Indonesian Sign Language (BISINDO) As Means to Visualize Basic Graphic Shapes Using Teachable Machine

Arya Harditya[1]

[1] *Sampoerna University*
Email:arya.harditya@sampoernauniversity.ac.id

ABSTRACT
In the contemporary digital era, it is arguable that the term Universal Design has rarely been discussed. Possibly, this is caused by the emergence of a new normal amongst the common society, the change of cultural communication, and behaviours. The shift of focus to technological development alone that hides a more imperative purpose such as integrating technology with universally designed objects. Hence, it broadened its accessibility to people with disabilities, especially deaf or people with muteness. Deaf people communicate differently, they are generally using hand movements to practice Sign Language, which this article is particularly highlighting IndoSL or BISINDO. Similarly, the most current technological discussion is image recognition by using a system called Machine Learning that can be considered as an interaction design's field of interest. In this article, I would like to take the opportunity to critically explore and integrate Sign Language hand tracking with Machine Learning System by using open-source JavaScript libraries, especially in terms of visualizing and/or generating basic graphic shapes.

Keywords: *Digital Media, Universal Design, Artificial Intelligence, Sign Language, Interaction Design.*

1. INTRODUCTION

According to Hurlbut, in 2009, there are more than 130 signed languages in the world. There are 2 signed languages listed in Indonesia, BISINDO (Bahasa Isyarat Indonesia) and SIBI (Sistem Isyarat Bahasa Indonesia). Hurlbut also pointed out that in 1987, approximately one out of every thousand children in developed countries are born deaf. Additionally, he also cited that more than 90% of deaf children were born from families with both parents are hearing. These children have no knowledge to learn to speak nor sign language. They simply do not learn to speak because they cannot hear [1].

Based on the above statistical facts, this article will aim at extending the functionality of signed language as a medium of digital interactivity but not limited to increasing people's interest in learning signed language through interactive media.

According to Hurlbut, many who learn American Sign Language (ASL) and works on projects in Indonesia have influenced the Deaf in Indonesia. In the Indonesian Sign Language (IndoSL) Dictionary (Kamus Sistem Isyarat Bahasa Indonesia, 2001) 50% of the wordlist items were derived from ASL. Generally, ASL has influenced IndoSL (also known as BISINDO), specifically for colour terms. In ASL several colours are initialized using English initial as well as hand movements. For example, "blue" in English is "biru" in the Indonesian Language, which resulted in false similarity because the initial consonant was "b". However, there is no semblance between ASL "b" and IndoSL "b" apart from the same articulation place [1].

Arguably, there has never been an issue for deaf people to access technology because it is a communication disability, not a physical challenge. Thus, the article would like to broaden the functionality of signed language to the level where it is a universally designed Human-Computer Interface (HCI), specifically categorizes as Machine Learning.

Universal Design is the design of a space and/or environment that can be accessed, understood, and used to the extent where all people of age, size, ability, or disability [2]. Upon developing the prototype, the article will apply the 7 principles of Universal Design: Equitable Use, Flexibility in Use, Simple and Intuitive Use, Perceptible Information, Tolerance for Error, Low

Physical Effort, and Size and Space for Approach and Use. According to the National Disability Authority, these principles were developed by a group of architects, product designers, engineers, and environmental design researchers at North Carolina State University in 1997. With the purpose to guide environment, product, and communication designs. The outcome of this article will primarily focus on the product (digital media technology) and communication aspects [2].

Concerning product and communication design, the article also includes Machine Learning as part of a developing Universal Design. Machine Learning is an artificial intelligence form that uses algorithms to allow a system to understand from rather than predetermined programming [3]. Machine Learning requires the correct dataset to be analysed, formatted, and accustomed to assembling a machine learning model. In the article's prototype, the data would be a signed language dictionary that involves visualization terms such as shapes and colours that eventually result in Artificial Intelligence arts or AI arts. According to Manovich, AI arts refer to humans programming computers to create with a significant degree of autonomy new artefacts or experiences that professional members of the art world recognize as belonging to "contemporary art." [4].

