ISSN : 2456-3307

International Conference on Artificial Intelligence and Machine Learning

Organized by
SECAB
Institute of Engineering and Technology,
Nauraspur, Bagalkot Road Vijayapur,
Karnataka, India

Volume 8, Issue 5, July-August-2021

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN COMPUTER SCIENCE, ENGINEERING AND INFORMATION TECHNOLOGY

Impact Factor Value 2020 = 6.135

Email: editor@ijsrcseit.com

International Conference on Artificial Intelligence and Machine Learning

On

25th and 26th Aug, 2021

Organized by

SECAB Group of Institute

SECAB ASSOCIATION

SOCIO ECONOMIC CULTURAL ASSOCIATION BIJAPUR

SECAB Institute of Engineering and Technology, Nauraspur, Bagalkot Road Vijayapur, Karnataka, India

In association with

IJSRCSEIT

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

Volume 8, Issue 5, July-August-2021

Published By

Techno Science Academy

The International Open Access Publisher

website : www.technoscienceacademy.com

CONTENTS

International Conference on Artificial Intelligence and Machine Learning
In association withInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

The Comprehensive Study of Facial Expression Recognition on Video

Ratnalata Gupta*[1], L. K. Vishwamitra[1]
[1]Department of Computer Science, Oriental University, Indore, Madhya Pradesh, India

ABSTRACT

In human life, facial expressions and emotions reveal external and internal responses. In human-computer interaction, the video clip plays an important part for extract the emotion of the end-user. In this type of system, it is necessary to observe rapid dynamic changes in the motion of the human face to give the essential response. A real-time application is based on identifying the face and recognizes his expressions such as happy, sad, disgruntled and tiredness, etc. for example Driver exhaustion detection to prevent road accidents on the road. We focus on the comprehensive study of different traditional and recent methodsfor preprocessing, features take out and emotion recognition is employed in facial expression and recognition systems (FER). This paper also describes various Terminologies are used in the Facial expression and recognition system (FER) system. The result compares with the Number of Expressions, algorithm accuracy, and implementation tools. Video-based facial emotion recognition is a very interesting and challenging problem hence the present study provides model complexity, implementation trends, and opportunities for researchers can consider as future research works.

Keywords: Facial Expression Recognition, Deep Learning, Classification.

I. INTRODUCTION

Today, Facial expressions are playing a prominent role in our day-to-day life for communication, human emotions, and intentions which play a vital role in decision-making perception such as physical fatigue detection [1]. Facial expression and emotions are the key features of nonverbal communication and this communication generally has eye gaze, gestures, postures, and body movements. Eyecontact is very important to interpersonal communication [2]. The face is also another important communicator. An expression is the type of emotion or feelings such as happiness, sorrow, anger, confusion, fear, disgust, and surprise. Facial expression and recognition systems (FER) have to contain feature extraction and classification (emotion detection). FER methods have two approaches first is conventional and the second one is deep learning-based approaches. We discuss methods and the Terminology used in the FER system [3].

II. RESEARCH BACKGROUND OFFER

Usually, the FER system uses a conventional method or deep learning method [3].The convention approach first detected the face and its components from the image, and then they extract the features and classify

emotions. This method is less dependent on data and hardware because it is based on manual feature extraction and it required small data for analysis. The deep convolutional neural network approach removes and minimizes the dependence onpre-processing techniques by employing "end-to-end" trainable learning. The input data train with a different source for classification [4]. The convolution neural network (ConvNet) is the most popular deep learning technique [2].

A. The facial expression recognition system

The Face-detection and recognition approaches include three key stages such as image pre-processing, feature extraction, and classification. Fig.1 shows the processes of the facial recognition system [3,4].

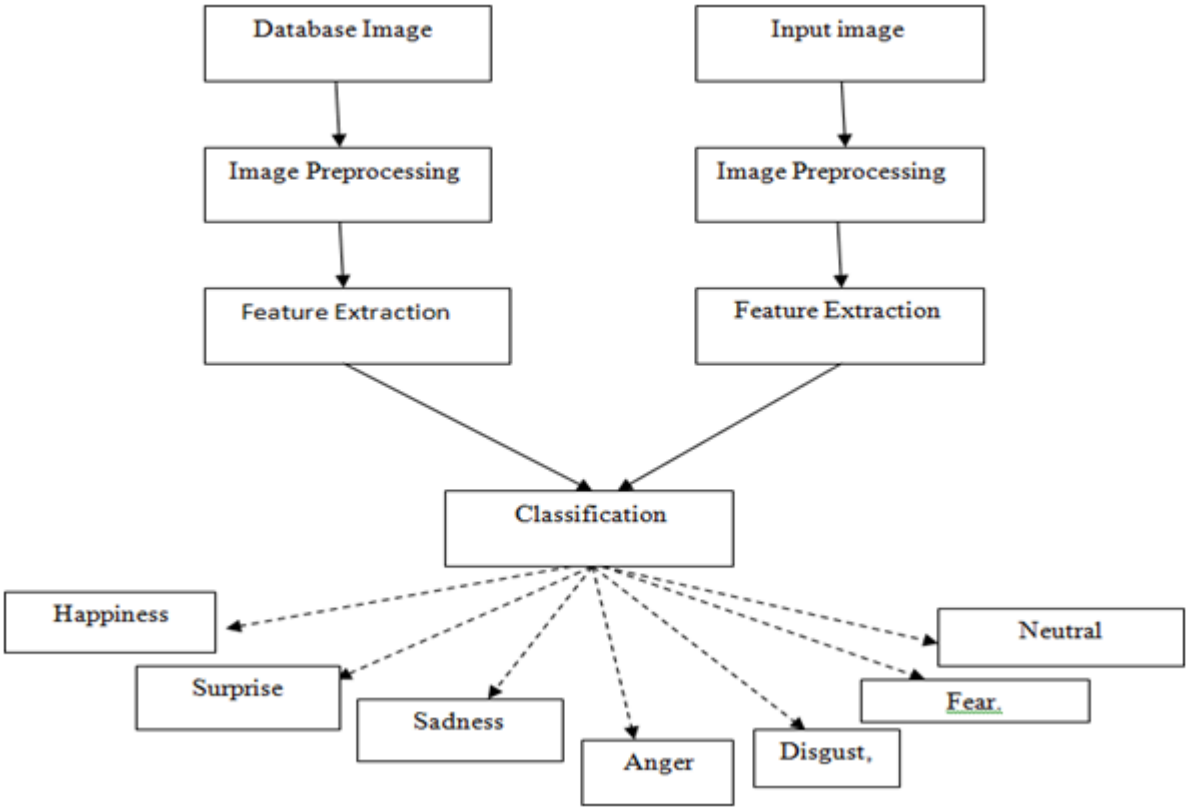


Figure1: Overview of FER system

B. Face image preprocessing

Image preprocessing is done before the feature extraction process to enhance the overall performance of the FER system. The different types of Processes included in Image preprocessing viz. image size and quality and scaling, the brightness of an image, normalization, and additional improvement processes to enhance the recognition rate. The cropping is performed to extract the detected face from the entire image, and image scaling is the process of increasing and decreasing the size of a face image, for this interpolation method is used [4, 5] and Gaussian filter (GF) is used for noisereduction and reducing the size of an image [6]. Normalization is a preprocessing method that reduces the illumination and variations of facial images, it also reduces redundant information such as background, hair, neck, shoulders, sunglasses, etc. The Viola-Jones algorithm is used for Localization and finding a face from an image. It is irrespective of its size, situation, and surroundings. The algorithm uses four key steps (1) haar-like feature (2) create an integral image (3) Adaboost training (4)

cascading classifiers for face identification [7]. The Adaptive-Boosting algorithm and a Haar-like feature are used for representing the rectangular region of an image at a particular location [7.8].The Scale-Invariant Feature Transform(SIFT) algorithm is used for feature detection and face alignment. It is a technique for detection and also describes local feature points in an image. These features must not change with Scale, Illumination, and rotation [9]. The histogram equalization method is an image processing technique that is used to increase contrast and adjust images by using histograms [7]. Region of interest (ROI) is the best preprocessing algorithm. It has based on image segmentation in which different types of functions are performed such as1.The settlement of the face dimension is based on color, thresholding, 2. Forehead, eye, and Mouth segmentation. This method is suitable for the findingof a face in an image[10].

C. Feature Extraction

The FER system is generallyclassified into various categories like Textureand shapedescriptor Methods, edgedetection techniques, Global and Local appearance-based approaches, analytic (geometric) approach, and small regions (patch) approach [3, 10]. Gabor filter and Local Binary Pattern(LBP)based method is used for extracting the features and it is based on the texture analysis technique [11,12]. Texture descriptors for images can be obtained by using thresholds its differences between neighborhood pixels and central pixels[13,14]. The Gaussian Laguerre (GL) [15] uses a single filter and there is no need for multiple filters for facial texture features extraction. Weber Local-Descriptor (WLD) [16] is a robust local descriptor and it depends on the actual intensity of the excitation. WLD consists of two portions 1. Differential excitation 2. Gradient orientation, by using this component a WLD histogram is made. Supervised Descent Method extracted main and related positions of the face and estimates the distance between various components [17]. Weighted Projection-based LBP (WP-LBP) extracts the instructive LBP features and that is weighted [18]. Discrete Wavelet Transform (DWT) is a method for converting image pixels to wavelets and is employed for a large image problem. It compresses the image and provides limited directional information. The feature which is not present in DWT provides a contourlet transform (CT) approach such as multi-scale and directional segregation a high degree of directionality and anisotropy which uses a double filter bank laplacian pyramid and Directional Filter Bank (DFB)[19]. The Local Curvelet Transform (LCT) is a geometric feature descriptor that uses average, entropy, and standard deviation to the local regions [20]. The edge-based methods combine with template matching and geometrical feature matching [21]. Active Shape Model(ASM) and Active Appearance Model(AAM) are used for Face alignment which is used as, statistical model. [7]. Histogram of oriented gradients (HOG) algorithm takes out features from an image which is done by the gradient and orientation of the edges. [21,22]. Features describe changes in facial texture when a particular action occurs viz crease(Jhurriyan), bulges, obverse, the space surrounding the mouth, and eyes. The feature vector is extracted by an image filter which is used to either the entire face or specific areas. Appearance-based algorithms are broad-based, it includes the principal component analysis (PCA) algorithm, which is used for the reduction of higher to lower dimensional space[2], independent component analysis(ICA)[23,24] is an extension to PCA it is the statistical and computational approach. Stepwise LinearDiscriminant Analysis (SWLDA) [25] algorithm takes out the small set of localized features by implementing (forward and backward) regression models, it is based on F-test values and is calculated by class labels. Texture feature-based descriptors are more batter than others so now day Local Directional Number Pattern (LDN), Local Ternary Pattern (LTP) [26], Karhunen-Loeve Transform Extended LBP (KELBP), and in recent years use Discrete wavelet transform (DWT) algorithm decompose input data for recognizing expression using time scale, in the FER system [7].

D. Classification

The last stage of the FER System is emotion detection and the output of the previous stage gives as an input to the classifier in which the classifier classifies expressions such as Fear, Happy, sadness, Surprise, Anger, and Disgust, etc. The Line segment Hausdorff Distance (LHD) is used for image matching and the Euclidean distance metric defines the distance between two sets of the points both methods used for classification purposes [3]. KNN (k-Nearest Neighbors) [17], SVM (Support Vector Machine) [24], Convolutional neural network (CNN) [21], Adaptive Boosting (Adaboost) [8], Bayesian, Sparse Representation based Classifier (SRC)[2]. KNN is a kind of lazy learning it takes more time for predicting. It works on instance-based learning techniques and compares test samples with training datasets, its classification decision is based on a small neighborhood of similar objects(analogy). The KNN algorithm is simple for all machine learning algorithms and easy to implement [17]. SVM (Support Vector Machine) [24] algorithm is suitable for linear and non-linear data for Classification and regression problems. It creates decision boundaries in high n-dimensional spaces and finds out the best decision boundary. This is called a hyperplane which is helpful for classification. Convolutional neural networks (CNNs) performances significantly high in classification [27]. CNNs consist of two steps 1. a convolution and pooling layer extract features and analyze image 2. the output of the first steps takes as input in a fully connected layer, this layer detects emotions from the image [28]. AdaBoost is an ensemble learning algorithm, so it builds one strong learner classifier such an algorithm may be less overfitting problem as compared to other classification algorithms [8]. Naive Bases classifiers algorithms based on Bayes' Theorem. It is highly scalable and required a set of linear parameters for the learning problems [21]. Deep Neural Network (DNN) [28] is a supervised learning method it contains multiple layers and Feed Forward Neural Networks. The Deep Belief Network (DBN) [2] is a robust learning network that contains multiple layers with values and defines as a stack of Restricted Boltzmann machines (RBM) layers that use the backpropagation layer (BP) for commutation. Which has two steps pre-training and fine. It optimizes the weights by minimizing the cross-entropy error [26].

SVM classifier provides better classification and recognition results as compared to other classifiers and CNN gives a better result than other neural network-based classifiers Table 2 shows FER techniques and algorithm that is used in preprocessing, feature extraction, and classification. It shows the different preprocessing methods like Localization, Normalization, Face acquisition, Histogram equalization, ROI segmentation, Haar-like features, Multitask cascaded convolutional network(MTCNN), Viola-Jones, Gabor filter, and Table 2 shows different feature extraction methods like Enhanced Independent Component Analysis (EICA), PCA, PCA-FLDA, ICA, local directional ternary pattern (LDTP) GL Wavelet, Local Binary Pattern(LBP), VTB (vertical time backward), and moments, Principal Component Analysis(PCA), Supervised Descent Method(SDM), WPLBP, Contourlet transforms (CNT) and curvelet transform(CLT), CNN, Attention convolutional network, DSN, DTN, BiLSTM, DBN, HOG, DWT, local directional rank histogram pattern(LDRHP), local directional strength pattern(LDSP), Kernel principal component analysis and generalized discriminant analysis to generate(KPCA-GDA)..

Nowadays preprocessing method recently used the histogram equalization and used for feature extraction LBP method give the best result and for classification SVM and CNN give best result than others

III. TERMINOLOGIES

In this paper, we explain some related terminologies which are used in the theory of the FER system. Facial Landmark points (FLs)[29], Action Units (AUs), Facial Action Coding System(FACS), Basic Emotions(BE), Compound Emotions (CE), and Micro Expressions(ME) [3,29] are basic categories of expression. Facial Landmark (FLs) is eyes & eyebrows, nose, and mouth, lips, jawline, etc. are the special landmark. FLs methods are divided into Shape- and Texture- based methods. The shape-based method is split into explicit and implicit methods. The Explicit-based statistical models which can detect landmarks are the Active ShapeModel (ASM) & Active Appearance Model (AAM)[28,29]. The deep neural network identifies multiple landmarks on the entire face without state information that are implicit-based techniques.

PCA & ICA [23]. Facial Action Units (AUs)[28] are related to the contraction of one or more specific facial muscles. Facial expressions are broken down into individual components of muscle movement [29]. The facial Action coding System (FACS) [3] was developed by Paul Ekman and Wallace Friesen, which describes each and every observable change in facial movement based on Action Units (AUs)[3,29].Compound emotions are those emotions that are formed by the combination of basic emotions such as Happily surprised and Happily disgusted there are two different emotional sense categories. In which 22 Compound Emotions are expressed by humans in 7 basic emotions and 12 compound emotions 3 additional emotions are also included appall, hate, and awe[3].

IV. DISCUSSION AND COMPARISON OF RESULTS

We have studied and compared FER techniques to findscope for improvement in video-based FER system.

Table 2 Study of FER Techniques.

Name of Authors and year	Preprocessing method	Feature extraction Methods	Classification methods	Methods	Datasets	Complexity	FER accuracy	Number of Expressions
Uddin et.al (2009)	Histogram equalization	EICA, PCA, PCA-FLDA, ICA	FLDA	EICA,FLDA,HMMs	CK	Average	93.23%	6
Poursaberi, et. al. (2012)	Histogram Equalization, Viola-Jones.	Gauss–Laguerre Wavelet	K-nearest Neighbor(KNN)	Gauss–Laguerre wavelet, KNN	JAFFE, CK,	Average	96.71% 92.2%	6
Yi and Khalid (2012)	Face acquisition	LBP, Vertical Time Backward (VTB)	SVM	LBP, VTB and Moments, SVM	CK, MMI	Average	95.83%	6
G. Fanelli et.al.(2012)	Localization, Normalization	Gabor filter	Hough transform voting methods	Hough forest voting, Gabor filters	CK MMI	Lower	76% 86.7%	6
SL Happy and A. Routray(2015)	ROIs, Histogram equalization	LBP	SVM	ROI+LBP, SVM	JAFFE, CK+	Less	92.22% 94.39%	6
M Abdulrahman, and A Eleyan(2015)	Localization, Size	PCA and LBP	SVM	PCA+LBP+SVM	JAFFE, MUFE	High	87% 77%,	7
Matamoros et. al. (2016)	Region of Interest(ROI) segmentation	Gabor filters, PCA	SVM	ROI+Gabor functions, SVM	KDEF	Less	99%	7
F. Salmam and M.Kissi (2016)	Viola-Jones	SDM	CART, Decision Tree	SDM+ CART+ Decision Tree	JAFFE, CK+	Less	90%	7
Sunil K. et al. (2016)	Localization, Holistic Based	Weighted Projection - based LBP (WP-LBP)	SVM	WP-LBP, SVM	CK+ JAFFE	Average	97.50% 98.51%	7
S. Biswas and J. Sil (2017)	Histogram equalization	CNT and CLT	SVM	HE+CNT+CLT+SVM	JAFFE	Average	97%	7
UDDIN et.al(2017)	Image acquisition	LDRHP | lLDSP KPCA-GDA	CNN	LDRHP | lLDSP KPCA-GDA +CNN	Ck	Average	95.42%	6
Byungyong et.al (2017)	Localization, Image filters	Local Directional Ternary Pattern (LDTP)	SVM	LDTP+SVM	CK+ JAFFE	Average	94.2 93:2	7
Ke Shan et.al(2017)	Haar-like features and Histogram Equalization	CNN	KNN	CNN+HE+KNN	CK+ JAFFE	Less	80.30% 76.74%	7
Fengping An (2019)	Image acquisition	CNN and LSTM	SVM	CNN and LSTM	CK FER2013	Less	98.9% 86.6	7
Shervin M et.al.(2019)	Localization	CNN, Attention convolutional network	CNN, visualization tech.	[RNN +CNN), visualization tech.	FER2013, CK+,	.Less	70.02% 98.0%	7
Dandan Li (2019)	MTCNN	DSN ,DTN ,BiLSTM	CNN	DSN+BiLSTM	CK+, MMI	High	99.6% 80.71%	6
A. R. Kurup. et.al. (2019)	Viola –Jones	DBN +HOG +DWT	CNN , SVM	DBN +HOG +DWT+CNN	CK+ ,	High	98.57%	7

We observed the precision and execution ofvarious strategies. Figure2 shows the Accuracyrate of different FER methods.

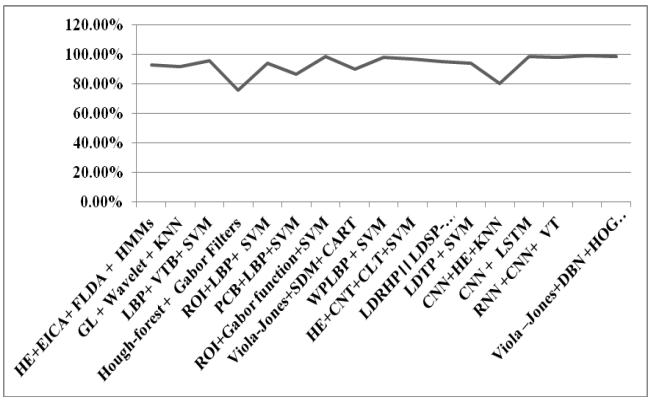


Figure 2 Accuracy rate of various techniques

V. CONCLUSION AND FUTURE WORK

Nowadays Facial Expression Recognition System (FER) is the centre of attraction and for this many algorithms have been developed using different types of parameters to show emotion in real-life applications viz driver monitoring, medicine, robotics Conversation, forensic section, and fraud detection. This paper is useful for analyzation and evolution of any algorithm so researchers can use it for further development of conventional methods and deep learning-based methods of the FER system and introduce some related FER terminology. We study and analyze all various possible methods for pre-processing, feature extraction, and classification. The end of the paper presents a performance analysis of FER techniques that require future research.

The goal of this survey paper is to find the power of algorithms for preprocessing, feature extraction, and classification. ROI segmentation, WPLBP, Histogram equalization with LBP method, CNT and CLT gives high accuracy of 99%, 98.51%, and 97% respectively for preprocessing and feature extraction and GF provides lower complexity which gives accuracy between 76 % to 86%, and the SVM and MTCNN classifier provided 98.9% and 99.62% accuracy respectively with some basic 7 universal emotion such as disgust, sadness, happiness, surprise, anger, fear, and neutral in which CK+ database gives the best result as compared to another database.

VI. REFERENCES

[1]. Chris Frith," Role of facial expressions in social interactions", Phil. Trans. R. Soc, pp 3453–3458, (2009).

[2]. S Zhang, X Pan, Y Cui, X Zhao, L Liu., "Learning Affective Video Features for Facial Expression Recognition via Hybrid Deep Learning", IEEE Access vol 7, pp.32297-32304, (2019).

[3]. F. Khan, "Facial Expression Recognition using FacialLandmark Detection and Feature Extraction via Neural Networks", Xiv:1812.04510v3 [cs.CV], pp.1-7,(2020).

[4]. M. Yeasin, B. Bullot, R. Sharma, "Recognition of Facial Expressions and Measurement of Levels of Interest from Video", IEEE transactions on multimedia, Vol. 8, pp.500- 508(2006).

[5]. S Bashyal, GK Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization", Engineering Applications of Artificial Intelligence Elsevier, pp 1056–1064, (2007).

[6]. S. Biswas, J. Sil, "An efficient face recognition method using contourlet and curvelet transform" Journal of King Saud University Elsevier, pp. 1-12, (2017).

[7]. Ebenezer Owusu, Yongzhao Zhan, Qi Rong Mao, "A neural- AdaBoost based facial expression recognition system" Expert Systems with Applications Elsevier, pp. 3383-3390(2014).

[8]. P Viola, M Jones, Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, pp. 511-518, (2001).

[9]. J. Yang, S.Liao, Stan Z. Li "Automatic Partial Face Alignment in NIR Video Sequences", International Conference on Biometrics Springer, pp. 249–258, (2009).

[10]. Hernandez- Matamoros, A. Bonarini, E. Hernandez, M Nakano- Miyatake, H.Meana " Facial Expression Recognition with Automatic Segmentation of Face Regions using a Fuzzy based Classification Approach" Knowledge-Based Systems, Vol- 110, Pages 1-14,(2016)

[11]. Ke Shan, Junqi Guo, Wenwan You, Di Lu, RongfangBie "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural-Network Structure", IEEE 15thInternational Conference on Software Engineering Research, Management and Applications (SERA) , pp: 123-128(2017).

[12]. Caifeng Shan, Shaogang Gong, Peter W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", Image and Vision Computing, pp. 803– 816, (2009).

[13]. Aliaa A. A. Youssif, Wesam A. A. Asker ", Automatic Facial Expression Recognition System Based on Geometric and Appearance Features", Canadian Center of Science and Education, Vol. 4, pp 115-124 (2011).

[14]. SL Happy, A. Routray," Automatic Facial Expression Recognition Using Features of Salient Facial Patches", IEEE Transactions on Affective Computing, pp. 1-12, (2013).

[15]. A. Poursaberi, A Noubari, M Gavrilova, and N Yanushkevich, "Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression Identification", EURASIP Journal on Image and Video Processing Springer, pp: 1-13, (2012).

[16]. J Chen, Shan, Chu, Zhao, Matti, Xilin, Wen, "WLD: A Robust Local Image Descriptor", ieee transactions on pattern analysis and machine intelligence,pp. 1-15, (2009).

[17]. Fatima Salmam, Mohamed KISSI, "Facial Expression Recognition using Decision Trees",13th International Conference on Computer Graphics, Imaging, and Visualization IEEE, pp. 125–130, (2016).

