2] como una lengua oficial nacional, fue desarrollada para incrementar la comunicación existente entre las personas con alguna discapacidad auditiva y las que no. La lengua de señas permite que personas sordas puedan comunicarse con otras utilizando un canal viso-gestual [3]. En muchas ocasiones las personas mudas también tienen problemas de audición [4]; los

alumnos sordos que tienen la oportunidad de ir a Centros de Atención Múltiple, no obtienen suficientes competencias en el español escrito, en comparación con los oyentes [5]; también les afecta que en muchas ocasiones los padres no cuentan con la preparación necesaria sobre la lengua de señas para poderse comunicar o enseñar a sus hijos de forma adecuada. Los sordos adultos presentan un nivel de lectura bajo, como un niño de cuarto grado de primaria, la mayoría de los niños sordos tienen los padres oyentes, quienes no dominan la lengua de señas, provocando que los niños aprendan fuera de su hogar [6]. Las personas que vivieron con familia de sordos no se ven como discapacitados, pero sin duda enfrentan barreras durante sus estudios, les cuesta trabajo encontrar intérpretes de Lengua de Señas en México, además los profesores no están capacitados [7]. Con base en la problemática existente de comunicación entre personas con discapacidad auditiva y los oyentes, es necesario implementar herramientas que ayuden en la comunicación. Se han desarrollado algunos dispositivos para el reconocimiento de la lengua de señas invasivos como guantes y no-invasivos que utilizan cámaras para posteriormente realizar pre-procesamiento con la finalidad de identificar los puntos de interés en este caso las manos. Cabe mencionar que la lectura de señas es difícil ya que presenta distintas posturas y configuraciones de la mano [8] así como la mayor capacidad de procesamiento de los dispositivos actuales, permiten el desarrollo de nuevas tecnologías capaces de detectar diferentes movimientos y actuar en tiempo real. A diferencia del reconocimiento de la voz hablada, que lleva más de 40 años de investigación, esta temática es relativamente nueva en el ambiente científico, y evoluciona de forma acelerada a medida que aparecen nuevos dispositivos así como nuevos algoritmos de visión por computador. La captura y reconocimiento de gestos dinámicos permite que sean utilizados en diversas áreas de aplicación como por ejemplo monitoreo de pacientes médicos, control en un entorno de videojuego, navegación y manipulación de entornos virtuales, traducción de léxicos de la lengua de señas, entre otras aplicaciones de interés. Particularmente la lengua de señas puede entenderse como un problema particular del reconocimiento de gestos dinámicos, el cual es sumamente apreciado en los últimos tiempos por distintas instituciones, ya que permite una ayuda directa a personas hipoacúsicas. Para poder utilizar un sistema de reconocimiento automático de lengua de señas para traducir los gestos de un intérprete, es necesario afrontar una serie de diversas tareas. En primer

lugar existen diferentes enfoques dependiendo el dispositivo de sensado a utilizar. Si bien existen dispositivos invasivos como guantes de datos, en esta Tesis se analizan sólo dispositivos no invasivos de dos tipos: las cámaras RGB convencionales, y las cámaras de profundidad (con particular interés en los nuevos dispositivos RGB-d. También Punsara et al. desarrollaron una aplicación móvil que se encarga de traducir la lengua de señas en voz digital y un guante portátil con tecnología IoT (Internet of Things) que permite reconocer dígitos del alfabeto y palabras en inglés [4]. En esta investigación se desarrolló un sistema que permite detectar los gestos del abecedario de la Lengua de Señas Mexicanas, utilizando como descriptor MediaPipe, además los clasificadores Random Forest y SVM.

El resto de este artículo está organizado de la siguiente manera: en la Sección 2 se explican los materiales y métodos que se utilizaron para desarrollar el sistema que aplica inteligencia artificial para detectar el abecedario de la Lengua de Señas Mexicana. En la Sección 3 se muestran las pruebas y resultados obtenidos; por último, en la Sección 4 se detallan las conclusiones.

