Image Processing Model for Sign Language Recognition

Nwachukwu, Chibuike, Ledisi G. Kabari

Computer Science Department, Ignatius Ajuru University of Education, Rivers State, Nigeria

Abstract: **The act of sign language dates back to the Stone Age when there was no formally defined language in the world. It is also a means of communicating with these special people (i.e. the deaf and dumb). Research shows that a variety of sign language recognition system is already in place presently. However, there are drawbacks associated with these systems. These drawbacks range from poor quality of training databases, un-robust recognition models and lack of GUI for good visualization of sign language recognition. In this work, an enhanced image processing model for sign language recognition was developed. The system was aimed at helping in the education and interaction with the speech impaired in the learning environment. Object-Oriented Analysis and Design Methodology were adopted in this approach. The system was implemented using MATLAB. The SURF algorithm was used to perform feature extraction on the input images in order to enhance detection of interest points on the images. The results show that using our model, sign language images can be recognized at the rate of 90% accuracy as compared to the existing model with performance accuracy of 62%. This study could be beneficial to speech-impaired individuals in the educational sector, to the academic communities that support speech impaired persons, to special schools for the impaired, to organizations that cater for speech impaired persons and to the research community.**

Key Terms: **Sign Language, Computer Vision, Recognition, Gesture Recognition, Speech Impaired.**

I. INTRODUCTION

Dumb people are very common in our society today. In fact they occupy a large space in the distribution of physically impaired persons and research has shown that their numbers are high in Africa. These people have the inability to speak or hear and thereby are very limited in communication and contribution of ideas in the community. Also, when it comes to access to education, these set are also under privileged as most institutions have reservations of admitting people with disabilities as they do not possess the right tools and techniques to accommodate them. Therefore, Deaf-Dumb people do not have normal opportunities for learning. The uneducated dumb people will therefore face serious problems in communicating or interacting with normal people in their society.

A deaf and dumb person is someone who cannot hear or neither hears and speaks. Only about a century ago did the education of the dumb become a thing. The act of sign language dates back to the Stone Age when there was no formally defined language in the world. It is also a means of communicating with these special people (i.e. the deaf and dumb). Finger and gesture spelling is the center of the sign language communication. The most common visual form of finger spelling is imitating the shape of letters in the air, or tactually, tracing letters on the hand. Finger spelling can use one hand such as in American Sign Language, French Sign Language and Irish Sign Language, or can use two hands such as in British Sign Language.

Image identification has in recent years become an essential step in most of the contemporary world problem solving system. Image identification which is a branch of image processing can be very useful in recognizing images such as sign language for better interpretation to the physically impaired persons. Image processing helps to train a given set of sign gestures either using the hands or facial expressions or any other gesture for that matter for classifying and recognizing these gestures for maximum accuracy.

Research shows that a variety of sign language recognition tools are already in place presently. However, there are drawbacks associated with these systems. These drawbacks range from poor quality of training databases, unrobust recognition models and lack of GUI for good visualization of sign language recognition.

The aim of this study is to develop an enhanced image processing model for sign language recognition. We develop a sign language recognition interface for speech impaired persons. We train and test our model using gesture datasets from open-source gesture databases and implemented using Matrix Laboratory (MATLAB). Finally, we compare our results with other existing system performance.

II. RELATED WORKS

Huang et al. [1] proposed a real time Chinese Sign Language translation research based on video processing. They performed a study of two different streams from each signing video, one with the signer and therefore the full movement, and another focused solely on the hand position on each frame. They fed the results of that to a CNN network for video processing, and that they were ready to achieve 82.7% accuracy with their vocabulary of 178 words. However, though considered a true time translation, they used specific gestures to sign the start and ending of every sentence, something that's impractical for real world use.

Rao et al. [2] proposed a continuous sign language recognition system using the front camera of a Smartphone to analyze single handed gestures of Indian Sign Language. Their approach extracted features and determined patterns based on the hand contour created by the hand shape in the picture at a specific frame using video processing techniques. The best score they got was 90.58 percent. Their vocabulary, on the other hand, was limited to only 18 words, all of which had to be performed in a controlled atmosphere with a simple background when shooting the video, which was a key limitation.

Sadek [3] proposed a study on SVD based image processing applications, its contributions and challenges. The study provided an investigational review for the SVD as an able transform in image processing applications. The study also provided some new insights that sprang from SVD properties analysis in different image processing. The main contributions in their study were novel perceptual image forensic technique, a new prospective vision in utilizing the SVD Properties, reviewing and experimental valuation of the developed SVD based application such as denoising, compression, a new block based roughness measure for application such as perceptual progressive compression as well as perceptual progressive data hiding. However, they were unable to use SVD algorithm to implement digital image compression.

Mozammel et al [4] proposed image compression using DWT. The model provided ample soaring compression ratios with no significant deprivation of the quality of the image. The value and sturdiness of this method was vindicated using a set of actual images. The images were taken with a digital camera called (OLYMPUS LI-40C). To display the accuracy of the proposed method, an evaluation between the proposed technique and other common compression techniques has been carried out. The algorithm was demonstrated using Visual C++ and tested on a Pentium Core 2 Duo 2.1 GHz PC with 1 GB RAM. However, they could not implement their compression using a machine learning algorithm such as Neural network (NN).

Peters et al. [5] implemented SVD to compress the microarray image. Huge amounts of DNA information for research purposes were stored as microarray images. These were of high resolution images which highlights minute details of the image. Because of the high resolution, these images tend to be larger in size, which means storage on the hard disk requires lot of space. However, they could not combine a hybrid algorithm for better compression ratio.

Pamar and Pancholi [6] propose image compression based on curvelet transform. They compressed the standard images with this algorithm and the simulation results showed that this compression algorithm was better performance compared to conventional transform techniques. However, they could not use proper threshold value and proper quantization method.

Kumar and Pamar [7] proposed a review on versatile approaches on medical image compression. They explained some of the techniques used for compression of medical images which were used to tackle problemsof storage and bandwidth. They conducted a review of related literatures and came to a conclusion that the compression technique was highly dependent on the ratio of the compression ratio. This implies that a higher compression ratio will determine how better the technique is. However, there was no experimental implementation to prove the hypothesis presented which makes it flawed and uncertain.

