Gurukul International Multidisciplinary Research Journal (GIMRJ)*with*
International Impact Factor 8.249
Peer Reviewed Journal
https://doi.org/10.69758/GIMRJ2406I8V12P092

e-ISSN No. 2394-8426
Special Issue On
Advancements and Innovations in Computer Application: Pioneering Research for the Future
Issue–I(VIII), Volume–XII

Sign Language Recognition (SLR)

Rutuja Burde,
PG Student
Department of Computer Science,
G.H. Raisoni University, Amravati, India
rutuja.burde282000@gmail.com

Received on: *11 May ,2024* ***Revised on:*** *18 June ,2024* ***Published on:*** *29 June ,2024*

***Abstract*—** Sign Language is mainly used by deaf (hard hearing) and dumb people to exchange information between their own community and with other people. It is a language where people use their hand gestures to communicate as they can't speak or hear. Sign Language Recognition (SLR) deals with recognizing the hand gestures acquisition and continues till text or speech is generated for corresponding hand gestures. Here hand gestures for sign language can be classified as static and dynamic. However, static hand gesture recognition is simpler than dynamic hand gesture recognition, but both recognition is important to the human community. We can use Deep Learning Computer Vision to recognize the hand gestures by building Deep Neural Network architectures (Convolution Neural Network Architectures.

Index Terms **– HTML, CSS, Python, Pip, Mediapip, opencv, NumPy**

I. INTRODUCTION

Deaf (hard hearing) and dumb people use Sign Language (SL) as their primary means toexpress their ideas and thoughtswith their own community and with other people with hand and body gestures.It has its own vocabulary, meaning, and syntax which is different from the spoken language or written language. Spoken language is a language produced by articulate sounds mapped against specific words and grammatical combinations to convey meaningful messages. Sign language uses visual hand and body gestures to convey meaningful messages.There are somewhere between 138 and 300 different types of Sign Language used around globally today.

V Using Deep Learning algorithms and Image Processing we can able to classify these hand gestures and able to produce corresponding text.An example of "A" alphabet in sign language notion to English "A" text or speech.. Sign language guides this part of the community and empowers smooth communication in the community of people with trouble talking and hearing (deaf and dumb). They use hand signals along with facial expressions and body activities to cooperate.

Key features of the SLT system include adaptability to different sign language dialects, user-friendly interfaces, and the ability to learn and adapt to new signs over time. The research also explores the integration of wearable devices and mobile applications to make the SLT system accessible and portable for users in various settings.

In development and implementation of the Sign Language Translation system, this research aims to empower the deaf community by breaking down communication barriers and fostering inclusivity in diverse social and professional environments. The potential impact of this technology extends beyond individual interactions, contributing to a more inclusive and equitable society where communication is accessible to everyone, regardless of their hearing abilities.

II. RELATED WORK

Sign Language Recognition (SLR) is an important part of the larger application field of Hand Gestures Recognition (HGR) in Human–Computer Interaction. Complete works describing applications and techniques in HGR.

Research Designing solutions that achieve good performance despite all the constraints and difficulties imposed by the complexity of the problem is the real challenge. Issues related to sign composition, interaction, relationships between hands, different classes of signs, lexicon complexity and non-standard sign translation

Propose an SLR system design with the following system architecture in three levels:

a) sign recognition,

b) sign selection and detection of relationships between signs.

Each level itself can involve a very complex system design.

The first issue to address when designing systems for SLR is related to the interaction between the user doing the signs and the computational interface. This interaction can employ glove-based systems or vision-based systems

On the other hand, vision-based systems can be much cheaper and more comfortable for users. Therefore, vision-based systems have received increased attention from researchers and developers. For these systems, general or special cameras capture images as data input. Papers describing system solutions using intensity (RGB), gray, and black and white images as their input can be found

