v. 4, n. 18, 2024

Journal of
Engineering
Research

DESIGN OF A SOLUTION FOR THE DEVELOPMENT OF AN APPLICATION THAT WILL INTERPRET THE RECOGNITION OF MEXICAN SIGN LANGUAGE THROUGH CONVOLUTIONAL NETWORKS

Gilberto Anduaga Márquez
International Institute of Aguascalientes: Rincón de Romos, Aguascalientes, Mexico, National Technological Institute of Mexico: Aguascalientes, Ags. Mexico
https://orcid.org/0009-0009-0875-7781

Abdel R. Dzul Bermejo
International Institute of Aguascalientes: Rincón de Romos, Aguascalientes, Mexico
https://orcid.org/0009-0003-1873-4109

Alma L. Esparza Maldonado
International Institute of Aguascalientes: Rincón de Romos, Aguascalientes, Mexico
https://orcid.org/0009-0007-3219-083X

Alfonso Recio Hernández
National Technological Institute of Mexico Aguascalientes, Aguascalientes, Mexico
https://orcid.org/0009-0006-7632-4601

Jorge Humberto Dzul Bermejo
National Technological Institute of Mexico Aguascalientes, Aguascalientes, Mexico
https://orcid.org/0009-0007-9010-7010

Abstract: This document addresses the problems faced by those people who have a hearing disability [1] in such a way that their main form of communication is through sign language, which for our work only considers Mexican and the interpretation of this by people who don't know him. The central objective of this document is to propose a design to generate an application for mobile devices and that, through the use of artificial intelligence using convolutional neural networks; be able to interpret personalized sign language transmitted by the hearing impaired person. Some previous works on the interpretation of sign language by other researchers are reflected, exposing its main characteristics and the differences with the one presented here. In the methodology part, the artificial intelligence techniques that can be applied to this topic are explained, justifying why convolutional neural networks are considered in this work through the Tensor Flow libraries. In conclusion, the design proposed in this article will allow the recognition of personalized Mexican Sign Language.
Keywords: Mexican Sign Language, artificial intelligence, neural network, Mobile device, Tensor Flow Lite.

BACKGROUND

The Convention on the Rights of Persons with Disabilities establishes that they are those who have long-term physical, mental, intellectual or sensory deficiencies, which when interacting with various barriers can prevent their full participation in society and on equal terms [1].

Consider now a Hearing Impaired Person [2] characterized by having significant audio loss, which can range from mild to profound, often using hearing aids such as hearing aids or cochlear implants. Relying heavily on visual and tactile stimuli, people with hearing disabilities can excel at lip reading, interpreting facial expressions and gestures. Many of them use sign language as their main means of communication, this being a visual and gestural way of expressing thoughts and emotions.

Sign language [3] plays a crucial role in the lives of people with hearing disabilities, providing an effective and natural means of communication [4]. Through facial expression, gestures and hand movements, sign language allows people with hearing impairments to communicate fully and richly, capturing emotional nuances and linguistic details. Furthermore, the use of sign language encourages inclusion and equality [5]. It allows people with hearing disabilities to actively participate in society, access information and relate fully to others. Sign language is not only a means of communication for people with hearing disabilities, but also a vital element of their identity and culture [6]. Its use is essential to build an inclusive and diverse society where each individual can express themselves and be understood, regardless of their hearing abilities.

In order for people with hearing disabilities who use sign language to be able to interact with other people in their environment, it is necessary that the latter be able to interpret sign language or, where appropriate, have some tool that allows them to communicate effectively using this sign language. language [7]. Artificial intelligence techniques can be very useful in this task, specifically those that allow images to be classified quickly and accurately.

On the other hand, the use of mobile devices among people with hearing disabilities has been transformative in terms of facilitating communication and accessibility [6]. Some ways mobile devices benefit these people are: Text messaging and instant messaging apps allow the hearing impaired to communicate effectively without relying on sound, video calling is essential for the hearing impaired as it gives them the ability to communicate through language of signs, allowing for richer and more expressive conversations, voice recognition technologies allow people with hearing disabilities to interact with their devices using voice commands, facilitating tasks such as sending messages, making calls and searching for information online, etc.

