JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

Development of Mobile Application for Arabic Sign Language based on Android Studio Software

[1]Noor Azlan Noor Azhar, [1]Shaharil Mad Saad, [1]Muhamad Zulhairi Mohd Nizam, [1]Muhamad Zulhelmi Mohd Nizam, [1]Wan Aliff Abdul Saad, [1]Muhammad Danial, [2]Mohd Azwarie Mat Dzahir

[1]*School of Mechanical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*

[2] *Faculty of Built Environment & Surveying, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*

shaharil@utm.my

ABSTRACT
Sign language is the main form of communication used by deaf people. Their daily activities such as speaking, learning and reading are generally involved in sign languages. Arabic sign language is usually being used for deaf people to read the ayat Al-Quran. It is become hard for them as the traditional method required a teacher that mastered in Arabic Sign Language (ArSL) in order to teach deaf people in learning Al-Quran. Currently, assistive technology for deaf people is considered to be relatively new and underdeveloped unlike visually impaired persons who have electronic braille to aid them in the process of learning and teaching Al-Quran. This project focuses on image-based devices since this project of developing applications involves mobile phone cameras that are specialized for image recognition. The trained dataset obtained for this project is obtained from a published journal and it is sufficient for training. The test dataset is gathered from training data that has been used for developing the model as well as through the experimental setup. An application is created by using Android Studio specialized to implement this model for classification of hand gesture image. The accuracy to predict the hand gesture is observed in order to indicate the model performance.

Index Terms — *Arabic Sign Language, Image based device, Image Recognition, Android Studio*.

I. INTRODUCTION

Sign language is a primary method of communication used by deaf people. Most of their daily activities such as speaking, reading and learning involved sign language. Learning and study have been difficult for them as the majority of the community where they live cannot understand any sign language. Even though most of the hearing-impaired have mastered the sign language, communication with them would be affected because only few "normal" individuals understand and/or can use sign language [1]. This is no exception for deaf Muslim who wanted to learn al-Quran. In Malaysia, there are 160,000 people with hearing problem and majority of them (62.61%) are Muslim [2]. Arabic sign language is important for deaf Muslim in Malaysia to read and learn the ayah Al-Quran. This has become a challenge for them as both of the parties need to understand and be familiar with this sign language in order to teach or learn Al-Quran between each other.


Figure 1: Teaching Al-Quran of deaf people [3]

Currently, there are many assistive technologies to aid impaired people in the process of learning and teaching Al-Quran. Unlike visually impaired people where they have devices that can electronically braille to aid them, but for deaf people, there are still lack of assistive technologies and devices that can be found to technically help them go through this learning process. K. O. Jimoh et al. stated it has been observed that people with hearing impairments do not always have access to systems that assist their process of communication [4].

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

Nowadays, mobile application seems to be the best way to aid deaf people in learning Al-Quran. This technology is suitable as mobile phones seem to be one of the reliable technologies in our daily life and people always bring mobile phones together with them everywhere. Since mobile devices are lightweight and highly portable, new possibilities can be created such as computer vision, pattern recognition including image processing [5]. So, this project can help and guide deaf people in learning Al-Quran with any person as they can use this mobile application to help people understand them. The application has an image-based classification that can help individuals understand the details generated by the representation of sign language through this mobile application. It also can reduce the gap between hearing impaired and people around them in order to communicate and aid them in learning Al-Quran by using Arabic Sign Languages.

II. LITERATURE REVIEW

Basically, sign languages are an important way for deaf people to communicate. Since sign language is a means of communication used by the deaf people to communicate in society, yet not many people know sign language plus there are over hundreds of different standards of sign language used over the world [6]. Besides, it is obvious that deaf people have difficult times to communicate with normal people since they do not understand sign language that is being used by deaf people. Usually, there will be an intermediate person (translator) to overcome this situation, however there is a lot of difficulty to find one because only few people can be a good interpreter for sign language. This gap can be reduced using a system that allows the translation of sign language automatically to text and vice versa [1].
Sign language that deaf people use to read al-Quran is Arabic Sign Language (ArSL). So, ArSL has been used for this project as it involves reading and learning al-Quran. Aforementioned, this project only focuses on sign language that is used in al-Quran which is Arabic Sign Language. ArSL uses one-handed fingerspelling alphabet to represent each letter. Figure 2 and figure 3 shows that only Arabic alphabet of the ayah Al-Quran is display— diacritical marks are ignored.
Figure 2. Arabic Sign Language [4]