Haddad et al. [8] proposed a unique joint watermarking scheme for the medical images. The proposed technique was made up of a combination of JPEG-LS and bit substitution watermarking modulation. From the experimentation, it was detected that the proposed technique was cable of offering the same watermarked images with great security services as compared to the other techniques. However, satellite images could not be compressed in a lossless manner using this technique.

Jiang et al. [9] presented a hybrid algorithm for the medical image compression. The key objective of the proposed approach was to wrap diagnostic related information with high compression ratio. Good PSNR was achieved by the proposed system after an experimentation was carried out and this was achieved in good time as compared to other approaches. However, they could not achieve digital image compression on other image formats or types except the medical image.

Katharotiya et al [10] proposed a comparative analysis between DCT and DWT techniques of image compression. They simulated two major techniques. The first method used the Discrete Cosine Transform (DCT), whereas the second used the Discrete Wavelet Transform (DWT) (DWT). The simulation results were compared using various quality settings and applied to several photographs. Benchmarking to determine performance at runtime was not included in the parameters used for the comparison.

Vasanth et al [11] proposed an efficient compound image compression using optimal discrete wavelet transform and run length encoding techniques. They focused primarily on the preprocessing and modifications that were used to shrink the size of a compound image in order to achieve a high compression ratio (CR), short compression time, and other benefits. The images were measured for preprocessing during the compression process, and DCT with an adaptive particle swarm optimization approach was utilized to preprocess the photos. The primary motivation for this optimization technology is in order to refine the CR value by increasing the wavelet coefficient in the Harr wavelet. To compress the complex images, Run Length encoding was used. Compound images with a low CR and quick computation time. In remote sensing images, however, this was not possible.

Saadi et al. [12] proposed a mixed software/hardware implementation of 2-D signals encoder/decoder utilizing dyadic DWT in light of quadrature mirror filters utilizing quick wavelet Mallat's algorithm. The embedded development kit EDK6.3i and the synthesis software ISE6.3i, both of which were accessible with the Xilinx Virtex-IIV2MB1000 FPGA, were used to compose and aggregate this work. The wavelet coefficients were encoded using the Huffman coding scheme in order to send them logically via an Ethernet transmission control protocol/Internet protocol based connection. The possible reorganization. The setup was implemented into the Es-Salem research reactor's neutron radiography framework. They could, however, use the SVD technique to do so.

Wang et al [13] proposed an approach to encrypt the compressed image by chaotic map and arithmetic coding. It improved the security of arithmetic coding as well as enhances the CR. With the express purpose of having high key and plain image sensitivities, which were utilized to make decisions about the parameter and beginning value of chaotic maps and were recognized with the plain image. The trial results validated the proposed plot's impact and showed that the compressed and encrypted image was safe and useful for transmission. They were, however, unable to attain lossless compression.

Asokan et al [14] presented an overview of image processing techniques for analysis of satellite images for historical maps classification. This research looked at the benefits and drawbacks of a variety of satellite image processing approaches. This study also included multiple comparative analyses to demonstrate the applicability of various approaches.

Barbhuiya et al [15] carried out a comparative study on image compression using DCT and DWT. A comparison is outlined to emphasize the results of this compression system between DCT and DWT using JPEG (Joint Photographic Experts Group) and PNG (Portable Network Graphics) color images. We have done conversion of color images into gray scale and also compression of gray scale image is shown after conversion using DWT method.

Sathik et al [16] proposed a hybrid image compression method which segments the image into background and foreground and compress them with different quality levels. The foreground of the image was given more importance than the background. An edge based segmentation method was used to segment the image into foreground area and background area. The proposed method highly preserved quality of the foreground image. JPEG compression is a widely used compression technique. The proposed method adapted variable quantization and threshold values corresponding to background and foreground. This ensured that the vital area of the image is highly preserved than the other areas of the image. This hybrid approach increased the compression ratio and produced a desired high quality compressed image.

Ponomarenko et al [17]presented an advanced discrete cosine transform (DCT)-based image compression method that combines advantages of several approaches. First, they used a strategy termed "rate-distortion-based modified horizontal-vertical partition scheme" to divide an image into blocks of various sizes. They used a bit-plane dynamical arithmetical coding with extensive context modeling to compress the image in the blocks and minimized the statistical severance of quantized DCT coefficients of each block. After that, another filtering was done. The results demonstrated better compression than JPEG and other DCT-based techniques.

Atish et al [18] proposed a novel contrast improvement approach for contrast improvement of a low-contrast satellite image using SVD and DCT. The information about intensity of the given image was illustrated using singular value matrix and any alteration on the singular values alters the strength of the input image. The proposed technique transformed the image into the SVD-DCT sphere and after standardizing the singular value matrix; the enhanced image was reassembled using inverse DCT. The visual and quantitative results proposed that the SVD-DCT approach evidently demonstrated higher efficiency and flexibility compared to exiting methods such as the histogram equalization, gamma correction and SVD-DWT based techniques.

Li et al [19] proposed a hybrid transform for 3D-DCT based video coding. In the proposed model, 3D-DCT and discrete Haar transform were repeatedly used to eliminate the empty blocks resulting in an adaptive scheme for realizing flexible sequential length of DCT implementations. When compared to other methods in the literature, the proposed model could intelligently choose the best DCT mode and eliminate the temporal correlations more effectively. Experimental results showed that the proposed approach had considerable enhancement than the traditional fixed-length 3D-DCT coding and other variable length 3D-DCT coding.

Beena and Namboodiri [20] proposed an automatic sign language finger spelling using convolution neural network (CNN). They main objective of their study was the recognition of gestures that were static in ASL which are sourced from Kinect sensor. The most stimulating aspect of their design was the development of a good classifier to classify the input static gestures with high accuracy. In their proposed system, design of classifier for sign languages recognition used CNN architecture from Kinect Depth images. The system trained CNNsm for the classification of 24 alphabets and 0-9 numbers using 33000 images. The classifier was trained using diverse parameter configurations and tabulated the results.

III. METHODOLOGY

We reviewed several related literatures related to our study; however, we are interested in the work done by Beena and Namboodiri (2017). They proposed an automatic sign language finger spelling using convolution neural network (CNN). They focused on the recognition of static gestures of ASL which are collected from Kinect sensor. The system trained CNN for the classification of 24 alphabets and 0-9 numbers using 33000 images. The system has trained the classifier with different parameter configurations and tabulated the results. However, they used low quality sign gestures dataset and could not develop a GUI for the display of finger spelling results. We intend to improve on the work carried out by Beena and Namboodiri [20] by developing an enhanced model for sign language recognition using image processing techniques.