III. PROPOSED WORK

The project will be structured into 3 distinct functional blocks, Data Processing, Training, Classify Ge
Data Processing:
The load data.py script contains functions to load the Raw Image Data and save the image data as numpy arrays into file storage. The process data.py script will load the image data from data.npy and preprocess the image by resizing/rescaling the image, and applying filters and ZCA whitening to enhance features. During training the processed image data was split into training, validation, and testing data and written to storage. Training also involves a load dataset.py script that loads the relevant data split into a Dataset class. For use of the trained model in classifying gestures, an individual image is loaded and processed from the filesystem.
Training:
The training loop for the model is contained in train model.py. The model is trained with hyperparameters obtained from a config file that lists the learning rate, batch size, image filtering, and number of epochs. The configuration used to train the model is saved along with the model architecture for future evaluation and tweaking for improved results. Within the training loop, the training and validation datasets are loaded as Dataloaders and the model is trained using Adam Optimizer with Cross Entropy Loss. The model is evaluated every epoch on the validation set and the model with best validation accuracy is saved to storage for further evaluation and use. Upon finishing training, the training and validation error and loss is saved to the disk, along with a plot of error and loss over training

Classify Gesture:
After a model has been trained, it can be used to classify a new ASL gesture that is available as a file on the filesystem. The user inputs the filepath of the gesture image and the test data.py script will pass the filepath to process data.py to load and preprocess the file the same way as the model has been trained.

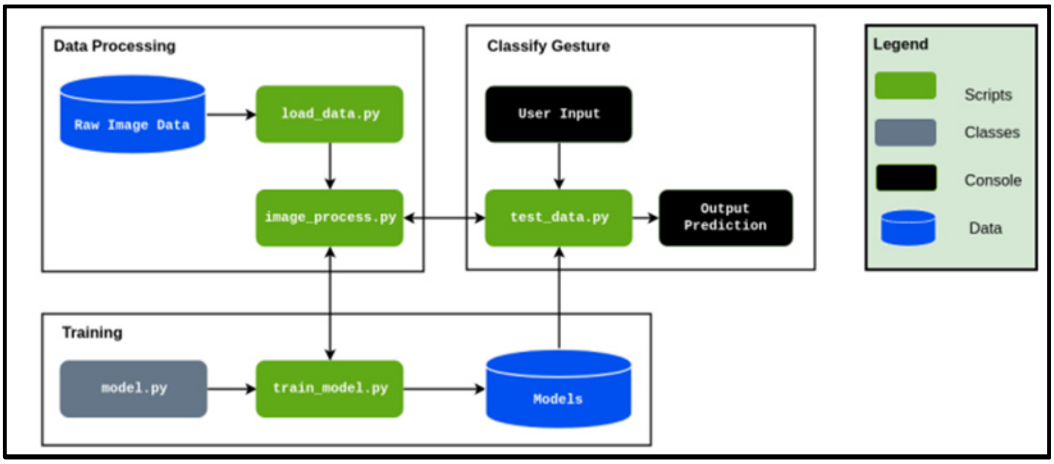


Fig: 1 Model Performance

IV. PROPOSED RESEARCH MODEL

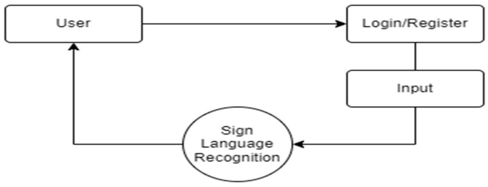
A proposed research model Sign language recognition is a topic of current research in Computer Science and Engineering field. This application will be a boost to the deaf and hard hearing people. They are not

Gurukul International Multidisciplinary Research Journal (GIMRJ)*with*
International Impact Factor 8.249
Peer Reviewed Journal
https://doi.org/10.69758/GIMRJ2406I8V12P092

e-ISSN No. 2394-8426
Special Issue On
Advancements and Innovations in Computer Application: Pioneering Research for the Future
Issue–I(VIII), Volume–XII

able to use the computers and other hand held devices as it is very difficult for them to interact with such devices. So, in this area a lot of research is going on to help them.

In India, either no standard database is available to carry research in this area and or no systems are available for them. In the proposed research we will try to develop a system for hard hearing and physically challenged persons.