MATERIALES Y MÉTODOS

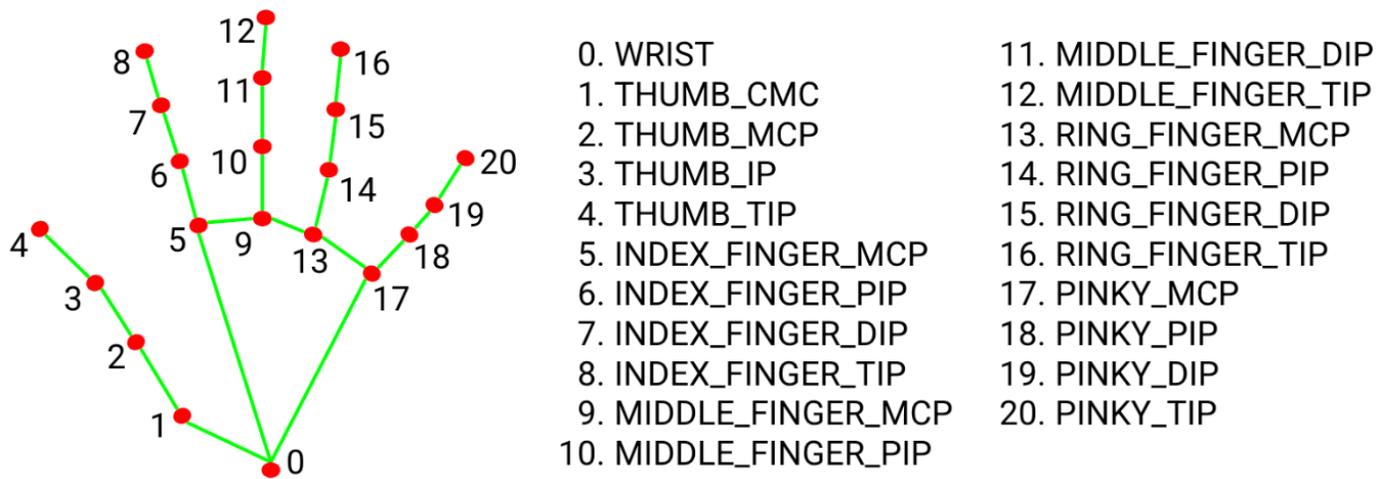
Se generó una base de datos propia, la cual contiene 5000 imágenes, 100 de cada letra o clase del abecedario con las 2 manos, omitiendo la ñ y z, se utilizó la seña final de las letras con movimiento; las fotos tienen formato BGR y se capturaron en ambientes controlados. Las fotografías se separaron dejando el 80% para la fase de aprendizaje y 20% para la fase de reconocimiento.

2.1 MediaPipe

Es una herramienta desarrollada por Google que implementa inteligencia artificial para identificar la lengua de señas, es un framework que da la posibilidad de procesar videos o audios; permite detectar puntos de interés tales como: caras, manos o alguna parte del cuerpo [9]. Para el desarrollo del sistema se utilizó la librería MediaPipe importando los descriptores de las manos “mp.solutions.hands.Hands”.

La Figura 1, muestra los 21 componentes que la librería Hand Landmark de MediaPipe detecta de la mano en cada fotografía y de los cuales entrega las coordenadas x, y de cada elemento.

Figura 1. Coordenadas generadas con MediaPipe hand [9].



2.2 SVM (Support Vector Machine)

Fue desarrollado inicialmente como un clasificador lineal que permite separar de un conjunto de datos $S = \{(x_1, y_1), \dots, (x_n, y_n)\}$, donde $x_i \in R^d$ and $y_i \in (-1, +1)$, x_1 corresponde al vector con las características, mientras que y_1 representa la clase. También se puede utilizar una función kernel para trabajar con más dimensiones, clases y mayor rapidez [10].