[18]. Sunil Kumar, M.K. Bhuyan, B. K. Chakraborty, " Extraction of informative regions of a face for facial expression recognition", IET Computer Vision Journals vol 10, pp. 567 – 576(2016).

[19]. S. Katsigiannis, G. Keramidas, D. Maroulis, "Contourlet Transform for Texture Representation of Ultrasound Thyroid Images" 6th IFIP WG 12.5 International Conf., AIAI 2010 Springer, pp. 138–145.,(2010).

[20]. Aysegu, Y. Demir, C. Guzelis, "A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering", Neural Comput&Applic Springer, pp. 131–142,(2016).

[21]. Md. Uddin, W Khaksar, J Torresen, "Facial Expression Recognition Using Salient Features and Convolutional Neural Network", a special section on visual surveillance and biometrics: practices, challenges, and possibilities IEEE Access Vol 5, pp: 26146 – 26161, (2017).

[22]. M. Nazir, Z. Jan, M. Sajjad," Facial expression recognition using histogram of oriented gradients based transformed features", Cluster Comput DOI 10.1007/s10586-017-0921-5 Springer, pp: 1-10, (2017).

[23]. Yi Ji, Khalid Idrissi," Automatic facial expression recognition based on spatiotemporal descriptors", Pattern Recognition Letters Elsevier, pp 1373-1380, (2012).

[24]. M. Abdulrahman, Alaa Eleyan,"Facial Expression Recognition Using Support Vector Machines", 23nd Signal Processing and Communications Appl. Conf. IEEE, pp.1-4(2015).

[25]. Md. Zia Uddin, J. J. Lee, T.-S. Kim," An enhanced independent component-based human facial expression recognition from the video", IEEE Transactions on Consumer Electronics, Vol. 55 IEEE Transactions, pp. 2216-2224,(2009).

[26]. Byungyong Ryu, A Rivera, J. Kim, O. Chae" Local Directional Ternary Pattern for Facial Expression Recognition" IEEE Transactions on Image Processing, pp: 6006–6018, (2017).

[27]. Shervin Minaee, AmiraliAbdolrashidi, "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network" arXiv:1902.01019v1 [cs.CV] ,pp. 1-8,4 Feb (2019).

[28]. Fengping An, Zhiwen Liu," Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM", International Journal of Computer Graphics Springer, pp.1-16, (2019).

[29]. S. Du, Y. Tao, A. M. Martinez, "Compound facial expressions of emotion". PNAS, pp.1454–1462, (2014).

International Conference on Artificial Intelligence and Machine Learning
In association withInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Face Mask and Social Distance Recognition using Deep Learning

Fathima Sholapur[1], S A Quadri[2]

[1]Department of Computer Science and Engineering, VTU/ Secab Institute of Engineering and Technology, Vijayapura, Karnataka, India

[2]Department of CNE, VTU / Secab Institute of Engineering and Technology, Karnataka, India

ABSTRACT

As the COVID-19 pandemic has caused a great despair in a variety of industries all over the ecosphere. The World Health Organization (WHO) certifies wearing a face mask and the practice of corporeal distances in order to reduce the virus's blowout. In this development, a workstation vision arrangement is developed for the automatic recognition of the violation of a mask to wear, and the corporeal distance among employees in the organization. For the face recognition, the broadside is collected and remarked on for over 1000 illustrations; set of data is obtained as an input, maximum up to 1853 photos. Then, it is trained and tested with a multi-Tensor Flow using state-of-the-art object recognition replicas on the aspect concealment to the set of data, and opt for the Nearer R-CNN Inception, ResNetV-2 to a system, which is supplied with an accuracy of up to 99.8%. The Euclidean distance is used to calculate remoteness among various objects under study. A barrier of six feet was kept as a safe distance between the objects. The corporeal distance between two or more than two objects is recognized using the R-CNN network. A real- time video of students entering the campus was shot in SECAB engineering campus and data is fed for learning and training of the proposed model. The proposed system is developed to monitor and improve safety measures by providing information about working masses in the organization by distinguishing them for wearing masks and having social distancing.

I. INTRODUCTION

The spread of COVID-19 has resulted in more than 1,841,000 global deaths and more than 3, 51,000 deaths in the US by Dec. 31, 2020. The spread of virus can be avoided by mitigating the effect of the virus in the environment or preventing the virus transfer from person to person by practicing corporeal distance and wearing façades. WHO defined corporal unsociability as keeping at least six feet or two meters distance from others and recommended that keeping the corporeal distance and impairment of a face mask can remarkably turn down transmission of the COVID-19 virus. Like other sectors, the construction industry has been affected, where unnecessary projects have been suspended or mitigated people's interaction. However, many infrastructure projects cannot be suspended due to their crucial role in people's life. Therefore, bridge maintenance, street widening, highway rehabilitation, and other essential infrastructure projects have been activated again to keep the transportation system's serviceability. Although infrastructure projects are activated, the safety of construction workers cannot be overlooked. Due to the high density of workers in construction projects, there is a high risk of the infection spread in construction sites. Therefore, systematic safety

monitoring in infrastructure projects that ensure maintaining the somatic detachment and wearing façade masks can enhance construction workers' safety. In some cases, safety officers can be assigned to infrastructure projects to inspect workers to detect cases that either social distancing or face mask wearing is not satisfied. However, once there are so many workers on a construction site, it is difficult for the officers to determine hazardous situations. Also, assigning safety officers increases the number of people on-site, raising the chance of transmission even more, and putting workers and officers in a more dangerous situation. Recently, online video capturing in construction sites has become very common.

Drones are used in construction projects to record online videos to manage worksites more efficiently. The current system of online video capturing can be used for safety purposes. An automatic system that uses computer vision techniques to capture real-time safety violations from online videos can enhance infrastructure project workers' safety. This study develops a model using Faster R-CNN to detect workers who either don't wear a face mask or don't maintain the corporeal distance in road projects. Once a safety violation occurs, the model highlights who violates the safety rules by a red box in the video.

II. METHODS AND MATERIAL

This research obtains a facemask dataset available on the internet and increases the number of data by adding more images. Then, the project trains multiple Faster R-CNN object recognition models to choose the most accurate model for face mask recognition. For the corporeal distance recognition, the paper uses a Faster R-CNN model to detect people and then uses the Euclidian distance to obtain the people's distance in reality based on the pixel numbers in the image. Transfer learning is used to increase accuracy. The exemplary was functional on multiple videos of students entering in SECAB institute of engineering and technology, to show the performance of the model.

A part of the dataset of face masks was obtained from Make ML website that contains 853 images that each image includes one or multiple normal faces with various illuminations and poses. The images are already explicated with faces with a mask, without mask, and incorrect mask wearing. To increase the training data 1,000 other images with their annotations were added to the database. The total of 1,853 images was used as the facemask dataset. Some samples of images with their annotations are illustrated in

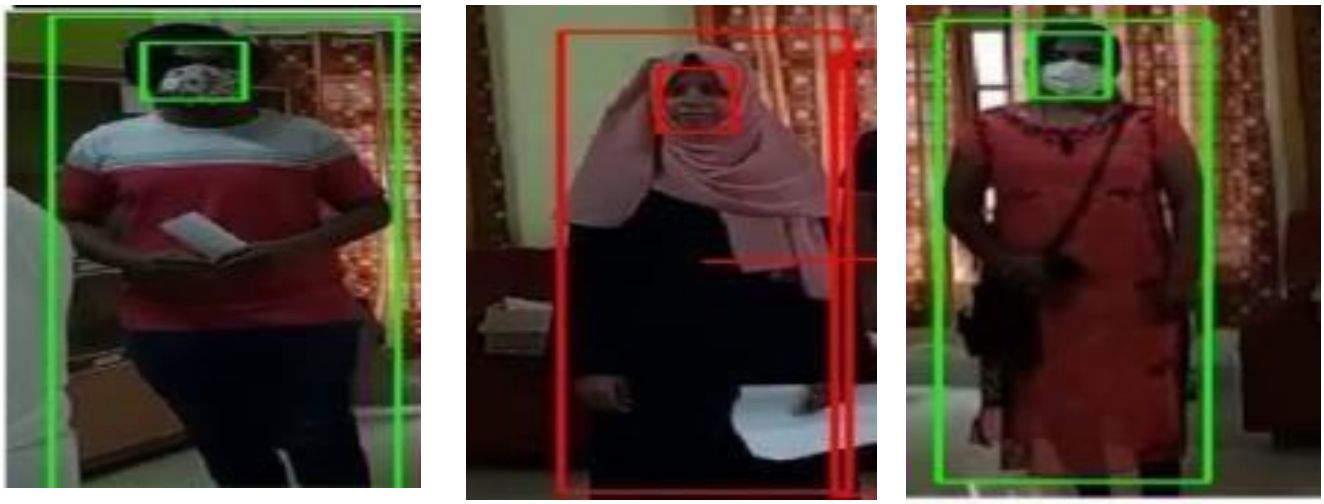


Figure1, where types of face mask wearing are with mask and without mask.

Figure-1 Corporeal distance recognition

The project used the corporeal distancing detector model developed by Roth. The model detects the corporeal distancing in three steps; people recognition, picture transformation, and distance measurement. Roth trained models available on the Tensor Flow object recognition model Zoo on the COCO data set that includes 1,20,000 images. Among all the models, the Faster R-CNN Inception V2 with coco weights was selected as people recognition through model evaluation due to its highest Camera from an arbitrary angle to the bird's eye view. Figure 2 shows the original image captured from a perspective to the vertical view of bird's eye, where the dimensions in the picture have a linear relationship with real dimensions. The rapport amid pixel of (x, y) in the bird's eye picture and pixel of (u, v) in the original picture is defined as:

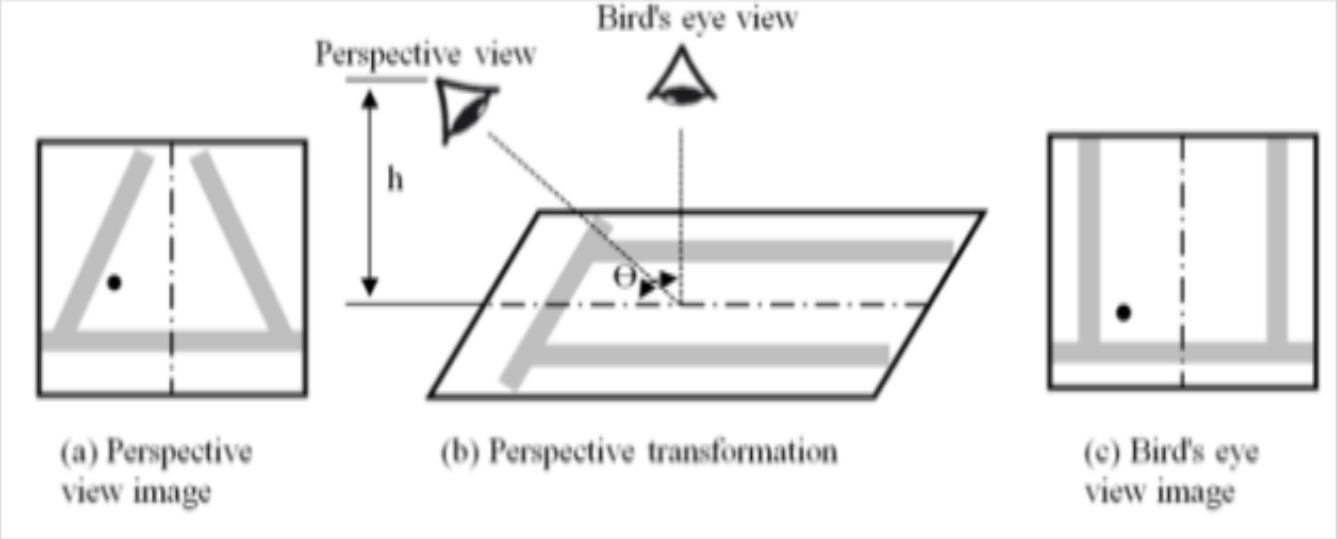


Figure 2. A perspective image transformation to a bird's eye image

Where x=x'/w' and y=y'/w'. For the transformation matrix, the Open CV library in Python was used. Finally, the distance between each pair of people is measured by estimating the distance between the bottom-center points of each boundary box in the bird's eye view. The actual corporeal distance, i.e., two feet, was approximated as 120 pixels in the image.

Faster R-CNN model

Figure 3 shows a schematic architecture of the Quicker R- CNN model. The Quicker R-CNN includes the Region Proposal Network (RPN) and the Fast R-CNN as detector network. The response image is passed through the Convolutional Neural Networks (CNN) Mainstay to extract the features. The RPN then suggests bounding boxes that are used in the Section of Interest (SOI) pooling layer to execute pooling on image's structures. Then, the network passes the output of the ROI pooling layer through two Fully Connected (FC) layers to provide the response of a pair of FC layers that one of them determines the session of each entity and the other one accomplishes a deterioration to improve the anticipated boundary line boxes.

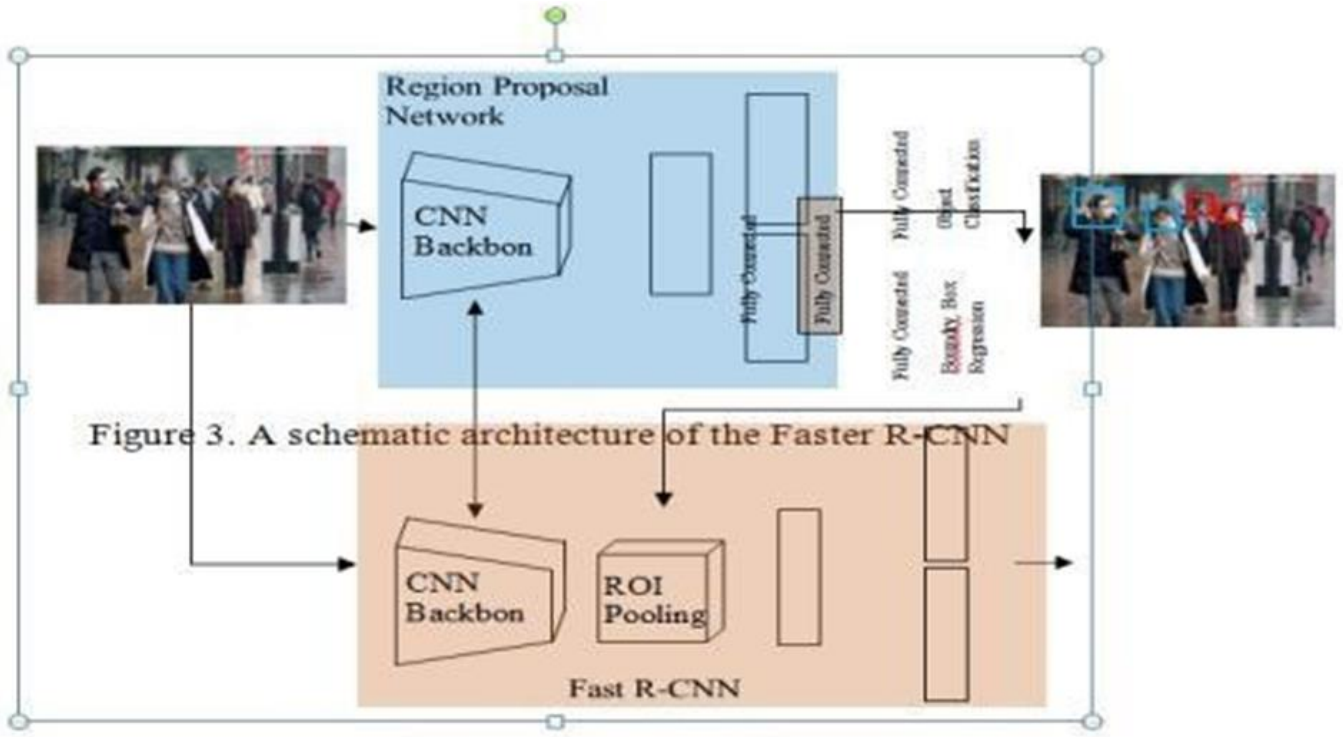


Figure-3.Faster R-CNN model

III. RESULTS AND DISCUSSION

For both networks of face mask recognition and corporeal distance recognition the Google Colaboratory was used. Google Colaboratory is a cloud service developed by Google Research that provides python programming environment for executing machine learning and data analysis codes and provides free access to different types of GPUs including Nvidia K80s, T4s, P4s and P100s and includes leading deep learning libraries. In this experiment, for the face recognition, we used the batch size of 1, the momentum optimizer of value 0.9 with cosine decay learning rate (learning rate base of 0.008), and the image size of 800*1333. The maximum number of steps was 200,000 and the training of the model was stopped when the classification loss reached below 0.07 that happened in near the 42,000th step (Figure 4). For the corporeal distance recognition model, we used the batch size of 1, the total number of steps of 200,000, and the momentum optimizer of value 0.9 with manual step learning rate. The first step was from zero to 90,000, where the learning rate was 2e-4, the second step was from 90,000 to 120,000, where the learning rate was 2e- 5, and the third step was from 120,000 to 200,000, where the learning rate was 2e-6.

Observed Result -1 (2 are safe and 3 are unsafe and 4 are wearing mask

- In this green color anchor box indicates the individuals is safe and individual is wearing mask and maintains social distance.
- Red color anchor box and RED color line indicates individuals are not wearing mask and not maintained social distance.

Observed Result -2(2 are safe and 0 are unsafe and 0 wearing mask)

- In this green colour anchor box indicates the individuals are safe and individual maintained social distance.
- Red colour anchor box indicates individuals are not wearing mask.

IV. CONCLUSION

This project developed a model to detect and differentiate two types of people; one with the face mask and safe distance and the other with without facemasks and social distance, this would be great information for security personnel to enforce safety measures in this COVID-19 pandemic. The project processed dataset including images of people with mask, without mask, and incorrect mask wearing. To increase the training dataset, 1,000 images with different types of mask wearisome were composed and additional to the dataset to create a dataset with 1,853 images. A Faster R-CNN Inception Res Net V2 network was chosen mid different models that generated the accuracy of 99.8% for face mask recognition. For corporeal detachment recognition, the Faster R-CNN

Inception V2 was used to detect people and a transformation matrix was used to remove the effect of the camera angle on actual distance. The Euclidian distance converted the pixels of the transformed image to people's real expanse. The classic set a threshold of six feet as a criterion for distance violation. Transfer learning was used for training models. Four videos of actual students entering in SECAB engineering campus Vijayapur, were used to evaluate the combined model. The output of the four gears indicated an average of more than 90% accuracy in detecting different types of mask tiring in the students. Also, the model accurately detected students/staff that were too close and didn't practice the corporeal distancing. The security personnel can use the model results to monitor students/staff to avoid infection and enhance campus safety. Future

studies can employ the model on other public domain organizations such as construction companies and factories. Furthermore, the accuracy of recognition can also be increased by tuning various object parameters.

V. REFERENCES

[1]. M. Afkhamiaghda and E. Elwakil, "Preliminary modeling of Coronavirus (COVID-19) spread in construction industry," J. Emerg. Manag., vol. 18, no. 7, pp. 9–17, Jul. 2020, doi: 10.5055/JEM.2020.0481.

[2]. M. Kamari and Y. Ham, "Automated Filtering Big Visual Data from Drones for Enhanced Visual Analytics in Construction," in Construction Research Congress 2018, Mar. 2018, vol. 2018- April, pp. 398–409, doi: 10.1061/9780784481264.039.

[3]. M. Kamari and Y. Ham, "Analyzing Potential Risk of Wind-Induced Damage in Construction Sites and Neighboring Communities Using Large-Scale Visual Data from Drones," in Construction Research Congress 2020: Computer Applications - Selected Papers from the Construction Research Congress 2020, 2020, pp. 915–923, doi: 10.1061/9780784482865.097.

[4]. Y. Ham and M. Kamari, "Automated content-based filtering for enhanced vision-based documentation in construction toward exploiting big visual data from drones," Autom. Constr., vol. 105, p. 102831, Sep. 2019, doi: 10.1016/j.autcon.2019.102831.

[5]. L. Liu et al., "Deep Learning for Generic Object Recognition: A Survey," Int. J. Comput. Vis., vol. 128, no. 2, pp. 261–318, Feb. 2020, doi: 10.1007/s11263-019-01247-4.

[6]. Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Recognition with Deep Learning: A Review," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11. Institute of Electrical and Electronics Engineers Inc., pp. 3212–3232, Nov. 01, 2019, doi: 10.1109/TNNLS.2018.2876865.

[7]. L. Jiao et al., "A survey of deep learning-based object recognition," IEEE Access, vol. 7, pp. 128837–128868, 2019, doi: 10.1109/ACCESS.2019.2939201.

[8]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object recognition and semantic segmentation," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Sep. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

[9]. Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Recognition in 20 Years: A Survey," 2019, Accessed: Aug. 02, 2020. [Online]. Available: http://arxiv.org/abs/1905.05055.

[10]. R. Girshick, "Fast R-CNN," 2015. Accessed: Aug. 03, 2020. [Online].Available: https://github.com/rbgirshick/.

International Conference on Artificial Intelligence and Machine Learning
In association withInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Smart Office Security System Using Face Recognition

Prajwal R M[1], Savitha C[2], M Z Kurian[3]

[1]M.Tech (VLSI & Embedded System) Student, Department of Electronics and Communication, SSIT, Tumkur, Karnataka, India

[2]Assistant Professor, Department of Electronics and Communication, ECE, SSIT, Tumkur, Karnataka, India

[3]HOD, Department of Electronics and Communication, SSIT, Tumkur, Karnataka, India

ABSTRACT

Nowadays, the security is the most important part of organized every day activities. There is an advanced awareness in the automated home technique by using internet of things. One of the major qualities in the automated home technique is the security. At present, in the prominent places like offices, banks, malls etc. the people are facing lot of security complications. To overcome this and to get needed security, especially trained laborers are required. The level of security is affected by the mistakes made by human beings as a laborer. So, these mistakes can be reduced by using face recognition security system, which can notice the robber to confined areas. There are two type of components are involved in this system. The hardware components consist of Arduino microcontroller, webcam, PIR sensor and LCD display. The face detection and face recognition algorithm are involved in the software component. In the face recognition system, webcam is used to capture the images rapidly, if any motion is noticed by PIR sensor in a specific area. After that in the capture image, the face is detected and recognized. At the end of the process, the images and notifications are sent to the smart phone of an authorized person by using IoT based application. On the other hand, the temperature sensor is used to measure the present temperature and the gas sensor is used to measure the present gas in the office. Finally, all the results are displayed on the LCD display.

Keywords - Face recognition, Security, IoT, webcam, LCD.

I. INTRODUCTION

Nowadays, the home security is the important thing in human life. At present, the security elements are needed in various regions. The privacy factor also influences the importance of security system. The important link in security chain is the detection of the client who will enter the office. One of the most advantageous electronic technologies is the IoT which enhance the security of the people. So, this technique is used in the smart offices, homes etc. The computer vision provides further security system in the IoT platform. Many years ago, the people were using traditional security system for the security of their homes. The traditional method includes the use of nonliving things such as ID card, password etc. The actual face recognition is included in the department of biometrics as a part. Biometrics is the capability of the computer to identify the human face by using particular features in the face. At present, biometrics is the rapidly evolving and most leading technique which is very useful to the people. The advanced improvements in the face recognition system is

used in many places for the confirmation of user for the security. Such a places are ATM security, attendance system, automated home security etc.

II. RELATED WORKS

In (1) N A Othman and I Aydin presented "A New IoT combined Face Detection of people by using computer vision for security application". The Raspberry PI3 together with PICAMERA are used in the operation of this system. Particularly in this method, the detection of face is carried out by using Haar cascade algorithm and Adaboost algorithm. In

(2) N A Othman and I Aydin presented "A New IoT combined Body Detection of people by using computer vision for security application". PIR sensor setup with the Raspberry PI is used in the implementation of this technique. The body detection is the chief target of this system and is achieved by using HOG algorithm and SVM algorithm. In (3) Amritha Nag et al. presented "IoT based door access control using Face recognition". The OpenCV based face recognition technique is used in the development of the system and the face recognition is carried out by using Haar classifier algorithm. In

(4) Moulik upala presented "IoT solution for smart library using facial recognition". The internet of things is improved by this method and reduction in the manual effort is the main aim of this system.