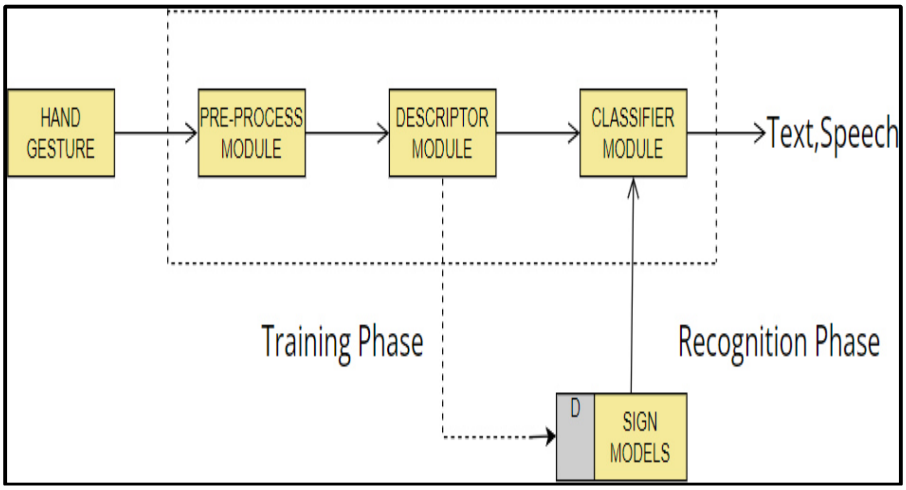


Fig. 2: The flow of proposed work

Data Collection

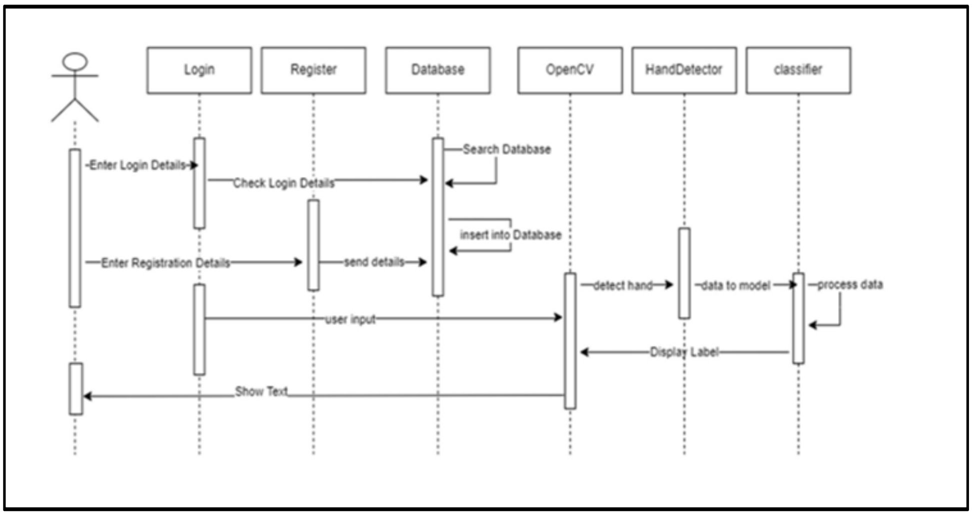


Fig2: Algorithm

Data Collection :

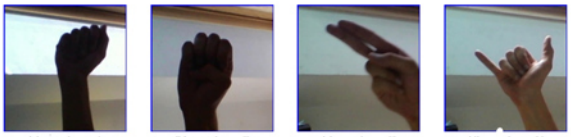
Gurukul International Multidisciplinary Research Journal (GIMRJ)*with*
International Impact Factor 8.249
Peer Reviewed Journal
https://doi.org/10.69758/GIMRJ2406I8V12P092

e-ISSN No. 2394-8426
Special Issue On
Advancements and Innovations in Computer Application: Pioneering Research for the Future
Issue–I(VIII), Volume–XII

The primary source of data for this project was the compiled dataset of Sign Language Recognitions called the SLR

The dataset is comprised of 87,000 images which are 200x200 pixels. There are 29 total classes, each with 3000 images, 26 for the letters A-Z and 3 for space, delete and nothing This data is solely of the user Akash gesturing in ASL, with the images taken from his laptop's webcam. These photos were then cropped, rescaled, and labelled for use.

A self-generated test set was created in order to to investigate the neural network's ability to generalize. Five different test sets of images were taken with a webcam under different lighting conditions, backgrounds, and use of dominant/non-dominant hand



(a) letter A (b) letter E (c) . . letter H (d) . letter Y

Fig 3: Examples of images from SLR dataset used for training

Data Pre-processing :

The data preprocessing was done using the PILLOW library, an image processing library, and sklearn .decomposition library, which is useful for its matrix optimization and decomposition functionality.