2. INDONESIAN SIGNED LANGUAGES (INDOSL)

As a digital citizen, we are not communicating the same language and it is not the case anymore when technology provides translation software with real-time speech and image recognition to obliterate the complex cultural boundary. It is not completely solving the problem, rather it aids users with a small fraction of vast multilinguistic knowledge. Critically, we must also aware that common literacy was not the only issue that required technological adaptation, but unique literacy such as sign language is also needed to be considered.

According to Hurlbut, many who learn American Sign Language (ASL) and works on projects in Indonesia have influenced the Deaf in Indonesia. In the Indonesian Sign Language (IndoSL) Dictionary (Kamus Sistem Isyarat Bahasa Indonesia, 2001) 50% of the wordlist items were derived from ASL. Generally, ASL has influenced IndoSL (also known as BISINDO), especially for colour terms. There are two types of Indonesian Sign Language: BISINDO and SIBI. The article chose BISINDO because it is the official Indonesian sign language agreed by a foundation called GERKATIN (Gerakan untuk Kesejahteraan Tuna Rungu Indonesia) or Movement for the Well-Being of Deaf Indonesians in English.

Figure 1 IndoSL (BISINDO) and ASL comparison [11].

In ASL many of the colours are initialized using English initials for several colours as well as have the same movement. For example, the English word "blue" is "biru" in the Indonesian Language, which resulted in a false similarity because the initial consonant was "b". However, there is no resemblance between ASL "b" and IndoSL "b" apart from the same place of articulation [1].

Theoretically, if sign languages adopted digital technology, we must well aware that there will be a new emergence of universal design notion and development, which is a new function of BISINDO in language education. The next utopian future when technology transforms communication, not only in a form of Google Translate but also an automated translator for the deaf and the mute. Continuously, the normal common knowledge needs to be upgraded, which consists of literacy, code literacy, and sign language literacy.

3. UNIVERSALLY DESIGNED HUMAN-COMPUTER INTERFACE (HCI)

Arguably, there has never been an issue for deaf people to access technology because it is a communication disability, not a physical challenge. Thus, the article would like to broaden the functionality of signed language to the level where it is a universally designed Human-Computer Interface (HCI), specifically categorizes as Teachable Machine.

Universal Design is the design of a space and/or environment that can be accessed, understood, and used to the extent where all people of age, size, ability, or disability [2]. Upon developing the prototype, the article will apply the 7 principles of Universal Design: Equitable Use, Flexibility in Use, Simple and Intuitive Use, Perceptible Information, Tolerance for Error, Low Physical Effort, and Size and Space for Approach and Use. These principles were developed by a group of architects, product designers, engineers, and environmental design researchers at North Carolina State University in 1997.

The best early development of using hand-tracking as a sign language translator is as shown below. Although the example shows an offline practice of machine learning using Mediapipe.

Figure 2 Sign language recognition with RNN and Mediapipe hand tracking DEMO [10]

It is time to realize that new technology is universally available, this includes the freedom of expression through arts. The outcome of this article will primarily explore and strictly focus on the product (digital media technology) and communication aspects.

4. ARTIFICIAL INTELLIGENCE (AI) ARTS

Machine Learning is an artificial intelligence form that uses algorithms to allow a system to understand from rather than predetermined programming [3]. Machine Learning requires the correct dataset to be analysed, formatted, and accustomed to assembling a machine learning model. In the article's hypothesis, the data is a sign language dictionary that involves visualization terms such as shapes and colours that eventually result in Artificial Intelligence arts or AI arts. AI arts refer to humans programming computers to create with a significant degree of autonomy new artefacts or experiences that professional members of the art world recognize as belonging to "contemporary art". One of the most popular methods for using computers in the arts and design is writing computer programs that generate objects in various media (text, image, video, 3D shapes, graphic designs, logos, urban plans, music, etc.) Such programs can take a variety of forms - simple instructions to draw a sequence of shapes, algorithms that generate fractals, cellular automata algorithms, genetic algorithms (Karl Sims), and so on [4].

AI can recognize moving objects in real-time, a similar feature is implemented in some mobile devices in-app cameras that enable a user to activate the shutter button by using a hand gesture. Hypothetically, a sign language that involves many interesting gestures and movements as the means to communicate, can also be applied to activate software features as trigger functions as AI arts described by Manovich.