III. DESIGN OF OFFICE SECURITY SYSTEM

A. Hardware Design:

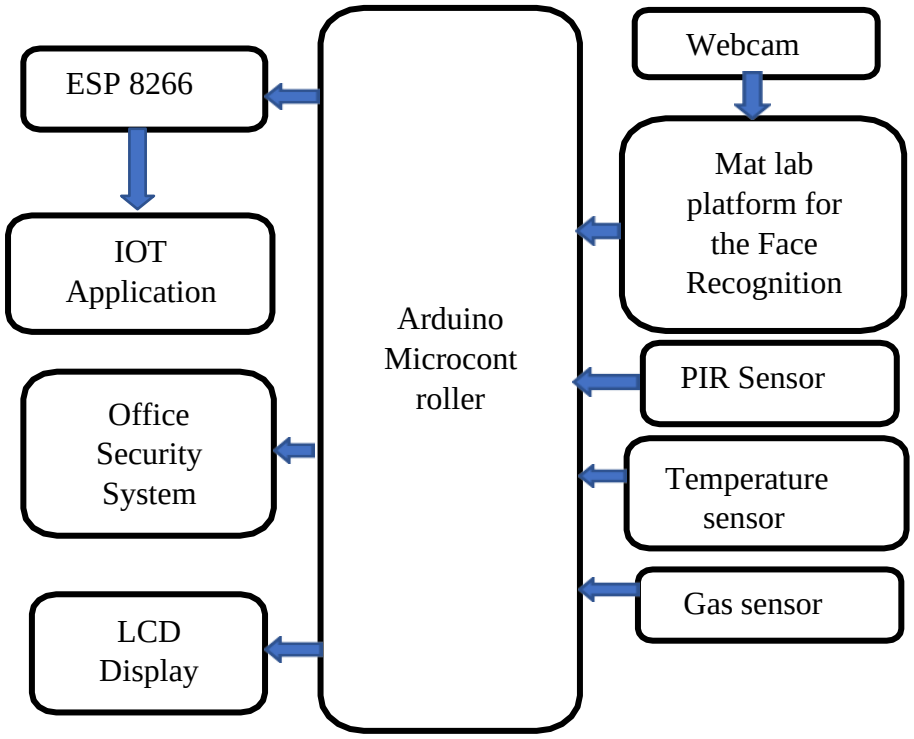


Fig.1: Proposed block diagram of office security system using face recognition.

Fig 1 shows the proposed block diagram of office security system. One of the major parts of this system is the Arduino microcontroller which is connected to the webcam. The Arduino microcontroller has the capacity to control the movement of the webcam. The images are captured by the webcam after any movement is noticed by the PIR sensor. Afterwards, the captured image is applied by the computer vision module, because it has the capability to recognize the face. If the detected face is recognized the door lock in the office is opened. Afterwards, by using IoT based application, the notification and images are sent to the authorized person's smart phone. If the detected face is not recognized the lock of the door in the office is not opened, but then also the images and notifications are sent to the smart phone. The temperature sensor is used to measure the present temperature of the office and gas sensor is used to measure the amount of gas present in the office. Finally, all the results are displayed on the LCD display.

B. Software Design:

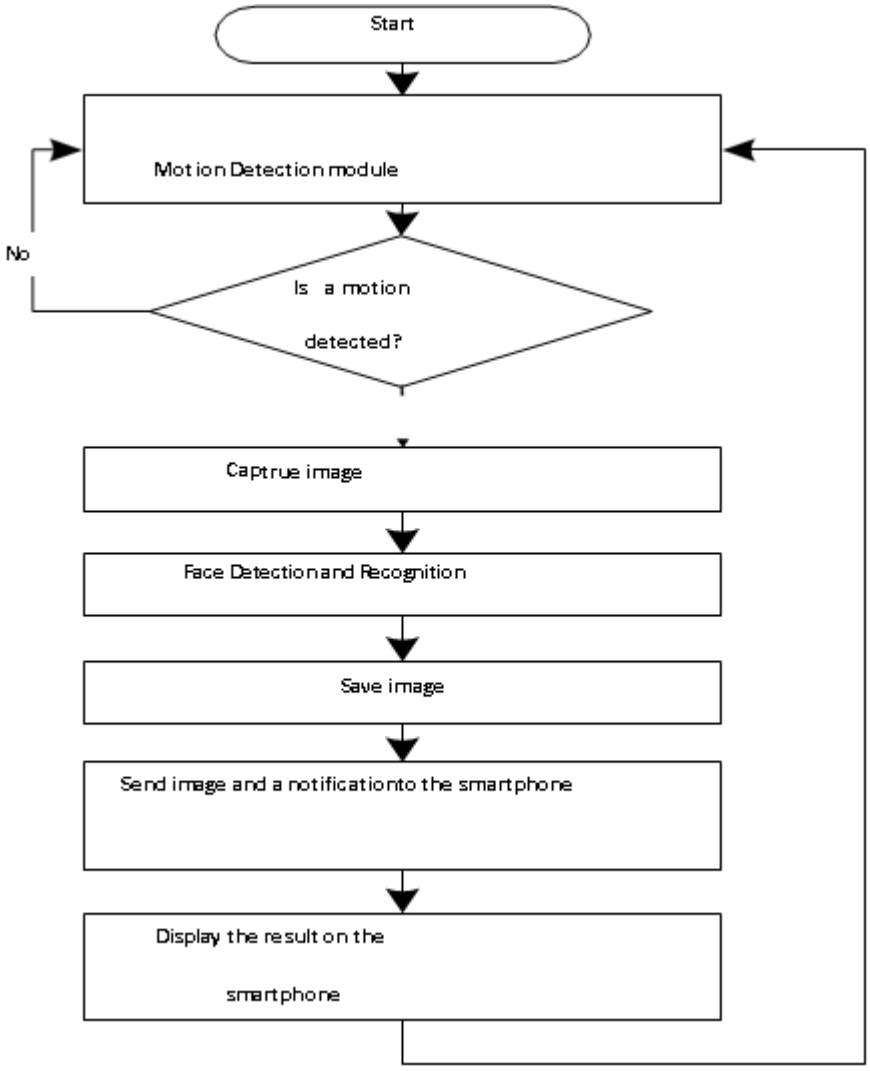


Fig.2: Flowchart for office security system using face recognition

Fig 2 shows the flow chart of the presented office security system using face recognition. If the PIR sensor does not detect any motion, the entire process which start from beginning. If PIR sensor detects any motion, then images are captured by webcam. After that many procedures are applied to the captured image. Whether the face is detected and recognized, the office door lock is opened. Then by using IoT application the notification

and images are sent to the authorized person's smart phone. Whether the face is not detected and recognized, the office door lock is not opened and the entire procedure will start from beginning. At the end of the process whatever the results are displayed on the LCD display.

IV. SYSTEM REQUIREMENTS

A. Hardware requirements:

1) **Arduino Nano:** Arduino Nano is a small microcontroller board based on the ATmega328p and it is bread board friendly. The length of the Arduino nano is 45mm and width is 18mm. The operating voltage of the Arduino nano is 5volt and input voltage is 6 to 20 volts. There are 14 digital Input/Output pins, 8 analog input pins in the Arduino nano and it has the clock speed of 16 MHz, flash memory of 32Kb. Out of the 32kb flash memory, 0.5 kb is used by bootloader.
2) **PIR sensor:** The commonly used PIR sensor in the face recognition system is HCSR-501. The main aims of utilizing this sensor are the detection of human being and movements. PIR sensor is one of the electronic sensors, so the presence of human is recognized by this sensor. The 10 to 15cm is the range of this sensor. Human body emit some high temperature radiations of the range 0.8 to 0.14nm, are captured by PIR sensor and provides 1or 0 for presence or absence of the human.
3) **Webcam:** Webcam is a small camera. In the face recognition system, the webcam is used to capture the images after any movement is noticed by PIR sensor. The Arduino microcontroller is connected to the webcam, because it controls the movement of the webcam.
4) **LCD display:** It is an electronic visual and flat paneled display developed from liquid crystal technology. The alphanumeric displays are used in extensive range of implementations. No light is emitted by liquid crystal immediately, so modulating techniques are used instead of this. The LCD display has the operating voltage of 4.7 to 5.3 volts and it has the 16x2 sharp alphanumeric dot matrix display. So, 224 numerous characteristics with their symbols are displaying on this display. There are two rows in this display and by each row 16 character are printed.

B. Software requirements:

1) **Arduino IDE:** For beginners the use of Arduino IDE is very easy and it is so flexible sufficient for modern users. A message area, a toolbar with button, text editor for writing code and a text consoles are some of the parts of Arduino IDE. It is linked to the Arduino and Genuino hardware to mail the scheme and circulate with them.
2) **IoT application:** IoT application is an android application which can be installed by using any android gadgets. The exchange of video files, photos, messages, audio files are takes place by this application. The complete conversion between convey of user with each other is contented by this application.

V. ADVANTAGES

1. To give high security using the images.
2. To improve public security.
3. More user friendly.

VI. APPLICATIONS

1. Identification of Face: It is used in the identification of people by their face images.
2. Access Control: It is used to get entry in the office and get login in the computer.
3. Security: Security is most important responsibility at airports for office of the airline employees and passengers.
4. General identity verification: Registration of Electoral, e-commerce, new-born identification, passports, worker IDs, banking, national IDs, Conformation of document, etc.

VII. RESULT

Our project result is that detection and recognition of face comparing with images which are stored in data collection. Whether the detected face is recognized then it is found to be known face. So, the image and notification are sent to the authorized person's smart phone by using IoT application and the office door lock is opened. Whether the detected face is not recognized, it is found to be unknown face. So, the image and notifications are sent to the authorized person's smart phone by using IoT application but the office door lock is not opened. On the other hand, the temperature sensor is used to measure the present temperature and the gas sensor is used to measure the present gas in the office. Ultimately, whatever the results are displaying on the LCD display.

VIII. CONCLUSION

The face recognition security system is a very easy and less expensive process. When compared with the other security systems, the face recognition security system provides better security. The homes and office security can be increased by using this system. Other security systems take lot of time and they use lot of features in the human body to identify a person But the proposed security system identifies a person by using only face in a very short time. So, the face recognition security system is user friendly and very useful process for the security.

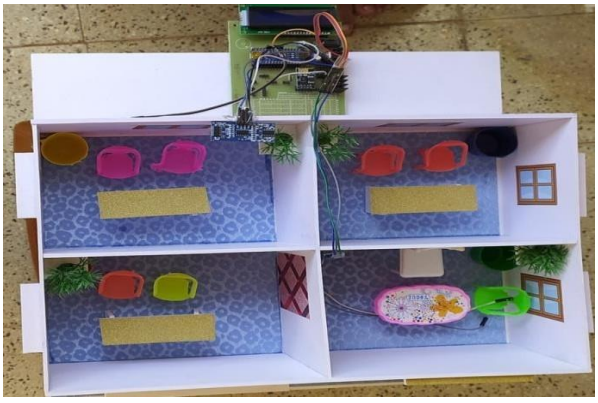


Fig.3: Prototype of office security system

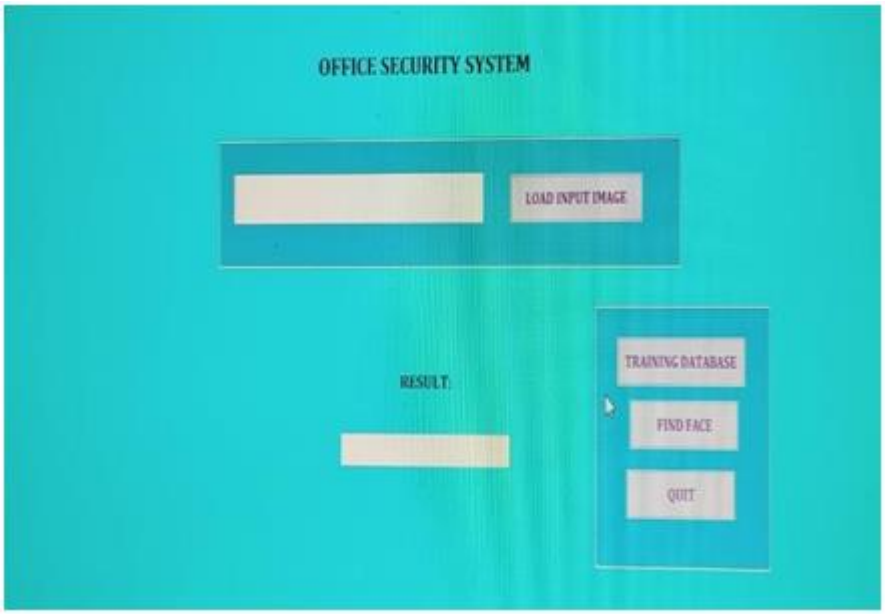


Fig.5: Created database

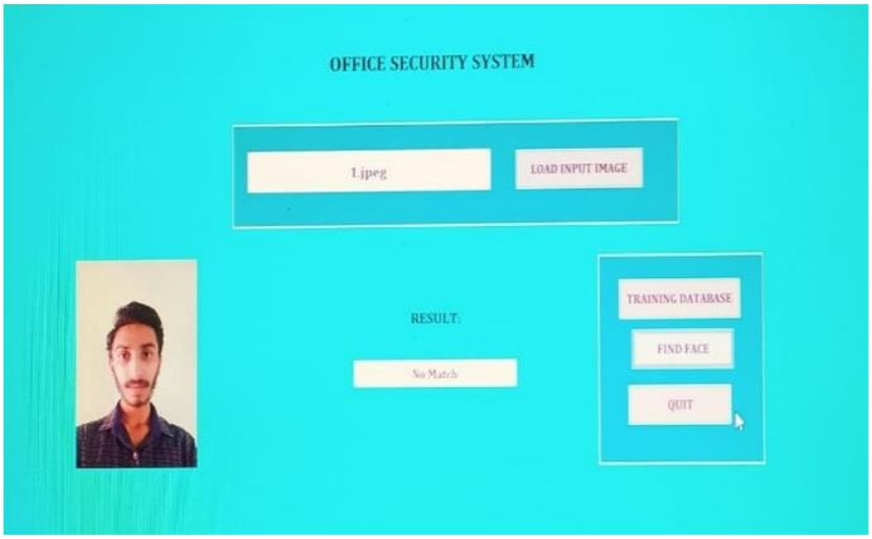


Fig.6: Unknown person captured image

Fig.7: Known person captured image

IX. REFERENCES

[1]. N. A. Othman, I. Aydin, "A New IoT Combined Face Detection of People by Using Computer Vision for Security Application," in Proc. 2017 IEEE International Conference on Artificial Intelligence and Data Processing (IDAP17), pp. 1-5.

[2]. N. A. Othman, I. Aydin, "A New IoT Combined Body Detection of People by Using Computer Vision for Security Application," in Proc. 2017 IEEE International Conference on Computational Intelligence and Communication Networks (CICN 2017), pp.1-5.

[3]. Amritha Nag, Nikhilendra J N, Mrutyunjay Kalmath, "IoT based door access control using face recognition", 2018 3rd International Conference for Convergence in Technology.

[4]. Maulik Upala, WK Wong, ''IoT Solution for Smart Library Using Facial Recognition'', IOP Conf: Materials science and Engineering 495(2019

[5]. Shrutika V. Deshmukh, Dr.U.A.Kshirsagar, "Implementation of Human Face Detection System for Door Security using Raspberry Pi", International journal of innovation Research in Eleectrical, Electronics, Instrumentation and control Engineering.

International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Fake Indian Currency Recognization

Ms. Zarina Begam*[1], Dr. S. A. Quadri[2], Mr. Sayed Md Abrar Qazi[3]

*[1]Department of CSE, VTU/Secab Institute of Engineering and Technology/ Vijayapura, Karnataka, India
[2]Department of CSE, VTU/Secab Institute of Engineering and Technology/ Vijayapura, Karnataka, India
[3]Department of CNE, VTU / Secab Institute of Engineering and Technology / Vijayapura, Karnataka, India

ABSTRACT

India is a developing country, Creation, and print of counterfeit currency of Rs.100, 500 and 1000 is by now there but after demonetization, the counterfeit notes of new Rs.50, 200, 500, 2000 have also come darkness in very tiny time and which achieve the country's monetary growth. From few years due to technological improvement in color printing, duplicating, and scanning, counterfeiting troubles are coming into the picture. In this a work, recognition and confirmation of paper notes with the help of digital image processing technique is described. The personality insertion is perform by picture of notes and it compared by the character of the real notes. The notes will be predictable and verified by using image processing techniques. The move toward consists of works including image processing, edge detection, image segmentation and characteristic extraction and comparing images. The preferred outcome will be text output of the notes recognized and confirmed.

Keywords: Currency Recognition, Image Processing Technique, Image Character Extraction, Segmentation Currency Verification.

I. INTRODUCTION

Technology is getting extremely speedy these days. Consequently, All banking department also modern day by day. Every day there is a drastically changes happening in technologies, The technologies give the techniques to identify the currency. Lot of searching of data for identifying fake note we researched and developed a automatic detection note machine, detect store notes is now widely using in accessing offered of modern categories like candies, soft munchies pot to automobile or railway ticket. These technologies is used for currency recognition essentially aim for identify and extract noticeable and hidden features of cash notes. From gasoline stations to the restricted vegetable seller, each one is cagey of accommodating stock annotations in Rs.50, 200,500 and Rs.2,000 (even be at large past the demonetization) a greater element of them are more or less impractical to tell beginning legitimate store notes. This usually effected of forge on the market is price rises. This allow resolution to is at the moment presented for general chap to discover make out imitation notes is "Fake Note Detector Machine". This device is by and large accessible in banks which can accessible all moment by the usual resident. every these scenario want a category of way out for general public to critic a copied collection note with to exhortation our coins from winning its price.

II. PROBLEM DEFINITION & PROPOSED SYSTEM

Now a days the technology is extremely fast growing in the globe. This increasing of technology the every year administration or bank sector faces the problem of fake currency. This problem is very serious problem in India now a day. Similarly the government is also improving day to day but using high printing skill counterfeit circulate the fake banknote in the Indian bazaar [1]. The Reserve Bank of India (RBI) in latest yearly report said that the during 2017-2018, 17,929 pieces of Rs 2,000 notes be detected in 2017-2018 while only 638 counterfeit notes of the similar value had been detected the year before. In the past, public detecting of fake banknote only guide or a hardware device which is not easy available in bazaar [3]. The technology of notes detection system basically use for recognition and extraction the features of bank note

The projected method of expose there of 2 parts

- Currency Recognition
- Currency Verification

- **Currency Recognition:**

In notes appreciation, we perceive also separate the quantity notes with the facilitate of figure meting out. now we be extract the skin tone of the acquire figure. We be pursue the follow ladder.

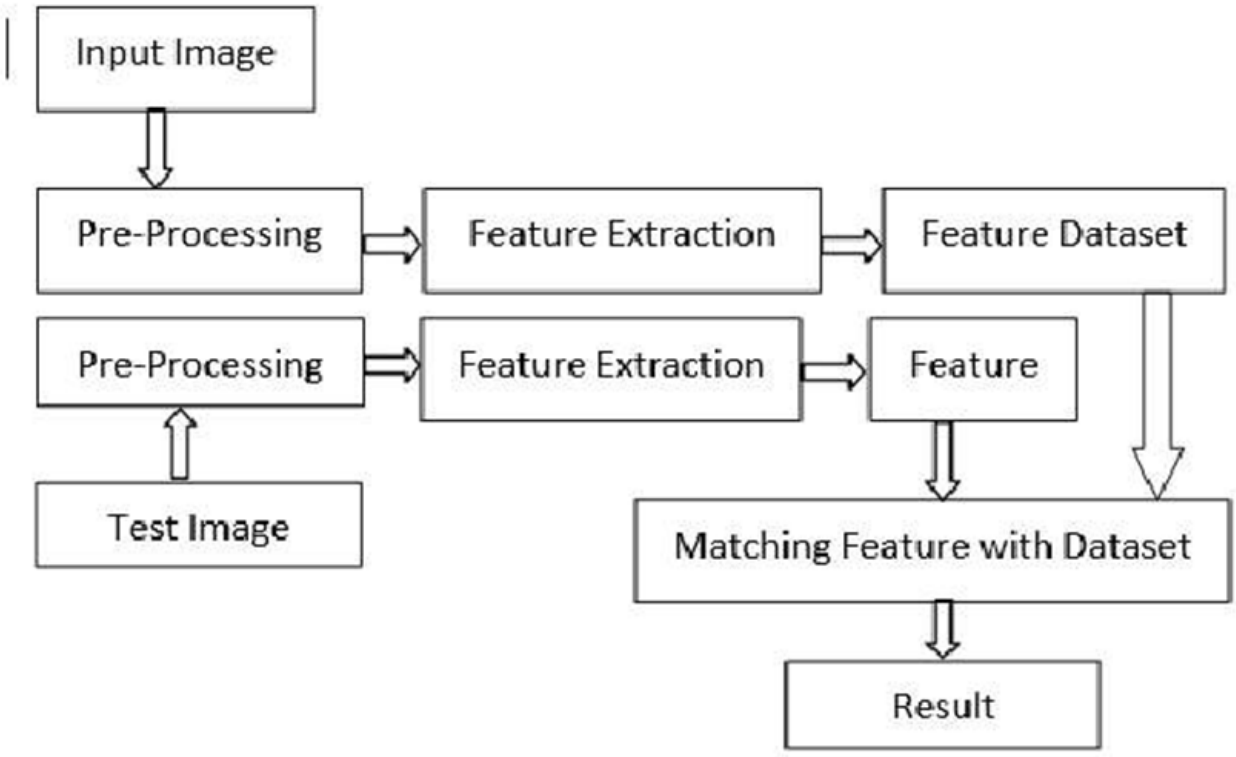


Fig -2: Proposed Detection System

1. **Image Acquisition:** The stage picture acquirement in picture giving out is always the initial footstep in the job flow run since, lacking an picture, no meting out is potential. After the figure has been obtained, a

mixture of method of dispensation can be functional to the illustration to present the countless diverse mental picture tasks.

2. **Pre-processing:** The foremost aim pre-processing to improve the diagram manifestation of descriptions and look up the handling of data sets. picture pre- processing, too call picture restitution, and involve the modification of deformation,

3. **Edge Detection:** border recognition be the names for a put of numerical method over to at identify point in a latest picture on which the figure brilliance change stridently or, new properly, have sum continuity.

4. **Picture Segmentation:** picture segmentation is the practice partition a digital picture keen on many segment (set of pixels, too famous as fantastic pixels). The target of segment is to abridge transform is depiction of access picture keen on great that be extra significant and easy to study.

- **Currency Verification:**

Image Acquisition: performing arts picture attainment in picture meting out is always the first stride in the work flow run as, exclusive of an likeness, refusal meting out is potential.

Pre-processing: The key aspiration of the pre- processed to improve the diagram form of similes and look up the exploitation of figures set.

Haar Skin Extraction: Aar-even skin are latest picture skin texture what are functioning just on picture intens (i.e) RGB standards of all and each pixel here an picture.

Feature Assessment: Even foremost aim is to evaluate the extract skin tone with the store skin tone even are kept in numbers sets which offer the outcome.

III. METHODOLOGY

The organization planned here job on the picture of Indian notes letter acquire by a digital camera. The system which is useful here is as follow.

a. Achievement of picture of Indian notes memo by easy latest handy or scan.
b. Picture want is RGB picture and transformed to Gray over picture.
c. Border discovery of whole aged scale picture.
d. Currently Indian notes skin tone of the essay exchange both view and turn around will be crop and segmented.
e. Spot from side to side inventory: The diminutive flowery devise written equally on the spectator side reverser surface (filled up) with note colour.

IV. RESULTS AND DISCUSSION

Figure 4.1: Image Acquisition of 2000 rs note Attribute drawing out of system capture the chart pleased of descriptions for indexing & recovery. ancient or low-level picture skin can be whichever broad skin tone, The consequences of the quality pulling out are planned and compare with standards give productivity of the Rectifies Fig [4.2].