To this point, these technological advances have significantly improved the quality of life of deaf people, providing them with powerful tools for communication, access to information and full participation in the digital society [7]. The national survey on availability and use of information technologies in homes (ENDUTIH) 2021 [8] registered 91.7 million mobile phone users. The figure is equivalent to 78.3% of the study population in Mexico.

On the contrary, sign language interpretation [9] plays a vital role in creating an inclusive and accessible environment for people with hearing disabilities. By providing a bridge between the hearing impaired community and those who use oral language, a Mexican Sign Language interpreter using mobile devices will facilitate effective communication in a variety of contexts, from professional meetings to social events and medical services. This skill not only breaks down communication barriers, but also promotes equal opportunity by allowing people with hearing disabilities to fully participate in everyday life. Ultimately, the role of a sign language interpreter contributes to building a more inclusive society that is respectful of diversity.

As stated, communication is essential for human interaction, but people who use sign language face unique challenges. This visual and gestural system is crucial for their social participation, but lack of general knowledge, access to education and/or employment, and limited availability of interpreters create obstacles. Furthermore, lack of access to information, associated costs and social exclusion are significant problems [10]. The inclusion of people with hearing disabilities [5] requires concrete measures, including the adoption of technology, such as mobile applications, and respect for the rights and use of sign language to eliminate communication barriers. Society must work towards accessible and equitable communication for all.

SOLUTION DESIGN

This article proposes the design of a Mexican sign language interpreter, which will be implemented through convolutional neural networks, and trained with user data, to be personalized to their specific form of expression according to their movements, through of your mobile device. To achieve this objective, it must be considered that, although current mobile devices have acceptable computing capacity, this is still generally lower than what desktop computing equipment can perform. Based on this premise, it is necessary to use techniques that allow its execution on mobile devices limited in calculation or processing capacity, therefore, we will have to focus on the use of Tensor Flow Lite.

Next, the developments and research related or similar to the proposed solution are presented, the Artificial Intelligence techniques that offer the best results to achieve a high percentage of effectiveness will be covered, such as Machine Learning techniques, the use of already developed libraries that implement these technologies and will facilitate the development of the application; These libraries are Tensor Flow Lite and MediaPipe. It will also be discussed how the design and training of the neural network is carried out, as well as the expected results.

INVESTIGATION OF THE STATE OF THE ART, AS WELL AS RESULTS OF SIMILAR DEVELOPMENTS

Analyzing the advancement of technology, it has been seen that there are computer applications that in some cases require a Leap Motion or the Microsoft Kinect or online applications (WEB) have also been developed that perform a sign language translation. However, these are not aimed at the objectives set out in this document since some only serve to teach the language, others were only in the measurement stage or require specific hardware that is not easy to transport. Some examples of these applications are: Classification of Mexican Sign Language with SVM generating artificial data, Recognition of words from Mexican Sign Language using RGB-D information, Android App for Learning Mexican Sign Language, Language Recognition System of Mexican Signs based on an RGB-D camera and machine learning, etc.

In comparison with the exposed works, our research proposal proposes that the majority of people use commonly used mobile devices, taking advantage of the hardware they have. The idea when using this type of device is that the interpretation is in real time without the need for special or additional hardware to what a person normally carries. The main difference that must be highlighted is adding personalized data to the training of the convolutional neural network [18] that increases the percentage of effectiveness of the application for each user. Based on the above, it is considered that artificial intelligence techniques are the best tool to achieve the objective, since they can be implemented with generic data and later personalized by adding data generated by the disabled person using the system. The tools that will be used are mentioned below:

CREATION, TRAINING AND EVALUATION OF CONVOLUTIONAL NEURAL NETWORK FOR GESTURE CLASSIFICATION

Image recognition is an artificial intelligence technique that uses complex mathematical algorithms to analyze and understand images. Image recognition tools can recognize, analyze and interpret images, and are more efficient than the human eye. These tools can search countless images quickly and return very accurate results. Among the Artificial Intelligence tools that could be adopted for this project, **Machine Learning** (ML) [19] must be considered, which is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without being explicitly programmed. These algorithms are used for various purposes such as data mining, image processing, predictive analytics, etc. to name a few. The main advantage of using Machine Learning is that once an algorithm learns what to do with the data, it can do its job automatically.

There are different Machine Learning implementations ready to be used in different types of projects. These implementations or libraries are constantly updated by various developers distributed throughout the world, so it is not necessary to start from scratch and, if necessary, use one of them; In this project we will focus on the following:

Tensor Flow which is an open source framework developed by Google that was first released in November 2015 [20]. However, its internal development began in 2011 as a research project in the Google Brain team. Tensor Flow was designed to make it easy to create and train deep neural networks and other machine learning models in an efficient and scalable way. Tensor Flow's development

Application	Characteristics
Classification of Mexican Sign Language with SVM generating artificial data [11]	It allows you to identify Mexican Sign Language with high precision by introducing very representative data generated artificially, which improves the generalization capacity of the classifier.
Recognition of Mexican Sign Language words using RGB-D information [12]	Construction of a system that can translate words from the Mexican Sign Language by recognizing them from the 3D trajectory of the movement of the signers' hands using a Kinect sensor.
Impact of the system for the teaching and translation of the Mexican sign language UAEMex in public institutions [13]	The objective is to test and measure the impact of the UAEMex Mexican Sign Language Teaching and Translation System software by creating artificial intelligence to facilitate communication in people with hearing or speech disabilities.
Android App for Learning Mexican Sign Language [14]	As its name indicates, the learning of Mexican sign language, in which many people or the vast majority do not know if it exists or have ever heard of it, therefore, through this tool, the aim is for young people to know and Learn this language, to be able to start a conversation with someone who suffers from this disability.
Mexican Sign Language recognition system based on an RGB-D camera and machine learning [15]	Sign language recognition based on the detection of characteristic points on the hands, body and face that are used when signing. To acquire the signs, a depth camera was used to obtain the 3D coordinates characterized by each sign.
Mexican Sign Language to Text Translator [16]	It proposes a software tool capable of recognizing static gestures of Mexican sign language, through the implementation of a neural network and the use of the leap motion sensor as a data entry device.
Vowel recognition system of Mexican Sign Language [17]	Interpret the vowels of Mexican Sign Language identified by static hand gestures.

Table 1: List of previous research on the interpretation of the LSM.

philosophy is based on several key principles: flexibility, scalability, portability, open source community, and extensibility

Tensor Flow has had a significant impact on the field of machine learning [20], artificial intelligence, image classification and has contributed to the advancement of computer vision in general.

The training of neural networks requires a lot of processing capacity, storage and time, so it is usually carried out on computing equipment with specialized hardware such as servers or desktop computers with enough processing cores (CPU and GPU)[23], main memory and secondary storage that allow rapid data transfer. This article has considered the use of mobile devices that do not have these characteristics, so the work has to be divided into training and implementation through the Tensor Flow Lite library.

Tensor Flow Lite [21] is a tool for running machine learning models on embedded, mobile, or IoT devices. Its features include optimization for offline execution, cross-platform support, support for multiple languages, high performance with hardware acceleration, and examples of common machine learning tasks. The workflow involves generating a model in Tensor Flow Lite format, optionally with metadata, and then running inference on the device. Specific APIs can be used depending on the model type and performance can be improved with hardware acceleration on iOS and Android devices. However, not all Tensor Flow models are compatible with Tensor Flow Lite and on-device training is not supported. For the rest, MediaPipe is a tool that uses Tensor Flow [22], it is a neural network already trained for the detection of gestures made with the hands and that will further facilitate the development of this project.