5. VISUAL RECOGNITION TECHNOLOGY

Contemporarily, Deep Learning has fuelled incredible development in the computer vision field, as neural networks constantly advancing the boundary of visual recognition technology. Most technologies such as object, landmark, logo, and text recognition are prepared for internet devices through Cloud Vision API, which Google believed that the rapidly growing mobile devices with computational power enable these technologies to be delivered into the hands of its users regardless of the time and space [5].

Figure 3 Example from Google AI including detection, fine-grain classification, attributes and geo-localization [6].

5.1. MobileNets

Google AI released MobileNets, a mobile computer vision models built for TensorFlow, which can be assembled for classification, detection, embeddings, and segmentation comparable to the common large-scale models. TensorFlow is an open-source platform for machine learning (ML). The tools have comprehensive, flexible ecosystem, modern libraries that lets researchers push the boundary of ML and ease developers building and deploying ML-driven applications. As seen in Figure 1, MobileNets being utilized on Google Doodles to create machine learning-based interactive games, hypothetically we can create a new output where sign

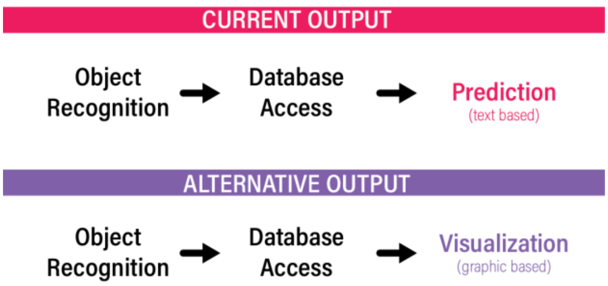

Figure 4 Workflow of Open Source JavaScript libraries that will be used in the visualization program. P5.js will be the main sketch to utilize ml5.js and TensorFlow.js libraries.

language is recognized by the ML, instead of giving text-based prediction and accuracy, alternatively, it draws graphic-based output on the screen. This section will provide proof of results that the hypothesis can be realized into further prototyping.

As suggested by Daniel Shiffman in his Machine Learning in JavaScript tutorial, p5.js will be the main sketch or host for the supplementary libraries such as ml5.js and TensorFlow.js [7]. The JavaScript libraries are executed in a code editor called Atom, it is used in this research because it has a built-in local server for instant live preview to how it operates online.

Figure 5 Workflow of Open Source JavaScript libraries that will be used in the visualization program. P5.js will be the main sketch to utilize ml5.js and TensorFlow.js libraries.

As seen in Figure 6, you see an image of a puffin bird predicted by MobileNets, which the result can be seen on the Developer Tools of Chrome browser. The result of the prediction consists of an array list and label or class name of the predicted image.

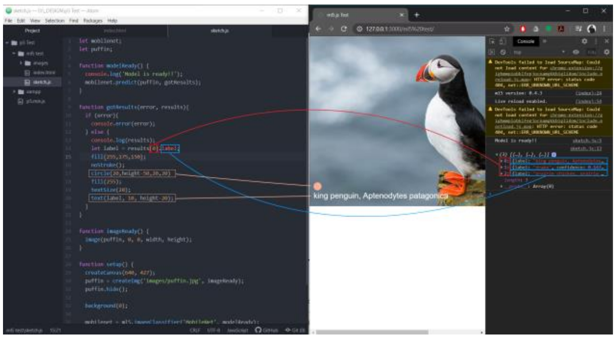

Figure 6 Proof of the result of image prediction using ml5.js by accessing data in MobileNets.

Furthermore, the predicted image can trigger variables programmed in p5.js, which in Figure 4 is exemplified to trigger a circle and the name of the array label on the screen.

5.2. PoseNet

As a static image being recognized by MobileNets, the next step of proving our hypothesis is to get it recognized objects in real-time, in other words, moving images such as using a camera to recognized objects. The visual recognition testing is using one of machine learning pre-trained models called PoseNet. It is a machine learning model that lets Real-time Human Pose Estimation, it also can be utilized for a single pose or multiple poses estimation, in other words, there is an algorithm version that can detect a person and another version that capable to detect multiple persons in an image/video.