Fig.[4.2]-Recognized Currency

The facial appearance are extracting with kept meant this evaluation present those confirmation outcome. Below diagram[4.3],[4.4]

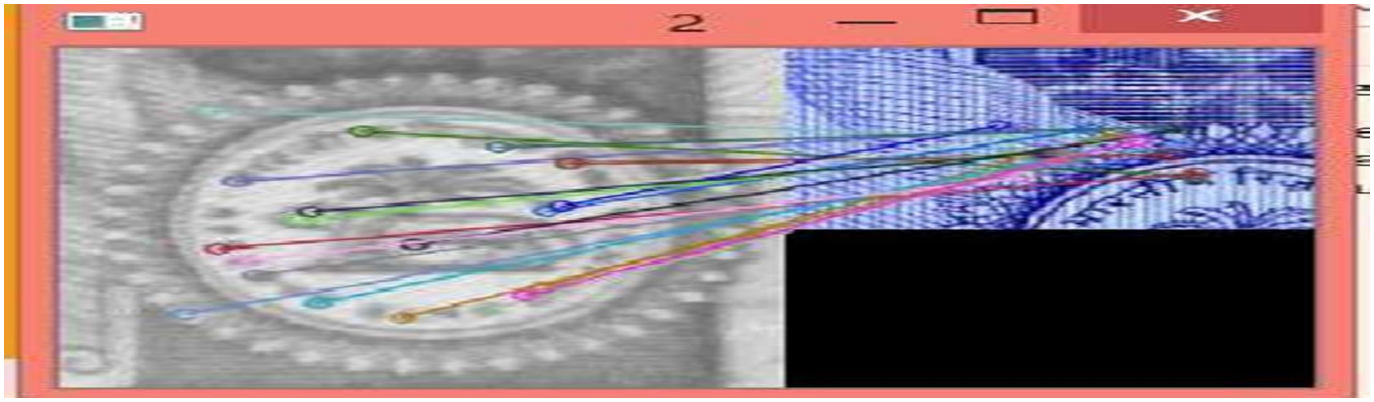


Fig [4.3]-Haar features extracted

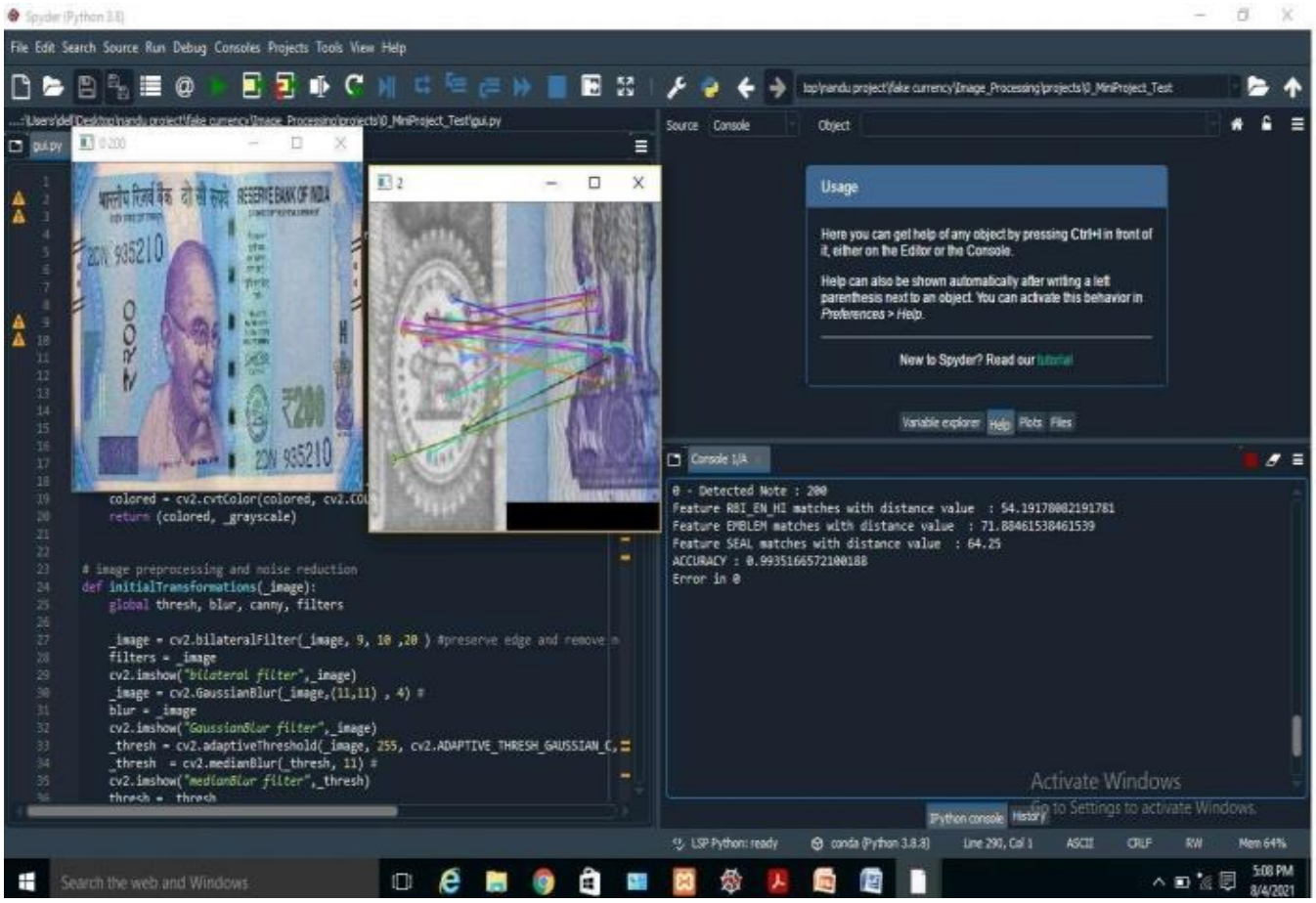


Fig [4.5]Output of Valid Currency

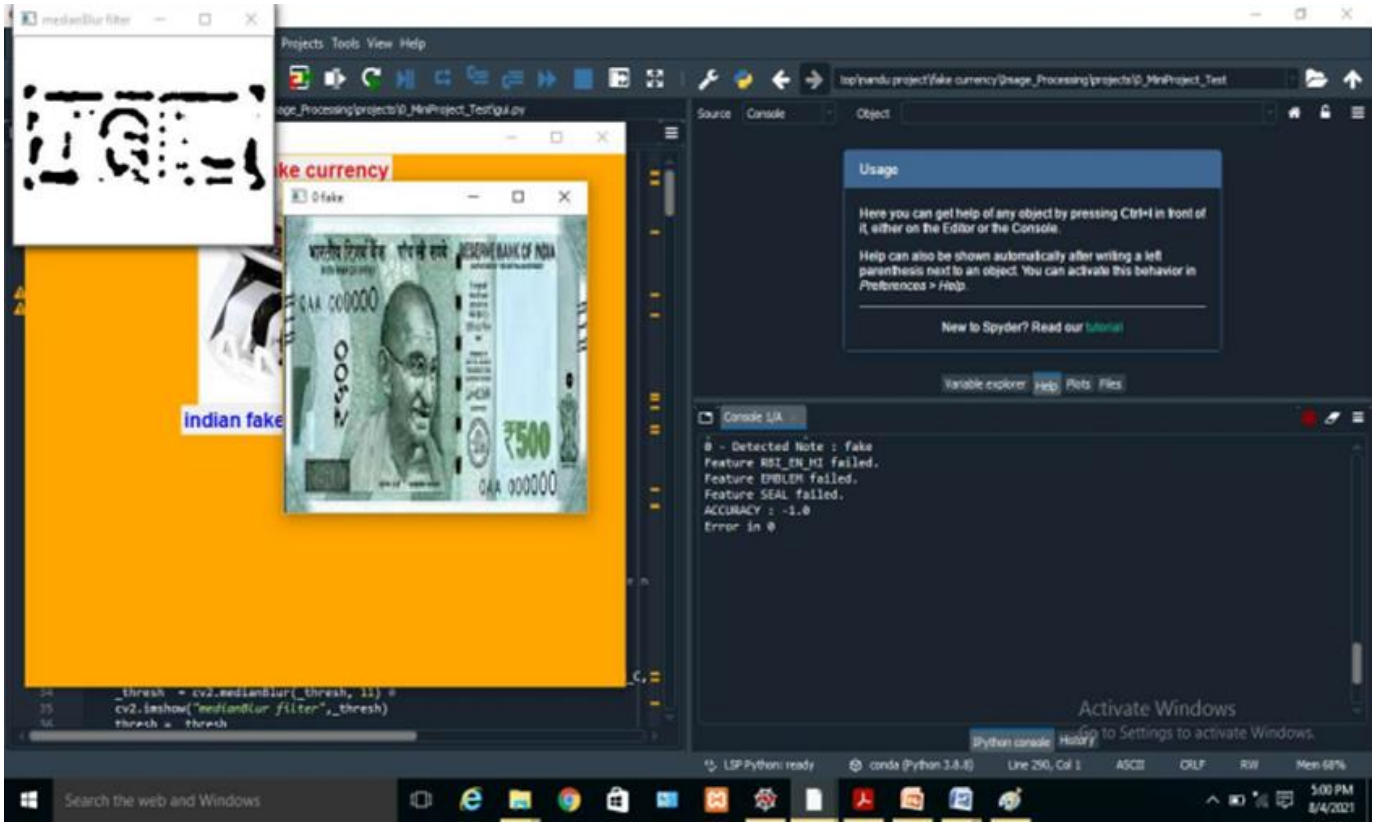


Fig [4.6] Output of Fake Currency

Attribute drawing out of system capture the chart pleased of descriptions for indexing & recovery.

V. CONCLUSION

By means of digital picture meting out, study of exchange picture is further correct as fighting fit as this system is proficient in language of fee and time unbearable compare to presented technique. Day by day study vocation is mounting in this ground and a variety of representation processing technique are implemented in array to find a supplementary correct consequence.

VI. REFERENCES

[1]. Fake currency detection using image processing TusharAgasti et al 2017 IOP Conf. Ser.: Mater. Sci. Eng. 263 052047
[2]. TruptiPathrabe G and SwapniliKarmore 2011 Int. J. CompTrends Tech 152-156
[3]. Tanaka M, Takeda F, Ohkouchi K and Michiyuk 1998 IEEE Tran on Neural Network 1748- 53.
[4]. Jahangir N, Ahsan Raja Chowdhury 2007 IEEE 10th Int. Conf. on Computer and Information Technology
[5]. Rubeena Mirza, Vinti Nanda 2012 IFRSA Int.J. Computing 2 375-80
[6]. JunfangGuo, Yanyun Zhao, and AnniCai 2010 Proc IEEE Int. Conf Network Infrastructure and Digital Content 359-363.
[7]. M.Deborah and SoniyaPrathap "Detection of Fake currency using Image Processing". IJISET-International Journal of Innovative Science, Engineering & Technology, Vol. 1, Issue 10, 2014
[8]. Ingulkar Ashwini Suresh," Indian Currency Recognition and verification Using Image Processing".IRJET-International Research Journal of Engineering and Technology, Vol.3, Issue-6, 2016
[9]. Brindha M "Object Detection using Haar-Like Feature Extraction "International Journal of Computer Applications (0975 – 8887) Volume 153 – No 9, November 2016
[10]. Srishti Roy1"Simulation of Iris Comparison and Recognition System Using MATLAB".SSRG International Journal of Computer Science .

International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques

Shanta Patil[1], S R Vinay[1], Kartik Kulkarni[1], Mehboobi Mulla[1], Amruta Naveen[2], Sahana Beelagi[2]

[1]B.E. Students, Final Year B.E, Department of Computer Science, VTU/Secab Institute of Engineering and Technology/ Vijayapura, Karnataka, India

[2]Professor, Final Year B.E, Department of Computer Science, VTU/Secab Institute of Engineering and Technology/ Vijayapura, Karnataka, India

ABSTRACT

Plant diseases affect the growth of their respective species, therefore their early identification is veryimportant. Many Machine Learning (ML) models have been employed for the detection and classification of plant diseases but, after the advancements in a subset of ML, that is, Deep Learning (DL), this area of research appears to have great potential in terms of increased accuracy. Many developed/modified DL architectures are implemented along with several visualization techniques todetect and classify the symptoms of plant diseases. Moreover, several performance metrics are used for the evaluation of these architectures/techniques. This review provides a comprehensive explanation of DL models used to visualize various plant diseases. In addition, some research gaps are identified from which to obtain greater transparency for detecting diseases in plants, even beforetheir symptoms appear clearly.

Keywords : plant disease detection, classification of diseases, image processing, Deep learning,disease of plant leaf

I. INTRODUCTION

Plant disease is defined as the state of local or systemic abnormal physiological functioning of a plant, resulting from the continuous, prolonged 'irritation' caused by phytopathogenic organisms.pests and diseases pose a threat to food security because they can damage crops, thus reducing the availability and access to food, increasing the cost of food. Plant pests and diseases may also negatively affect the palatability of foods resulting in changes to the traditional food preferences of populations.

In general, pathogens that infect plants do not specialize in infecting people. You are not likely to catch a disease from working with diseased plants in your garden, but it is a potential risk (depending on the infection), and consideration should be taken.The most direct economic impact of a transboundary pest or disease is the loss or reduced efficiency of agricultural production - whether it be of crops or animals - which reduces farm income. The severity of the economic effect will depend on the specific circumstances.

Insects can vector or infect a plant with a pathogen when they feed on an infected host plant, and then move and feed on an uninfected plant. Pathogens can also spread through infected seeds, transplants, or

contaminated equipment, irrigation water, and humans.Plant diseases can affect plants by interfering with several processes such as the absorbance and translocation of water and nutrients, photosynthesis, flower and fruit development, plant growth and development and cell division and enlargement.

Plant diseases are caused by different micro-organisms such as viruses, bacteria and fungi. In addition, various soil-borne and above ground insect pests also affect crop production It causes the reduction of available resources for plants, which fail to produce enough biomass, seeds, and thus yield.Common methods for the diagnosis and detection of plant diseases include visual plant disease estimation by human raters, microscopic evaluation of morphology features to identify pathogens, as well as molecular, serological, and microbiological diagnostic techniques.

Inconsistency and delay in the identification of plant diseases cause a reduction in the quantity and quality of yield. Losses due to plant diseases or other pest accounts for 20 to 40% of global annual productivity. Studies have been carried out to assess the estimated loss caused by different diseases. Yield loss also contributes toward increased consumer prices and a drop in the earnings for crop producers. Accurate and timely identification of plant diseases is crucial for ensuring maximum yield and is beneficial for farmlands in remote areas.

The agricultural land mass is more than just being a feeding sourcing in today's world. Indian economy is highly dependent of agricultural productivity[9]. Therefore in field of agriculture, detection of disease in plants plays an important role.The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops.

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas[1][2]. Whereas if automatic detection technique is used itwill take less efforts, less time and become more accurate. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases. Image processing is used for measuring affected area of disease and to determine the difference in the color of the affected area.Deep learning techniques used to detect plant diseases are more accurate and less time consuming compared to the traditional image processing techniques. Researchers are facing major issues in the field of plant disease like unavailability of data set for each and every disease, background noise in captured images, sometimes texture property of plant leaf varies during the change of environment[4].

II. LITERATURE REVIEW

S.N	Title of the paper Year	Algorithm	Methodology	Result
1.	Plant leaf detection and disease recognition usingdeep learning	Deep learning convolutional neural network model.	Involves 3 key stages: acquisition of data, preprocessing of data and imageclassification	A 96.5% accuracy rate was achieved using 75 epochs during the training of the model. The model also achieved a maximum accuracy rate

	IEEE 2019			of 100% whentesting andom images of plant varieties and diseases.
2.	Disease detection on the leaves of the tomato plantsby using deep learning	One of the most used optimization algorithm is Stochastic Gradient Descent (SGD) algorithm. Briefly, SGD minimizes the loss through iterations by updating means according to gradient.	Deep learning is used to detect diseases from the leaves of the tomato plants. Two different deep learning network architectures AlexNet [9] and SqueezeNet [10] are trained and tested on the tomato images of the Plantvillage dataset. Both training and testing are done on the mobile supercomputer NvidiaJetson Tx1.	Accuracy results are the results from the Caffe tests.AlexNet performed slightlybetter than SqueezeNet. Accuracy on Test Set:AlexNet: 0.9565 SqueezeNet: 0.943
3.	Tomato Leaf Disease Detection Using Deep Learning Techniques ICCES 2020	A convolutional Neural Network algorithm used in this proposed method is a hierarchical feature extraction that maps the pixel values and evaluates the same withthe trained dataset image.	This has been implemented effectively using image processing technique, Feature extraction, segmentation and convolutional neuaral and the latest niche algorithms and usingthe open source programming language Python.	The proposed method is achieved an accuracy level of98%. The proposed method is used as a CNN algorithm for hierarchical feature extraction that map input image pixel intensities and compare the same with the trained dataset image.
4.	Sugarcane Disease Recognition using Deep Learning. IEEE 2019	This paper provides an idea of helping farmers with the aid of deep learning algorithm in detecting and classifying sugarcane diseases.	1. Image Dataset Acquisition 2. Pre-processing of Images 3. Feature Extraction 4. Classification	The highest recorded validation accuracy duringthe training was 95% with60 epochs. shows of detection and recognition of a sugarcane plant with 35.52% accuracy that it is infected with smut diseaseshow an accuracy rate of 98.98% infected with grassy shoot disease. infected with a yellow leaf

				disease on the right image..
5.	Plant Leaf Disease Classification Using Grid Search Based SVM. IEEE 2020	SVM is used as an algorithm to develop a plant disease identification and classification system.	[1] Image acquisition [2] Noise removal [3] Image segmentation [4] Feature extraction [5] Classification(SVM and grid search basedSVM) detection	The accuracy of SVM classifier was 80% and whenapplied with Grid Search hyper parameter tuning accuracy increased to 84%.

III. PROPOSED METHODOLOGY

Diseases in Plants are a major concern to the farmers these days. Many a times, the farmers are not sure which pesticide or insecticide is needed to treat a particular diseased plant because they are not sure of the type of disease. This results in spraying wrong pesticides, damaging the plants which affect the plant yield. To overcome with this problem, we have come up with a solution of developing a system that easily identifies some common diseases that occur in the plants.

Through image processing and Deep learning algorithms, we aim to classify diseases and generate a model that would provide an easy and accurate way of determining the plant disease through on click of an image of the affected plant leaf. This system is not only beneficial to the farmers in saving the crops, but also in saving money by buying only right kind of pesticides suitable to treat the particular disease. As the system does not involve any heavy machineries and electricity, the system proves to not only be a cost-effective solution, but also an environment-friendly one.

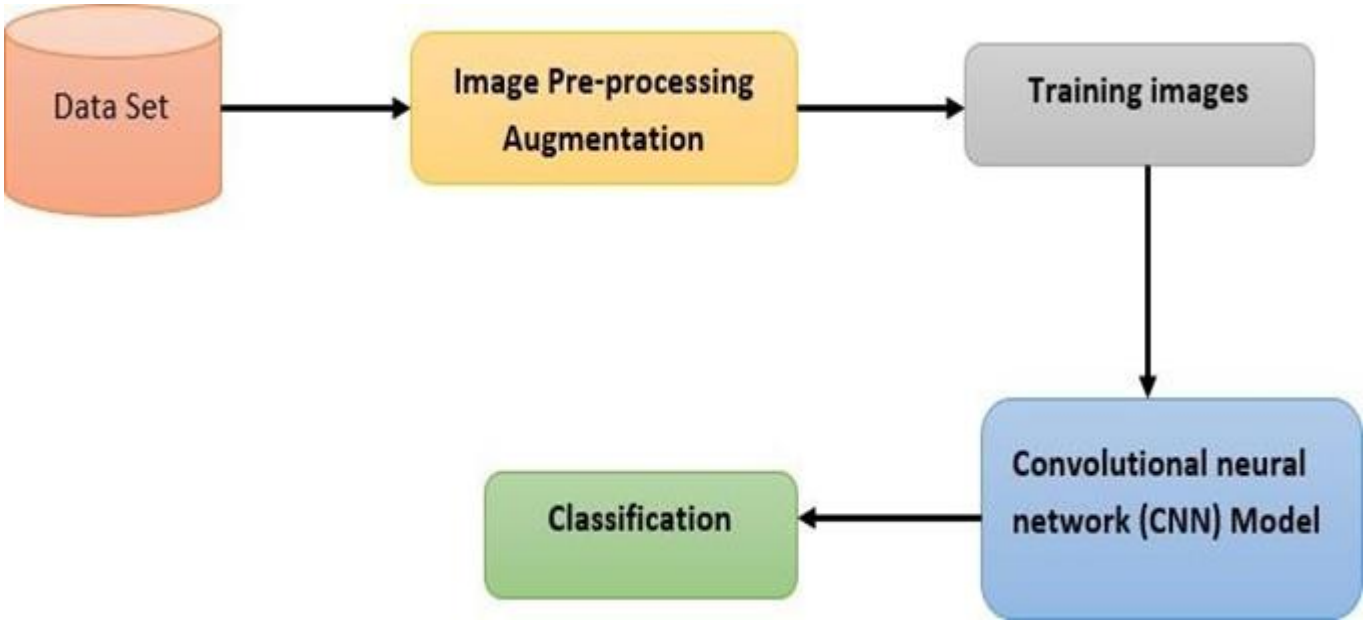


Fig 3.1: Flowchart diagram

Image Database: The next point in the project is creation of the image database with all the images that would be used for training and testing. The construction of an image database is clearly dependent on the application. The image database in the proposed approach consists of 28 different classes of image samples. The image database itself is responsible for the better efficiency of the classifier as it is that which decides the algorithm.

Image Pre-processing: Image pre-processing is the name for operations on images at the lowest level of abstraction whose aim is an improvement of the image data that suppress undesired distortions or enhances some image features important for further processing and analysis task. It does not increase image information content. Its methods use the considerable redundancy in images. Neighbouring pixels corresponding to one real object have the same or similar brightness value. If a distorted pixel can be picked out from the image, it can be restored as an average value of neighbouring pixels. In the proposed approach image pre- processing methods are applied to the captured image which are stored in image database.

Augmentation: Image augmentation is a technique of altering the existing data to create some more data for the model training process. In other words, it is the process of artificially expanding the available dataset for training a deep learning model.

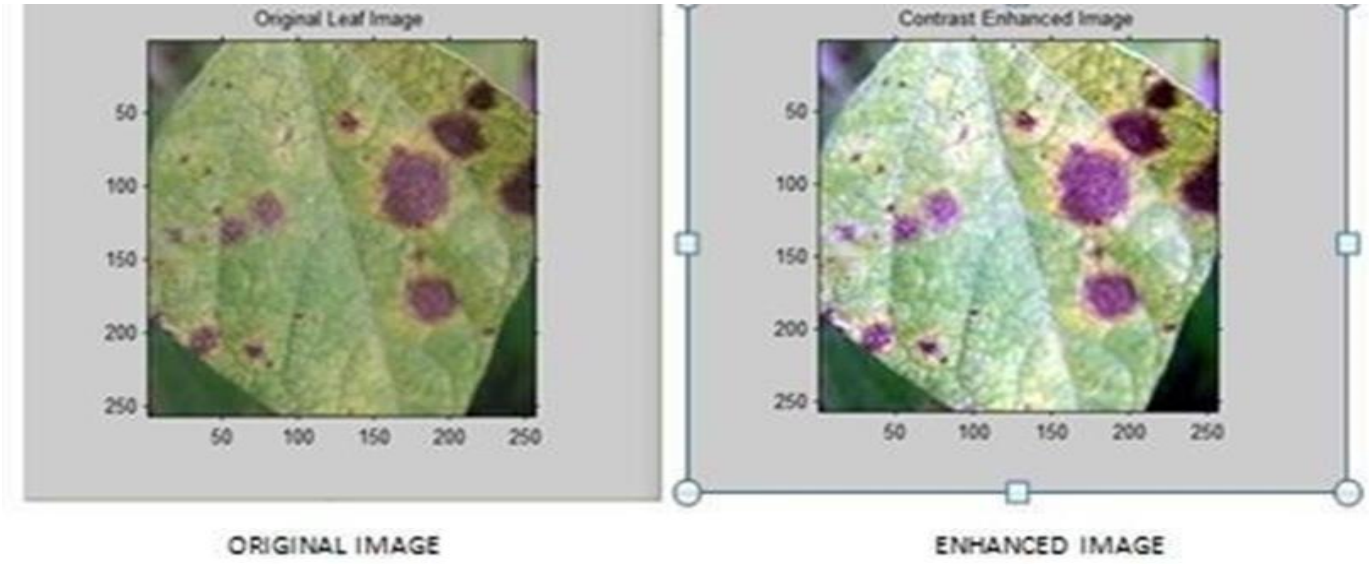


Figure.4.1 Enhanced Image

Classification: Classification is a process of categorizing a given set of data into different types of classes. Classification of image is done using CNN (Convolutional Neural Network).