Object Oriented Analysis and Design (OOAD) methodology was used during the software development. Object Oriented Analysis and Design methodology is a software engineering approach that modules a system as a group of interacting object. OOAD models are pictures that illustrate the system's objects from various perspectives such as structures and behaviors. Two stages are involved in the approach; Object Oriented Analysis and Object Oriented Design.

Unified Modeling Language (UML) notation is the design tool used for modeling in this research. UML used in the research includes: Use Case diagrams, Sequence diagrams and Class.

3.1 Analysis of the Existing System

The existing system was proposed by Beena and Namboodiri[20]. In the existing system, they implemented PDNN CNN. PDNN is a Python deep learning toolkit developed under the Theano environment. The existing system architecture consisted of a single convolutional layer with 20 feature maps, local filter having a size of 5x5, and a pooling size of 2x2. On feature map with dimension of 28x28 was the input to the architecture. The outputs of the network were flattened. The number of targets or output classes was 33. They trained their network with FC hidden layers ranging from 1 to 4.

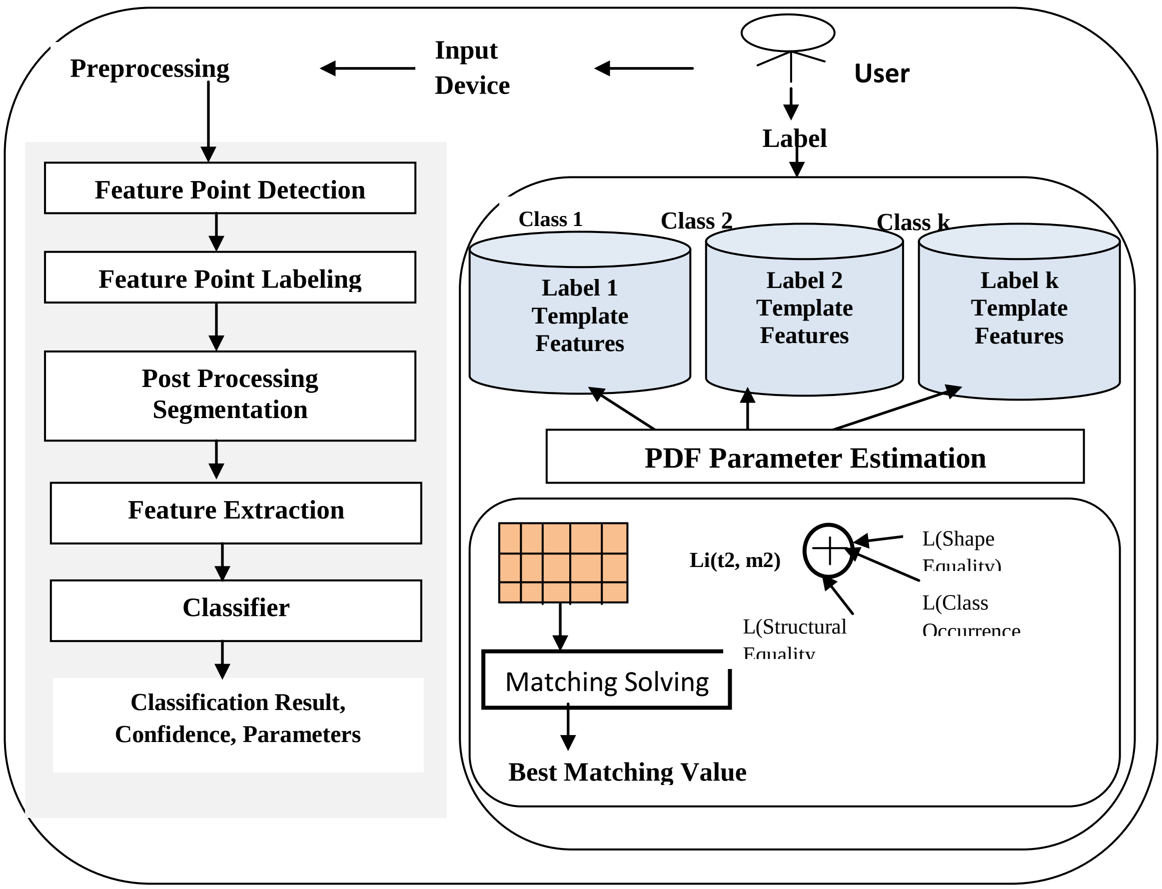

Fig. 1: Architecture of the Existing System (Source: Beena and Namboodiri [20].