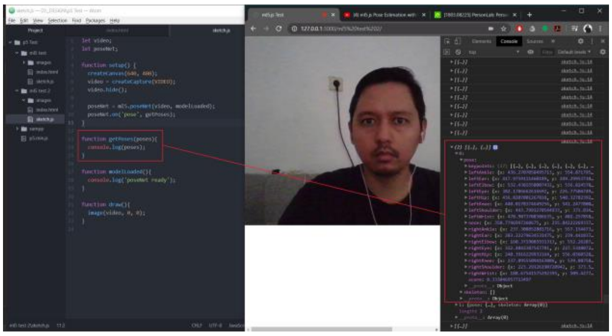

Figure 7 ml5.js syntax that reveals PoseNet body coordinates of Pose and Skeletons.

Similar access method to MobileNets, PoseNet also acquired data from the array list generated by the ml5 library. As seen in Figure 5, ml5.js can list the body's x and y coordinates based on 2 types of labels, Pose, and Skeleton, which are detected by the camera. The Pose coordinates are based on Skeleton points that consist of:

Table 1 Skeleton points inside the Pose label. Inside of these points is the x and y coordinates.

POSES	
leftAnkle	rightAnkle
leftEar	rightEar
leftElbow	rightElbow
LeftEye	rightEye
LeftHip	rightHip
leftKnee	rightKnee
leftShoulder	rightShoulder
leftWrist	rightWrist

nose

Based on Table 1, it is possible to trigger visuals by using PoseNet poses x and y coordinates. As seen in Figure 6, where the code is accessing the nose coordinates, nested in if loops to visualize a white triangle object.

5.2. PoseNet: Pose Classification

It is a complex mechanism to prepare data collection for Pose Classification, the most important step is the inputs, outputs, task, and debug initialization:

- **Inputs** : specifies the number of skeleton points that will be captured.
- **Outputs** : specifies how much content is being recorded.

