e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:06/June-2023 Impact Factor- 7.868 www.irjmets.com

SOCIAL BEHAVIOR OF AUTONOMOUS VEHICLE

Mr. Hrishikesh Bhave*1, Mr. Swanand Joshi*2

*1,2Assistant Professor VPM's MPCOE, Velneshwar M.Tech In Mechanical Engineering, India.

DOI : https://www.doi.org/10.56726/IRJMETS41294

ABSTRACT

It is interesting to hear about your use of automatic facial expression recognition in the context of autonomous vehicles. The automotive industry is indeed heading towards autonomous vehicles, and AI plays a crucial role in achieving that goal. Facial expression recognition can be a valuable tool for understanding the social behaviour of drivers and enhancing human-computer interaction. Convolutional Neural Networks (CNNs) have proven to be highly effective in various image-related tasks, including facial expression identification. CNNs are used for feature extraction and inference, and the specific architectures and approaches can vary significantly among different works. The performance impact of these elements is typically assessed through ongoing research and experimentation .In this paper, we reviewed the state of the art in image-based facial expression identification using CNNs and highlighted algorithmic variations and their performance impact. By combining several contemporary deep CNNs, you were able to achieve an accuracy of 70.47% on the FER2013 test without requiring additional training data or face registration. This result suggests that your approach has outperformed prior studies in this area. Using facial expression recognition to control autonomous vehicles based on the driver's emotions introduces an intriguing concept. It opens up possibilities for creating more personalized and intuitive driving experiences. However, it is important to consider various factors, such as the reliability and real-time nature of the emotion recognition system, ensuring driver safety, and addressing potential challenges or limitations associated with this technology. Continued research and development in this field will likely contribute to further advancements in autonomous vehicles and human-computer interaction, ultimately shaping the future of transportation

Keywords: Face Recognition, Image Processing, Face Detection, Opencv, Tensor Flow.

I. INTRODUCTION

Using a coding scheme, we address the social behaviour of the driver and the autonomous vehicle in this paper. Understanding how to read facial expressions. Recognizing facial expressions is essential for non-verbal human communication, and the creation, reception, and interpretation of face expressions have all received extensive research. [1]. Given the significance of facial emotions in human contact, computer vision's ability to automatically perform Facial Expression Recognition (FER) opens up a variety of unique applications in areas like data analytics and human-computer interaction[2].Consequently, FER has been widely studied and significant progress has been made in this field. In fact, recognizing basic expressions under controlled conditions (e.g. frontal faces and posed expressions) can now be considered a solved problem [1]. The term basic expression refers to a set of expressions that convey universal emotions, usually anger, disgust, fear, happiness, sadness, and surprise. Recognizing such expressions under naturalistic conditions is, however, more challenging. This is due to variations in head pose and illumination, occlusions, and the fact that unopposed expressions are often subtle, illustrates. Reliable FER under naturalistic conditions is mandatory in the aforementioned applications, yet still an unsolved problem Convolutional Neural Networks (CNNs) have the potential to overcome these challenges. CNNs have enabled significant performance improvements in related tasks, and several recent works on FER successfully utilize CNNs for feature extraction and inference. These works differ significantly in terms of CNN architecture, preprocessing, as well as training and test protocols, factors that all affect performance. It is therefore not possible to assess the impact of the CNN architecture and other factors based on the reported results alone. Being able to do so is, however, required in order to be able to identify existing bottlenecks in CNN-based FER, and consequently for improving FER performance. The aim of this paper is to shed light on this matter by reviewing existing CNN-based FER methods and highlighting their differences (Section II), as well as comparing the utilized CNN architectures empirically under consistent settings(Section III). On this basis, we identify existing bottlenecks and directions for improving FER performance. Finally, we confirm empirically that overcoming one such bottleneck improves performance