IV. RESULTS

All the experiments are performed in Deep Learning(DL). For input data disease, samples of plant leaves Apple leaf, Apple rust leaf Bell_pepper leaf, Bell_pepper leaf spot, Corn Gray leaf spot, Corn leaf blight, Squash Powdery mildew leaf, Tomato Early blight leaf , Tomato leaf late blight, Tomato mold leaf and so on are considered.

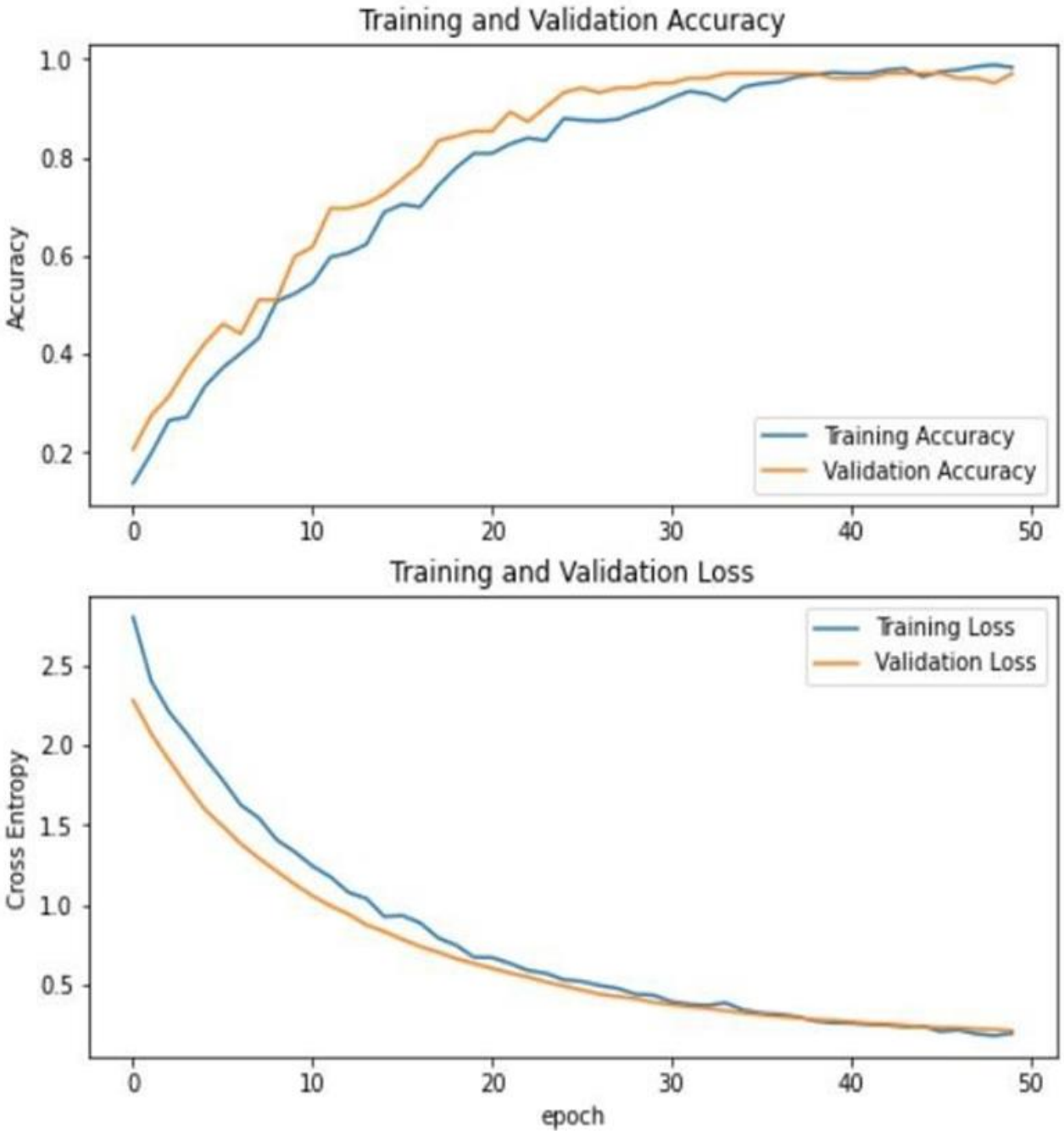


Figure 5.1: graph for epoch 50

Here in figure 5.1 for only 50 times training is happened for the dataset. In the figure 5.1 we can see that the line of training and validation are not a straight line so it says that the trained model is less accurate.

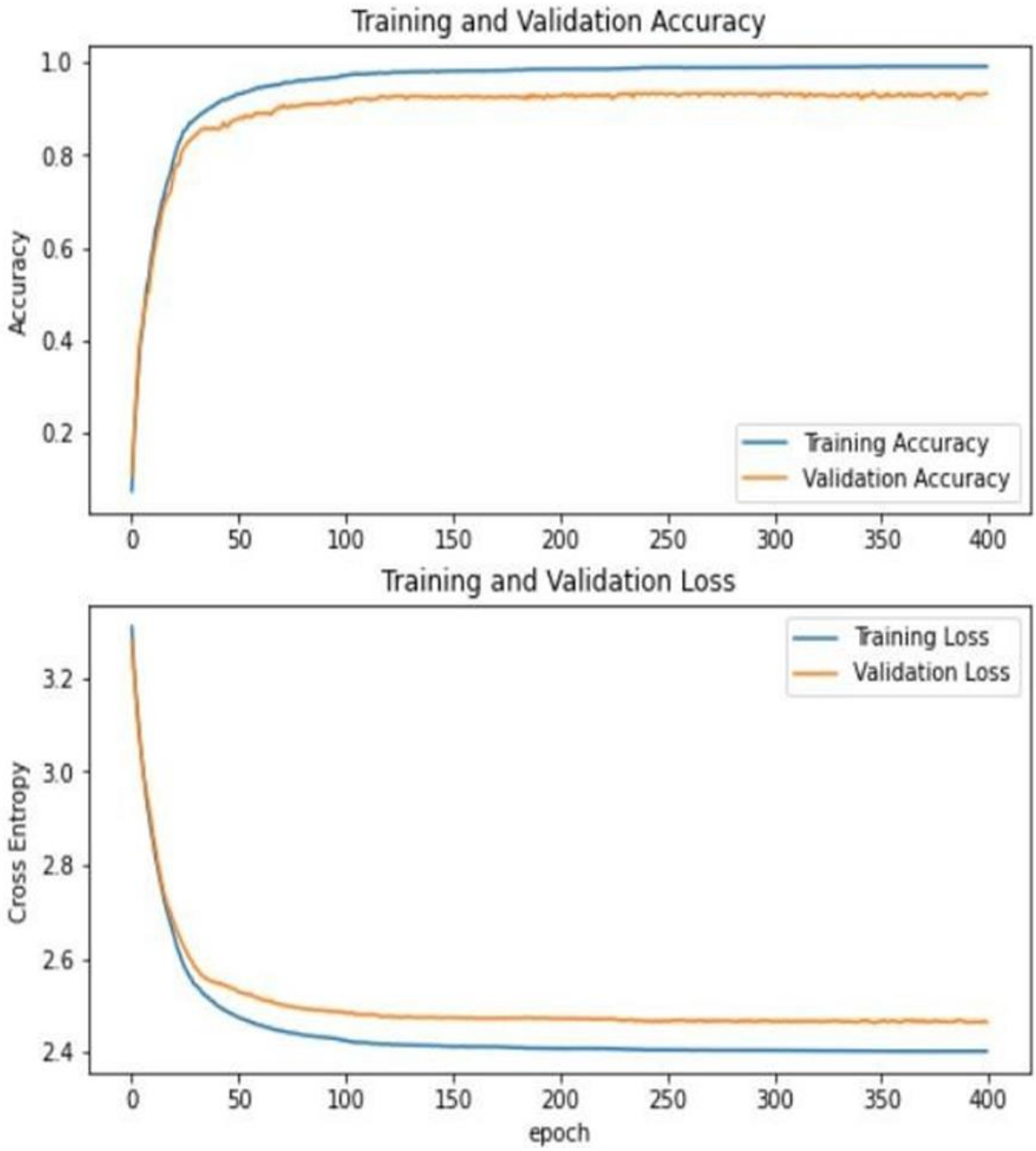


Figure 5.2 graph for epoch 400

Here in figure 5.2 for only 400 times training is happened for the dataset. In the figure 5.2 we can see that the line of training and validation are straight line so it says that the trained model is of more accurate than the first one.

The model was trained using images of disease leaves. The model could classify with approximate 90 percent accuracy. By comparing figure 5.1,5.2 we can say that accuracy can be increased when trained withmore number of epochs.

V. CONCLUSION

The objective of this project is to recognize abnormalities that occur on plants in their greenhouses or natural environment. The image captured is usually taken with a plain background to eliminate occlusion. The algorithm was contrasted with other machine learning models for accuracy. Using CNN classifier, the model was trained using images of disease leaves. The model could classify with approximate 90 percent accuracy. The accuracy can be increased when trained with vast number of images.

VI. REFERENCES

[1]. Jasim, M. A., & AL-Tuwaijari, J. M. (2020, April). Plant leaf diseases detection and classification using image processing and deep learning techniques. In 2020 International Conference on Computer Science and Software Engineering (CSASE) (pp. 259-265). IEEE.

[2]. Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2010, December). A framework for detection and classification of plant leaf and stem diseases. In 2010 international conference on signal and image processing (pp. 113- 118). IEEE.

[3]. Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 382-385). IEEE.

[4]. Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2011). Detection and classification of leaf diseases using K-means-based segmentation and. Information technology journal, 10(2), 267-275.

[5]. Swain, S., Nayak, S. K., & Barik, S. S. (2020). A Review on Plant Leaf Diseases Detection and Classification Based on Machine Learning Models. Mukt shabd, 9(6), 5195-5205.

[6]. Tripathi, G., & Save, J. (2015). An image processing and neural network based approach for detection and classification of plant leaf diseases. Journal Impact Factor, 6(4), 14-20.

[7]. N. S. Sapare and S. M. Beelagi, "Comparison Study of Python in Scientific Computing Using Mathematical Problems", vol. 7, no. 9, pp. 1-10, 2020.

[8]. Prajapati, B. S., Dabhi, V. K., & Prajapati, H. B. (2016, March). A survey on detection and classification of cotton leaf diseases. In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) (pp. 2499-2506). IEEE.

[9]. Oo, Y. M., & Htun, N. C. (2018). Plant leaf disease detection and classification using image processing. International Journal of Research and Engineering, 5(9), 516-523.

[10]. S. Sapare and S. M. Beelagi, "Comparison study of Regression Models for the prediction of post-Graduation admissions using Machine Learning Techniques," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 822-828, doi: 10.1109/Confluence51648.2021.9377162.

[11]. Singh, M. K., Chetia, S., & Singh, M. (2017). Detection and classification of plant leaf diseases in image processing using MATLAB. International journal of life sciences Research, 5(4), 120-124.

[12]. Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., & Alrahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. International Journal of Computer Applications, 17(1), 31-38.

International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Deaf Sign Interpreter Service

Shivanand Manjaragi*[1], Revati Patil[1], Chaitra Shende[1], Aruna Gadakeri[1], Sourab Nagare[1]
*[1]Department of Computer Science and Engineering, Hirasugar Institute of Technology, Nidasoshi, Belgaum, Karnataka, India

ABSTRACT

Sign language is the way of communication for hearing impaired people. There is a challenge for common people to communicate with deaf people which makes this system helpful in assisting them. This paper aims at implementing computer vision which can take the sign from the users and convert them into text in real time. The proposed methodology contains four modules such as: image capturing, pre-processing classification and prediction. This system uses a camera, which captures various gestures of the hand. Next, the captured image is pre-processed, the edges are determined an edge detection algorithm. Finally, a template-matching algorithm identifies the sign and display the text. As the output is text, one can easily interpret the meaning of a particular sign. This also curtails the difficulty to communicate with the deaf. The system is implemented by using OpenCV-Python. The system uses various libraries and finally translates the sign gestures to text with accuracy.
Keywords: Deaf Sign Interpreter, Deep learning, template-matching algorithm, RCNN

I. INTRODUCTION

The deaf and dumb people find it difficult to communicate their needs in their day to day life. The Sign language is developed so that they can communicate with other people using hand movements. Sign languages are developed primarily to aid deaf and dumb people. They use a concurrent and specific combination of hand movements, hand shapes and orientation in order to convey particular information. One such set of language is the Indian Sign Language (ISL) system which is predominantly used in south Asian countries. Certain aspect that distinguishes ISL from other sign languages is that ISL devoid of any temporal inflections in its finger spelling chart and also the usage of both the hands. Many people in India are speech and/or hearing impaired, and they thus use hand gestures to communicate with other people. However, apart from a handful number of people, not everyone is aware of this sign language and they may require an interpreter which can be inconvenient and expensive. This paper aims to narrow this communication gap by developing software which can predict the ISL alphanumeric hand gestures in real time.

The objective of this system is to elevate people with hearing disability and help them socialize with common people. It is a form of non-verbal communication.

- Reorganization of sign language automatically from video stream to help deaf and dumb people, this system helps them to communicate more effectively with each other or normal people.

- Reorganization of hand gesture from the video stream and converting gesture to text form which helps deaf and dumb people to commutate effectively.
- This reorganization system provides us an innovative, natural, user friendly way of communication with the computer which is more familiar to the human beings.

The organization of this paper is as follows. The Section 2 is literature review, Section 3 presents the details of the proposed methodology, Section 4 consists of result analysis and section 5 is conclusion and future work.

II. LITERATURE REVIEW

[1]. Deaf Mute Communication Interpreter- A Review: This paper aims to cover the various prevailing methods of deaf-mute communication interpreter system. The two broad classifications of the communication methodologies used by the deaf – mute people are - Wearable Communication Device and Online Learning System. Under Wearable communication method, there are Glove based systems, Keypad method and Handicom Touch- screen. All the above mentioned three sub-divided methods make use of various sensors, accelerometer, a suitable micro-controller, a text to speech conversion module, a keypad and a touch- screen. The need for an external device to interpret the message between a deaf –mute and non-deaf- mute people can be overcome by the second method i.e. online learning system. The Online Learning System has different methods. The five subdivided methods are- SLIM module, TESSA, Wi-See Technology, SWI_PELE System and Web- Sign Technology.

[2]. An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform: The proposed ISLR system is considered as a pattern recognition technique that has two important modules: feature extraction and classification. The joint use of Discrete Wavelet Transform (DWT) based feature extraction and nearest neighbor classifier is used to recognize the sign language. The experimental results show that the proposed hand gesture recognition system achieves maximum 99.23% classification accuracy while using cosine distance classifier.

[3]. Hand Gesture Recognition Using PCA in: In this paper authors presented a scheme using a database driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching with can be effectively used for human robotics applications and similar other applications.. Initially, hand region is segmented by applying skin color model in YCbCr color space. In the next stage thresholding is applied to separate foreground and background. Finally, template based matching technique is developed using Principal Component Analysis (PCA) for recognition.

[4]. Hand Gesture Recognition System for Dumb People: Authors presented the static hand gesture recognition system using digital image processing. For hand gesture feature vector SIFT algorithm is used. The SIFT features have been computed at the edges which are invariant to scaling, rotation, addition of noise.

[5]. An Automated System for Indian Sign Language Recognition in: In this paper a method for automatic recognition of signs on the basis of shape based features is presented. For segmentation of hand region from the images, Otsu's thresholding algorithm is used, that chooses an optimal threshold to minimize the within-class variance of threshold black and white pixels. Features of segmented hand region are calculated using Hu's invariant moments that are fed to Artificial Neural Network for classification. Performance of the system is evaluated on the basis of Accuracy, Sensitivity and Specificity.

[6]. Hand Gesture Recognition for Sign Language Recognition: A Review in: Authors presented various methods of hand gesture and sign language recognition proposed in the past by various researchers. For deaf and dumb people, Sign language is the only way of communication. With the help of sign language, these physical impaired people express their emotions and thoughts to other person.

[7]. Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People in: In this paper author proposed a system to aid communication of deaf and dumb people communication using Indian sign language (ISL) with normal people where hand gestures will be converted into appropriate text message. Main objective is to design an algorithm to convert dynamic gesture to text at real time. Finally after testing is done the system will be implemented on android platform and will be available as an application for smart phone and tablet pc.

[8]. Real Time Detection and Recognition Of Indian And American Sign Language Using Sift In: Author proposed a real time vision based system for hand gesture recognition for human computer interaction in many applications. The system can recognize 35 different hand gestures given by Indian and American Sign Language or ISL and ASL at faster rate with virtuous accuracy. RGB-to- GRAY segmentation technique was used to minimize the chances of false detection. Authors proposed a method of improvised Scale Invariant Feature Transform (SIFT) and same was used to extract features. The system is model using MATLAB. To design and efficient user friendly hand gesture recognition system, a GUI model has been implemented

[9]. A Review on Feature Extraction for Indian and American Sign Language in: Paper presented the recent research and development of sign language based on manual communication and body language. Sign language recognition systems typically elaborate three steps pre processing, feature extraction and classification. Classification methods used for recognition are Neural Network(NN), Support Vector Machine(SVM), Hidden Markov Models(HMM), Scale Invariant Feature Transform(SIFT),etc.

[10]. SignPro- An Application Suite for Deaf and Dumb. In: Author presented application that helps the deaf and dumb person to communicate with the rest of the world using sign language. The key feature in this system is the real time gesture to text conversion. The processing steps include: gesture extraction, gesture matching and conversion to speech. Gesture extraction involves use of various image processing techniques such as histogram matching, bounding box computation, skin color segmentation and region growing. Techniques applicable for Gesture matching include feature point matching and correlation based matching.

The other features in the application include voicing out of text and text to gesture conversion.

From the literature review, we have come across multiple solutions and study on detection and interpretation of sign language for deaf and dumb. From the literature review studied, it can be noted that the majority of solutions proposed by the research scholars is hardware based solution which consists of sensory gloves which can track the gestures to determine the sign language using MEMS sensors mounted on the system. The comparative study can be performed concluded that the sensor based solutions require extra hardware to be attached to the deaf and dumb person's body to interpret the sign language. This is not only hectic to wear but also has accuracy issues. Other Computer Vision based algorithms use SURF and SIFT for Detection of hand gestures. This required more optimization to get accurate results. Thus our proposed solution using deep learning RCNN neural network is more accurate and by using this we train the model. This model is user friendly to operate using the developed python GUI application and can be scaled anytime by altering the data.

III. PROPOSED METHODOLOGY

The proposed system consists of development of Deaf sign Interpreter service using Python and Deep learning. The system consists of training a machine learning model capable of detection and interpreter of sign languages for deaf and dumb using deep learning and Python. The proposed system consists of a python GUI application which can accept the image data feed form the Camera. The system then uses deep learning to detect the sign language in the image and then interprets it using trained deep learning RCNN neural network to detect the sign language. The sign interpreted is showed as text in GUI application. Thus proposed system not only can help the deaf and dumb to communicate easily but also avoid the necessity of person with the knowledge of sign language who works as an interpreter between the deaf and dumb people and common people.

Sign language is evolving as an inevitable communication method for the hearing impaired persons. The basic element of the sign language is the sign language alphabets. By using these alphabets the system is built by using the Deep learning and CNN technologies. This system is basically built to help impaired people for communication purpose.

The proposed work consists of development of sign language interpreter for deaf and dumb using deep learning. As shown in block diagram the system consists of python GUI application, Python backend with OpenCV and a deep learning model capable of detecting or recognizing the deaf Sign language. The GUI application developed will have the option of loading the image for the interpretation of the sign language or starting a live video stream to detect the sign language in the live video. The Camera captures the video and feeds the frame by frame data to the python backend where OpenCV performs image preprocessing to feed it to the deep learning model. The Deep learning neural network is trained on sign language data using transfer learning approach to detect and interpret sign language using the provided dataset and to predict the sign message which is conveyed by the deaf and dumb people. The preprocessed images are fed to the trained neural network which will display the text as output regarding the detected sign. The system thus forms and better and accurate solution of sign language interpretation which is not only accurate but also can be scaled in future since it is based on deep learning.

A. System Architecture

The System architecture is shown in Fig. 1. In our system we are using a pre-trained model. The first process is Data acquisition. Here we are collecting data in terms of images. Lots of data is collected, it is further used for preprocessing. In preprocessing, the LabelImg software is used for generating XML files. Which were further converted into TF-Record files. By using these TF-Record files the RCNN deep learning model is trained. This trained model further used for creating GUI application. Here the images will be feed on the top of this model. Then the inference session collects the features of the image, where capabilities are learned from the deep learning trained model for the sign detection. After detecting the gesture, python Tkinter GUI is used to provide fast and easy way to create GUI application. And finally detected gesture will be displayed as an output in text format.

RCNN Algorithm:

Faster RCNN is the modified version of Fast RCNN. The major difference between them is that Fast RCNN uses selective search for generating Regions of Interest, while Faster RCNN uses “Region Proposal Network”, aka RPN. RPN takes image feature maps as an input and generates a set of object proposals, each with an objectness score as output.