Figure 3. Illustration of deaf people read ayat Al-Quran [5]

At present, there are two main approaches available to aid deaf peoples in translating their Arabic Sign Language (ArSL) to languages that understand by vocal people. These approaches are sensor-based approach and image-based approach [8]. Sensor based recognition requires devices that imply sensors to detect the sign language exhibited by the deaf people. There are many types of sensor-based device but when it comes to detection of hand gestures, commonly resistive bend-sensing technology is used on a glove. A Glove-based system as shown in Figure 4 requires the user to put on a glove that has various sensors fitted on it. These sensors are used to detect the motion, position, and velocity of the hand and then analyze to display the appropriate output. This type of technology has a major disadvantage where complex gestures cannot be portrayed due several computer-connected cables that being used for the recognition process [3]. Unlike Image based device, sensor-based device do not require High definition camera, however the communication can be unnatural since it required the user to wear the glove in order to communicate. Next, Electromyography (sEMG) sensor also one of the sensor-based recognition system that being used to detect sign language exhibited by deaf people. sFMG sensor is an electrical device that can measure the electrical pulse produced by the body muscle and can detect finger movement by

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

integrating all the electrical signals [11]. This sensor has proved its ability to distinguish various hand shape and finger movement of the user in the SLR System that is being developed for detection of American Sign Language [12].

Kinect-based systems use Kinect motion sensors to capture the hand signs of the users. This type of sensor can capture 3 dimensional dynamic gestures [13] while capturing the user's hand signs. Moreover, the Kinect sensor has the ability to extract features of dynamic gestures from the sign and converts it to the intended textual form. However, this type of sensor lacks accuracy as the sensor cannot detect fully the hand gesture of the user since it only a sensor detection, not an image recognition. As a result of this, the image- based approaches for recognizing gestures are preferred [10].

The image-based recognition collects images as data input from the camera and then processes them for gesture recognition. This method focuses on gesture of the captured image and main features extraction for the recognition process [3]. Image-based recognition has a major advantage where it can distinguish various patterns of different gestures for the recognition process. Usually, this type of recognition involves computer machine learning and easier to use compared to sensor-based recognition. In addition, when it comes to communication between people, image-based recognition is more normal due to the free movement of user hands from computer-connected cables that need to be attached to hands. Most of the image-based recognition involves Machine Learning (ML). Machine Learning is the study of computer algorithms which, through practice, it will develop automatically. Different from traditional programming where users need to learn and identify patterns which then add it to the algorithm manually, machine learning can learn from the data that exist and update the algorithm automatically [15]. Machine learning can be categorized into four categories which are supervised, unsupervised, semi supervised machine learning and Reinforcement learning [15]. This project is limited to supervised machine learning only that required human supervision to provide training data with solutions called 'labels'. Usually, this type of machine learning was used for image processing, prediction, and classification. For this project, supervised machine learning uses Convolutional Neural network since it is powerful enough to be able detect shapes and edges even input of RGB color images.

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most widely used for visual imagery processing. Convolutional Neural Networks (CNN) is a type of neural network model which allows extraction of higher representations for the image content [16]. They are made up of neurons with learnable weights and biases. Each particular neuron receives multiple inputs and then takes a weighted sum over them, passing through an activation function and responding with an output. Convolutional Neural Networks, or CNNs, were designed to map image data to an output variable as in Figure 5. They have proven so effective that they can predict problems involving image data as an input. Typical architecture of Convolution neural network includes Convolutional/padding Layer, Pooling Layer, and Fully-Connected Layer. Padding is important for reducing image size while retaining its features. Pooling layer, is important for removing noise in the images. lastly, Fully Connected layer is a regular feed forward neural network.