The MediaPipe gesture detector [22] allows the recognition of hand gestures in real time and provides the results of the recognized gestures along with the reference points of the detected hands. This library can be used to recognize specific hand gestures of a user and invoke functions of an application that correspond to these.

This process operates on image data with an ML model and accepts static data or a continuous stream. The library generates hand landmarks in image coordinates, hand landmarks in world coordinates, dexterity (left/right hand), and multi-hand gesture categories. The gesture detector uses a model package with two pre-packaged model packages: a hand landmark model package and a gesture classification model package. The reference model detects the presence of hands and their geometry; and the gesture recognition model recognizes gestures based on the geometry of the hands.

The hand landmark model package detects the 21-coordinate keypoint location of the hand knuckles within the detected hand regions. The model was trained with approximately 30,000 real-world images, as well as several rendered synthetic hand models imposed on various backgrounds. See the definition of the 21 benchmarks below:

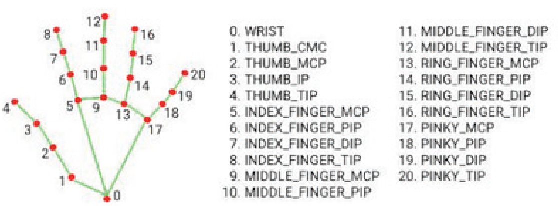

Image 1 [22]. Landmarks on the hands, detected by MediaPipe.

Note: Adapted from Google AI for Developers [Imagen], 2024, Google (https://ai.google.dev/edge/mediapipe/solutions/vision/gesture_recognizer?hl=es-419#configurations_options.). CC BY 4.0

The hand landmark models package contains two models one for palm detection and the second for hand landmark detection. The palm detection model localizes the hand region from the entire input image, and the second detects the landmarks on the cropped hand image defined by the palm detection model.

The gesture classification model package contains a two-step neural network with a gesture embedding model followed by a gesture classification model. The gesture embedding model encodes image features into a feature vector, and the classification model is a lightweight gesture classifier that takes the feature vector as input.

Once the tools to be used have been defined, only the core part remains to be carried out, which is the training of the gestures corresponding to Mexican sign language. For this, the following process was followed:

1. Data collection: A set of images containing the gestures to be recognized was collected.
2. Data preprocessing: Before feeding the images to the neural network, it was necessary to perform some transformations, such as resizing them to a specific size, normalizing the pixel values, or applying data augmentation techniques to increase the variability of the data set.
3. Division of the data set: Three parts were generated: training, validation and testing.
4. Model training
5. Model validation: After each training epoch, the performance of the model was validated using the validation set.
6. Model evaluation: The final performance of the model on the test set was evaluated to obtain an objective estimate of its ability to generalize.

DESIGN CREATION

Taking into account that most of the tools necessary for the development of the application have already been developed, the central objective will be how to use or, where appropriate, adapt these tools in the Mexican sign language interpreter. For this, the steps to follow are first outlined in a flow chart:

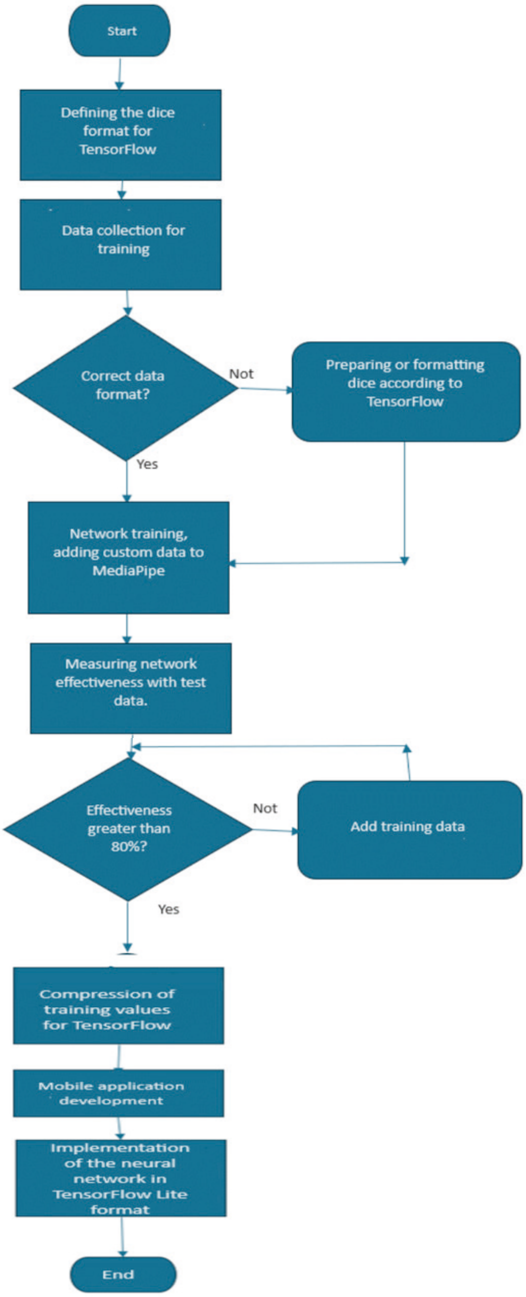

Image 2. Design of the LSM interpreter.

Designing a neural network using Tensor Flow Lite involves following a few steps:

I)Define the neural network model, that is, its architecture, including the number and type of layers (convolutional, recurrent, fully connected, etc.), and the activation functions.

II) Train the model using custom data to adjust the weights and biases of the neural network.

III) Optimize and convert the model to Tensor Flow Lite Once the model is trained, it must be optimized and converted to a format compatible with Tensor Flow Lite. This involves removing any unsupported operations and quantifying the weights to reduce model size and improve execution efficiency on mobile devices.

IV) Integrate the model into a mobile application using the Tensor Flow Lite API. This will allow inferences to be run using the model directly on mobile devices.

V) Test and debug the application.

RESULTS

The design will allow the development of an application that interprets sign language in real time through the use of a mobile device which would have several significant contributions in the area of computer science and in the lives of people with hearing disabilities:

Improved Accessibility: This app will increase accessibility for the hearing impaired by allowing them to communicate more easily with people who do not know sign language. In addition, it will expand their participation in society and improve their opportunities for education, employment and social life.

Image recognition technology: The development of this application by using advanced image recognition and signal processing algorithms. This will boost research and development in the field of computer vision and machine learning to improve the accuracy and speed of gesture recognition.

Integration of emerging technologies: By taking advantage of emerging technologies such as artificial intelligence, natural language processing and augmented reality, the application will be able to improve its functionality and user experience. Additionally, it will stimulate innovation in the integration of multiple technologies to address complex problems effectively.

Sensitization and awareness: The development and dissemination of this application will increase awareness of the challenges that hearing impaired people face in daily communication. Thus promoting greater support and resources for the research and development of personalized assistive technologies for people with disabilities.

CONCLUSIONS

This article has explained a revolutionary tool that has the potential to transform communication for people with hearing disabilities: the ability of mobile devices, such as cell phones, to interpret Mexican sign language. Through the combination of gesture recognition and image processing technologies, a new world of possibilities opens up for inclusion and accessibility.

Through the development of advanced gesture recognition and image processing algorithms, artificial intelligence is providing a new avenue for the inclusion and accessibility of people with hearing disabilities.

By harnessing the power of artificial intelligence, we break down language barriers and provide people with hearing disabilities with a tool that allows them to communicate more easily in environments where they may have previously experienced difficulties. This technology not only facilitates communication between people, but also opens opportunities for more active and meaningful participation in society.