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:06/June-2023 Impact Factor- 7.868 www.irjmets.com

substantially, demonstrating that modern deep CNNs achieve competitive results without auxiliary data or face registration (Section IV). An ensemble of such CNNs obtains a FER2013 test accuracy of 75.2%, outperforming existing CNN-based FER methods. Drowsiness of the drivers is one of the key issues for majority of road accidents. Drowsiness threatens the road safety and causes severe injuries sometimes,resulting in fatality of the victim and economical losses. Drowsiness implies feeling lethargic, lack of concentration, and tired eyes of the drivers while driving vehicles. Most of the accidents happen in India due to the lack of concentration of the driver. We have developed a system to detect driver drowsiness and alert them promptly. Drowsiness is a critical factor that can significantly impact driver performance and safety. By incorporating image-based methods and adapting them to support image sequences, we can effectively monitor and analyze the driver's state over time. Integrating per-frame results using graphical models is a suitable approach for sequence-based Facial Expression Recognition (FER). By considering the temporal information and the context provided by image sequences, you can gain a more comprehensive understanding of the driver's drowsiness level. This enables you to take appropriate actions to prevent potential accidents or hazards. By reviewing the state of the art in image-based facial expression identification using CNNs and highlighting algorithmic variations and their performance impact, we are provide valuable insights and guidelines for developing efficient and accurate systems for detecting drowsiness in drivers. Continued research in this area will contribute to improving the performance and reliability of such systems, ultimately enhancing driver safety on the roads. Additionally, integrating these findings and techniques into autonomous vehicles can further augment the capabilities and responsiveness of the vehicles to ensure a safe and comfortable driving experience for everyone.

II. METHODOLOGY

For this study, we used the FER-2013 dataset, which is made up of around 37,000 well-structured grayscale photographs of faces measuring 48 by 48 pixels. The faces in the photographs are processed so that they are almost balanced and take up approximately the same amount of space in each picture. Every picture must be assigned to one of the seven classes that represent various face expressions of emotion. These facial expressions have been divided into six categories: 0 represents anger, 1 disgust, 2 fear, 3 happiness, 4 sadness, 5 surprise, and 6 neutrality. Each facial expression type is represented by a single example in Figure 1.. In addition to the image class number (a number between 0 and 6), the given images are divided into three different sets which are training, validation, and test sets. There are about 29,000 training images, 4,000 validation images, and 4,000 images for testing.