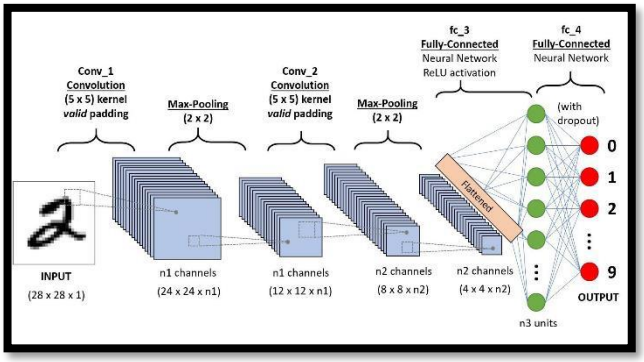

Figure 5. Convolutional Neural Network (CNN) Architecture [7]

In most Machine Learning (ML) projects, the hyperparameters is tuned until the good machine learning accuracy is obtained. Before tuning the ML model, it is crucial to understand what a good ML model entails. Training a data is required but it must have a limitation on the process of training. Based on the loss function graph, machine learning models can be classified as overfitting, underfitting, or optimal. ML models can be overfitted if there is too much data when the training process is undergoing. Thus, it starts learning the noise and unimportant features [19] in the dataset instead learning to generalize the data. In order words, the ML memorizes the data instead of understanding the data. Furthermore, ML models also can be underfitted if it is trained with less data or limited degree of freedom [15]. This underfitting result can cause the ML to have limitations on learning the required data. Besides, the learning process would be generalized and thus produces less error. The optimum model implies that the model has the optimum degree of freedom to sufficiently predict performance rather than match the training data. To make it simple, instead of memorizing the training data, it learned and generalized the output so that it produces less error when a new image is introduced and tested.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

III. METHODOLOGY

This study started by setting up workspace for development of the project, developing a GUI for interface, coding for classification process, testing the accuracy of developed model and coming out with higher precision results. Moreover, the model also will be tested by varying the distance of the input image from the camera. Accuracy of the result will be recorded in order to determine the performance of the model in classifying input images of hand gesture. This project focused on testing the model performance of interpretinghand gesture of user into Arabic Sign Language using Tensorflow and Tensorflow Lite APIs that is compatible for mobile application.

A. Setting up Workspace for Project Development

This project requires workspace for development of the application. Android Studio software is used to develop this Arabic Sign Language Interpreter for the used of deaf people to study and learn Al-Quran. It has many functions and features that have been included in this software such as Gradle Scripts and Manifests files that ease developers while using this software. Besides, programming language that are used for this project is Java programming languange. It is usually used to create applications on computers or mobile devices [23]. Java programming language is simple, efficient, and can be called a general-purpose language. Coding to create GUI and interpret images of hand gesture to Arabic sign language can be done through this programming language.

B. Development of Graphical User Interface (GUI)

A graphical user interface (GUI) is a type of user interface through which users interact with electronic devices via visual indicator representations. A GUI allows the user to communicate with the device by clicking any button that exists on the screen.

Figure 6: User Interface (UI) of the Application

This mobile application will be finalized with a simple design of User Interface (UI) to make it user-friendly and easy to be use. The application is created to have two options whether the user can select the input images from their internal storage or the user can click a button to access their mobile phone camera before taking a picture of any hand gesture that need to be translated.

C. Working Process of Input Images

Coding by using Java Programming Language in Android Studio software is important as it helps in development of the application. After the h5 file has been converted to a tflite file by using Java coding, it is ready to be implemented into a mobile application. Besides, dimensions of received input images will be extracted to 64 x 64 pixels, so that all of the dataset of input images will be uniform throughout the classification process. Moreover, the working process of the image classification is important throughout this application. When the input image is received, it will undergo a conversion from RGB image to grayscale image for classification process. The input image will be compared to trained model file and maximum probability will be classified. Specific labels will be chosen and print out the result on the mobile screen. Figure 7 shows the flow of classification process of input images.