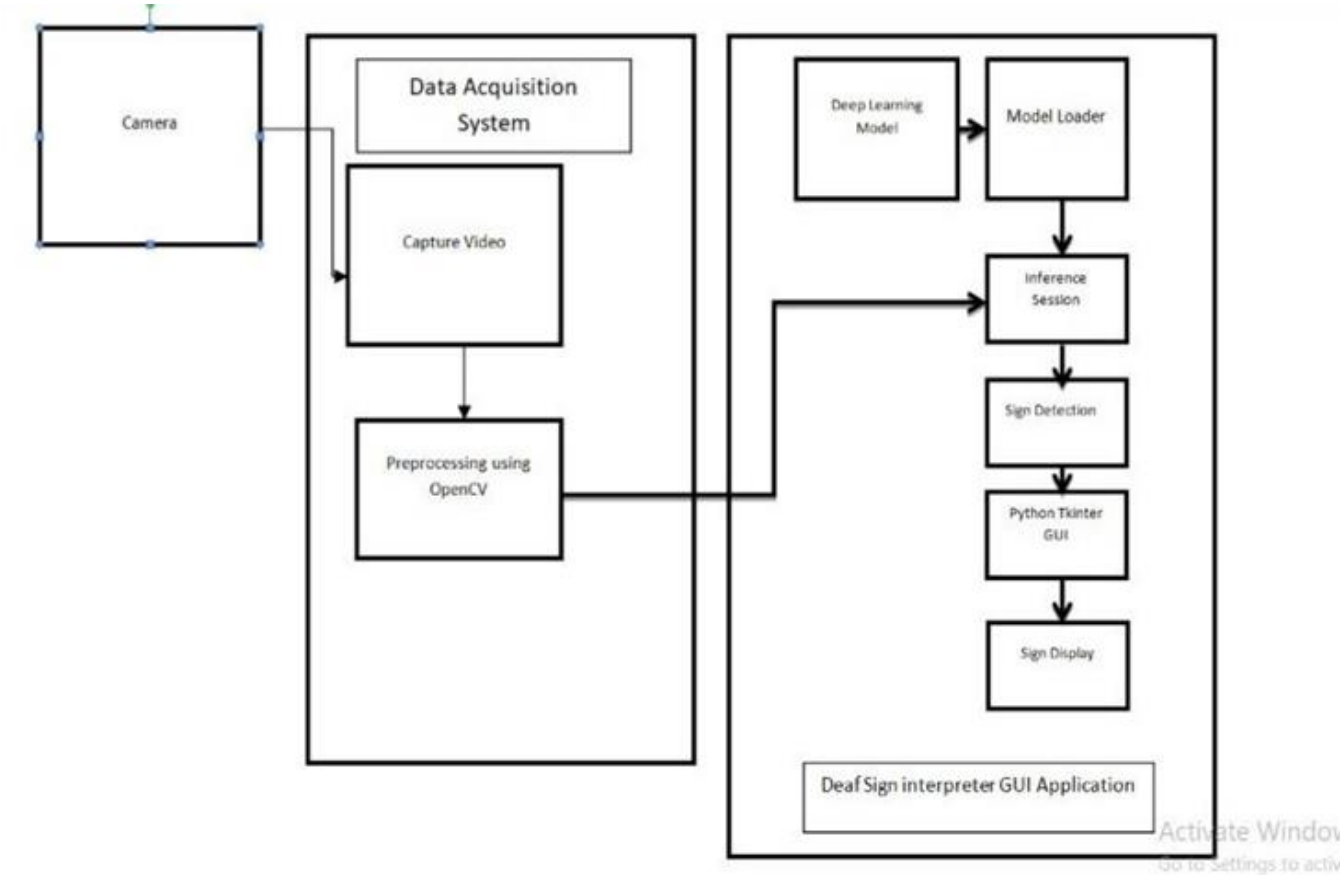


Fig.1 System Architecture

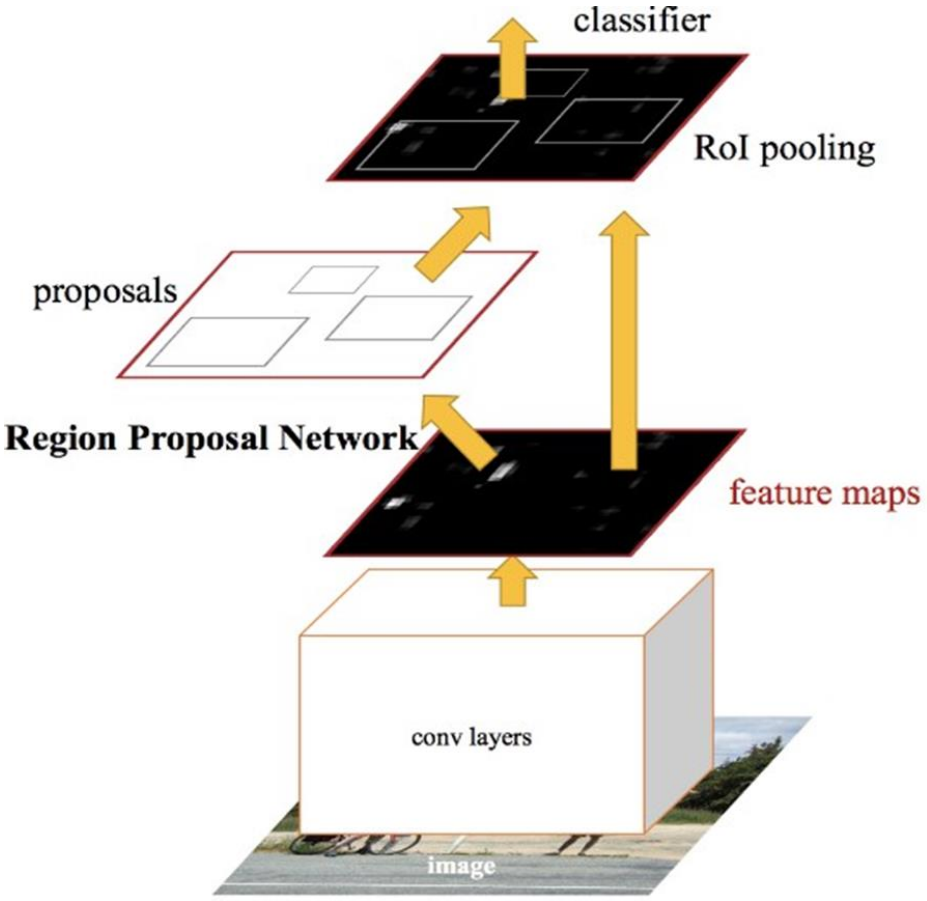


Fig.2 Faster RCNN Working

The below steps are typically followed in a Faster RCNN approach:

- We take an image as input and pass it to the ConvNet which returns the feature map for that image.
- Region proposal network is applied on these feature maps. This returns the object proposals along with their objectness score.
- A ROI pooling layer is applied on these proposals to bring down all the proposals to the same size.
- Finally, the proposals are passed to a fully connected layer which has a softmax layer and a linear regression layer at its top, to classify and output the bounding boxes for objects.

B. Implementation

The below diagram shows the detailed components of the system model.

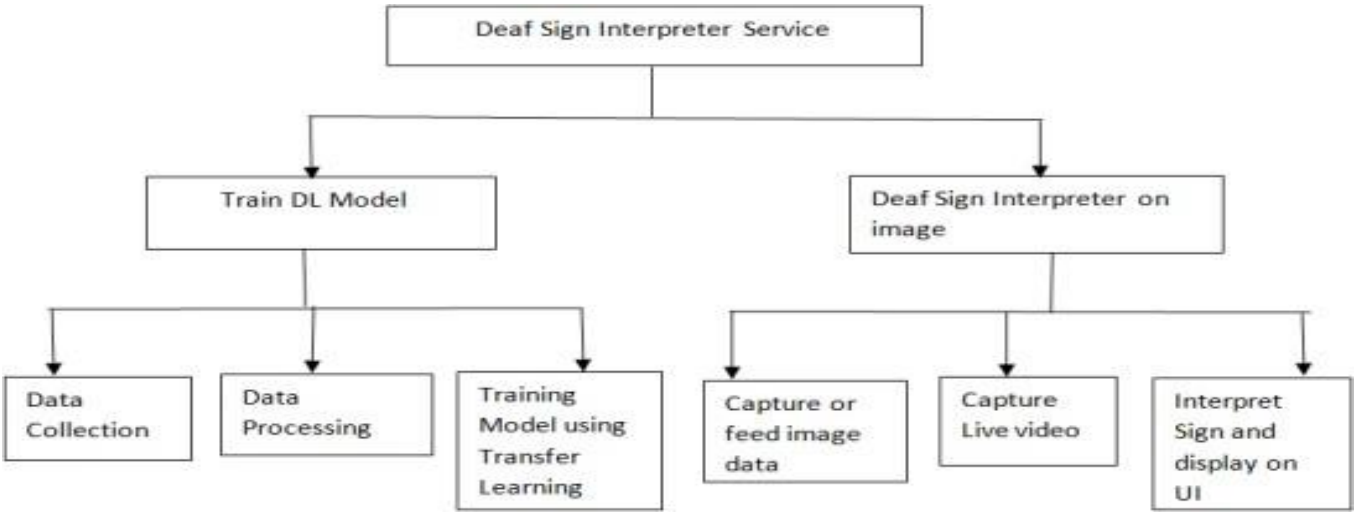


Fig.3 Data Decomposition Diagram

The different components of the system are:

Train DL model

The first step in sign language detection is the training of DL model. The pre-trained model is trained and then once the model is trained the successful inference on the model can be performed by interacting with the deep learning model to display the results. The training of DL model involves following functions:

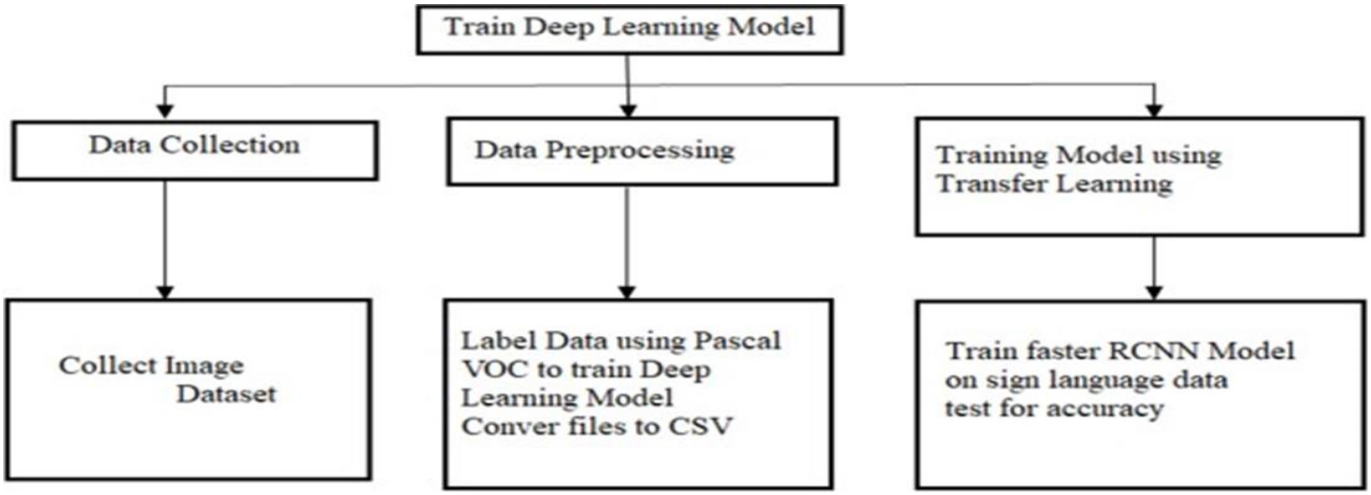


Fig.4 Training Deep Learning Model

- Data collection: Since we are using the transfer learning approach to train the existing model, we require data. The data collection involves image data of the number of sign language samples to train the model. We collect the data samples of each sign language instance and then used the same data for training the model. Once the data is collected the data is split into train and test prior to preparing the data for training.
- Data Pre-processing: The data preprocessing involves resizing and labeling the image data based on the deep learning network chosen for transfer learning. All the images in the test and train are resized and the sign are labeled properly using the Labeling Software. The labeled files are then converted into CSV which will be used generate the train record and test record files used for training the DL model
- Training the model using transfer learning approach: Once the dataset required for training is ready we need to train the model using transfer learning approach. The RCNN network architecture is chosen and the model is retrained on the deaf sign data for detecting of inference the deaf signs. Once the model training is complete the results are tested for accuracy. If accuracy is missing the parameters for training are fine tuned and the model training is repeated till we get accurate results or till the loss function remains constant.
- Deaf Sign interpretation APP: Once the trained model is ready the python app for deaf sign interpretation will be designed. The python app will load the image of video data and perform deaf sign detection and plot the results on image or video feed. The functions of the deaf sign interpreter app are.

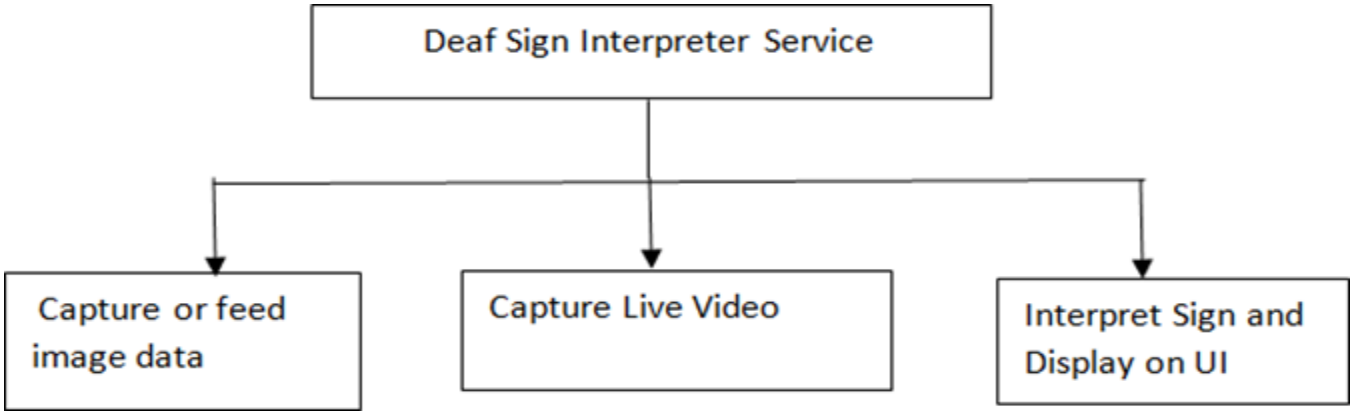


Fig.5 Deaf Sign interpretation APP

- Capture Image or feed image: The image to be interpreted can be fed using this module.
- Capture Live video: The video can be fed or live video can be directly processed using deaf sign interpreter service.
- Interpret the Sign and display on UI: The input image or video split into frame by frame will be fed to the trained model. If sign is recognized the model returns detected sign, occurrence, confidence and dictionary of coordinates of occurrence. This will be plotted on the Output UI image or video feed using Python OpenCV.

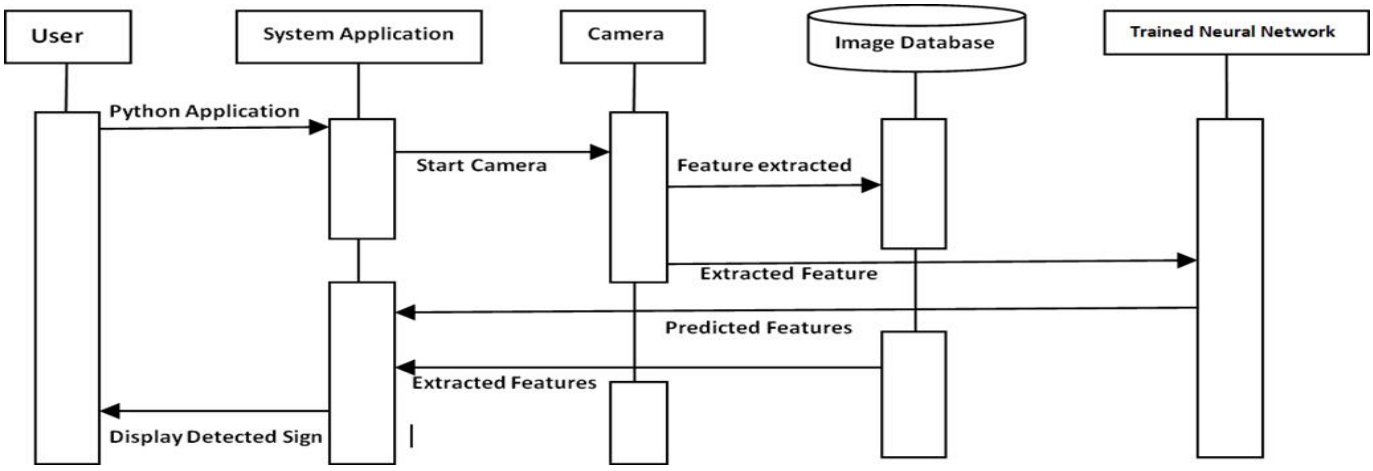


Fig. 6 Sequence Diagram

User first opens the python application, starts the camera and captures the image of hand gestures made. Then the hand gesture features are scaled and extracted by the algorithm. The extracted features are taken by the trained model to predict the output. After predicting the output the sign text will be displayed on the screen.
User first opens the python application, starts the camera and captures the image of hand gestures made. Then the hand gesture features are scaled and extracted by the algorithm. The extracted features are taken by the trained model to predict the output. After predicting the output the sign text will be displayed on the screen.
The algorithms of the each module are given below:

Data Collection:

Step 1 : Start
Step 2 : Collect real-time images as Dataset.
Step 3 : Split the Dataset into 80% training set and 20% testing set.
Step 4 : End

Image labeling:

Step 1 : Start
Step 2 : Label data using Pascal VOC to Train Deep Learning Model.
Step 3 : Convert files to CSV.
Step 4 : End

Symbol Detection:

Step 1 : Start
Step 2 : We take an image as input and pass it to the ConvNet which returns the feature map for that image.
Step 3 : Region proposal network is applied on these feature maps. This returns the object proposals along with their objectness score
Step 4 : A ROI pooling layer is applied on these proposals to bring down all the proposals to the same size.
Step 5 : Finally, the proposals are passed to a fully connected layer which has a softmax layer and a linear regression layer at its top, to classify and output the bounding boxes for objects.

User Interaction:

Step 1 : Start

Step 2 : Capture or feed image data.

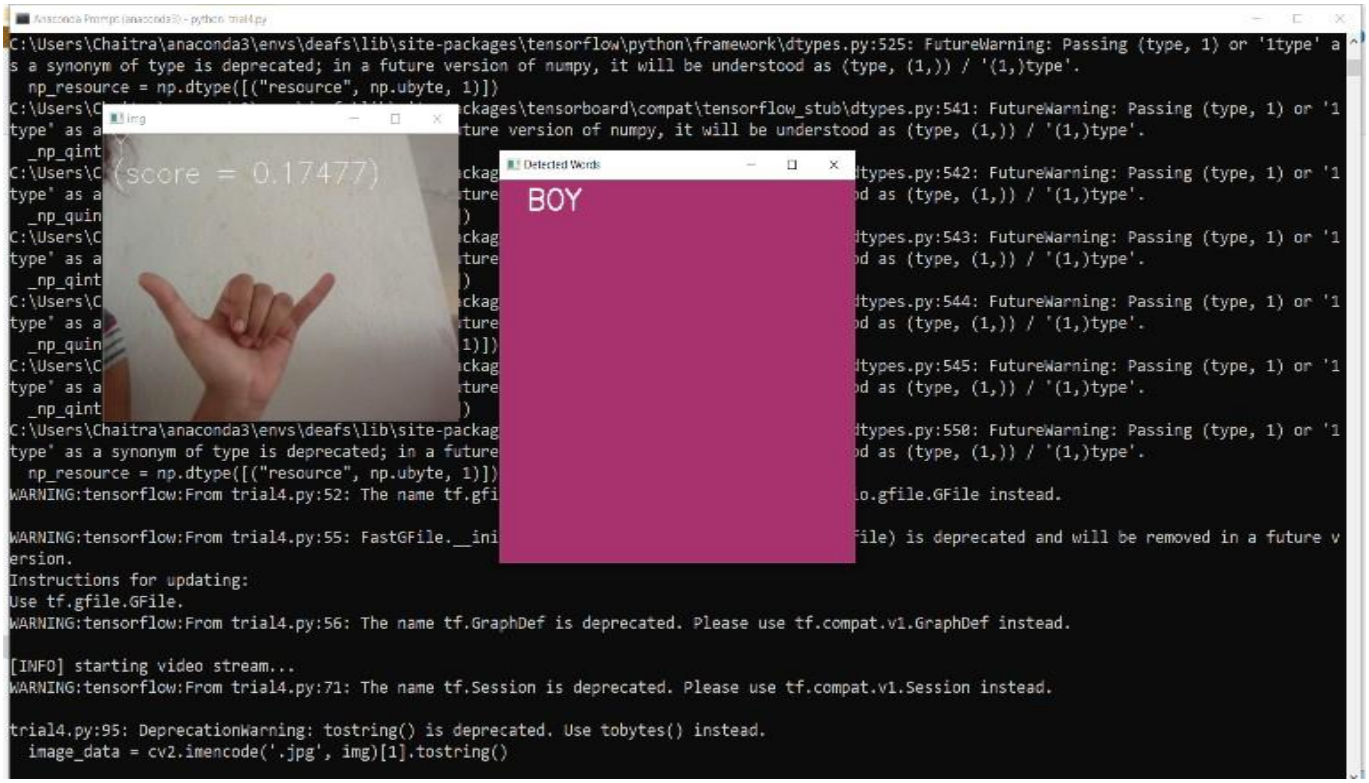
Step 3 : Capture Live Video.

Step 4 : Interpret Sign and Display on UI.

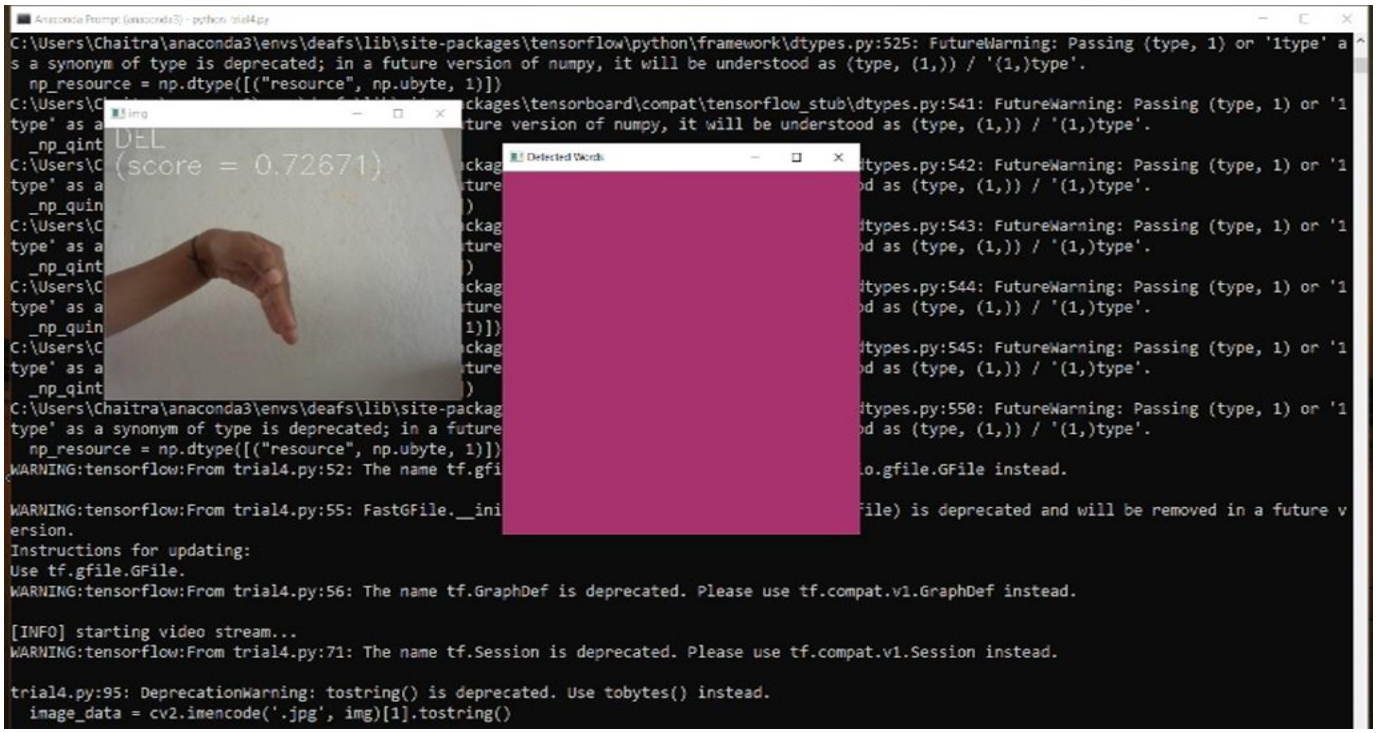
Step 5 : End

IV. RESULTS

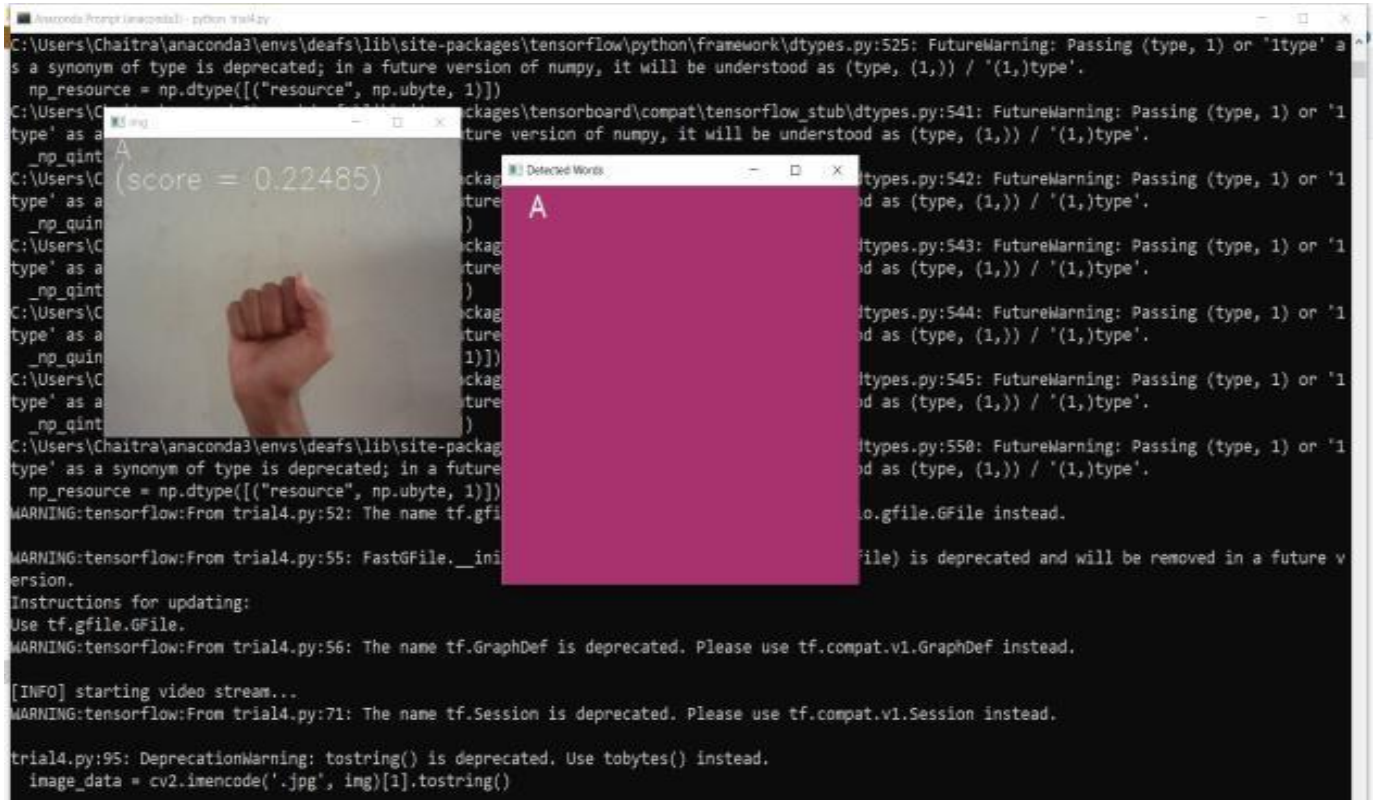
Below are some snapshots which include the human interface design with the computer.



