IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Facial Emotion Detection using Convolution Neural Network

Aniruddha Gaikwad, Ojas Verma, Prof. Priyanka Jadhav
Department of Computer Engineering
Smt. Kashibai Navle College of Engineering Pune, India

Abstract: *Human emotion detection through images is one of the most daunting challenge in the field of artificial intelligence. Machine Learning and Deep Learning are the two major technologies which are used for detecting the facial emotions. Deep Learning based neural networks provides better performance and accuracy. Face detection is a major step in our system that allows us to identify people or objects that are facing the camera. We then use feature extraction to separate individual images into different parts of multidimensional spaces. Finally, we use emotion classification to detect what emotions you're experiencing. Deep-learning approaches are taking part in a vital role in classification tasks. This paper deals with facial emotion recognition by applying transfer learning approaches. During this work pre-trained networks of VGG16 and MobileNet are used. The fully connected layers of the pre-trained ConvNets are eliminated, and that we add our fully connected layers that are appropriate for the number of directions in our task. Finally, the recent added layers are solely trainable to update the weights. This paper proposes a Convolutional Neural Network (CNN) based on Deep Learning architecture which will extract the features and at the same time classify the emotions into seven categories. The model will use three datasets: Japanese Female Facial Expressions (JAFFE), Extended Cohn-Kanade Dataset (CK+) and Facial Expression Recognition 2013 (FER2013) to train itself and increase its accuracy.*

Keywords: Artificial Intelligence (AI), Facial Emotion Recognition (FER), Convolutional neural networks (CNN), Japanese Female Facial Expression (JAFFE) Dataset, Cohn Kanade Dataset (CK+), Facial Emotion Recognition Dataset (FER2013), Deep learning (DL).

I. INTRODUCTION

Emotions are known to play a very dominant role in our life. It is not possible to imagine our life without expressions. Emotions can be expressed through verbal and non-verbal ways. Emotions govern the human body, and it is said that Facial Expressions and the body language is a key feature of human interaction. Facial expression is a form of non-verbal method of emotion expression. Research has shown that 55%

of the information during communication is passed on by Facial Expressions. Facial expression is also found to play a crucial role in the area of human-machine interaction. Currently

Artificial Intelligence (AI) is used for face detection using neural networks. Machine Learning and Deep Learning have been the major technologies which are used to recognise.

these facial emotions. In this review we find a brief description of various types of pre-processing and classification methods which are used to identify facial emotions using Machine Learning and Deep Learning. We propose a system using Convolution Neural Network (CNN) based on deep learning to identify seven facial emotions. The seven emotions include emotions such as anger, sadness, disgust, happiness, fear, surprise and neutral. Previously researchers have used both machine learning as well as deep learning for emotion recognition. This paper also dives into the various types of datasets available to train the neural network as well as compares various algorithms used for classifying the emotions. It discusses the CNN and the various layers used for extracting features from the facial image. The classic human computer interaction does not take into account the emotional state of the user and hence is likely to lose a lot of valuable data. The emotions captured by facial expressions can be used in various researches as well as other fields such as marketing, security, educational activities, etc. Due to the ever-increasing availability of computational power and the training datasets, the performance and the accuracy of these systems is bound to increase.

II. RELATED WORKS

2.1 Human Facial Expressions

Many attempts have been made previously to make an automatic facial expression tool [1] as it has applications in many fields such as driving assist systems, robotics, medicine, and lie detector [2,3,4]. The key features of human interaction include facial expression and body language. In the nineteenth century, [5] Charles Darwin published about the Globally common facial expressions. Ekman Friesen in 1971 showed how facial behaviours are correlated uniformly with specific emotions in humans [6], he defined seven expressions (anger, feared, happy, sad, contempt [7], disgust, and surprise). Not only humans but also animals produce certain specific muscle movements which are related to a mental state.

2.2 Image Classification Techniques

Image classification generally consists of feature extraction and classification. Various techniques have been inquired for FER in recent times. The traditional pioneer methods start with extracting features from the image and then use the feature values to classify the emotion. The recent deep learning-based methods combine both tasks in a single process. Several studies were reviewed and compared with the existing FER models [8,9,10,11]. The recent ones include the deep learning-based FER methods [10,11].