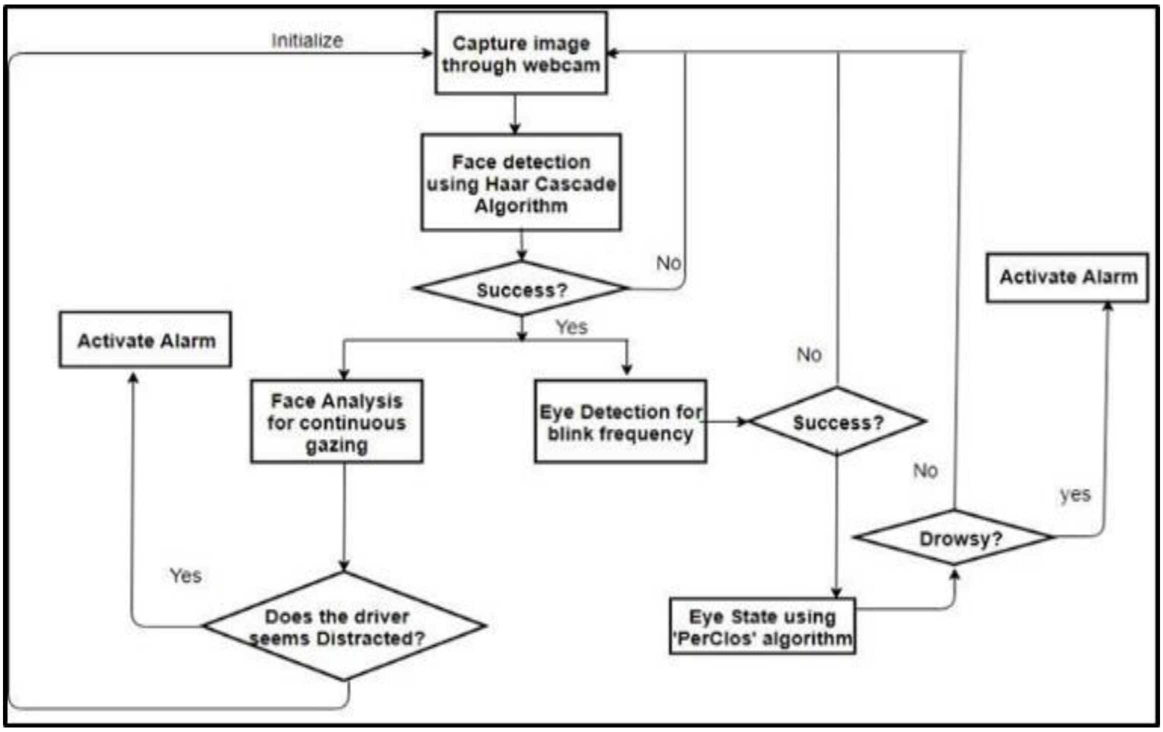

Fig.1 Architecture[5]

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:06/June-2023 Impact Factor- 7.868 www.irjmets.com

Architecture

This is the architecture for detecting the drowsiness and facial detection of the driver First of all the system captures images through the webcam and after capturing it detects the face through haar cascade algorithm. . It uses haar features which can detect the face. If the system founds it as face then it will proceed for next phase i.e. eye detection. The eye is also detected using haar cascade features and it is used for blink frequency. The state of eye will be detected using perclos algorithm. Through this algorithm we can find the percentage of time the eye lids remains closed. If it found eyes in closed state then it detects driver in drowsy state and alerts him by an alarm. In some cases distraction can be measured by continuous gazing. The drivers face is analyzed continuously to detect anydistraction. If found then alarm is activated by the system.

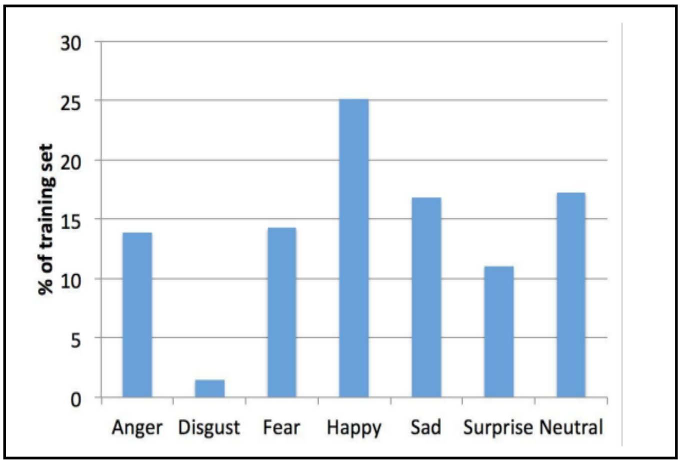

Figure 2: Distribution of different emotions across the FER-2013 dataset [2]

Recognize- Facial-Expressions for all image, depicting facial expressions face dataset do convert the image to grey-scale detect frontal face in and crop only the face extract POOL5 (256×6×6) features for cropped face using DCNN copy predefined facial expression label for each image as per the input dataset use the resulting POOL5 vector of dimension 9216(256×6×6) for tenfold and leave-one-out cross validation with SVM classifier to recognize facial expression [4].

A. CNN

We evaluated both a variety of preprocessing techniques as well as several model architectures, ultimately developing a custom CNN model capable of attaining near-state-of-the-art accuracy of 70.47% on the FER-2013 test set. For preprocessing, we experimented with centering (i.e., subtracting mean) and scaling data. We found it generally helpful to subtract the mean found in the train distribution from all sets before training/evaluating. We also implemented data augmentation: we randomly rotate, shift, flip, crop, and sheer our training images. This yielded about a 10 p.p. increase in accuracies. We implemented several CNN architectures from papers applying emotion recognition to these and other datasets. Ultimately, what yielded the best performance was our custom developed CNN architecture (left). Analyzing error in neural networks is infamously difficult. We analyzed our error across different classes, as well as by visual inspection of images we classified correctly and incorrectly. One early observation was that we fail much more at certain emotions, and that we were failing to classify images where it was necessary to rely on fine details in the images (e.g., small facial features or curves). Due to this, we increased the number of layers and decreased filter sizes to increase the number of parameters in our network, which had a clear effect in allowing us to fit the dataset better. This led to some over fitting, which we addressed by using dropout, early stopping around 100 epochs, and augmenting our training set. Given this, we only start learning training set noise after achieving approx. 70% dev set accuracy; this is clear from plotting accuracy during training. This leaves us with some suggestions for future work, which largely focus on enabling increased parameterization of the network.