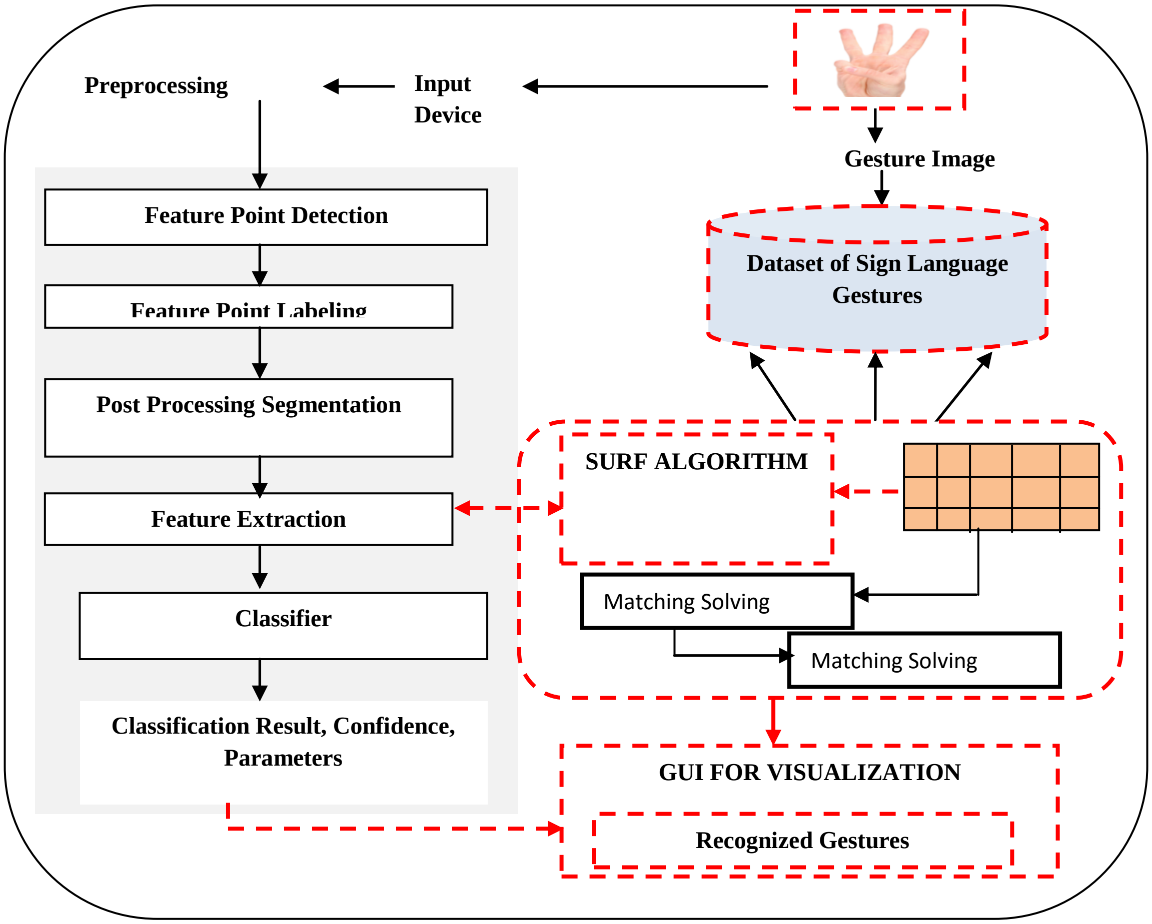

Figure 2: Architecture of the Proposed System

A learning rate of 0.1 was used for their model. For the system implementation, they collected 33000 images from the accessible dataset, using Creative Senz3D depth camera of the resolution of 320x240.

Depth images were collected using this approach Compared to RGB images, more information can collect from depth images. So in the existing system, they took benefit of Kinect images for attaining extreme efficiency.

The dataset consisted of 1,000 images for each of the 33 different hand signs from five subjects. 33 hand signs include all the finger spellings of both alphabets and numbers except J and Z which require temporal information for classification. Since (2/V) and (6/W) were differentiated based on context, only one class was used to represent both one alphabet and one number. Their aim was to create a light weight CNN classifier that can be used with resource constrained embedded devices.

So they down-scaled the dataset image to 28x28 gray scale images. This helped to reduce the number of input nodes in the first layer. They used only 784 features from each image for both training and testing. For simplicity they pickled all the images using pythons pickle function.

3.2 Disadvantages of the Existing System

i. The existing system had a low quality dataset in which the sign images were poorly defined and made the task of interpreting them tedious.
ii. The existing system model was unrobust which implies that lots syntaxes were not included in the sign language.
iii. The existing system did not have a user friendly interface that could be used to visualize the interpreted language.

3.3 Analysis of the Proposed System

The proposed system is an enhancement of the existing system. The proposed system was developed to help the teachers of the speech and hearing impaired to communicate efficiently with them using sigh language or hand gesture recognition techniques. In this proposed system the SURF algorithm was used to perform feature extraction and feature detection of the input image in the image processing process. SURF is an acronym for Speed up Robust Features.

The images used for the system's training were gotten from an open-source sign language repository. The dataset contained about 50, 000 sign language still images. This still image was converted from video frame format to RGB colour Model Format for further processing. After the RGB images were obtained, it was further converted into gray scale. Certain interest points are detected in the image using SURF algorithm after conversion from gray. Captured images are used to test the recognition of the model. The interest points of the captured image & reference image stored in database are also detected. After feature detection, features were extracted from both the captured image and the reference image. Feature extraction here can also be referred to as feature description. The final step is matching. Matching reference feature and reference points between reference images and input image i.e. captured image with closest match using Minimum Euclidean distance. The matching process links the interest point to symbolize certain alphabets or numbers which they represent in the sign language dictionary. The sign language dictionary was defined in the dataset used for the training of the model. The output of the recognition process is numbers and alphabets of the English Language.

3.4 Advantages of the Proposed System

i. The proposed system has a high quality dataset in which the sign images were well defined and thus the task of interpreting them is made easy.
ii. The proposed system model is robust which implies that a lot of syntaxes were included in the sign language definition dataset.
iii. The proposed system has a user friendly interface that could be used to visualize the recognized language.

IV. SYSTEM IMPLEMENTATION

The model was implemented using Matrix Laboratory (MATLAB). MATLAB is used for simulation and image processing programming. The MATLAB software requires that the system it is installed is has at least 2GB of RAM i.e. primary storage and at least 500MB of secondary storage space. The software also requires a processor speed of 2.00GHz to efficient execution results. Figure 3 shows the coding environment of the MATLAB software.

A high quality Android Camera was used to capture hand gesture images which will serve as input into the recognition model. The images are stored within the MATLAB projects folder for easy directory access. The images were preprocessed to remove noise from the image so that the interest points of the image can be clearly seen and easily identified. The SURF algorithm was imported using MATLAB import functions and used for the feature extraction on the image. The images were processed one at a time. Multiple processing and recognition was not carried out in this model to avoid any misinterpretation of signs. This is important because there are certain words that can be represented by hand gestures using both hands, but the model is designed to recognize the Basic English alphabets which are 26 in number and the digits 0-9 in the numbers section.

Figure 4, 5, 6, 7, 8 are used to display the input image and the recognized output of the model using MATLAB software.