The following sections briefly explain the techniques employed in the FER methods: *Machine Learning-Based FER Approaches:*

In the Artificial Intelligence (AI) domain, Automatic Facial Emotion Recognition is a challenging task chiefly in the Machine Learning subdomain. Several conventionally used machine learning methods (e.g., neural network, Knearest neighbor (KNN)) are used for the evolution of the FER methods. Xiao-Xu and Wei [12] gave the initiated FER method by adding wavelet energy feature (WEF) to the facial image, then extracted the features using Fischer's linear discriminants (FLD) and classified the images finally using the K-nearest neighbor (KNN) method.

Zhao et al. [13] also used KNN but for feature extraction, they employed Principal component analysis (PCA) and nonnegative matrix factorization.

Support Vector Machine (SVM) was used to classify the emotions extracted in several models:

Liew and Yairi [8], did a comparative study of SVM and other methods (e.g., KNN, LDA, etc.) for feature classification by engaging different methods (e.g., Gabor, LBP, and Haar).

The crucial limitation of the traditional aforementioned method is that it only considers frontal view for the FER model. The features extracted by conventional methods are different for profile and frontal views.

Deep Learning-Based FER Approaches:

A relatively newer approach in machine learning is Deep Learning, and to date, various CNN-based studies have been conducted.

A deep belief network (DBN) was integrated with the neural network which was used for unsupervised feature learning by Zhao and Zhang [14].

A standard CNN architecture was considered by Pranav et al.

[15] with two convolutional pooling layers for FER.

A recent model developed by Akriti et al. [33] has proposed a deep learning-based CNN model carried out in terms of validation accuracy, computational complexity, detection rate, learning rate, validation loss, computational time per step.

To handle the occlusions and pose variations Sreelakshmi et al[32] presented an emotion recognition system by using MobileNet architecture. The model is sampled on real- time clotted images and achieves an precision of 92.5% . It runs on a CPU core in seconds and can be trained from scratch or with quantized datasets.

Aravind Ravi[33] proposed pre-trained CNN features based on facial emotion recognition. In this work, a pre-trained VGG19 network is used to extract the features and the support vector machine is used to predict the expressions in an attempt of obtaining a model that can be utilized for classification of expressions.

A recent model developed by Shengbin Wu [35], has proposed a model based on the improved VGG16 model which has insufficient feature data and a low recognition rate in face expression recognition. The improved VGG16 network is used to extract the features of facial expression gray image, and the high-level features are input into the CNN model. As results show that the improved VGG16 network model is effective in training and testing expression recognition.

I Gede Putra Kusuma Negara, Andreas Pangestu Lim, Jonathan Jonathan [36] has proposed network is classified as SBNN with VGG-16 as the base model, which was modified into using 13 convolutional layers and GAP as the last pooling layer.

Recent Innovations

Recent studies on the facial recognition techniques: The impact of facial asymmetry as a marker of age estimation was studied by Sajid et al. [16]. The right face asymmetry was found to be better when compared with the left face asymmetry. Still one of the big questions remains with the face pose appearance. The solution to the problem of variability in the facial pose was provided by Ratyal et al. [17]. He used a threedimensional approach for facial poses. Many issues persist such as make-up [19], lighting conditions, and poses which are solved using CNN [17,18].

Recently, outstanding accomplishments have been made in Facial Expression detection [20,21,22] which has led to advancements in neuroscience [23] and cognitive science [24]. Facial emotion recognition is growing quickly as a subdomain of image processing and AI. Given the availability of computational power and training datasets, the performance of the FER models can be increased substantially.

III. DATASET

The database helps us to build a facial emotion recognition system that can produce results that are even more contrasting with related works. Most databases are based on 2D still images. Some basic emotions are categorized by databases namely (happiness, disgust, fear, sadness, surprise, and anger) and additionally the neutral expression. Several informationbased subjects for recognizing facial emotions have been approached to represent the emotions of a particular reference, while others attempted to animate unconstrained and real facial emotions. We have incorporated images from the following datasets and we have created our own unique datasets:

- *Extended Cohn-Kanade Dataset (CK+):* The Extended Cohn-Kanade Dataset (CK+) is a public benchmark dataset for action units and emotion recognition. The dataset comprises a sum of 5,876 marked images of 123 beings, where the sequences distinguish from neutral to peak expression. Images in the CK+ dataset are all posed with similar backgrounds, mostly grayscale, and 640 x 490 pixels. Mostly, there are eight emotions with which each sequence is marked: neutrality, anger, happiness, sadness, surprise, fear, disgust, and contempt.
- *Japanese Female Facial Expressions (JAFFE)*: This dataset has a collection of 213 images of seven fundamental feelings (6 basic facial expressions and neutral), in addition to the neutral emotion presented by 10 Japanese female models. 60 Japanese subjects have labelled each of these images and a resolution of 256 x 256 pixels.
- *Karolinska Directed Emotional Faces (KDEF)*: Includes a set of 4900 photographs depicting the emotions of the human face. This set includes 70 consisting of half men and half women, showing six basic emotions and including neutral expressions with the resolution of 762 x 562 pixels. Every emotion is seen from five distinct angles and was captured in two meetings.
- *Facial Expression Recognition 2013 (FER-2013)*: This is created with Google's image search API to search for photos of people that coordinate a bunch of 184 emotions related to terms such as "happiness", "madness" and so on. These terms, combined with words identified by age, gender, and nationality, resulted in a grayscale image of 35,887 with a resolution of 48 x 48 planned for six major emotions, in addition to neutral emotions.
- *Real-Time Dataset:* When the user runs our application the emotions are recognized and the images captured, then the captured images are used to update the existing dataset, thereby creating a realtime dataset and increasing the accuracy.

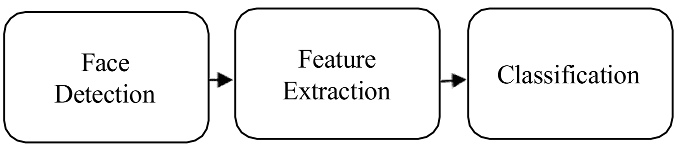

Figure 1. Model

III. METHODOLOGY

Emotion detection by facial expressions is a booming investigation in the field of Computer Vision. The progress made in this research tells us that there are three main steps to identify facial emotions, i.e., face detection, feature extraction and emotion classification. In our proposed model we used these three steps using deep learning as we found with our research that the computational time decreases, performance accuracy increases with this approach.

Face Detection

Face detection is an image pre-processing stage that is used to identify the facial expressions of humans. Image processing also advances the performance of the FER system. It has several types of processes such as adjustment of contrast, image clarity, cropping and scaling of the image. There are numerous methods for detection and image processing.

Haar Cascade

Haar Cascade is an object detection algorithm which is used to identify facial features in a real time video or image, the algorithm uses edge or line detection features. Here the darker areas are pixels with value 1, and lighter with value 0. Each of these values is in charge for finding one feature in the image.

OpenCV

OpenCV is an open-source computer vision and machine learning software system library. The library has over 2500 optimized algorithms, which has a comprehensive set of each classic and progressive computer vision and machine learning algorithms. These algorithms will be used to identify and acknowledge faces, detect objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, build 3D point clouds from stereo cameras, repair images along to generate a high-resolution image of a complete scene, appear similar images from an image database, etc. OpenCV leans mostly towards real-time vision applications.

Adaboost

Adaboost is implemented to eliminate the redundant feature of Haar. A precise number of these features can be combined to create a good classifier. The primary challenge is to search out these features. A variant of AdaBoost is used for selecting features and train the classifier. Using Adaboost we can determine the relevant feature and if it is present, it would have the binary value as 1, otherwise binary value 0.

Image Data Generator

Extraction of features

It is a crucial step for the FER system. It is the finding and identifying the features within the input image. Feature extraction is an important step for the better categorization of the emotion.

Image data generator is implemented for the generation of the batches containing the data of tensor images and is utilized within the domain of real-time data augmentation. We are able to loop over the data in batches once we utilize the image data generator in Keras. There are different techniques and arguments of the image data generator class that helps to outline the behavior of the data generation.

Convolutional Neural Network (CNN)

A convolutional neural network, or CNN, is a deep learning neural network aimed for processing structured array of data such as images or videos. CNN is widely used in computer vision and has become the most commonly used for many visual applications such as image classification, and has also found success in natural language processing for text classification.

A convolutional neural network is usually used to investigate visual images through processing data with gridlike topology. In additionally, referred to as a ConvNet. It is used to detect and classify objects in an image. A CNN often with up to 20 to 30 layers. The power of a CNN comes from the special kind of layer is known as convolutional layer. The usage of convolutional layer in an exceedingly CNN mirrors the structure of the human cortical region, wherever a series of layers method an incoming image associated determine a lot of more complex options.