Additionally, it is essential to continue working to improve and refine these systems, ensuring they are accurate, reliable, and accessible to everyone. This will require continued collaboration between developers, sign linguists, people with hearing impairments, and other experts in the field.

Thanks. We thank Dr. Alma Laura Esparza Maldonado for her support and direction in writing this article.

REFERENCES

1. United Nations. (febrero 2007). Convención sobre los Derechos de las Personas con Discapacidad Recuperado de: https://www.un.org/esa/socdev/enable/documents/tccconvs.pdf

2. García L. (17 de septiembre de 2022), Inclusión desde la lengua de señas [articulo], https://www.informador.mx/cultura/Derechos-Humanos-Inclusion-desde-la-lengua-de-senas--20220917-0038.html

3. Arellano A. (diciembre de 2021). ¿Qué es el lenguaje de señas?, [articulo]. https://www.ucv.edu.mx/uncategorized/que-es-el-lenguaje-de-senas/

4. Rodríguez, M. I., & del Pilar Velásquez, R. (2000). Historia y gramática de la lengua de señas. Pedagogía y saberes, (14), 91-104. Recuperado de: https://revistas.pedagogica.edu.co/index.php/PYS/article/view/6242/5695

5. De Diputados, C. (2012). Constitución política de los estados unidos mexicanos. México: Cámara de diputados. Recuperado de: https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf

6. Patiño O., Patiño J., Fernandez A., Jimenez A., (25 de agosto de2020), Tendencias investigativas en el estudio de tecnologías inclusivas para población sorda, recuperado de, https://www.redalyc.org/journal/1942/194264514019/html/

7. Ruíz Villa, A. (2021). El lenguaje de señas en un mundo globalizado. IJD: Blog Digital Universitario, 1-12. Chihuahua: IJD, Educación Superior / Universidad Autónoma de Chihuahua. Recuperado de https://edu.ijd.org.mx/data/files/El-lenguaje-de-se-as-en-un-mundo-globalizado_Alejandra-Ruiz-Villa_VBLOG_f_3.pdf

8. INEGI, (4 de julio de 2022), COMUNICADO DE PRENSA NÚM. 350/22 ENCUESTA NACIONAL SOBRE DISPONIBILIDAD Y USO DE TECNOLOGÍAS DE LA INFORMACIÓN EN LOS HOGARES (ENDUTIH) 2021, (PDF recuperado de), https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2022/OtrTemEcon/ENDUTIH_21.pdf

9. ONU, Convención sobre los Derechos de las Personas con Discapacidad y Protocolo Facultativo, [PDF recuperado] https://www.un.org/disabilities/documents/convention/convoptprot-s.pdf

10. Aranda, B. E. (2008). La vulneración de los derechos humanos de las personas Sordas en México. Derechos Humanos Mexico. Revista del Centro Nacional de Derechos Humanos, 3, 105-127. Recueprado de http://historico.juridicas.unam.mx/publica/librev/rev/derhumex/cont/8/art/art5.pdf

11. Cervantes, J., Lamont, F. G., Santiago, J. H., Cabrera, J. E., y Trueba, A. (2013). CLASIFICACIÓN DEL LENGUAJE DE SEÑAS MEXICANO CON SVM GENERANDO DATOS ARTIFICIALES. Revista Vínculos, 10(1), 328–341. https://doi.org/10.14483/2322939X.4684

12. Trujillo-Romero, F., & García Bautista, G. (2021). Reconocimiento de palabras de la Lengua de Señas Mexicana utilizando información RGB-D. ReCIBE, Revista electrónica De Computación, Informática, Biomédica Y Electrónica, 10(2), C2–23. https://doi.org/10.32870/recibe.v10i2.209

13. Rojas Hernández, R., González Jaimes, E. I., Trujillo Mora, V., López Chau, A., & González Morán, C. O. (2023). Impacto del sistema para la enseñanza y traducción de la lengua de señas mexicana UAEMex en instituciones públicas. Ciencia Latina Revista Científica Multidisciplinar, 7(1), 822-838. https://doi.org/10.37811/cl_rcm.v7i1.4434