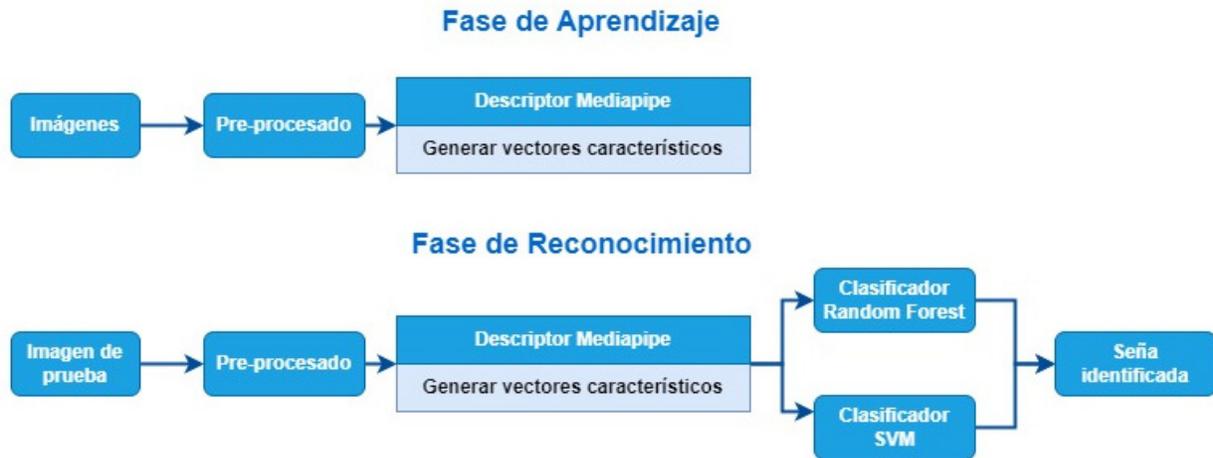
2.3 Clasificador Random Forest

Es un método de clasificación estadístico que contiene varios árboles de decisión, es decir, cuando el árbol tiene una entrada, esta se somete a varias pruebas en los diferentes nodos, hasta encontrar la respuesta [11] [12], donde un vector aleatorio de p -dimensiones $x = (x_1, \dots, x_p)^T$, representa las variables de entradas predictoras, mientras que Y corresponde a una respuesta real de una distribución $P_{x,y}(X, Y)$ [13]. El objetivo principal es encontrar una función de predicción $f(x)$ para detectar Y .

3. PRUEBAS Y RESULTADOS

En la Figura 2, se muestra el procedimiento general que se utilizó para la identificación de las señas.

Figura 2. Diagrama de la metodología.



Fase de aprendizaje: se tomaron como entradas las imágenes con formato BGR, las cuales posteriormente durante la etapa de pre-procesado se cambiaron como RGB y se utilizó como descriptor MediaPipe para detectar las señas y extraer la información del espectro de textura de cada una de las imágenes, con lo cual se generó su respectivo vector con las características.

Fase de Reconocimiento: se introdujeron las fotos para realizar la validación, después se realiza el pre-procesado, enseguida se implementa el descriptor MediaPipe para generar los vectores característicos y posteriormente se usaron los clasificadores Random Forest y SVM para determinar la clase o letra a la que corresponde cada una de las señas que aparecen en las fotografías.

3.1 Matriz Confusión

Con esta matriz se pueden mostrar los resultados de las pruebas realizadas para determinar el nivel de acertación de los clasificadores: Random Forest y SVM; el valor entero positivo en la diagonal principal indica los casos en los que se realizó de forma correcta la clasificación, mientras que fuera de la diagonal se indican el número de clasificaciones incorrectas.

Las matrices confusión de las Figuras 3 y 4, se generaron con una base de datos propia de fotografías con una sola mano, para el entrenamiento se usaron 80 fotografías y 20 para la validación o reconocimiento. En total se tomaron 2500 fotografías, 100 para cada seña o clase, de las cuales 2493 se procesaron de forma correcta y 7 no fueron detectadas por landmark, de cada foto se obtuvieron vectores de dos dimensiones con 21 coordenadas [x-xmin, y-ymin]. En las matrices confusión, algunos resultados de clasificación no presentan la acertación total de 20, debido a los 7 problemas en la detección de la mano, los cuales quedaron distribuidos en las imágenes utilizadas en la fase de aprendizaje y reconocimiento; también se debe a imágenes que se clasificaron incorrectamente.

Figura 3. Matriz confusión obtenida con el clasificador Random Forest, una mano.

		SEÑA PREDICHA																								
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
SEÑA REAL	A	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	18	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0
	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
	S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0
	U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	

Figura 4. Matriz confusión generada con el clasificador SVM, una mano.