Structure of CNN contains:

- Convolution Layer: Convolutional layer is also known for feature extraction. The filter is placed on top of the input image and the dot product is taken. The output is known as a feature map and is further fed to other layers. Pooling Layer: Pooling layer is used to reduce the size of the feature map. This layer is placed in between convolutional and Fully connected layer.
- ReLU Layer: The Rectified Linear Unit is a linear function that will produce a positive output for the input, otherwise if there is no input then the output would be zero. It has become a default function as it is easy to train and gives better performance.
- Fully Connected L*ayer*: The goal of the Fully Connected Layer is to classify the input image into various segments based on the trained dataset.

After extraction of all features a Softmax classifier is used for classification.
There are several transfer learning models we selected a couple to analyse our dataset namely MobileNet and VGG16.

Classification

This is the last step for the FER system, the classifier is used to extract expressions. The classification is done in 7 emotions. The facial emotions that can be detected and classified by this system are Happy, Sad, Anger, Fear, Disgust Surprise and Neutral. We have used OpenCV for image processing tasks where we identify a face from a live webcam feed and then process that video with the dataset we have already trained.

Transfer Learning Models

Transfer learning is a machine learning technique that enables data scientists to benefit from the knowledge gained from a previously used machine learning model for a similar task. There are several models include:

MobileNet

The MobileNet model is meant to be utilized in mobile applications, and it's TensorFlow's first mobile computer vision model. MobileNet uses depthwise
separable convolutions. It considerably reduces the quantity of parameters when contrast to the network with regular convolutions with identical depth within the nets. This leads to light-weight deep neural networks [26,27].
MobileNet is a model that will constant convolution as done by CNN to filter images however in an exceedingly totally different method than those done by the previous CNN. It uses the concept of Depth convolution and Pointwise convolution that is totally different from the commonly convolution as done by normal CNNs [26]. This will increase the potency of CNN to predict images and hence they will be able to contend within the mobile systems in addition. Since these ways of convolution bring to the comparison and recognition time plenty, therefore it provides a higher response in an exceedingly very short time and hence we tend to utilize them as our image recognition model.

Architecture of MobileNet:

MobileNets for mobile and embedded vision applications is planned, which supported an efficient design that uses depthwise separable convolutions to create light-weight deep neural networks.
Two straightforward global hyper-parameters that expeditiously tradeoff between latency and accuracy take place to introduce.

Advantages of MobileNet:

- Reduced network size - 17MB.
- Reduced range of parameters - 4.2 million.
- Faster in performance and are helpful for mobile applications.
- Small, low-latency convolutional neural network.

Disadvantages of MobileNet:

- It is less correct than alternative progressive networks. There is just a small reduction in accuracy in comparison to alternative networks

VGG16

VGG may well be a classical convolutional neural network (CNN) design. The sixteen in VGG16 refers that it has sixteen layers that have weights. It's 3*3 convolutions but immeasurable filters. The thought of this model was projected in 2013, but the actual model was submitted throughout the ILSVRC ImageNet Challenge in 2014. VGG16 is one in each of the numerous innovations that sealed the approach for several innovations that followed throughout this field.

VGG16 was known to be the most effective performing arts model on the ImageNet dataset. It's free a series of convolutional network models starting with VGG, which might be applied to face recognition and image classification, from VGG16 to VGG19. The initial purpose of VGG's analysis of convolutional networks is to know however the depth of convolutional networks affects the accuracy and accuracy of enormous scale image classification and recognition

Architecture of VGG16:

The number 16 within the name VGG refers to the certainty that it is 16 layers deep neural network (VGGnet). This suggests that VGG16 may be a pretty intensive network and includes a total of around 138 million parameters. Even as per the current standards, it's an enormous network. However, VGGNet16 architecture's simplicity is what makes the network a lot of appealing. Simply by observing its design, it will be same that it's quite uniform.

There are rare convolution layers followed by a pooling layer that reduces the height and the width. If we glance at the number of filters that we are able to use, around 64 filters are convenient that we are able to double to concerning 128 and then to 256 filters. Within the last layers, we are able to use 512 filters.

Advantages of VGG16:

- It is a very good architecture for bench marking on a particular task.
- similarly, pre-trained networks for VGG are obtainable freely on the internet, so that it's generally exercised out of the box for multiple uses
- Disadvantages of VGG16:
- It's bitterly slow to train.
- The network architecture loads themselves are relatively bulky

Parameters	VGG16	MobileNet
ImageNet Accuracy	71.5%	70.6%
Learnable Parameters (Millions)	138	4.2
Using our dataset Accuracy	93.8%	94.4%
Layers	16	28
Size (MB)	53	17

Figure 2. Comparison of VGG16 vs MobileNet

According to the implementation, we have arrived at the conclusion that the accuracy of the model depends not only on the number of layers but also on the types of layers as well as the data provided to the model. Since we are using a real-time dataset which stores the images of the users while the application is running, the data generated can be used to increase the accuracy of the model. Overall, we have achieved higher accuracy and tried to implement the system using transfer learning models.