14. Jiménez García, Erick D, (29 de marzo de 2022), App Android para el Aprendizaje del lenguaje de señas mexicano https://rinacional.tecnm.mx/jspui/handle/TecNM/4527

15. Mejía K., (1 de septiembre de 2022), Sistema de reconocimiento del Lenguaje de Señas Mexicano basado en una cámara RGB-D y aprendizaje automático, https://ri-ng.uaq.mx/handle/123456789/3799

16. Díaz G, (diciembre de 2016), Traductor del lenguaje de señas mexicano a texto, https://repositorio.cinvestav.mx/bitstream/handle/cinvestav/2314/SSIT0014424.pdf?sequence=1

17. Cuecuecha E., Martínez J., Méndez D., Zambrano A., Barreto A., Bautista V., Ayala S. (febrero de 2018), Sistema de reconocimiento de vocales de la lengua de señas mexicana, https://pistaseducativas.celaya.tecnm.mx/index.php/pistas/article/view/1152

18. García P. (2013), reconocimiento de imágenes utilizando redes neuronales artificiales, recuperado de, https://d1wqtxts1xzle7.cloudfront.net/61283296/ProyectoFinMasterPedroPablo20191120-128137-13v2s7i-libre.pdf?1574303392=&response-content-disposition=inline%3B+filename%3DRECONOCIMIENTO_DE_IMAGENES_UTILIZANDO_RE.pdf&Expires=1699386338&Signature=a4j1T2QQD3mmac5tnoKnVkJRBAFuTa7Ayt~QsTgz37bNGAKk4hK0XEBeofZXUNJdtTYZAa763Y-OKzZWPJ0cChP9IDKJgmwzE5U5Ed4-tr1fPFgj3Y3J6NoWjnq-Sr9tZR94cewT1-4tleMqa1zDJMkJu5prosjas7oCDVslj1arwnBAlXxoO~sUe-C1v8EwJy8OguqPlZ5NLQ-t-~Lkas2f8QR20CTr63hcfg35Q-X7MbPh9D2-192kLkn4iSoVksJCQLrgtFCu4ROkV9sk7Lyy3d0IbZhXJE39lJ5fJ3XDcW6dStis3qdpV81nV5vHEaM7e6L8J7JgWTm4i2yFCg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

19. Artola A. (2019), Clasificación de imágenes usando redes neuronales convolucionales en Python, recuperado de, https://idus.us.es/bitstream/handle/11441/89506/TFG-2402-ARTOLA.pdf?sequence=1&isAllowed=y

20. Haykin S. (2005). «Neural Networks - A Comprehensive Foundation». Editorial: Pearson.

21. Larranaga P., Inza I, Moujahid A. «Tema 8. Redes Neuronales» http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t8neuronales.pdf. Tensor Flow, (2023), Por qué Tensor Flow, recuperado de, https://www.Tensor Flow.org/about?hl=es-419

22. Tensor flow (2021) Tensor Flow Lite. Recuperado de https://www.Tensor Flow.org/lite/guide?hl=es-419

23. Google for Developers (02/noviembre/2023), Gesture recognition task guide, https://developers.google.com/mediapipe/solutions/vision/gesture_recognizer#configurations_options.

24. Soto-Orozco O., Corral-Saenz A., Rojo-Gonzalez C., Ramirez-Quintana J. (13/noviembre/2019), Análisis del desempeño de redes neuronales profundas para segmentación semántica en hardware limitado, Recuperado de: https://www.redalyc.org/journal/5122/512261374010/html/

25. Dukhan M., Barchard F. (29/noviembre/2023), Half-precision Inference Doubles On-Device Inference Performance, recuperado de: https://blog.Tensor Flow.org/2023/11/half-precision-inference-doubles-on-device-inference-performance.html?hl=es-419