(a) Sample from non-uniform background test set (b) Sample from plain white background test set (c) Sample from a darker test set (d) Sample from plain white background test set

Image Enhancement :

A combination of brightness, contrast, sharpness, and color enhancement was used on the images. For example, the contrast and brightness were changed such that fingers could be distinguished when the image was very dark

Edge Enhancement:

Edge enhancement is an image filtering techniques that makes edges more defined. This is achieved by the increase of contrast in a local region of the image that is detected as an edge. This has the effect of making the border of the hand and fingers, versus the background, much more clear and distinct. This can potentially help the neural network identify the hand and its boundaries.

Image Whitening:

Image whitening, is a technique that uses the singular value decomposition of a matrix. This algorithm decorrelates the data, and removes the redundant, or obvious, information out of the data. This allows for the neural network to look for more complex and sophisticated relationships, and to uncover the underlying structure of the patterns it is being trained on. The covariance matrix of the image is set to identity, and the mean to zero

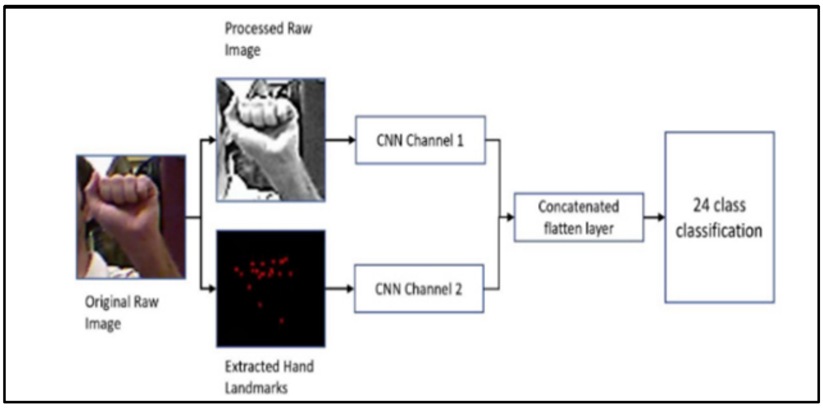


Fig 4: Examples of image preprocessing.

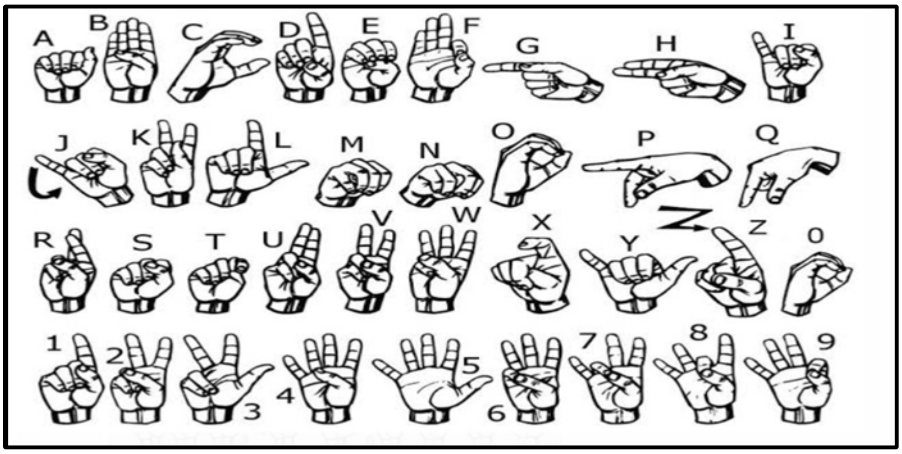


Fig 5: **Sign of hand language recognition**

V. PERFORMANCE EVALUATION

For overall performance measurement, a confusion matrix and classification file are computed.

The method for evaluation metrics is as follows: The frequency with which the classifier plays an accurate vaticination is referred to as accuracy.