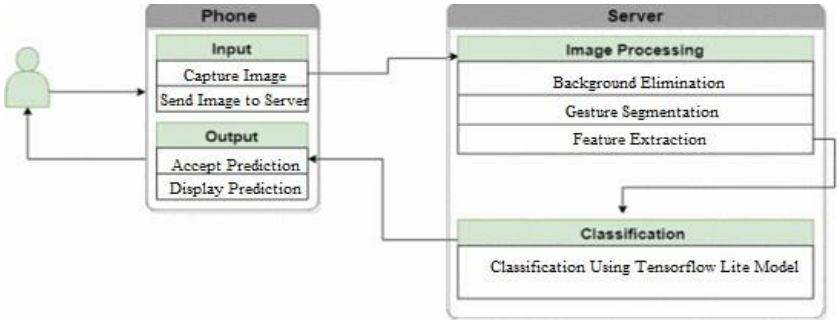

Figure 7: The flow of Classification Process of Input Images

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

D. Prediction Test on Different Distance and Background

The ability of the trained model to classify the input image will be measured so that the accuracy of this mobile application can be determined. Experiment by varying the distance (30cm, 60cm, 90cm) while capturing input image will be done to verify this trained model accuracy. This can ensure how accurate this mobile application with different distance of input images from the user. Besides, different backgrounds of hand gesture image also can affect the classification process of the hand gesture to predict the sign language output. So, model testing of different backgrounds such as white and black background should be done to indicate the performance of the model that has been developed.

E. Performance Evaluation Measures (Average Accuracy)

Model accuracy can be calculated by using accuracy formula. Precision is the ratio of the positives that are correctly identified by the model over total positive records. For example, if there are 10 'alif' images, the model has to predict correctly whether there are 'alif' images by undergoing the classification process. If the model recognized 8 'alif' images correctly out of 10 'alif' images, then the precision of the model for alphabet 'alif' is 8/10, i.e., 80%. Five different tests will be conducted throughout this project. All of the test will be measures for its precision to test the performance model. Sample of 20 images hand gesture of each alphabet will be taken for every type test. The prediction result of true or false will be note down into a table and precision of the model to classify each of the hand gesture image will be calculated by using formulae below. This measure helps to understand the model's performance to print out the correct result of Arabic Sign Language.

$$Accuracy = \frac{true}{true + false} \quad (1)$$

IV. RESULT & ANALYSIS

The result is based on the performance and ability of the model that has been developed to receive captured sign gestures from the user for classification process and display the predicted sign language through mobile application. The model will be tested with various conditions such as different background of the hand gesture and varying distance of the hand gesture from the camera. The accuracy of the result will be measured to indicate the performance of the model, thus recommendations on what to improve need to be considered in order to make sure the model is reliable and ready to be implemented into a mobile phone.

A. Functionality Test

Before the model can be tested by using mobile application, the functionality test of the application will be run first to make sure the application can be used without any issues.

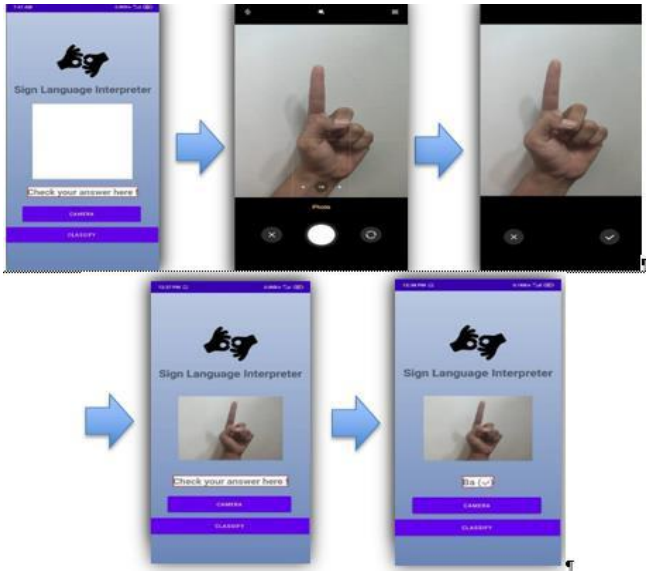

Figure 9: Flow of the Camera Access Process for Image Classification

The application has been developed for the user to choose whether to use an already taken picture of a hand gesture by

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

accessing their mobile phone internal storage or by taking a new image of hand gesture from the mobile phone camera before going through the classification process. Figure 9 shows flow of the process on how the user use the application by accessing their mobile phone camera.