B. Real-Time Classification

There is a possibility that our training set may not represent the true distribution of lighting conditions in a true manner. As such, this suggests that our training set may not be a true representation of the distribution of lighting conditions in real-life situations. We might not be representing the distribution of lighting conditions in our training set as it is likely to be inaccurate. In light of the above, it might be suggested that our training set may not really represent the distribution of lighting conditions in the real world. In light of this, it seems that our training set may not adequately represent the distribution of lighting conditions in practice. Accordingly, it might be that the distribution of lighting conditions observed in our training set isn't accurate representation of the true distribution. It seems that out training set may not accurately reflect the distribution of lighting conditions in the environment. As a result, it may be that our training set does not truly represent the distribution of lighting conditions within each category of lighting. Accordingly, our training set of lighting conditions may not be a valid representation of the distribution of conditions in the real world. This suggests that the training set may not, in truth, provide us with a true representation of the distribution of lighting conditions in real life.

C. Python

Python is a high-level, general purpose programming language. It is developed by Guido van Rossum in 1991. It is supported in Wind0ws, Linux and macOS. It is simple and easy to understand language. It is an object-oriented programming language and having dynamic semantics. It is known for its simplicity, readability and having less maintenance costs. It is an interpreted language. Since there is no compilation, It is faster than other programming languages. It throws an exception when an error is detected due to which it is easy to catch the exception. It supports program modularity and code reusability. Large number of inbuilt modules, packages and libraries are already present which make it more advanced. Nowadays the developers are choosing python as their coding language because of all the advantages it have. Python libraries are free and already available on all the major platforms for use. Python is used to make web-applications, game development, machine learning and artificial intelligence. There are large number of python libraries and frameworks which are already present and can be used directly for development purpose. Happy, Sad, Angry, Surprise, Clam. These are five categories in which this model is train to judge the human facial emotion. If the expression is likely to be happy the output is happy or it may be depending upon the input expression of the webcam input value [6]. It will distinguish the constant articulation no need of any outside equipment. It runs on a straightforward PC and doesn't need any high computational force. It is straightforward and standard prepared model to foresee the outward appearance. This is simply the enhanced work since it doesn't need any sort of outer source. It will utilize a straight forward standard web cam to take input and for yield it will utilize the screen of PC. The yield will show in a hued framed square having feeling composed on its advantage right corner. It utilizes Open CV to perceive the face first then it will find a next way to change over the RGB picture to a dim scale picture. With this progression the picture is practically prepared to perceive the facial feeling.

2. Face detection using python

There were several factors that affected these algorithms, such as extreme head postures and different lighting conditions in the image, which are based on machine learning. According to the algorithms, they are based on machine learning and they were affected by factors such as extreme head postures as well as different lighting conditions in the image used in the analysis. There were a number of factors affecting these algorithms, such as extreme head postures and lighting conditions in the image, which were affected by factors such as machine learning and the methods used to calculate them. Several factors, including extreme head postures, different lighting conditions in the image, and different lighting conditions in the scene have affected these algorithms which are based on machine learning algorithms.

It is worth mentioning that these algorithms are based on machine learning; however, they have been adversely affected by factors such as extreme head postures and the lighting conditions present in the image. Machine learning algorithms were used in the development of these algorithms, which was affected by factors such as extreme head postures and changing lighting conditions in the image over the course of the learning process. A set of algorithms based on machine learning was used to analyze the images and use different factors such as extreme head postures and variations in lighting conditions in order to determine the results. Based on a machine learning approach, these algorithms were developed that are able to address factors like extreme head postures and different lighting conditions in an image that may affect their performance. Based on machine learning, these algorithms were affected by the fact that the images were taken under different lighting conditions and whether there were extreme head postures involved with the image as well. As a result of this

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:06/June-2023 Impact Factor- 7.868 www.irjmets.com

algorithms were based on machine learning and were affected by factors such as extreme head postures and different lighting conditions in the image.