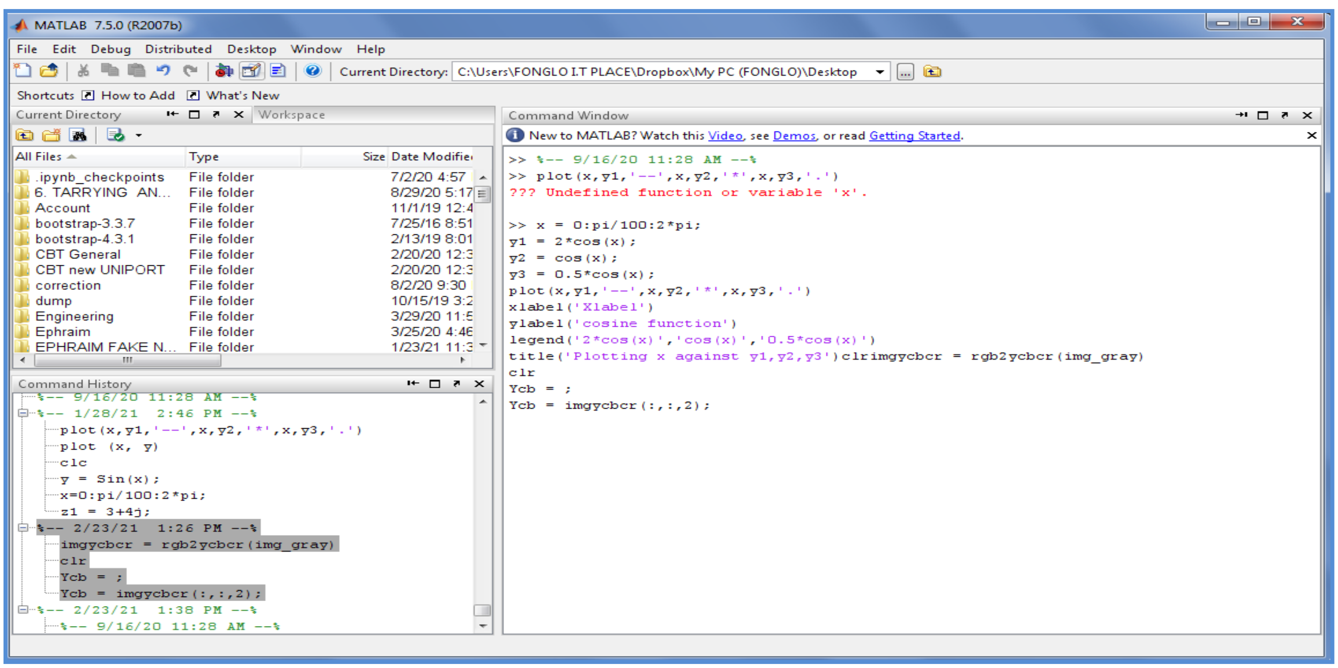

Figure 3: MATLAB Coding Environment

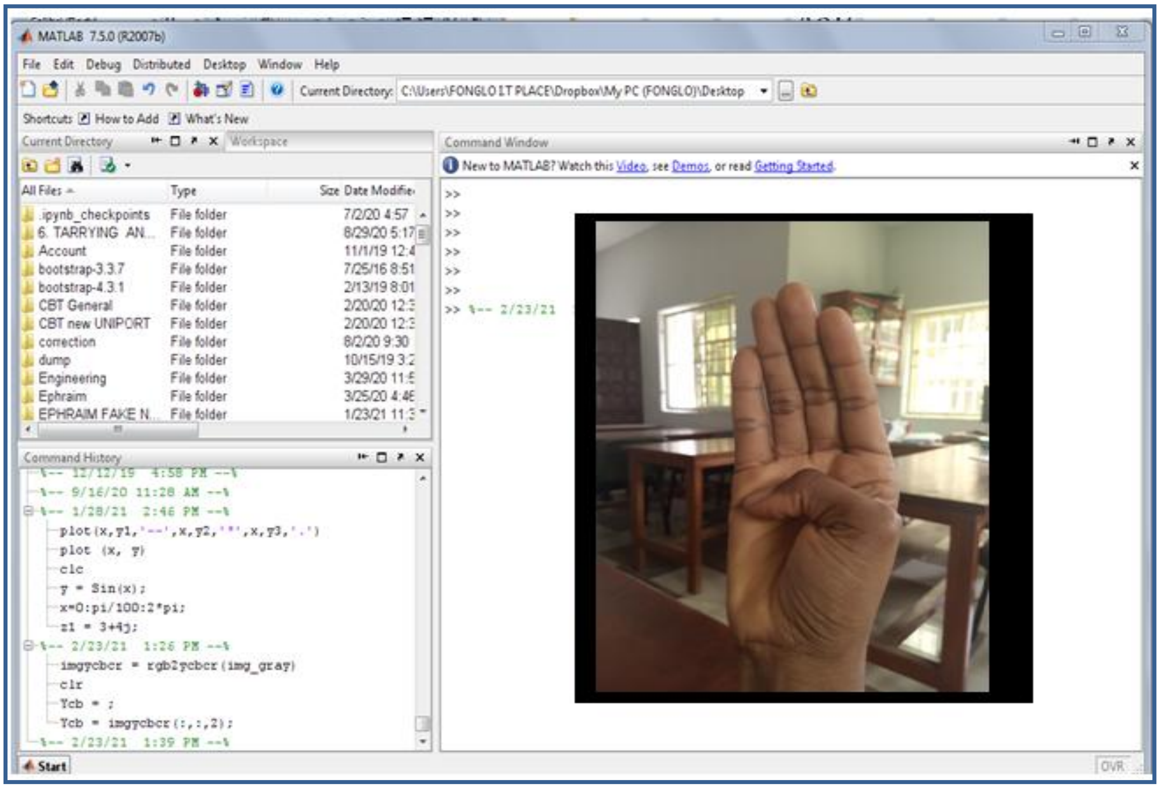

Figure 4: First Image Input of Hand Gesture

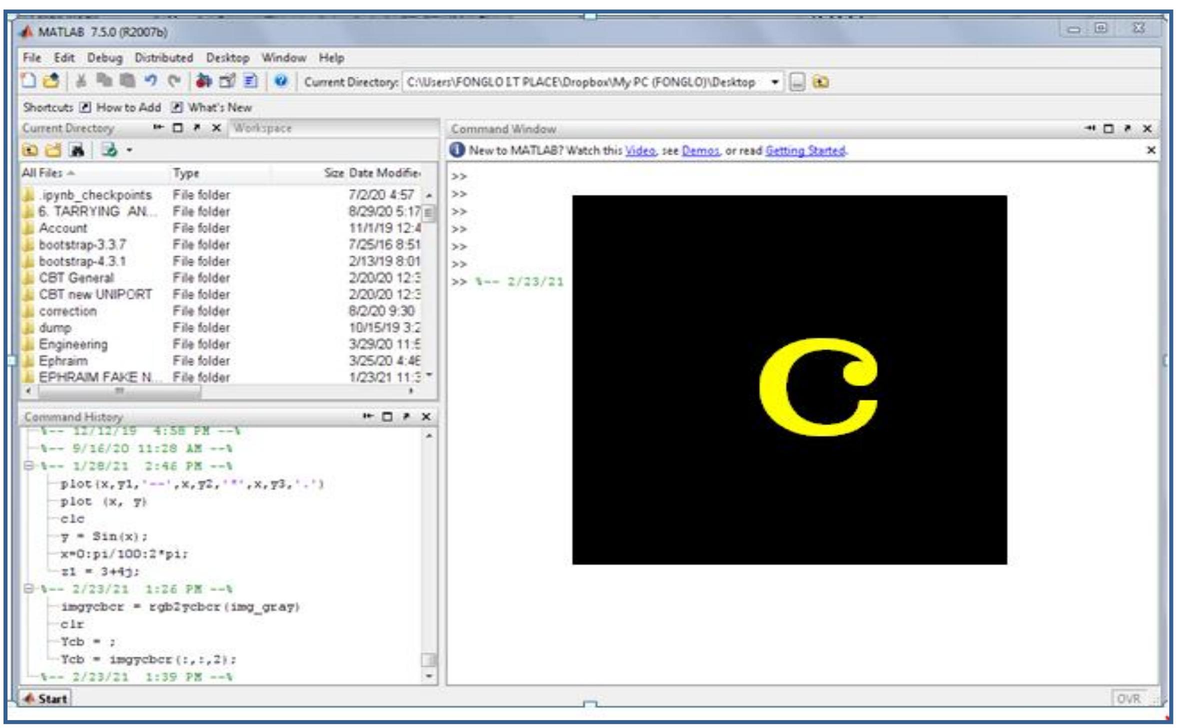

Figure 5: Second Image Input of Hand Gesture

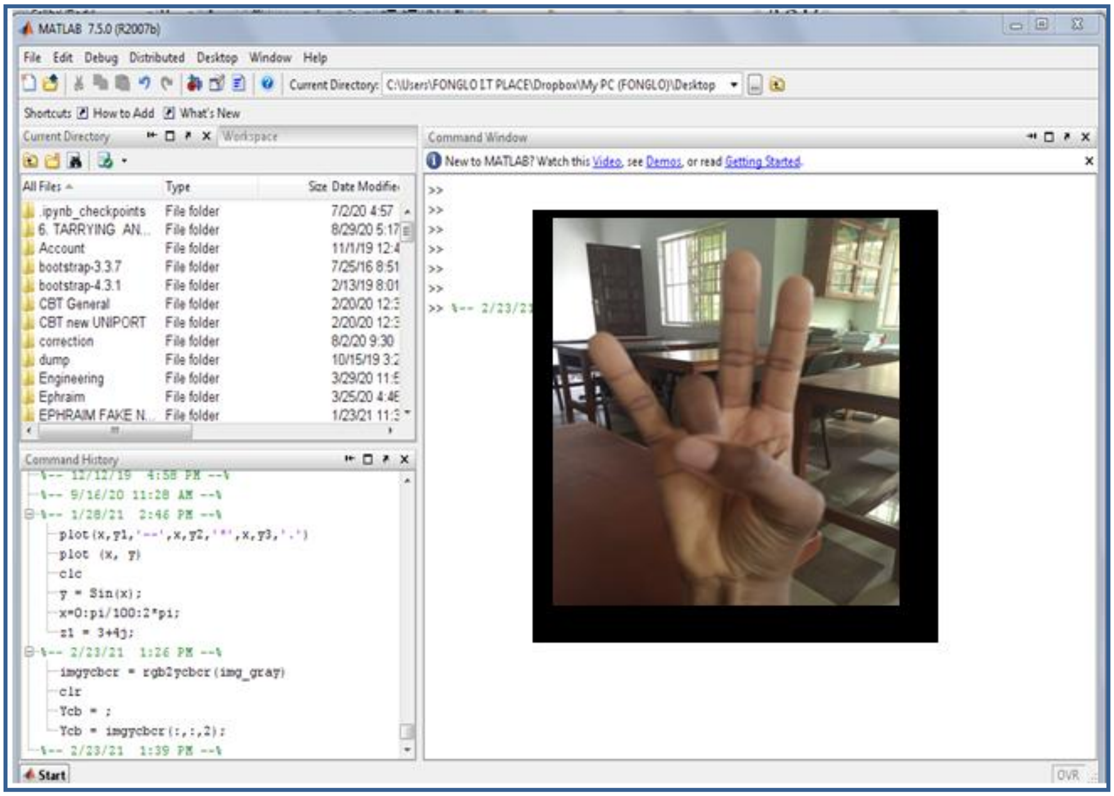

Figure 6: Third Image Input of Hand Gesture

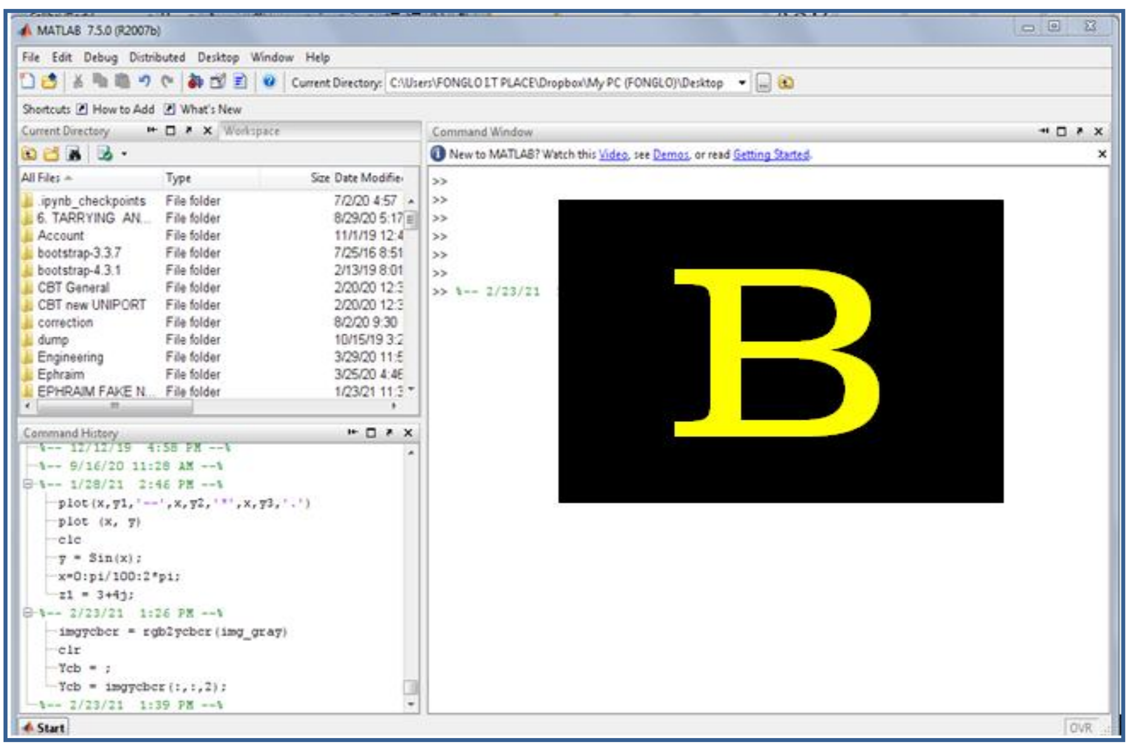

Figure 7: Third Image Input of Hand Gesture

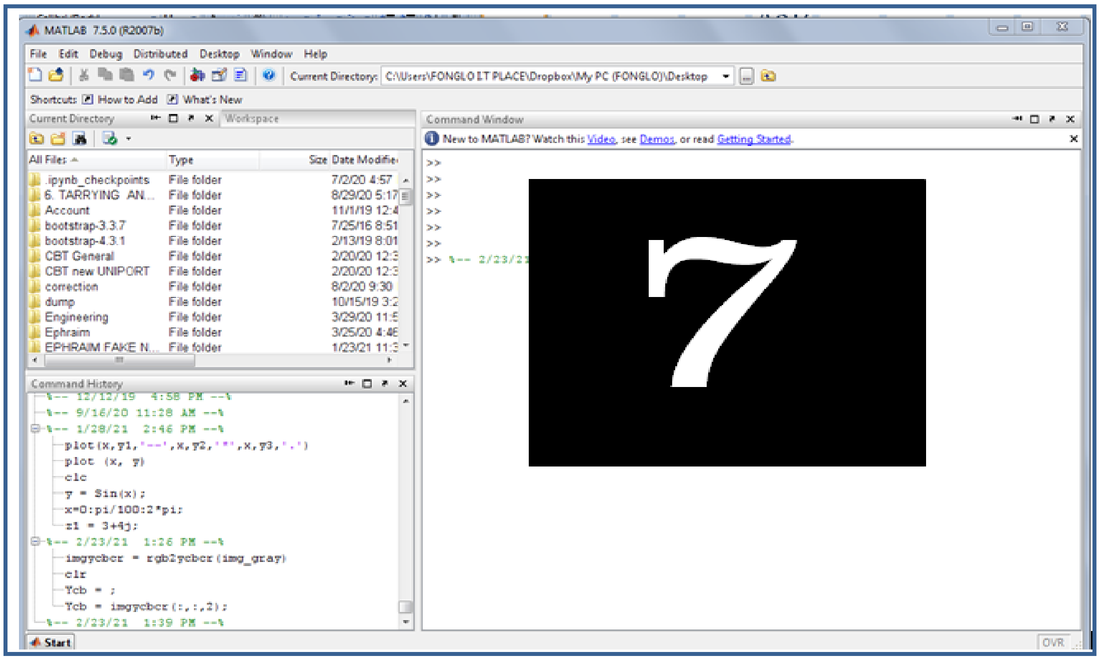

Figure 8: Recognition Output of the Third Input Image

The Image input which is shown in Figure 6 represents the letter "**C**" of the English alphabet. This output will enlighten the impaired person that whenever that hand gesture is made by the teacher or any other person, they are trying to communicate the letter ton them. The model also allows for a clear recognition output to avoid any misunderstanding in the interpretation or representation of the sign.