V. CONCLUSION

It is important to note that there is no specific way to build a neural network that would work well. Different problems would require different network architecture and a lot of trial and errors to produce desirable validation accuracy. This is the reason why neural nets are often called as "black-box algorithms.". From the research conducted till now from research papers and articles, we have decided to use the CNN algorithm for training our deep learning model which will extract features from the images and will predict the emotions. As with the database, we have decided to use the JAFFE, CK+ and FER2013 datasets to train our model as these datasets provide higher accuracy and also have an adequate quality of images which will give us a better accuracy overall. This research and software implementation will be beneficial for future works related to facial feature extraction.

ACKNOWLEDGMENT

We'd like to express our sincere gratefulness to Prof.Priyanka Jadhav, whose role as project guide was invaluable for the project. We are extremely thankful for the keen interest she took in advising us, for the reference materials provided, and for the moral support extended to us. We express our deep sense of gratitude and humble thanks, for her valuable guidance throughout the project work. Furthermore, we are indebted to Prof. R.H. Borhade, HOD Computer, Dr. A. V. Deshpande, Principal whose constant encouragement and motivation inspired us to do our best

REFERENCES

[1] Zafar B, Ashraf R, Ali N, Iqbal M, Sajid M, Dar S, Ratyal N (2018) A novel discriminating and relative global spatial image representation with applications in CBIR. Appl Sci 8(11):2242

[2] Ali N, Zafar B, Riaz F, Dar SH, Ratyal NI, Bajwa KB, Iqbal MK, Sajid M (2018) A hybrid geometric spatial image representation for scene classification. PLoS ONE 13(9):e0203339 Ali N, Zafar B, Iqbal MK, Sajid M, Younis MY, Dar SH, Mahmood MT Lee IH (2019) Modeling global geometric spatial information for rotation invariant classification of satellite images. PLoS ONE 14:7

[1] Ali N, Bajwa KB, Sablatnig R, Chatzichristofis SA, Iqbal Z, Rashid M, Habib HA (2016) A novel image retrieval based on visual words integration of SIFT and SURF. PLoS ONE 11(6):e0157428

[2] C. R. Darwin. The expression of the emotions in man and animals. John Murray, London, 1872.

[3] Ali N, Zafar B, Iqbal MK, Sajid M, Younis MY, Dar SH, Mahmood MT, Lee IH (2019) Modeling global geometric spatial information for rotation invariant classification of satellite images. PLoS ONE 14:7

[4] Ali N, Bajwa KB, Sablatnig R, Chatzichristofis SA, Iqbal Z, Rashid M, Habib HA (2016) A novel image retrieval based on visual words integration of SIFT and SURF. PLoS ONE 11(6):e0157428

[5] C. R. Darwin. The expression of the emotions in man and animals. John Murray, London, 1872.

[6] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2):124, 1971.

[7] Matsumoto D (1992) More evidence for the universality of a contempt expression. Motiv Emot 16(4):363

[8] Liew, C.F.; Yairi, T. Facial Expression Recognition and Analysis: A Comparison Study of Feature Descriptors. IPSJ Trans. Comput. Vis. Appl. 2015, 7, 104–120.

[9] Ko, B.C. A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors 2018, 18, 401.

[10] Huang, Y.; Chen, F.; Lv, S.; Wang, X. Facial Expression Recognition: A Survey. Symmetry 2019, 11, 1189.

[11] Li, S.; Deng, W. Deep Facial Expression Recognition: A Survey. IEEE Trans. Affect. Comput. 2020.

[12] Xiao, X.Q.; Wei, J. Application of wavelet energy feature in facial expression recognition. In Proceedings of the 2007 International Workshop on Anti-Counterfeiting, Security and Identification (ASID), Xiamen, China, 16–18 April 2007; pp. 169–174.

[13] Zhao, L.; Zhuang, G.; Xu, X. Facial expression recognition based on PCA and NMF. In Proceedings of the 2008 7th World Congress on Intelligent Control and Automation, Chongqing, China, 25–27 June 2008; pp. 6826–6829.