		SEÑA PREDICHA																								
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
SEÑA REAL	A	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	18	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	12	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0
	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
	S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0
	U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0
	V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0
	W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0
	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0
	Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20

Las matrices confusión de las Figuras 5 y 6, resultaron al probar los clasificadores Random Forest y SVM, usando 25 clases con fotografías capturadas con las 2 manos, por lo tanto, cada clase cuenta con 200 fotos, es decir, 5000 imágenes totales; de las cuales 160 de las fotografías de cada clase se usaron para el entrenamiento y 40 para la validación. 15 imágenes de las 5000 no fueron detectadas con landmark, por lo tanto, en las matrices confusión de las Figuras 5 y 6, nuevamente algunos resultados de clasificación no presentan la acertación total de 40, ya que se tuvieron problemas en la detección de la mano y otras fueron clasificadas de forma incorrecta.

Figura 5. Matriz de confusión del clasificador Random Forest para ambas manos.

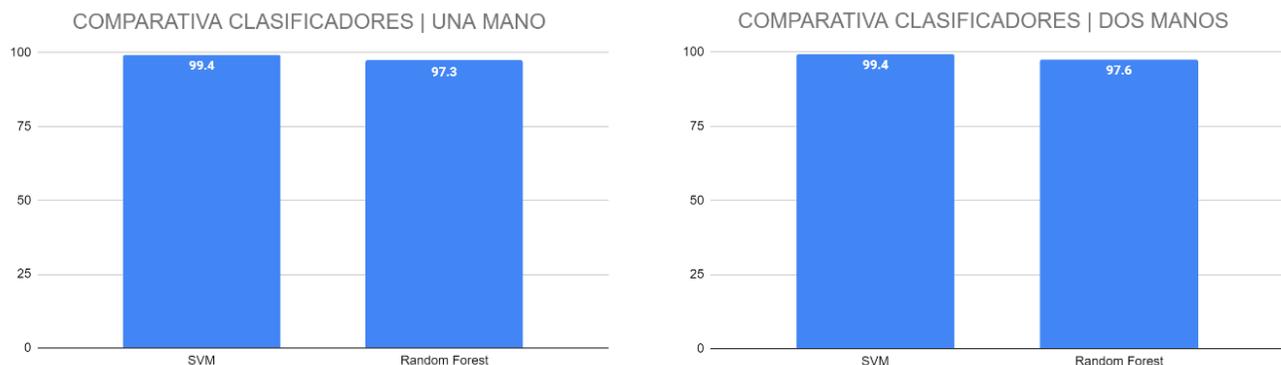
		SEÑA PREDICHA																										
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y		
SEÑA REAL	A	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	B	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	1	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	0	0	0	0	1	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0
	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0
	S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0
	U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0
	V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0
	W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0
	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0
	Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0

Figura 6. Matriz de confusión del clasificador SVM para ambas manos.

		SEÑA PREDICHA																										
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y		
SEÑA REAL	A	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	B	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
	D	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	E	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	F	0	0	0	0	3	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	G	0	0	0	0	0	0	39	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
	H	0	0	0	0	0	0	0	39	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
	I	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	J	0	0	0	0	0	0	0	0	0	26	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	
	K	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	L	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	M	0	0	0	0	0	0	1	0	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	
	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	
	S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	
	U	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	
	V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	
	W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	
	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	
	Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	

La gráfica de la Figura 7, muestra los resultados de acertación con los clasificadores Random Forest y SVM, con clases de una mano y dos.

Figura 7. Resultados de acertación con los clasificadores Random Forest y SVM con una y dos manos.



4. CONCLUSIONES

Con el sistema se logró realizar la detección de señas utilizando MediaPipe para obtener información sobre las características de las imágenes, y para clasificar se usaron Random Forest y SVM. Se tuvieron fallas al detectar las imágenes, ocasionadas por la inclinación de la mano, lo cual ocasionó que no se pudieran identificar los signos de forma correcta, afectando los resultados de las matrices confusión, ya que algunas de las imágenes erróneas se utilizaron en la fase de aprendizaje y otras en la de reconocimiento. Los clasificadores se utilizaron con imágenes de una y dos manos, logrando los mejores resultados de 99.4% de efectividad con SVM en una y dos manos, en comparación con Random Forest que alcanzó 97.3% y 97.6% de efectividad con imágenes de una y dos manos respectivamente.