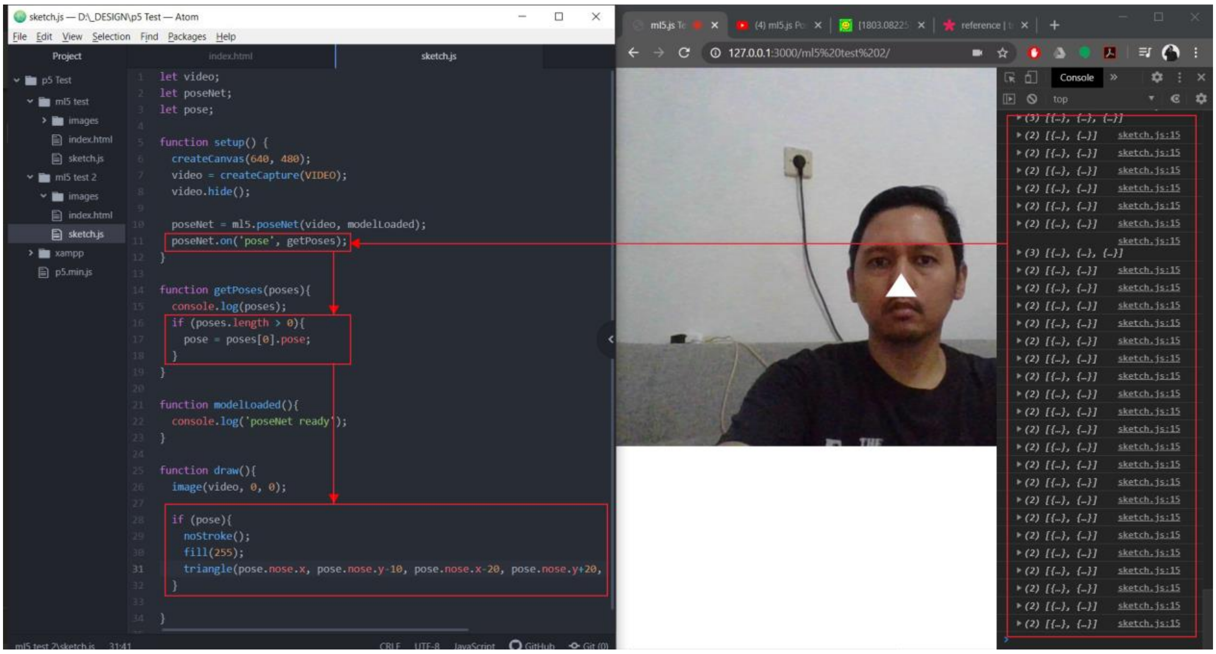

Figure 8 The result of using x and y coordinates of nose point to draw a triangle on the screen, with an addition of simple mathematical calculation.

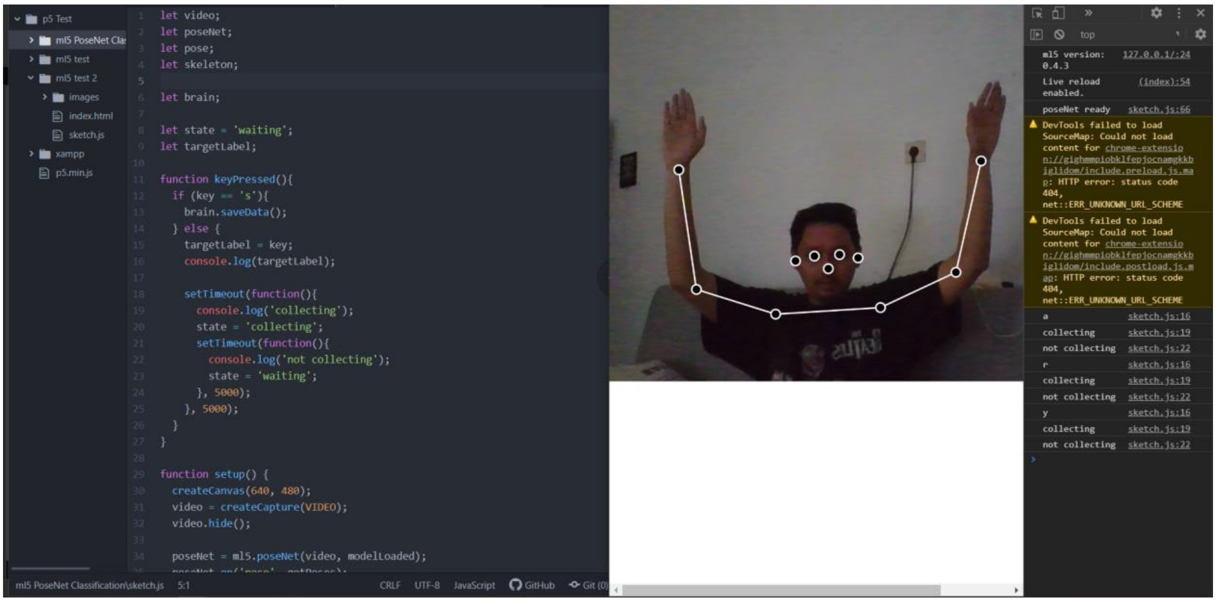

Figure 9 Result of Pose Classification with PoseNet. "Teaching" PoseNet to memorize specific alphabets

- **Task** : specifies the activity given to the machine.
- **Debug** : specifies the boolean state of the activity.

The output of Pose Classification is a JSON file with the skeleton's x and y input coordinates, which then will be used to "teach" PoseNet to recognize the poses. Unfortunately, PoseNet does not have Hand Tracking capability built into the library, although there are many developments in exploring an accurate 3D hand tracking library. One of the many promising developments is Google's MediaPipe.

5.3. Google's MediaPipe Hand Tracking

As claimed by Google that MediaPipe is the simplest way for researchers and developers to build Machine Learning (ML) applications for mobile, edge, cloud, and the web. It is a node-based framework for building multiple modes (video, audio, and sensor) machine learning pipelines that is designed for researchers, students, and software developers, that requires production-ready ML applications, research work that accompanies publishing code, and building technology prototypes [9].