B. First Prediction Test Result (Internal Storage & Camera Access)

The model started to be tested by using trained images (image that has been used to develop the model) in order to test the performance of the model to classify the images that it trained itself. Then, input image that being captured from mobile phone camera has been tested one by one to proceed with the classification process of the image. Sample of 20 trained images & captured images from each of the alphabet have been used to measure the precision of the output result.

Table 1: Accuracy Result for Both of the Test

Sources of Input Images	**Type of Image**	**Average Accuracy (%)**
Internal Storage	Trained Image	91.1
Captured from Camera	New Image	72.5

C. Second Prediction Test Result (Different Distance)

Input images that are taken from mobile phone cameras can vary by distance. The further the input images from the camera, the blurry input images would be obtained. It shows that the accuracy of the images would be affected by the distance that the hand gesture (input image) is being taken. For this test, the hand gesture (input images) will be taken by using a mobile phone camera from three different distances which are 30cm, 60 cm and 90 cm.

Table 2: The Accuracy Result for Distance of 30cm, 60cm and 90cm

Distance (cm)	**Average Accuracy (%)**	**Result**
30	72.5	
60	32.0	
90	7.12	

D. Third Prediction Test Result (Dark & Random Background)

Background of the image is important as it will take part in classifying the input images of the hand gesture. Hence, the model has been tested several times with hand gesture images that have different backgrounds. In this test, black and random background images are tested to look up the performance and ability of the model in the classifying process.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

Background	Average Accuracy (%)	Result
Dark	70.9	
Random	39.2	 

CONCLUSION

This project has achieved its objective which is to develop a mobile application that convert input images of sign language into Arabic characters. The ability and performance of the model also has been tested by conducting several image classification test. Thus, the highest accuracy of classification result in predicting input images of hand gesture is 91.1% where the trained images of Arabic Sign Language dataset is used for classification process. It shows that this model already has higher precision results for classification images. However, the accuracy of the result drops to 72.5% when the input image is taken from a mobile phone camera. This proved that the conversion of RGB image to grayscale image affects the classification process of the input image of hand gesture. Moreover, the further the image that is taken from the user camera, the lower the accuracy of the classification process. It can be inferred that the model is capable of performing effectively for images obtained at a distance of less than 50 cm while it performs poorly for images obtained from a distance of more than 50 cm. Furthermore, random background of input images also affects the accuracy of the prediction result as objects that exist behind the hand gesture interrupt the classification process of the input images. Last but not least, it can be concluded that this model is not reliable enough and not ready to be implemented into mobile application as most of the accuracy of the prediction result are below 80% accuracy.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial sponsorship provided by Universiti Teknologi Malaysia (UTM) through university research grant, No: 4C489, 16J14 and 20J83.

REFERENCES

[1] R. Alzohairi, R. Alghonaim, W. Alshehri, S. Aloqeely, M. Alzaidan, and O. Bchir,"Image based Arabic Sign Language recognition system," Int. J. Adv. Comput. Sci.Appl., vol. 9, no. 3, pp. 185–194, 2018, doi: 10.14569/IJACSA.2018.090327.
[2] T. Malaysian and S. Language, "ARTICLE JOSHUA The Deaf of Malaysia ," no. 2010, 2012.
[3] AL-ITIJAH 2018, https://www.youtube.com/watch?v=QrGQ8yP1yPA
[4] Life and Lexis, (2013). "How I Learned Arabic Sign Language,", pp.1-6.

[5] Muhamad Zulhairi Mohd Nizam, Shaharil Mad Saad, Mohd Azlan Suhaimi, Mohd Azuwan Mat Dzahir, Shayfull Zamree Abd Rahim, and Mohd Azwarie Mat Dzahir, "Development of Al-Quran sign language classification based on convolutional neural network," AIP Conference Proceedings, 347, 020205 (2021) https://doi.org/10.1063/5.0051490

[6] K. O. Jimoh, A. O. Ajayi, and I. K. Ogundoyin, "Template Matching Based Sign Language Recognition System for Android Devices," FUOYE J. Eng. Technol., vol. 5, no. 1, 2020, doi: 10.46792/fuoyejet. v5i1.465.