3. Equations and Mathematical Expressions

PERcentage of eye CLOSure (PERCLOS) is defined as the proportion of time for which the eyelid remains closed more than 70-80% within a predefined time period. Level of drowsiness can be judged based on the PERCLOS threshold value.

$$EAR = (A + B) / (2.0 * C) \ldots\ldots(i)$$

where

A is the distance between the 2-points (p2 and p6)

B is the distance between the 2-points (p3 and p5)

C is the distance between 2-points (p1 and p4)

The accuracy formula is detailed below. It simply counts the number of samples where the model correctly predicted the emotion and divides it by the total number of samples in the testing set. Here, the testing set consists of about 7,178 images.

$$\frac{\text{Num. Correctly Predicted Emotions}}{\text{Total Num. Samples}}$$

III. RESULTS AND DISCUSSION

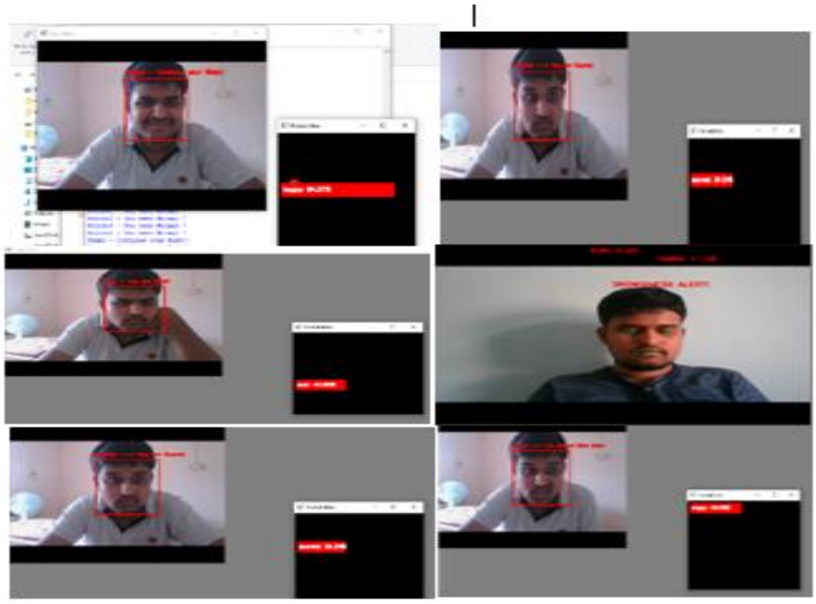

Fig.3 Results after program run

The five facial emotions shown in Figure 4 are:

1. Happy: The driver is smiling and has a relaxed expression.
2. Sad: The driver has a downturned mouth and furrowed brows.
3. Angry: The driver has a furrowed brow, narrowed eyes, and a clenched jaw.
4. Disgusted: The driver has a wrinkled nose and a protruded tongue.
5. Fearful: The driver has wide eyes, a furrowed brow, and a gaping mouth

The drowsiness of the driver is also shown in Figure 4. The driver's eyes are closed and their head is tilted back. After recognition of facial emotions, a message will appear on the screen indicating the driver's emotional state. The message will also provide probabilities of driver's behaviour. For example, if the driver is happy, the message might say "Driver is happy. Probability of driver speeding is 10%."The probabilities of driver's

IRJMETS

e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:05/Issue:06/June-2023 Impact Factor- 7.868 www.irjmets.com

behaviour are based on a number of factors, including the driver's emotional state, the time of day, the weather conditions, and the traffic conditions. The probabilities are used to help the driver stay safe and avoid accidents.