This is the case with the other outputs represented in the figures. This sign language recognition technique is the best that has been developed in recent years. This can be applied in Nigerian schools for special people for better communication of ideas between the teachers and the students, which will lead to ease in the education process of the speech and hearing impaired.

4.1 Performance Evaluation

This section evaluates the performance of our proposed model as compared to the performance of the existing system. Certain performance parameters such as model efficiency, recognition speed, training data and quality of test data among other were used for the evaluation. Table 1 shows the performance values based on the parameters used. And Figure 12 shows the performance evaluation chart of the proposed and existing models. At the end of the performance evaluation, our system had a model efficiency value of 10% which is a very good rating for image processing models. The existing system efficiency value was 7%.Quality of the test data was 11% while that of the existing system was 5%. This is as a result of the preprocessing done on the image to remove noise and other impurities. Our model also recognized the images in 5 Seconds while the existing system took about 1 minute due to large amount of noise and redundant data contained in the images.

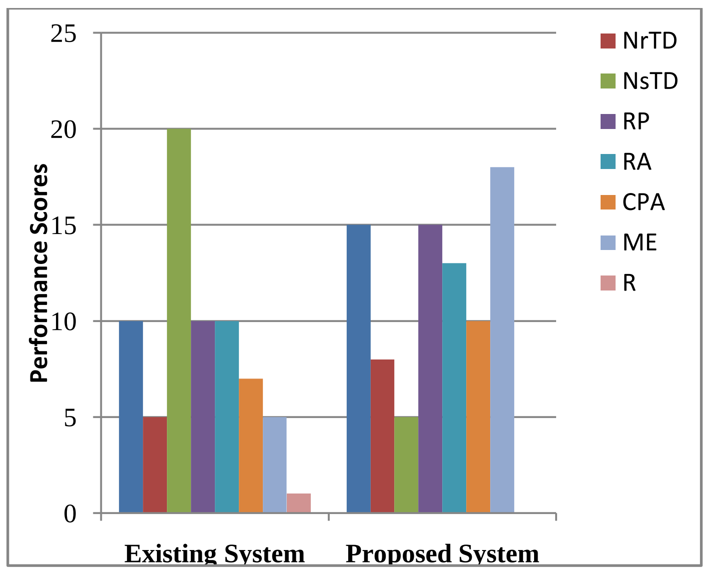

Figure 9: Performance Evaluation Chart

Table 1: Performance Evaluation Table

S/N	Parameter	Existing Model	Proposed Model	Parameter
1	No. of Training Data (NTrD)	33,000	50,000	No. of Training Data (NTrD)
2	No. of Testing Data (NTsD)	5	8	Mean Square Error (MSE) (NTsD)
3	Recognition Speed (RP)	1 Minutes	5 Seconds	Recognition Speed (RP)