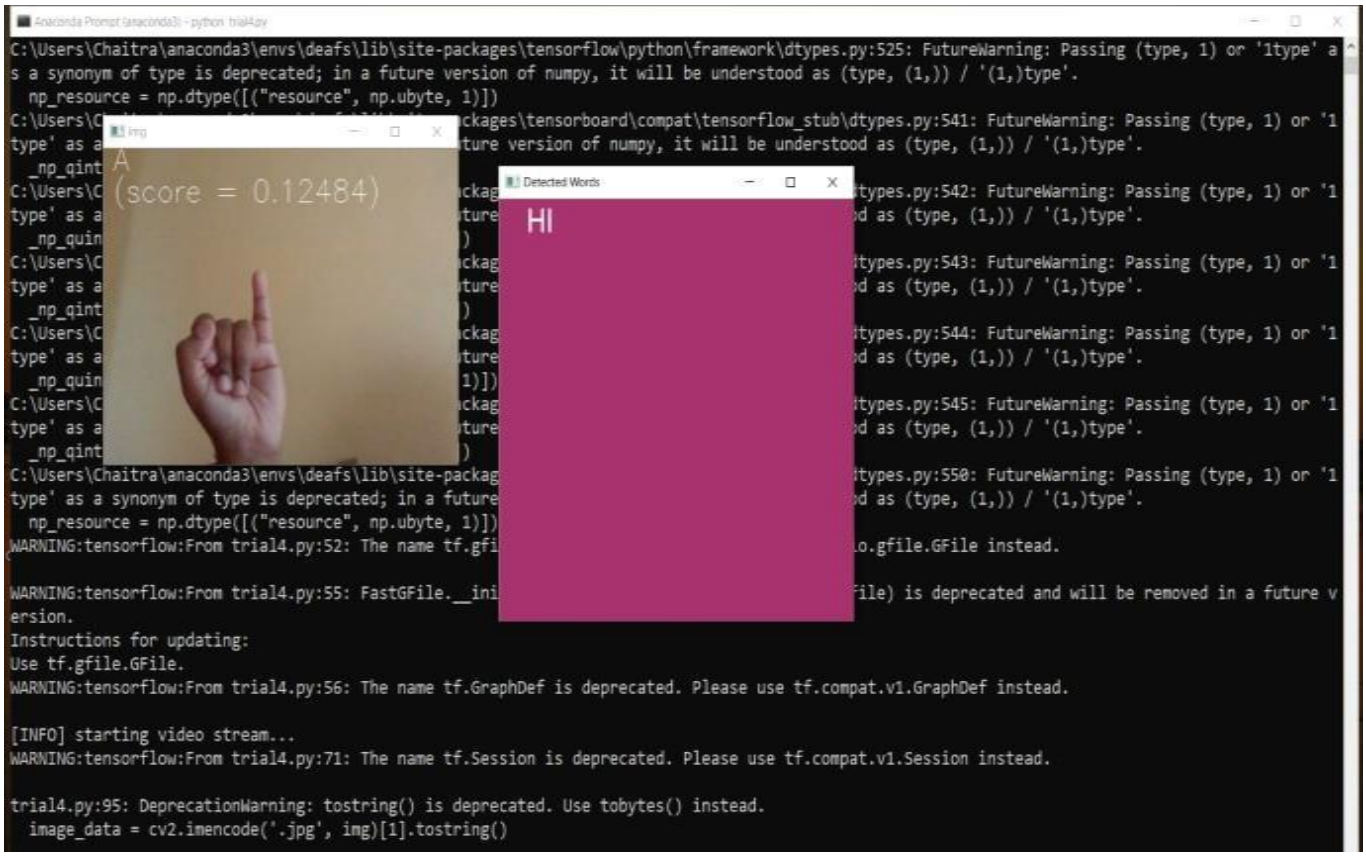
In above snapshot we want to display the word BOY so we used the signs and after some trail we displayed the word BOY successfully.



In the above snapshot the Delete sign will be detected.



In the above snapshot we can see that the letter A will be displayed by using the sign language.



In the above snapshot the word HI will be displayed by using the sign Language.

The accuracy of recognizing the symbol is 85% and in future this can be improved by training the model with larger dataset.

V. CONCLUSION AND FUTURE WORK

The proposed system not only can help the deaf and dumb to communicate easily but also avoid the necessity of person with the knowledge of sign language who works as an interpreter between the deaf and dumb people and common people. Thus our proposed solution using deep learning is user friendly to operate using the developed python application and can be scaled anytime by altering the training data.

This work has wide scope for future modification. In Future this work can be modified for accuracy by training on huge datasets as machine learning is data hungry. More the data, more the accuracy. At present we have worked on only basic sign letters. It can be further used to train on different sign languages and can be modified.

VI. REFERENCES

[1]. Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research ,Volume 11, pp 290-296 , 2016.

[2]. Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874-1883, 2016.

[3]. Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.
[4]. Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people",
[5]. International Journal of Science and Research (IJSR)
[6]. Chandandeep Kaur, Nivit Gill, "An Automated System for Indian Sign Language Recognition", International Journal of Advanced Research in Computer Science and Software Engineering.
[7]. Pratibha Pandey, Vinay Jain, "Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015 .
[8]. Nakul Nagpal,Dr. Arun Mitra.,Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 5,pp- 147 – 149.
[9]. Neelam K. Gilorkar, Manisha M. Ingle, "Real Time Detection And Recognition Of Indian And American Sign Language Using Sift", International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 5, Issue 5, pp. 11-18 , May 2014
[10]. Neelam K. Gilorkar, Manisha M. Ingle, "A Review on Feature Extraction for Indian and American Sign Language", International Journal of Computer Science and Information Technologies, Volume 5 (1) , pp314-318, 2014.
[11]. Ashish Sethi, Hemanth , Kuldeep Kumar,Bhaskara Rao ,Krishnan R, "Sign Pro-An Application Suite for Deaf and Dumb", IJCSET , Volume 2, Issue 5, pp-1203-1206, May 2012.
[12]. Priyanka Sharma,"Offline Signature Verification Using Surf Feature Extraction and Neural Networks Approach", International Journal of Computer Science and Information Technologies, Volume 5 (3) , pp 3539-3541, 2014.

International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Orchestrating Dynamic Big Data End to End ETL Pipeline

Syed Azimuddin Inamdar*[1], Sayyid Abrar[1], Gayatri Bajantri[1]
[1]Department of Computer Science Engineering, VTU, SECAB Institute of Engineering and Technology, Vijayapura, Karnataka, India

ABSTRACT
Now a days data is said to be the new currency and key to triumph. Gathering a rich quality information from numerous dispersed sources across the world necessitates abundant struggle and time. There stand quite a lot of other challenges that consists while transferring information from its start point to its end point. Data ETL pipelines are employed to extend the complete effectiveness of flow of data from its source to the final destination. In the meantime it is automated and decreases the involvement of humans. In spite of prevailing study on ETL pipelines, the study on this topic is limited. ETL pipelines are intellectual representations of end to end data pipelines. To make use of the full possible of the data pipeline, we need to recognize the events that are going in it and the way they're associated in an end to end pipeline. This thesis gives an summary of designing a conceptual model of data pipeline which may be further used as means of communication among various data teams.

Keywords : Bigdata, ETL pipeline.

I. INTRODUCTION

The impartial of this presented thesis is to implement the ETL data pipeline in cloud by optimizing cost and improving the overall performance of the data pipeline. It is a cloud based data integration service that allows us to create data driven workflows for orchestrating data movement and converting data at scale. Using Azure Data Factory we can create and schedule data driven workflows called pipelines that can ingest data from disparate data stores.

ETL stands for Extract, Transform and Load, it is a process of extracting data from one or multiple data sources distributed across the world, then, transforming the data as our business requirements and loading the data into data warehouse. Now a days data is charming more prevalent within the commerce world do to the importance of knowledge crops like APIs, consoles, standards and reports. The data plays important role in policymaking and in the expansion of ML and DL models. Hence, all processes related to data starting from generation of data to data response needs to be watched. The fault detection, reporting and justifying the effect of faults are very multifaceted but unavoidable while construction of effectual data products.

II. EXISTING SYSTEM

Here I have taken two separate cases which existing data pipelines are used, the cases are from telecommunication firm, Each case is separate from other and there is no other interaction between them.

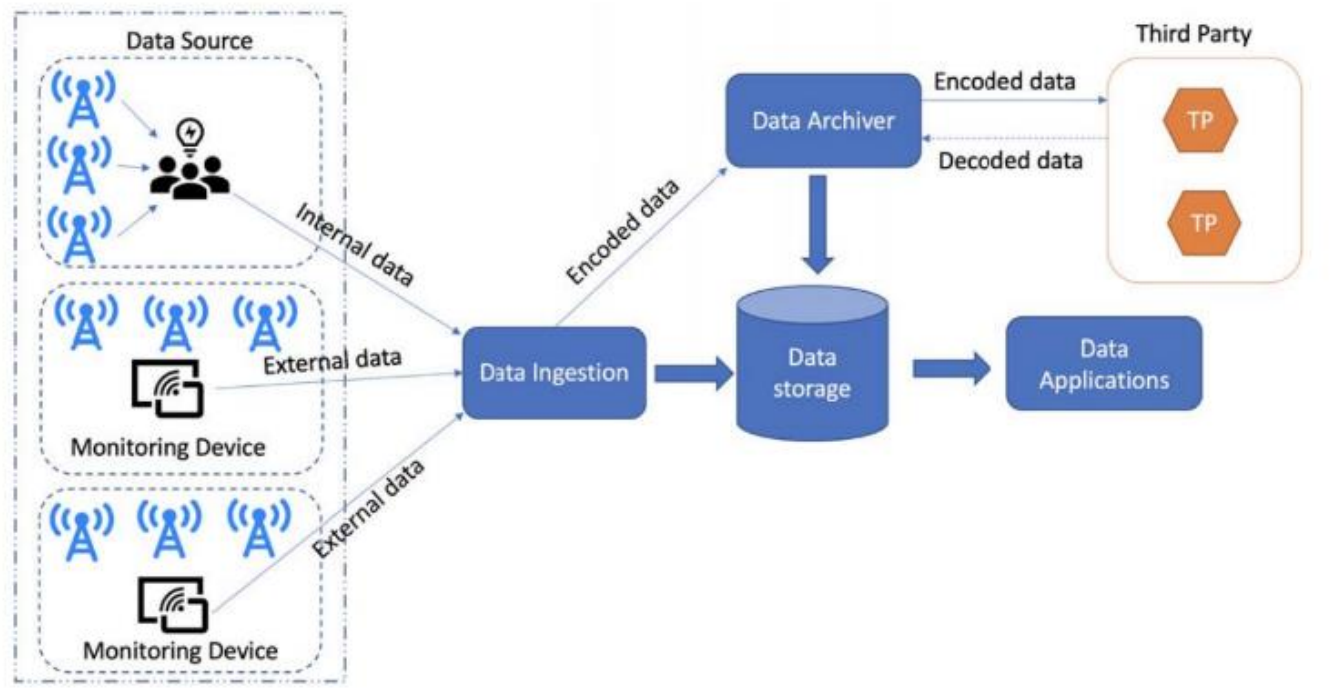


Fig1: Existing pipeline

A. The Data collection process.

The business gathers data from wide spread manifold sources dispersed across the world which are stimulating continuously. Data is composed from the device which is situated in another nation or from the client network. Also the subtle info in the data should be held correctly, it should not leak, Furthermore, the data gathering should deliberate that data generated from different data resources are in different incidences and arrangements. The data can be composed continuously or as batches. The data assortment mechanism should be proficient to adjust with numerous strengths of data-flow. These trials should be addressed religiously when data collection is automated. Fig. 1.5 shows the automatic data collection from the devices. In this case the device is placed inside a piece of equipment owned by customers, the device data is mined without the customer's sensitive information. Base stations have got nodes as well as a device for tracking and managing the nodes. Data assembly equipment are situated on the customer's premises that is physical place which can relate with the nodes straight or with the device to collect the data.

B. Data Governance Pipeline

The data pipeline shown in below fig. 1 is established to aid the data squads in the corporation who are employed with data whensoever it is needed. It is nothing but the connection from where the original data can be copied. There are two types of dumps that data pipeline gets, internal and external. The internal data dump

is the data that is consumed by the squads employed on it inside the corporation and external data dump is the data gathered from the devices in the fields directly. The technique of data absorption is different for different sources and the consumed data is kept in the data storage for future use. The data might have been encrypted connections that need to be decrypted before storing it. There will be a vendor party service to decrypt the encrypted data dump, the data archiver unit sends it to third-party services for decryption. The decoded links are then stored in central storage. Thus the data is made available at any point of the process.

Teams which are working on the data pipeline can appeal data from any phase of the pipeline. The pipeline is tracked manually by flow custodian who is accountable for recognizing the faults in the pipeline and solving them.

C. Problem Statement

Taking into the considerations of my cross case analysis of my data management, I have identified some of the main encounters with the data administration and current data pipeline used in the companies. These are listed below.

Availability of Data: For the successful development of a information product the availability of the accurate data at the accurate time and in the accurate format is very significant. Collection of data is a very tough task and sometimes it fails due to verification failure, climatic factors or even the failure of devices used for collection of data. Even after the gathering of vast data sets from the equipment, it may not reach the defined end point. Data composed can be left unfinished. i.e. unfinished data will be accessible in the data warehouse, it can be due to software letdowns, slices of the data can be vanished. It is hard to identify the loss unless we have a tracking mechanism. Availability of distinct data is vital as input to the model, to scale up the performance.

Quality of Data for systems like ML and DL: Quality of data is vital for systems like Machine Learning and Deep learning. When poor-quality data is nursed to the algorithms of ML and DL, the systems will yield poor-quality output. There should be a clear difference amid faults due to ecological factors and liabilities due to device letdown while collecting the fault logs from the field devices. The key test is when the information is processed to suitable a predefined structure, redundant parts of the data is detached from it. Consequently, we should have a technique to save the original data file. It is not possible to transform the data on the fly and store it, hence it is constantly decent to have the raw data file stored so that it can be retrieved whenever needed by the data squads or when the processed data becomes inadequate to fill the necessities.

Low Storage Capacity: Every team when building their own data pipeline, stores the same data in various forms in the data storage leading to lack of storage space. An amplified quantity of data pipelines will lead to lack of space. When the available storage is separated between different squads, each one will get only a minor share of the real obtainable storage space.

High Cost: To set up a data pipe line in a traditional way, requires high infrastructure cost. Traditionally the pipeline is designed in an on premises system. The system requires high performance processor, RAM and other hardware which make the cost to rise exponentially.

Location Based: Suppose the server is installed in Asia the teams accessing in US will face latency due to geo location of the server. The server responds much faster to the teams accessing the pipeline near to the geo location of the server.

These were the few challenges in data organization and Current data pipelines. Further my thesis I have designed the data pipelines which further reduces or overcomes some of the challenges listed above.

III. PROPOSED SYSTEM DESIGN

The below figure shows the conceptual model, it is a set of ideas used to create data pipelines. The basic fundamentals used to build the data pipeline here are Nodes and Connectors. Connectors are used to interlock the Nodes.

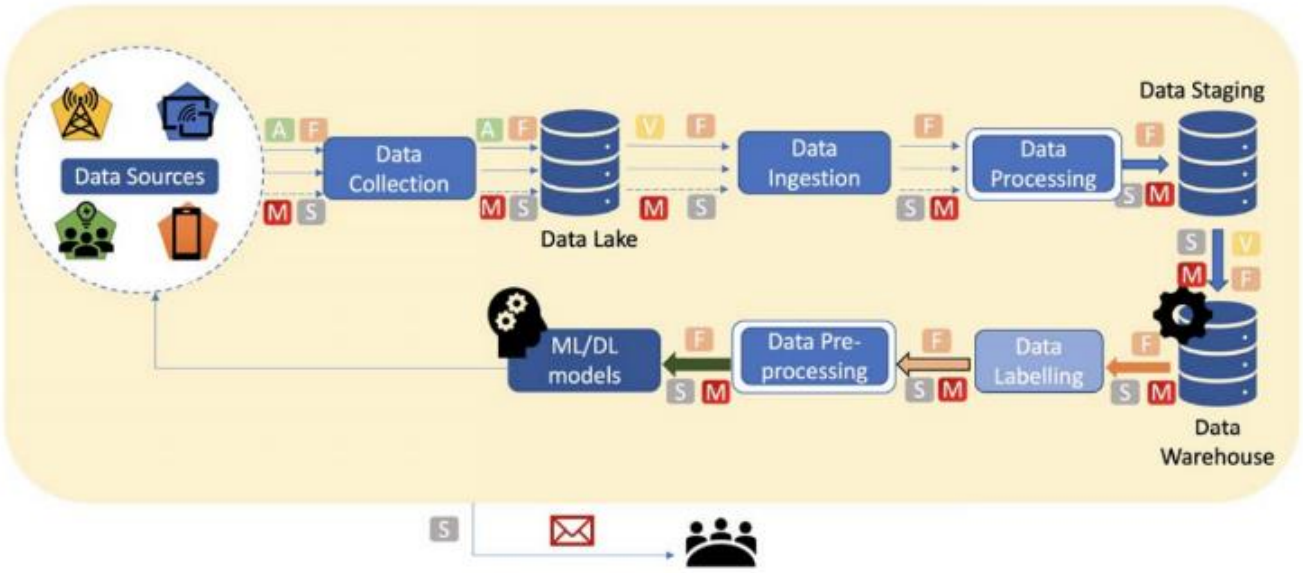


Fig: 2

Data Collection: The information generated from the source could be in the form of batch, intermittent or nonstop type. The three dissimilar flairs of arrows opening from data sources indicate that the connector can carry all the three types of data flow between the data source and data absorption. The data gathering node can gather information from the sources, it can display the consent to gather data from the data sources.

Data Lake: The composed information from the data source will be fresh and it should be kept so that the original files can be repossessed in the future for future use. The data gathering node has to display its right to consume data into the data pipeline. This verification will be approved by connectors among data collection and Data Lake.

Data Processing: Data processing is a complex process which comprises of multiple stages such as data combination, data analysing, data transformation, etc. Data combination is a process in which raw information is translated in a suitable form for arithmetical analysis. The data transformation, is a procedure in which the unstructured combined data is changed into a structured format or semi-structured format. Therefore, we can say that the data processing step translates all diverse kinds of information into a sole format and is stored in data staging area. This is represented in the fig. 1.9 with three dissimilar arrows at the input to data processing showing batch, intermittent and continuous data. The output from the data processing stage is a sole thick arrow as shown.

Data Warehouse: The data staging is a momentary storing area in which the information is stored for authentication. Once the authentication of structured or semi structured information is done, the authenticated information is then located to the data warehouse. This data warehouse acts as a point of admittance from where the information can be reserved for several data applications like formation of report, Machine Learning or Deep Learning applications, etc.

Data Labelling: The data pipeline shows the essential phases to mechanize the data pipeline for Machine Learning applications. Machine Learning algorithms can be of administered, unverified or strengthening. The data labelling phase is done to the administered algorithms, whereas data labelling stage is bounced for unsupervised algorithm. As the most of the businesses are using a administered algorithm method for their Machine Learning applications, the emphasis is high on data labelling.

IV. DISCUSSION AND COMPARISON OF RESULTS

The implementation stage of the mission is where the thorough design is essentially converted into working model. Goal of the stage is to decode the project into a finest conceivable solution in a appropriate programming language. This section covers the application of features of the project, giving particulars of the technology and development atmosphere used. It also gives an impression of the core components of the mission with their phase by phase flow.

The implementation phase requires the following tasks.

1. Cautious planning.
2. Examination of system and limitations.
3. Design of approaches to attain the changeover.
4. Assessment of the changeover technique.
5. Right decisions concerning assortment of the platform.
6. Suitable assortment of the language for application development

The result stage of the project is where the system is assessed and verified in terms of performance and whether or not the goals set in the commencement of the project are attained. Goal of the stage is to attain appropriate data that can be strategized and checked for performance authentication. This section covers the outcomes aspect of the project, giving details of the various complexity stages of the project and relating them.