IV. CONCLUSION

For continued work on this project, we believe there are two major areas of focus that would improve our real-time emotion recognition system. First, we suggest fine tuning the architecture of the CNN used for the model to fit perfectly with the problem at hand. Some examples of this fine tuning include finding and removing redundant parameters, adding new parameters in more useful places in the CNN's structure, adjusting the learning rate decay schedule, adapting the location and probability of dropout and experimenting to find ideal stride sizes. A second area of focus lies in adapting the datasets to more closely reflect real-time recognition conditions. For example, simulating low light conditions and "noisy" image backgrounds, could help the model become more accurate in real-time recognition. Additionally making sure that the distribution of models in the training dataset accurately reflects the distribution of subjects that the system will see when running in real-time.

V. REFERENCES

[1] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: A survey of registration, representation, and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI),vol. 37, no. 6, pp. 1113–1133, 2015.

[2] M. V. B. Martinez, "Advances, Challenges, and Opportunities in Auto-matic Facial Expression Recognition," in Advances in Face Detection and Facial Image Analysis. Springer, 2016, pp. 63–100.

[3] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza,B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, Y. Zhou,C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor,M. Milakov, J. Park, R. Ionescu, M. Popescu, C. Grozea, J. Bergstra,J. Xie, L. Romaszko, B. Xu, Z. Chuang, and Y. Bengio, "Challenges inrepresentation learning: A report on three machine learning contests,"Neural Networks, vol. 64, pp. 59–63, 2015.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.

[5] K. He, X. Zhang, haoqing Ren, and J. Sun, "Deep Residual Learning for Image Recognition," CoRR, vol. 1512, 2015.

[6] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[7] Y. Tang, "Deep Learning using Support Vector Machines," in International Conference on Machine Learning (ICML) Workshops, 2013.

[8] Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," in ACM International Conference on Multimodal Interaction (MMI), 2015, pp. 435–442.

[9] B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, "Fusing Aligned and Non-Aligned Face Information for Automatic Affect Recognition in the Wild: A Deep Learning Approach," in IEEE Conf. Computer Vision and Pattern Recognition (CVPR) Workshops, 2016, pp. 48–57.

[10] A. Dhall, O. Ramana Murthy, R. Goecke, J. Joshi, and T. Gedeon, "Video and Image Based Emotion Recognition Challenges in the Wild: EmotiW 2015," in ACM International Conference on Multimodal Interaction (ICMI), 2015, pp. 423–426.

[11] B.-K. Kim, J. Roh, S.-Y. Dong, and S.-Y. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," Journal on Multimodal User Interfaces, pp. 1–17, 2016.

[12] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going Deeper in Facial Expression Recognition using Deep Neural Networks," CoRR, vol. 1511, 2015.

[13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.

[15] K. He, X. Zhang, haoqing Ren, and J. Sun, "Deep Residual Learningfor Image Recognition," CoRR, vol. 1512, 2015.

[16] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in IEEE Conferenceon Computer Vision and Pattern Recognition (CVPR), 2015.

[17] Y. Tang, "Deep Learning using Support Vector Machines," in International Conference on Machine Learning (ICML) Workshops, 2013.

[18] Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," in ACM International Conference on Multimodal Interaction (MMI), 2015, pp. 435–442.

[19] B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, "Fusing Aligned and Non-Aligned Face Information for Automatic Affect Recognition inthe Wild: A Deep Learning Approach," in IEEE Conf. Computer Vision and Pattern Recognition (CVPR) Workshops, 2016, pp. 48–57.

[20] A. Dhall, O. Ramana Murthy, R. Goecke, J. Joshi, and T. Gedeon, "Videoand Image Based Emotion Recognition Challenges in the Wild: EmotiW2015," in ACM International Conference on Multimodal Interaction(ICMI), 2015, pp. 423–426.

[21] B.-K. Kim, J. Roh, S.-Y. Dong, and S.-Y. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expressionrecognition," Journal on Multimodal User Interfaces, pp. 1–17, 2016.

[22] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going Deeper in FacialExpression Recognition using Deep Neural Networks," CoRR, vol. 1511,2015.

[23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan,V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2015.