4	Recognition Accuracy (RA)	10%	15%	Recognition Accuracy (RA)
5	Cross Platform Adaptability (CPA)	10%	13%	Cross Platform Adaptability (CPA)
6	Model Efficiency (ME)	7%	10%	Model Efficiency (ME)
7	Robustness (R)	5%	18%	Robustness (R)
8	Quality of Test Data (QTD)	5%	11%	Quality of Test Data (QTD)

IV. CONCLUSION

The study concludes on the importance of image processing in sign language recognition to assist the impaired in gaining education and integration into the social communities where they too can make their contributions towards building a better world. It is import to note that the camera that was used for the capture of the images used for testing the model has a very high resolution, as blurry images will make the recognition process inefficient. The proposed model can be used to recognize the 26 alphabets and the ten basic numeric values using Image processing technique in which feature detection and feature extraction of hand gesture is done with the help of SURF (Speed up Robust Features) algorithm. The system resolves the difficulties of disabled person to learn sign language and identify the characters, thus improving the quality of teaching and learning in deaf and dumb institutes.

SUGGESTION FOR FUTURE WORK

For future work, we suggest that an advanced recognition model which recognizes sign words instead of characters be developed. This model will make use of two hands gesture images to communicate words to its audience.

REFERENCES

[1] J. Huang, W. Zhou, Q. Zhang, H. Li, and W. Li, W. “Video-based Sign Language Recognition without Temporal Segmentation. “2018.

[2] T. Rao, P. Escudeiro, N. Escudeiro, E. Rocha, and F. M. Barbosa. “Continuous Sign Language Recognition Using Smartphone Camera.” *Researchgate.* Pp. 1-6. 2019

[3] R. A. Sadek. R.A. SVD Based Image Processing Applications: State of the Art, Contributions and Research Challenges. *International Journal of Advanced Computer Science and Applications (IJACSA)*. Vol. 3, No. 4, pp.26-34. 2012

[4] M. Mozammel, H. Chowdhury, and A. Khatun. “Image Compression Using Discrete Wavelet Transform.” *International Journal of Computer Science Issues*, vol. 6 No. 4, pp. 327-330.

[5] Z. Peters, H. Chowdhury, and A. Khatun. “Compression of Microarray Images Using SVD” *International Journal of Computer Science Issues*, vol. 2 No. 3, pp. 300-308.

[6] K. C Parmar, and K. Pancholi. “Image Compression Based On Curvelet Transform.” *International Journal of Engineering Research & Technology (IJERT).* Vol. 3, No. 4. 2013.

[7] P. Kumar, and A. Pamar. “Versatile approaches for Medical Image Compression: A Review.” International Conference on Computing Intelligence and Data Science (ICCIDS). 2019.

[8] C. Haddad, T.J. Peters, R. Smolíková-wachowiak, and M. P. Wachowiak.. Microarray Image Compression Using a Variation of Singular Value Decomposition. Pp. 1176–1179. 2007.

[9] H. Jiang, M. Zhiyuan, H. Yang. B. Yang and L. Zhang. “Medical image compression based on vector quantization with variable block sizes in wavelet domain.” *Computational Intelligence and Neuroscience.* Vol. 5. 2012

[10] A. Katharotiya, S. Patel, and M. Goyani. “Comparative Analysis between DCT and DWT Techniques of Image compression.” *Journal of Information Engineering and Applications.* Vol. 1, No. 2, pp. 9-17. 2011.

[11] P. Vasanth, S. Rajan, and A. L. Fred. “An Efficient Compound Image Compression Using Optimal Discrete Wavelet Transform and Run Length Encoding Techniques” *Journal of Intelligent Systems.* Vol. 28, No. 1, pp. 87-101. 2019

[12] S. Saadi, M. Touiza, F. Kharfi,.and A. Guessoum. “Dyadic wavelet for image coding implementation on a Xilinx MicroBlaze processor: application to neutron radiography.” *J. Appl. Radiat. Isotopes.* Vol. 82, pp. 200–210. 2013.

[13] B. Wang, X. Zheng, S. Zhou, C. Zhou, X. Wei, Q. Zhang and C. Che. “Encrypting the compressed image by chaotic map and arithmetic coding.” *J. Optik*. Vol. 125, pp. 6117–6122. 2014.

[14] A. Asokan, J. Anitha, M. Ciobanu, A. Gabor, A. Naaji, and D. J. Hemanth, “Image Processing Techniques for Analysis of Satellite Images for Historical Maps Classification: An Overview.” *Applied Sciences.* Vol. 10, No. 4207, pp. 1-21. 2020.

[15] A. H. M. Barbhuiya, T. A. Laskar, and K. Hemachandran. “An Approach for Color Image Compression of JPEG and PNG Images using DCT and DWT.” 2014 Sixth International Conference on Computational Intelligence and Communication Networks. Pp. 129-133/

[16] M. M. Sathik, K. S. Kannan, and Y. J. V. Raj. “Hybrid JPEG Compression using Edge Based Segmentation.” *Signal & Image Processing: An International Journal (SIPIJ).* Vol. 2, No. 1, pp. 165-176. 2011

[17] N. N. Ponomarenko, K. O. Egiazarian, V. V. Lukin, and J. T. Astola. “High-Quality DCT-Based Image Compression Using Partition Schemes.” *IEEE Signal Processing Letters.* vol. 14, No. 2, pp. 105-108. 2018.

[18] S. Ponomarenko and. H. C. Hsu, (2016). “Advanced Discrete Cosine Functions (DCT).” *J. ACM SIGCSE Bull***.** Vol.32. Pp. 73–76.

[19] B. K. Ashish, A. Kumar and P. K. Padhy. “Satellite Image Processing Using Discrete Cosine Transform and Singular Value Decomposition.” *Communications in Computer and Information Science.* Pp. 277-290. 2011.

[20] H. R. Li, S. S. Surbhi and. G. Gopichand. “Hybrid Transform for 3D-DCT based Video Coding”. IOP Conf. Series: Materials Science and Engineering. Vol. 263, pp. 1-9. 2017. doi:10.1088/1757-899X/263/4/042082

[21] M. V. Beena and M. N. A. Namboodiri. “Automatic Sign Language Finger Spelling Uisng Convolution Neural Network: Analysis” *International Journal of Pure and Applied Mathematics*. Vol. 117, No. 2. Pp. 9-15. 2017.