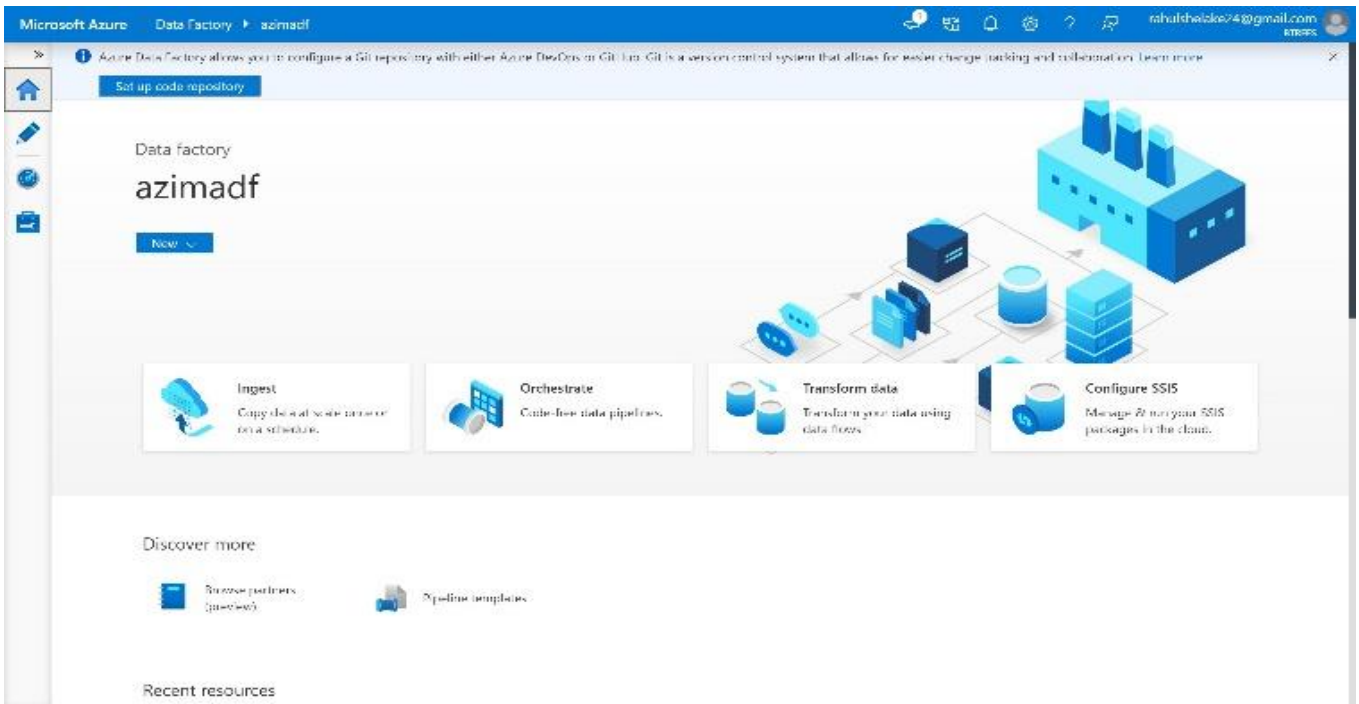


Fig:3 Azure Data Factory

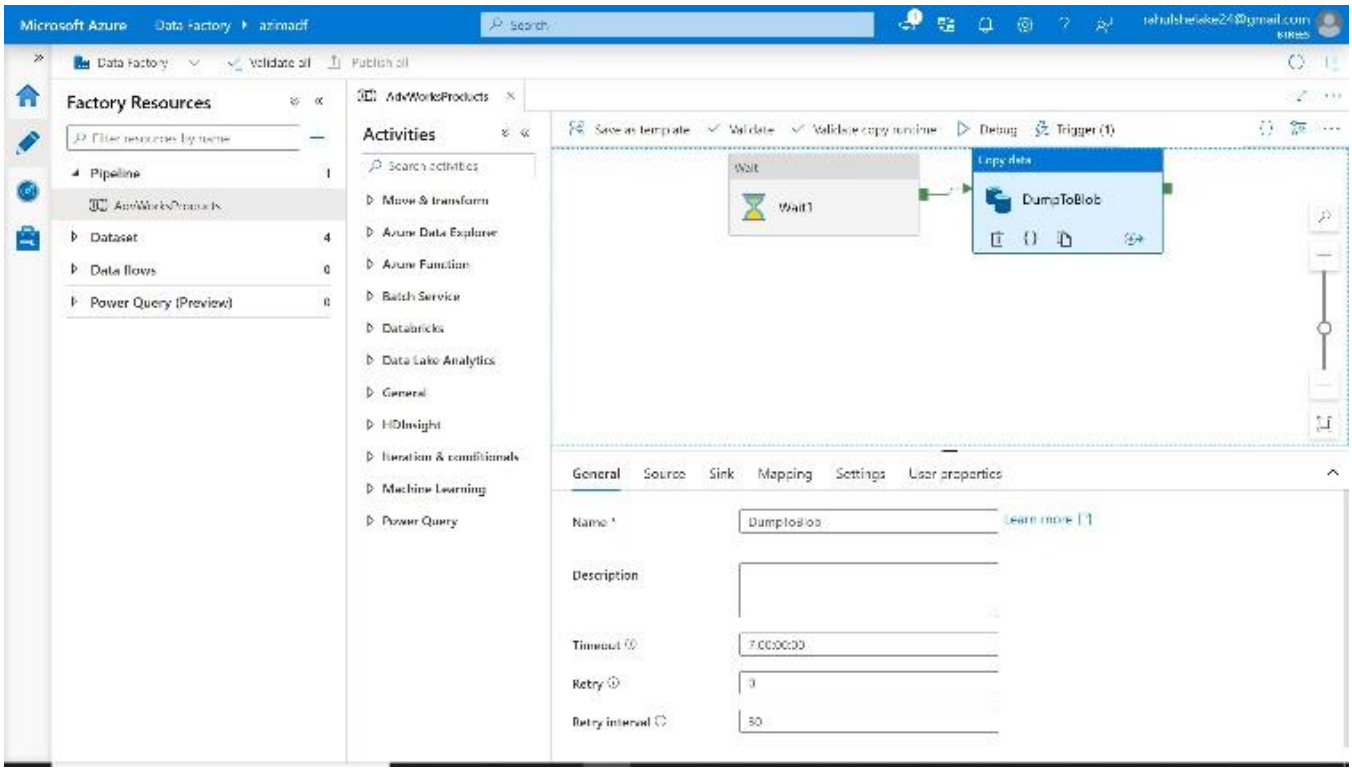


Fig 4: Pipeline

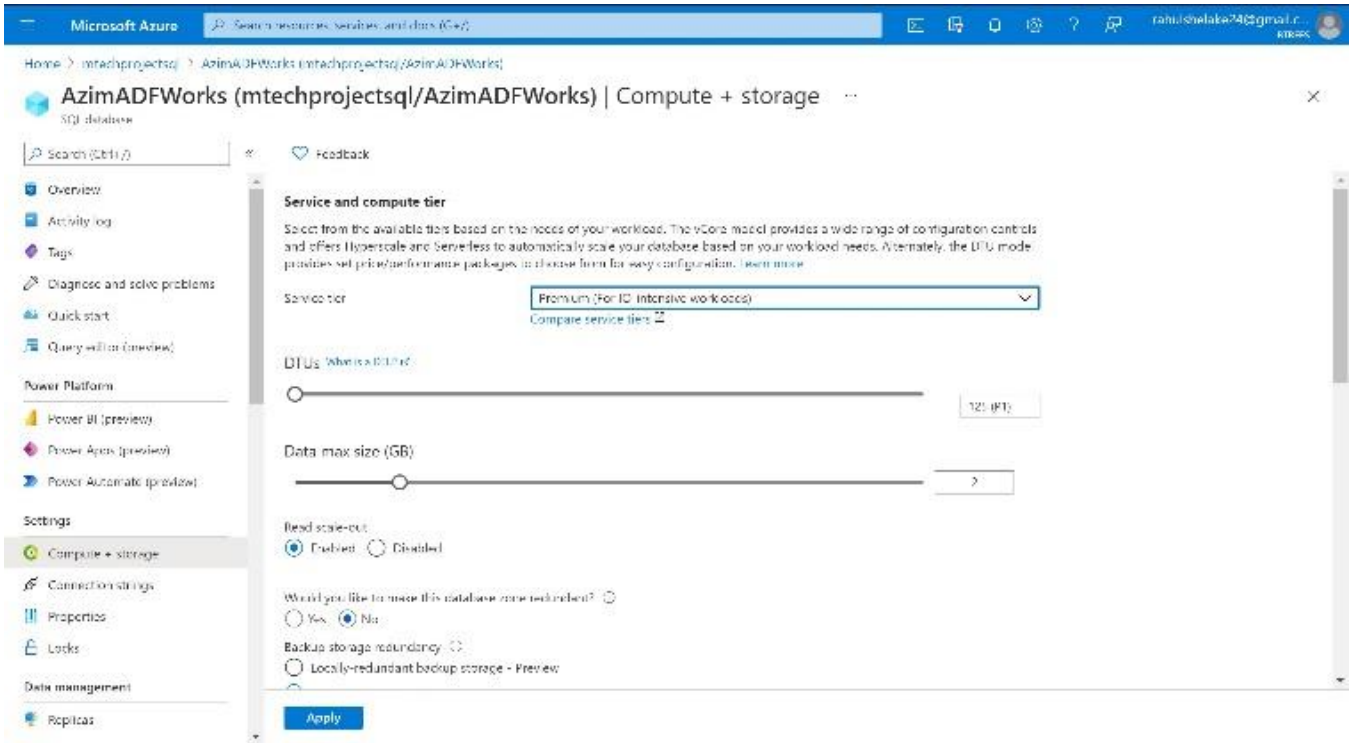


Fig 7:Scaling

Monitoring in Azure Data Factory or any other cloud service like Amazon Web Services (AWS) or Google Cloud Platform(GCP). Orchestration to external services with less or cost optimised way.

Data is new oil, it can manipulate quickly as our need or requirement to achieve much better result in less time using cloud or my design implementation. Using my design, we can extract new value from it as per our future need.

V. CONCLUSION AND FUTURE WORK

It guarantees maximum availability for both single region and multi region databases. It also provides a maximum read availability SLA on multi region databases. To make the automatic failover process more efficient, set a preferred region list for each region. User can access ETL pipeline across the globe without being limited to specific location.

Two other services that can run Spark jobs are Azure Databricks and HDInsight, which can be implemented with more efficient and cost optimized with built-in security provided by cloud services.

VI. REFERENCES

[1]. K. Goodhope, J. Koshy, J. Kreps, N. Narkhede, R. Park, J. Rao, and V. Y. Ye, "Building linkedin's real-time activity data pipeline." IEEE Data Eng. Bull., vol. 35, no. 2, pp. 33–45, 2012.

[2]. E. Deelman and A. Chervenak, "Data management challenges of dataintensive scientific workflows," in 2008 Eighth IEEE International Symposium on Cluster Computing and the Grid (CCGRID). IEEE, 2008, pp. 687–692.

[3]. P. Vassiliadis, "A survey of extract–transform–load technology," International Journal of Data Warehousing and Mining (IJDWM), vol. 5, no. 3, pp. 1–27, 2009.

[4]. J. Trujillo and S. Lujan-Mora, "A uml based approach for modeling ´ etl processes in data warehouses," in International Conference on Conceptual Modeling. Springer, 2003, pp. 307– 320.

[5]. Alkis Simitsis, Kevin Wilkinson, Umeshwar Dayal, Malu Castellanos HP Labs Palo Alto, CA, USA, Optimizing ETL Workflows for Fault-Tolerance, Conference: Proceedings of the 26th International Conference on Data Engineering, ICDE 2010, March 1-6, 2010, Long Beach, California, USA.

International Conference on Artificial Intelligence and Machine Learning
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Detection of Lung Inflammation using Transfer Learning

Javed Akkalkot*[1], Dr. S. A. Quadri[2]
*[1]Department of Computer Science, Secab Institute of Technology, Bijapur, Karnataka, India
[2]Professor, Department of Computer Science, Secab Institute of Technology, Bijapur, Karnataka, India

ABSTRACT

Covid19 is a destructive infection that has killed millions of people across the world. Elderly people and individuals with genuine fundamental ailments or past instances of pneumonia seem, by all accounts, to be at more danger of growing more genuine intricacies from the infection. Serious types of the infection can cause lung inflammation, prompting danger of death. The challenge here is to speed up the determination of the onset or the ongoing pneumonia infection.This paper proposes a Transfer learning based Deep Learning model for automated way of using X-ray imaging for early diagnosis of COVID19/pneumonia infection. We initially train the system using an 8-layer CNN model and a 13-layer CNN model and then compare it with ResNet model which is based on transfer learning for higher accuracy.

Keywords: Covid19, Lung inflammation, Deep Learning, X-ray, CNN model, ResNet model.

I. INTRODUCTION

Deep learning has reached a stage where we can use it to have human-level accuracy in segmenting and analyzing an image. Due to this, the deep learning has wide usage in a range of industries and medical industry in not far behind in utilizing its potential. A lot of diagnosis in medical industry happens though the means of imaging be it X-Rays, Sonography, CT-Scans, MRI Scans etc. It can be widely used in the detection of tumors and lesions in patients. [1] [2]

In this paper, we detect inflammation of lungs by giving the X-Ray images as input. The dataset is taken from the Gangzhou Women and Children's Medical center. These images are transformed to make them legible for training using Python Transforms. This dataset is then used to create a 8 layer CNN model and a 13 layer CNN model. Then we use transfer learning-based technique of reusing a pretrained model where we use ResNetmodel (Residual Network) and train it on the data set available.

Once these models are trained, we evaluate the models for loss and accuracy. We compare the loss and accuracy of these 3 models and determine which is the most efficient one. The models trained can be used for detecting lung inflammation/pneumonia by giving the X-Rays in the test data set. Even though deep learning still cannot replace doctors in medical diagnosis, it can provide support to experts in the medical domain, including examining chest X-Rays for signs of pneumonia. Inflammation of the lungs which is commonly known as Pneumonia can be caused by bacteria, viruses, and fungi. It can strike people of any age, even healthy

individuals. It can be life-threatening to infants, people with certain diseases, and people with impaired immune systems. [3] Researchers have proposed different artificial intelligence (AI)-based solutions for different medical problems recently. With convolutional neural networks (CNNs), researchers have been able to achieve successful results in a wide range of medical problems, like breast cancer detection, brain tumor detection & segmentation and disease classification in X-ray images [4] A deep CNN model such as ResNet requires a lot of data to train from scratch, since it contains millions of trainable parameters, so a small dataset wouldn't suffice to generalize the model. In Transfer Learning [5], a pre-trained CNN model is reused to take advantage of its weights as initialization for another CNN model tailored to a different purpose using the model's weights.

II. METHODOLOGY

The methodology used in this Paper is depicted below:Fig.1
We use the algorithm provided under Pytorch based package called TORCH.

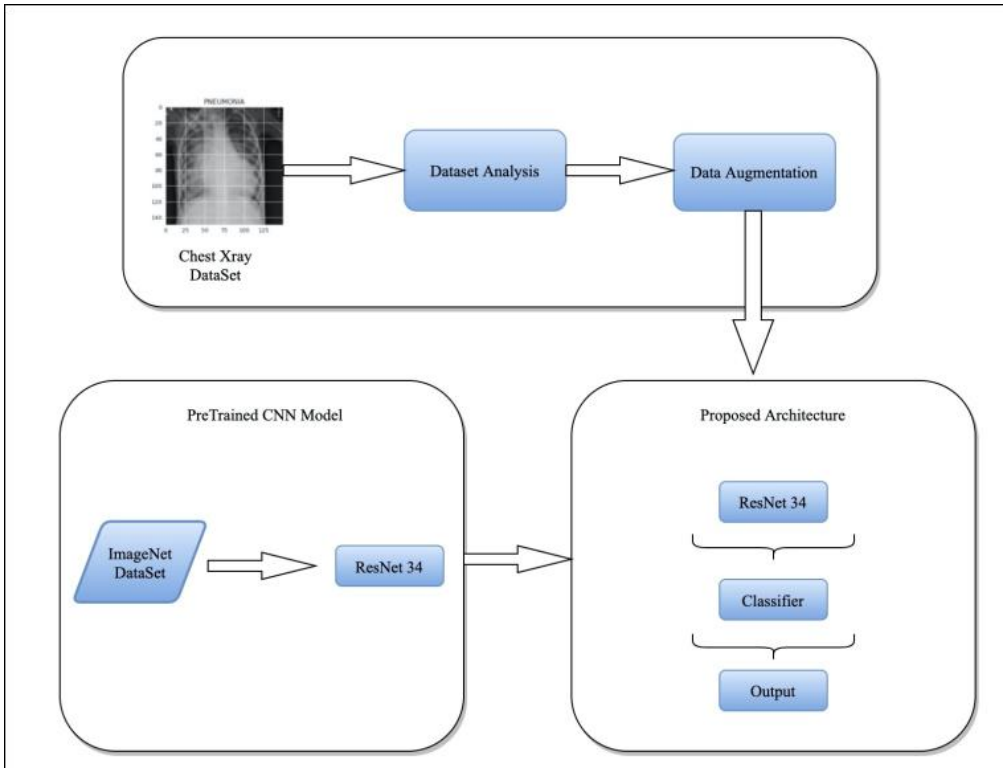


Fig:1. Methodology Used

The torch package contains data structures for multi-dimensional tensors and defines mathematical operations over these tensors. Additionally, it provides many utilities for efficient serializing of Tensors and arbitrary types, and other useful utilities. Class Used: torch.nn. Conv2d

We use Maxpooling and ReLu in between the layers. Maxpooling reduces the dimensions of the image, due to this the network will look at expanded areas of the image and this reduces the number of parameters associated which reduces the computational load. ReLu, Rectified Linear Unit is the most commonly used activation function in deep learning models. The function returns 0 if it receives any negative input, but for any positive value xx it returns that value back. So it can be written as f(x)=max(0,x)f(x)=max(0,x).

```
evaluate(model2, test_dl)
```

```
{'val_loss': 0.6960327625274658, 'val_score': 0.375}
```

Fig 2.CNN 8 Layer – Loss & Accuracy values

We add two additional functionalities to the above 8 layer model called Batch Normalization & Dropout

```
evaluate(model3, test_dl)
```

```
{'val_loss': 0.6847389340400696, 'val_score': 0.625}
```

Fig 3.CNN 13 Layer – Loss & Accuracy values

Because of batch normalization and dropout layers along with a few more layers, this 13-layer CNN model proves to be a better alternative than the CNN 8-layer model. It is now able to categorize more than 50% of the images in the testing dataset properly. However, this is still not as good it has to be for using in the real world comfortably.

We solve this problem by leveraging the optimization technique of transfer learning. Here we repurpose complex models trained for one purpose for another. Here we will use a pretrained residual network ResNet that is trained on the ImageNet dataset & repurpose that to our chest x-rays dataset. We use a 34-layer residual network with dropout in the final layer and check how well it compares to the previous models.

Transfer Learning is a supervised learning technique. It works by reusing parts of an earlier trained model on a fresh network which is tasked for altogether a different problem. Transfer learning significantly reduces the time required for feature based specific engineering and its corresponding training. First, a source model needs to be selected. Ideally the one which can train with a large set of data. Many researchers or labs release their pre-trained models for further use in the industry and research [10]

Second, we need to decide which layers of the pretrained model to be used in our system. The idea is to build a framework which is bare minimum better than an amateur model which gives way for new feature learning. Top layers tend to focus on more finely tuned to a particular problem whereas deeper layers are reused for more typical problems.

Finally, now we trained the newly built model on the new dataset. The benefit is that the model tends to converge faster and with less data and requires lesser computational cost.

We utilize ResNet(Residual Network) in our paper. Let's us understand the architecture of ResNet34 at high level. Deeper neural networks are harder to train which gave rise to ResNet. It is basically a residual learning framework which facilitate the training of networks that are significantly deeper than those used previously. [11]

```
{'val_loss': 0.7321296334266663, 'val_score': 0.7948718070983887}
```

Fig 4. ResNet34 Transfer Learning – Loss & Accuracy Values on CPU

III. RESULTS AND DISCUSSION

A model evaluation accuracy of about ~80% on a CPU based computing which is definitely an improvement from the prior two models. This shows the power & efficiency of using transfer learning to solve common image problems. Although the accuracy of the test has increased significantly, it is debatable whether this is

sufficient for real-world implementation. The efficiency further increases when the model is run on a GPU based machine.

We tested for 20 X-Ray images and out of 20, it successfully identified 19 of them and 1 was identified incorrectly. This leaves us with an accuracy of 95% on the test data.

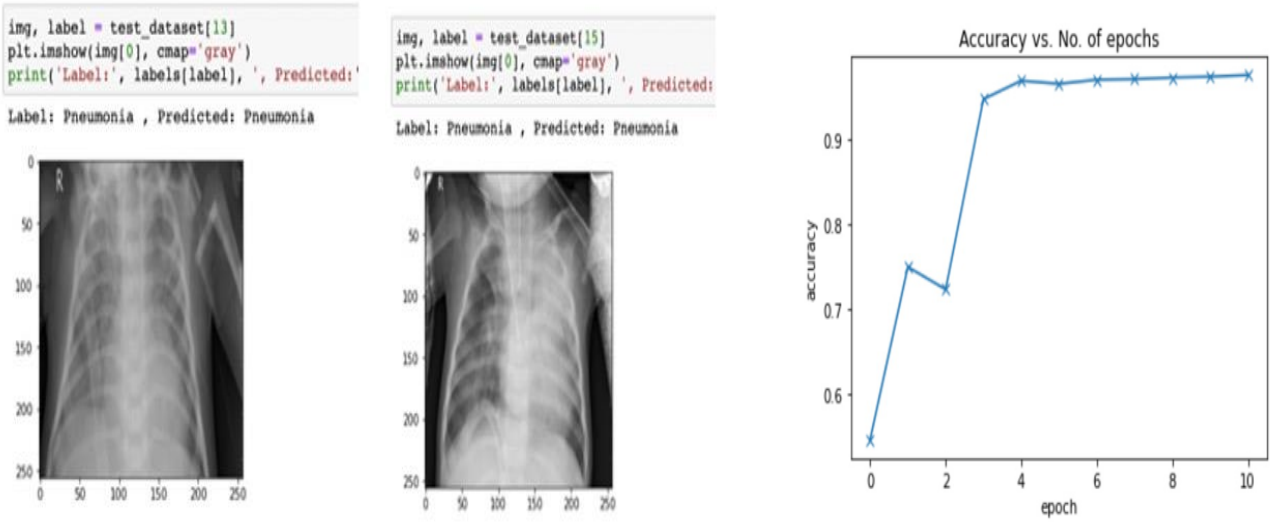


Fig 5: Test Results & Accuracy vs No. of Epochs for ResNet34

IV. CONCLUSSION

The proposed paper speed up the determination of the onset or the ongoing pneumonia infection.This paper proposes a Transfer learning based Deep Learning model for automated way of using X-ray imaging for early diagnosis of COVID19/pneumonia infection.

V. REFERENCES

[1]. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. Med. Image Anal. 2017, 42, 60–88. [CrossRef] [PubMed]

[2]. Brunetti, A.; Carnimeo, L.; Trotta, G.F.; Bevilacqua, V. Computer-assisted frameworks for classification of liver, breast and blood neoplasias via neural networks: A survey based on medical images. Neurocomputing 2019, 335, 274–298. [CrossRef]

[3]. Gilani, Z.; Kwong, Y.D.; Levine, O.S.; Deloria-Knoll, M.; Scott, J.A.G.; O'Brien, K.L.; Feikin, D.R. A literature review and survey of childhood pneumonia etiology studies: 2000–2010. Clin. Infect. Dis. 2012, 54 (Suppl. 2), S102–S108. [CrossRef] [PubMed]

[4]. Kallianos, K.; Mongan, J.; Antani, S.; Henry, T.; Taylor, A.; Abuya, J.; Kohli, M. How far have we come? Artificial intelligence for chest radiograph interpretation. Clin. Radiol. 2019, 74, 338–345. [CrossRef]

[5]. Tsiakmaki, M.; Kostopoulos, G.; Kotsiantis, S.; Ragos, O. Transfer Learning from Deep Neural Networks for Predicting Student Performance. Appl. Sci. 2020, 10, 2145. [CrossRef]

[6]. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax

Diseases. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition CVPR 2017, Honolulu, HI, USA, 21–26 July 2017; IEEE: Danvers, MA, USA, 2017; pp. 3462–3471.

[7]. Blumenfeld, Greenspan, and Konen presented a CNN for Pneumothorax detection using CXRAY images.

[8]. Roth, H.R.; Lu, L.; Seff, A.; Cherry, K.M.; Hoffman, J.; Wang, S.; Liu, J.; Turkbey, E.; Summers, R.M. A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations. In Proceedings of the 17th International Conference on Medical Image Computing and Computer-Assisted Intervention, Boston, MA, USA, 14–18 September 2014; Springer International Publishing: New York, NY, USA; pp. 520–527.

[9]. Rajpurkar, P.; Irvin, J.; Ball, R.L.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.P.; et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med. 2018, 15, e1002686. [CrossRef]

[10]. Shin, H.-C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging 2016, 35, 1285–1298. [CrossRef]

[11]. Liang, G.; Zheng, L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Comput. Methods Programs Biomed. 2019, 104964. [CrossRef] [PubMed]

International Conference on Artificial Intelligence and Machine Learning

Organized by

SECAB Institute of Engineering and Technology,
Nauraspur, Bagalkot Road Vijayapur, Karnataka, India

Publisher

Technoscience Academy

Website : www.technoscienceacademy.com
Email: info@technoscienceacademy.com

Email : editor@ijsrcseit.com, ijsrcseit@gmail.com