It is decided via partitioning the amount of nicely grouped instances by means of the whole wide variety of instances. Precision is a measure of how often the classifier accurately predicts a effective instance.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad ,$$

Here TP is the real +ve, TN is the real -ve, FP is the fake +ve, and FN is the fake -ve. It's computed through dividing the entire of TP and FP via the overall quantity of real positives.

Recall is a degree of how often the classifier effectively predicts a +ve example out of all +ve instances. $Precision = \frac{TP}{(TP+FP)}$ It's decided through isolating the amount of actual up-sides by means of the quantity of TP

$$Evaluation = \frac{TP}{(TP+FN)}$$

The F1 rating is the balanced means of perfection and recall. it's for a share of the classifier's exactness.

F1 rating is identical to $\frac{(2 \times precision \times recall)}{(precision + recall)}$.

VI. RESULT ANALYSIS

In human action recognition tasks, sign language has an extra advantage as it can be used to communicate efciently. Many techniques have been developed using image processing, sensor data processing, and motion detection by applying diferent dynamic algorithms and methods like machine learning and deep learning. Depending on methodologies, researchers have proposed their way of classifying sign languages.

As technologies develop, we can explore the limitations of previous works and improve accuracy. proposes a technique for acknowledging hand motions, which is an excellent part of gesture-based communication Because of a proficient profound deep convolutional neural network (CNN) architecture. Te proposed CNN design disposes of the requirement for recognition and division of hands from the captured images, decreasing the computational weight looked at during hand pose recognition with classical approaches.

In our method, we used two input channels for the images and hand landmarks to get more robust data, making the process more efcient with a dynamic learning rate adjustment.the presented results were acquired by retraining and testing the sign language gestures dataset on a convolutional neural organization model utilizing Inception

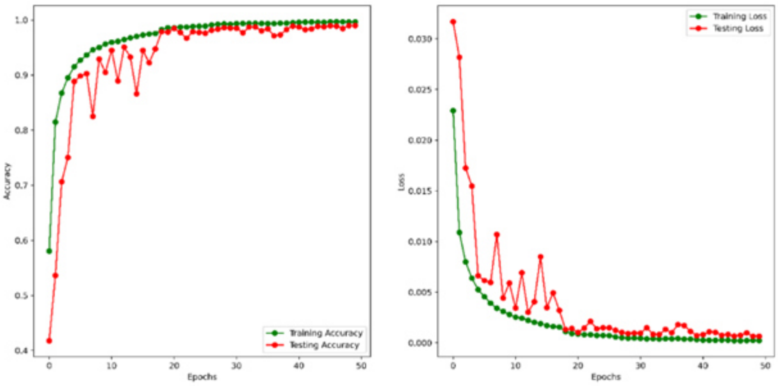


Figure 6. Training versus testing accuracy and loss for 50 epochs.

VII. CONCLUSION

This work proposes a methodology for perceiving the classification of sign language recognition. Sign language is the core medium of communication between deaf-mute and everyday people. It is highly implacable in real-world scenarios like communication, human–computer interaction, security, advanced AI, and much more.

For a long time, researchers have been working in this feld to make a reliable, low cost and publicly available SRL system using different sensors, images, videos, and many more techniques.

Many datasets have been used, including numeric sensory, motion, and image datasets. Most datasets are prepared in a good lab condition to do experiments, but in the real world, it may not be a practical case. Tat's why, looking into the real-world situation, the Fingerspelling dataset has been used, which contains real-world scenarios like complex backgrounds

VIII. FUTURE SCOPE

- The future scope of SLR systems may incorporate multimodal data sources, such as video, depth sensors, infrared sensors, and wearable devices, to capture a more comprehensive representation of sign language gestures.
- There will be a continued emphasis on achieving real-time performance in SLR systems to enable seamless communication and interaction between signers and non-signers. Optimization of algorithms, parallel processing techniques, and hardware acceleration will be explored to reduce latency and improve responsiveness.
- In Future research will focus on enhancing the accuracy and robustness of SLR systems, especially in handling variations in sign language gestures, lighting conditions, occlusions, and background clutter. Advanced machine learning techniques, including deep learning and reinforcement learning, will be employed to develop more accurate and adaptable recognition models.
- Advancements in natural language processing (NLP) and context-aware computing will enable SLR systems to not only recognize individual signs but also understand the semantic meaning and contextual nuances of sign language

IX. REFERENCES

[1] Farnaz D. Notash and Elahe Elhamki. "Comparing loneliness, depression and stress in students with hearingimpaired and normal students studying in secondary schools of Tabriz". In: International Journal of Humanities and Cultural Studies February 2016 Special Issue (2016). issn: 2356-5926.