[7] A. Huskanović, D. A. Macan, Z. Antolović, B. Tomaš, and M. Mijač, "Sustav prepoznavanja slikovnih uzoraka mobilnim uređajem Image pattern recognition using mobile devices," Image pattern Recognit. using Mob. devices, no. January, 2014, doi: 10.13140/2.1.2520.1926.

[8] N. I. on D. and O. C. Disorders, "American Sign Language," No. 11- 4756, 2019. https://www.nidcd.nih.gov/health/american-sign-language (accessed Jan. 03, 2020).

[9] ai Media, "Sign Language Alphabets From Around The World," 2020. https://www.ai-media.tv/sign-language-alphabets-from-around-the- world/ (accessed Feb. 01, 2021).

[10] G. R. S. M. & R. S. Jadon and With, "A Review of Vision Based Dynamic Hand Gestures Recognition," Comput. Sci. Appl., vol. 10, no. 05, pp. 990–1001, 2020, doi: 10.12677/csa.2020.105102.

[11] Kour K. P. and Mathew L. (2017) Literature Survey on Hand Gesture Techniques for Sign Language Recognition. International Journal of Technical Research and Science. 2 (7), pp. 431-434.

[12] S. A. Mehdi and Y. N. Khan, "Sign language recognition using sensor gloves," ICONIP 2002 - Proc. 9th Int. Conf. Neural Inf. Process. Comput. Intell. E-Age, vol. 5, no. May 2014, pp. 2204–2206, 2002, doi: 10.1109/ICONIP.2002.1201884.

[13] K. A. and A. A. S. Mitra, "Evolution of Hand Gesture Recognition," Int. J. Res. Comput. Commun. Technol., pp. 108–120.

[14] M. J. Cheok, Z. Omar, and M. H. Jaward, "A review of hand gesture and sign language recognition techniques," Int. J. Mach. Learn. Cybern., vol. 10, no. 1, pp. 131–153, 2019.

[15] J. Wu, L. Sun, and R. Jafari, "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors," IEEE J. Biomed. Heal. Informatics, vol. 20, no. 5, pp. 1281– 1290, 2016.

[16] A. Potgantwar and P. Bachchhav, "Sign language interpreter using kinect motion sensor using machine learning.," Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 12, pp. 3151–3156, 2019, doi: 10.35940/ijitee.L2645.1081219.

[17] Mandeep, K. A. and Amardeep S. (2015) Hand Gesture Recognition using PCA. International Journal of Computer Science and Engineering Technology 5(7): 267- 285.

[18] A. Geron, "Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow," in Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow, O'reilly, 2019, pp. 3–15.

[19] V. Tatan, "Understanding CNN (Convolutional Neural Network)," 2019. https://towardsdatascience.com/understanding-cnn- convolutional-neural-network- 69fd626ee7d4 (accessed Jan. 19, 2021).

[20] P. Veličković, "2D Convolution," 2016. https://github.com/PetarV-/TikZ/tree/master/2D Convolution (accessed Jan. 01, 2021).

[21] Missinglink.ai, "Fully Connected Layers in Convolutional Neural Networks: The Complete Guide," 2019. https://missinglink.ai/guides/convolutional-neuralnetworks/fully-connected-layers-convolutional-neural-networks-complete-guide/ (accessed Jan. 20, 2021).

[22] GeeksforGeeks, "Underfitting and Overfitting in Machine Learning," 2020. https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/ (accessed Feb. 02, 2021).

[23] M. Petrou, Image processing: the fundamentals, 2nd ed. Chichester, U.K: Wiley, 2010.

[24] H. H. O. Nasereddin, "MMLSL : Modelling Mobile Learning for Sign Language", Int. J. Eng. Comput. Sci., vol. 6, no. 2, pp. 20268–20272, 2017, doi: 10.18535/ijecs/v6i2.20.

[25] Techopedia, "Android Studio," 2019. https://www.techopedia.com/definition/33631/android-studio (accessed Feb. 02, 2021).

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3152-3160
https://publishoa.com
ISSN: 1309-3452

[26] Clearbridgemobile, “Kotlin vs. Java: Which is the Better Option for Android App Development?,” 2020. https://clearbridgemobile.com/java-vs-kotlin-which-is-thebetter- option-for-android-app-development/ (accessed Jan. 01, 2021).