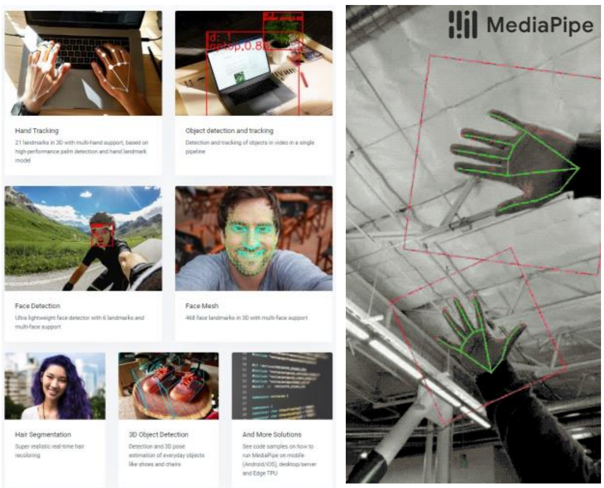

Figure 10 MediaPipe's output examples [8].

It means that there are more reasons to learn sign language not only as communication but also as a human-computer interaction method. As shown in Figure 8, MediaPipe does not only do Hand Tracking specifically but also Face Mesh, which is as important to predict facial expression in practising sign language.

Unfortunately, MediaPipe is still a recent development as this article is written, so with limited resources, it is difficult to provide proof of the result that sign language can be applied by using this contemporary method. As stated in MediaPipe's website that installation status on Windows platform is still Experimental. As the researcher was trying to build the desktop app, according to MediaPipe's Github users report it is failed due to incompatible Bazel version (3.5.0, and author have tried the updated 3.6.0 version as well but still returns android_sdk_repository path undefined while the build was for desktop, not Android). This is also an issue on macOS and Ubuntu operating systems.

To complete the article below is an example of MediaPipe's web-based app and how it can track our hand.

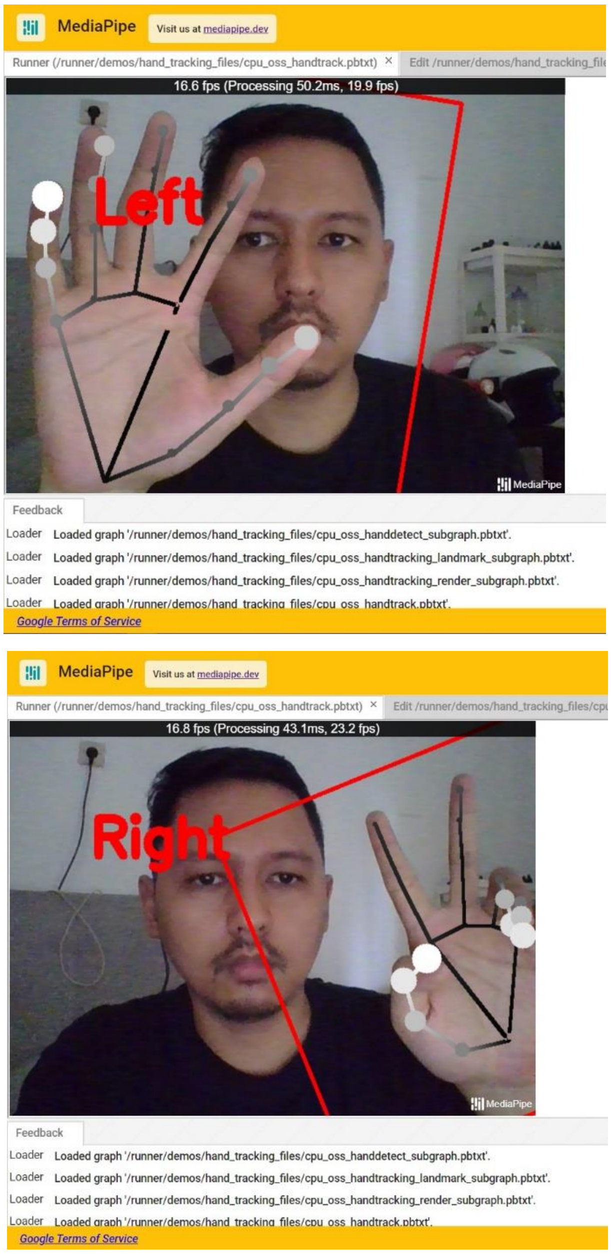

Figure 11 Mediapipe as Web App

The most current update is that Google has made a dedicated website to Machine Learning called Teachable Machine (teachablemachine.withgoogle.com) that is using ml5.js and Tensorflow.js to track body poses.