[14] Zhao, X.; Shi, X.; Zhang, S. Facial Expression Recognition via Deep Learning. IETE Tech. Rev. 2015, 32, 347–355.

[15] Pranav, E.; Kamal, S.; Chandran, C.S.; Supriya, M. Facial emotion recognition using deep convolutional neural network. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; pp. 317–320.

[16] Sajid M, Iqbal Ratyal N, Ali N, Zafar B, Dar SH, Mahmood MT, Joo YB (2019) The impact of asymmetric left and asymmetric right face images on accurate age estimation. Math Probl Eng 2019:1–10

[17] Ratyal NI, Taj IA, Sajid M, Ali N, Mahmood A, Razzaq S (2019) Threedimensional face recognition using variance-based registration and subject-specific descriptors. Int J Adv Robot Syst 16(3):1729881419851716

[18] Ratyal N, Taj IA, Sajid M, Mahmood A, Razzaq S, Dar SH, Ali N, Usman M, Baig MJA, Mussadiq U (2019) Deeply learned pose invariant image analysis with applications in 3D face recognition. Math Probl Eng 2019:1–21

[19] Sajid M, Ali N, Dar SH, Iqbal Ratyal N, Butt AR, Zafar B, Shafique T, Baig MJA, Riaz I, Baig S (2018) Data augmentation-assisted makeupinvariant face recognition. Math Probl Eng 2018:1–10

[20] Xie S, Hu H (2018) Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks. IEEE Trans Multimedia 21(1):211

[21] Danisman T, Bilasco M, Ihaddadene N, Djeraba C (2010) Automatic facial feature detection for facial expression recognition. In: Proceedings of the International conference on computer vision theory and applications, pp 407– 412. https://doi.org/10.5220/0002838404070412

[22] Mal HP, Swarnalatha P (2017) Facial expression detection using facial expression model. In: 2017 International conference on energy, communication, data analytics, and soft computing (ICECDS). IEEE, pp 1259–1262

[23] Parr LA, Waller BM (2006) Understanding chimpanzee facial expression: insights into the evolution of communication. Soc Cogn Affect Neurosci 1(3):221

[24] Dols JMF, Russell JA (2017) The science of facial expression. Oxford University Press, Oxford.

[25] A. Jaiswal, A. Krishnama Raju and S. Deb, "Facial Emotion Detection Using Deep Learning," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154121.

[26] F. Chollet. Xception: Deep learning with depthwise separable convolutions. arXiv preprint arXiv:1610.02357v2, 2016.

[27] M. Courbariaux, J.-P. David, and Y. Bengio. Training deep neural networks with low precision multiplications. arXiv preprint arXiv:1412.7024, 2014.

[28] Yundong Zhang, Haomin Peng haomin and Pan Hu, ―Towards Real time Detection and Camera Triggering, ‖ CS341.

[29] Ibai Gorordo Fernandez and Chikamune Wada, ―Shoe Detection Using SSD-MobileNet Architecture, ‖ 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech 2020).

[30] Yu-Chen Chiu, Chi-Yi Tsai, Mind-Da Ruan, Guan-Yu Shen and TsuTian Lee, ―Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems, ‖ ©2020 IEEE.

[31] Andres Heredia and Gabriel Barros-Gavilanes, ‖ Video processing inside embedded devices using SSD-Mobilenet to count mobility actors, ‖ 978-1-7281-1614-3/19 ©2019 IEEE.

[32] Sreelakshmi P, Sumithra (2019) Facial expression recognition to robust to partial occlusion using MobileNet. Int J Eng Res Technol (IJERT) Vol. 8. No. 06.

[33] Ravi A (2018) Pre-trained convolutional neural network features for facial expression recognition. arXiv preprint arXiv:1812.06387

[34] K. Simonyan, A. Zisserman, ―Very deep convolutional networks for largescale image recognition, ‖ 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 1–14, 2015.

[35] Shengbin Wu ―Expression Recognition Method Using Improved VGG16 Network Model in Robot Interaction, ‖ Volume 2021 |Article ID 9326695 | https://doi.org/10.1155/2021/9326695

[36] I Gede Putra Kusuma Negara, Andreas Pangestu Lim, Jonathan Jonathan ―EmotionRecognitiononFER-2013FaceImagesUsingFine- TunedVGG-16, ‖ Article • January 2020 DOI: 10.25046/aj050638