5. REFERENCIAS

- [1] Gobierno de México, «¿Qué es la Discapacidad Auditiva?» Accedido: 13 de noviembre de 2023. [En línea]. Disponible en: <https://www.gob.mx/difnacional/articulos/que-es-la-discapacidad-auditiva>
- [2] C. E. Escobedo Delgado, C. A. Mercader Flores, L. Escobar Dellamary, M. del R. Ramírez Barba, M. Pool Westgaard, y M. Cruz Aldrete, *Diccionario de lengua de señas Mexicana*. Ciudad de México, 2017.
- [3] L. A. Tovar, «LA IMPORTANCIA DEL ESTUDIO DE LAS LENGUAS DE SEÑAS», 2001.
- [4] K. K. T. Punsara, H. H. R. C. Premachandra, A. W. A. D. Chanaka, R. V. Wijayawickrama, A. Nimsiri, y S. Rajitha De, «IoT Based Sign Language Recognition System», en *2020 2nd International Conference on Advancements in Computing (ICAC)*, Malabe, Sri Lanka: IEEE, dic. 2020, pp. 162-167. doi: 10.1109/ICAC51239.2020.9357267.
- [5] M. Cruz-Aldrete, «Towards the construction of a dictionary of Mexican», vol. 38, 2014.
- [6] V. Figueroa y M. R. Lissi, «LA LECTURA EN PERSONAS SORDAS: CONSIDERACIONES SOBRE EL ROL DEL PROCESAMIENTO FONOLÓGICO Y LA UTILIZACIÓN DEL LENGUAJE DE SEÑAS», *Estud. pedagóg.*, vol. 31, n.º 2, 2005, doi: 10.4067/S0718-07052005000200007.
- [7] J. Pérez-Castro y J. C. Cruz-Cruz, «Experiencias de Inclusión-Exclusión de un Grupo de Sordos Usuarios de la Lengua de Señas Mexicana», *Rev. latinoam. educ. inclusiva*, vol. 15, n.º 1, pp. 39-54, 2021, doi: 10.4067/S0718-73782021000100039.
- [8] F. Ronchetti, «Reconocimiento de gestos dinámicos y su aplicación al lenguaje de señas», Doctor en Ciencias Informáticas, Universidad Nacional de La Plata, 2017. doi: 10.35537/10915/59330.
- [9] A. S. Reig, «Interfaz hombre-máquina mediante estimación de pose».
- [10] E. J. C. Suárez, «Tutorial sobre Máquinas de Vectores Soporte (SVM)».
- [11] R. F. Medina-Merino y C. I. Ñique-Chacón, «Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python», *Interfases*, vol. 0, n.º 010, p.

165, dic. 2017, doi: 10.26439/interfases2017.n10.1775.

[12] G. Sun, «Cervical Cancer Diagnosis based on Random Forest», *IJPE*, 2017, doi: 10.23940/ijpe.17.04.p12.446457.

[13] C. Zhang y Y. Ma, Eds., *Ensemble Machine Learning: Methods and Applications*. New York, NY: Springer New York, 2012. doi: 10.1007/978-1-4419-9326-7.

—

EDGARD JOSÉ MACIEL CATAÑO. Departamento de Ciencias Tecnológicas, Centro Universitario de la Ciénega, Universidad de Guadalajara

JUAN CARLOS ESTRADA GUTIÉRREZ. Departamento de Ciencias Tecnológicas, Centro Universitario de la Ciénega, Universidad de Guadalajara

ARIADNA BERENICE FLORES JIMÉNEZ. Departamento de Ciencias Tecnológicas, Centro Universitario de la Ciénega, Universidad de Guadalajara

MIROSLAVA CANO LARA. Departamento de Ingeniería Mecatrónica, Tecnológico Nacional de México

MARICELA JIMÉNEZ RODRÍGUEZ. Departamento de Ciencias Básicas, Centro Universitario de la Ciénega, Universidad de Guadalajara