[2] "The Cognitive, Psychological and Cultural Impact of Communication Barrier on Deaf Adults". In: Journal of Communication Disorders, Deaf Studies Hearing Aids 4 (2 2016). doi: 10.4172/2375-4427.1000164.

[3] Akash. ASL Alphabet. url: https://www.kaggle.com/grassknoted/asl-alphabet. (accessed: 24.10.2018). [4] Vivek Bheda and Dianna Radpour. "Using Deep Convolutional Networks for Gesture Recognition in American Sign Language". In: CoRR abs/1710.06836 (2017). arXiv: 1710.06836. url: http://arxiv.org/abs/1710. 06836.

[4]. Anderson, R., Wiryana, F., Ariesta, M. C. & Kusuma, G. P. Sign language recognition application systems for deaf-mute people: A review based on input-process-output. Proced. Comput. Sci. 116, 441–448. https://doi.org/10.1016/j.procs.2017.10.028 (2017).

[5]. Mummadi, C. et al. Real-time and embedded detection of hand gestures with an IMU-based glove. Informatics 5(2), 28. https:// doi.org/10.3390/informatics5020028 (2018).

[6] Hickeys Kinect for Windows - Windows apps. (2022). Accessed 01 January 2023. https://learn.microsof.com/en-us/windows/ apps/design/devices/kinect-for-windows

[7] Rivera-Acosta, M., Ortega-Cisneros, S., Rivera, J. & Sandoval-Ibarra, F. American sign language alphabet recognition using a neuromorphic sensor and an artifcial neural network. Sensors 17(10), 2176. https://doi.org/10.3390/s17102176 (2017).

[8] Ye, Y., Tian, Y., Huenerfauth, M., & Liu, J. Recognizing American Sign Language Gestures from Within Continuous Videos. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2145–214509 (IEEE, 2018). https:// doi.org/10.1109/CVPRW.2018.00280.

[9] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection", *1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA),* 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30, https://doi.org/10.46335/IJIES.2022.7.8.5

[10] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "Revealing and Classification of Deepfakes Videos Images using a Customize Convolution Neural Network Model", *International Conference on Machine Learning and Data Engineering (ICMLDE)*, 7th & 8th September 2022, 2636-2652, Volume 218, PP. 2636-2652, https://doi.org/10.1016/j.procs.2023.01.237

[11] Usha Kosarkar, Gopal Sakarkar (2023), "Unmasking Deep Fakes: Advancements, Challenges, and Ethical Considerations", *4th International Conference on Electrical and Electronics Engineering (ICEEE)*,19th & 20th August 2023, 978-981-99-8661-3, Volume 1115, PP. 249-262, https://doi.org/10.1007/978-981-99-8661-3_19

[12] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2021), "Deepfakes, a threat to society", *International Journal of Scientific Research in Science and Technology (IJSRST)*, 13th October 2021, 2395-602X, Volume 9, Issue 6, PP. 1132-1140, https://ijsrst.com/IJSRST219682

[13] Usha Kosarkar, Prachi Sasankar(2021), " A study for Face Recognition using techniques PCA and KNN", Journal of Computer Engineering (IOSR-JCE), 2278-0661,PP 2-5,

[14] Usha Kosarkar, Gopal Sakarkar (2024), "Design an efficient VARMA LSTM GRU model for identification of deep-fake images via dynamic window-based spatio-temporal analysis", Journal of Multimedia Tools and Applications, 1380-7501, https://doi.org/10.1007/s11042-024-19220-w

[15] Usha Kosarkar, Dipali Bhende, " Employing Artificial Intelligence Techniques in Mental Health Diagnostic Expert System", International Journal of Computer Engineering (IOSR-JCE),2278-0661, PP-40-45, https://www.iosrjournals.org/iosr-jce/papers/conf.15013/Volume%202/9.%2040-45.pdf?id=7557