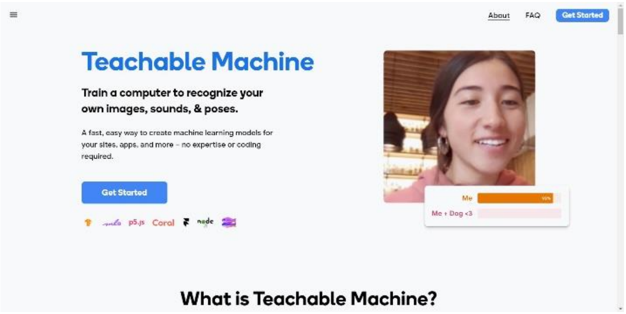

Figure 12 Teachable Machine Web App by Google

6. CONCLUSION AND FUTURE WORK

Finally, with the brief experience of creating a teachable machine by practising Neural Network training using ml5.js, it is highly recommended to acknowledge the bigger illustration of what Artificial Intelligence has expanded to many fields of research. I would still strongly point out the hybridity circumstances as a multimedia designer to adapt to rapid and optimistic technological development. Specifically, in the creative area of interest, we should explore the opportunity in arts and design as a social problem-solving tool despite the ubiquitous entertainment values of it.

The proof of results provided above; it is only a hint of what Machine Learning can do. I will pursue this development further in a bigger research topic in the future by also involving IndoSL (BISINDO) experts from GERKATIN and PUSBISINDO, to encourage and remind the society to involve and continue expanding Universal Design opportunities in the digital era.

REFERENCES

[1] Hurlburt, H. (2014). The Signed Languages of Indonesia: An Enigma (5th ed.). SIL International.

[2] National Disability Authority (2020). What is Universal Design | Centre for Excellence in Universal Design., Available at: http://universaldesign.ie/What-is-Universal-Design [Accessed 3 June 2020].

[3] *Machine Learning*. ibm.com. (2020). Retrieved 17 November 2020, from https://www.ibm.com/analytics/machine-learning.

[4] Manovich, L. (2019). Defining AI Arts: Three Proposals. [online], Available at: http://manovich.net/index.php/projects/defining-ai-arts-three-proposals [Accessed 27 May 2020].

[5] Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. [online] Available at: https://arxiv.org/abs/1704.04861v1 [Accessed 27 May 2020].

[6] Google. (2017). Example use cases include detection, fine-grain classification, attributes and geo-localization. [Image]. Retrieved 8 June 2020, from https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html

[7] [The Coding Train]. (2018, July 31). Beginners Guide to Machine Learning in JavaScript [Video Playlist]. Available at: https://www.youtube.com/playlist?list=PLRqwX-V7Uu6YPSwT06y_AEYTqIwbeam3y [Accessed on 27 May 2020].

[8] Google. (2020). Explore what is possible with MediaPipe today [Image]. Retrieved 8 June 2020, from https://mediapipe.dev/

[9] MediaPipe. (2020).MediaPipe: Cross-platform ML Solution Made Easy, Available at: https://mediapipe.dev/ [Accessed 5 June 2020]

[10] Youtube. (2019). Sign language recognition with RNN and Mediapipe hand tracking DEMO [Video]. Retrieved 8 October 2020, from https://www.youtube.com/watch?v=5epWNiv5EKk

[11] Yasmin, R. (2018). Mengenal Bahasa Isyarat: SIBI dan BISINDO [Image]. Retrieved 8 June 2020, from http://rembulanss.blogspot.com/2018/07/mengenalkan-bahasa-isyarat-sibi-dan.html