



Wielding Neural Networks to Interpret Facial Emotions in Photographs with Fragmentary Occlusions

K. Anji Reddy^{1,*}, K. Sivarama Krishna², Bhanu Prakash Battula³, Bajjuri Usha Rani⁴, P. V. V. S. Srinivas⁵

¹Senior Assistant Professor, Department of Computer Applications, V.R.Siddhartha Engineering College, Vijayawada, India

²Associate Professor, Dept. of CSE, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India

³Professor & Head, Department of CSD, KKR & KSR Institute of Technology and Sciences, Guntur, India

⁴Sr. Assistant Professor, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, India

⁵Associate Professor, Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram Guntur, Andhra Pradesh, India

Emails: anjireddy5558@gmail.com; sivaramkosuru@gmail.com; Prakashbattula33@gmail.com; bajjuri.usharani2022@gmail.com; cnu.pvv@kluniversity.in

Abstract

For many years, scientists have studied the way people express their emotions through body language and facial expressions. However, it is extremely difficult to accurately interpret the emotions of a person from just a single image. Interpreting facial emotions in photographs is a complex task. It is challenging to accurately detect facial emotions with the help of neural networks when the face is occluded with fragmentary blocks. With the advent of technology, emotion detection has become more accurate and reliable. It is now possible to use facial expression recognition in images to detect emotions such as happiness, sadness, anger, fear, surprise, and more. This research discusses the effectiveness of using neural networks to identify facial emotions in photographs with occlusions present. The datasets like Fer2013 dataset, CREMA-D and RAVDESS were used to train the model and the datasets were altered by implanting occlusions randomly in the images. The altered datasets were also used to evaluate the model. The challenges and opportunities that arise when neural networks are used in this context are explored. Additionally, insight is also provided into the best approach to accomplish the task.

Keywords: Neural Networks, Deep Learning; Occlusions; Emotion Interpretation; Human-Computer Interaction

1. Introduction

Facial emotion recognition is an important area of research, particularly in the field of computer vision. With the help of recent advancements in deep learning algorithms and artificial neural networks, it is now possible to detect facial emotions in photographs even when occlusions are present. This research will discuss the current state of neural networks to detect facial emotions in photographs with occlusions, exploring their advantages and limitations. In recent years, deep learning algorithms have become increasingly powerful for facial emotion recognition. These algorithms are capable of extracting subtle features from photographs and predicting the facial emotion of the subject. For example, a neural network could recognize a slight furrow in the brow of a subject as indicating a feeling of anger. By training a neural network on a large dataset of labeled images, it can be used to classify facial emotions in photographs. However, the presence of occlusions in the image can make it more difficult for a neural network to accurately detect facial emotions. Occlusions can range from glasses and hats to hands covering parts of the face or even facial hair. In this case, a code has been designed to overlap certain areas of the image like eyes and mouth. To train a neural network to recognize emotions in photographs with occlusions, the dataset must include images with occlusions and the datasets are tailored according to the needs. Specialized occlusion-aware neural networks can also be used to better recognize facial emotions in photographs with

occlusions. Overall, neural networks are effective for detecting facial emotions in photographs with occlusions, given the right training dataset. By training neural networks on large datasets of images with occlusions, they can be used to accurately predict emotions in photographs with occlusions. However, the presence of occlusions can still limit the accuracy of a neural network's predictions. Recent advancements in the field of emotion detection from facial expressions have allowed for faster and more accurate identification of emotional states. This has enabled the development of innovative applications such as emotion-based chatbots, interactive robots, and more. These advances are mainly due to the development of advanced computer vision algorithms. For example, convolutional neural networks (CNNs) can be trained to detect different facial expressions from an image. By analyzing the facial features from the image, the system is able to accurately determine if a person is happy, sad, or angry. Another recent development is the use of 3D facial models to improve the accuracy of emotion detection. This type of model is capable of providing the system with a 3D representation of a person's face, allowing the system to recognize subtle expressions and emotions. Finally, transfer learning techniques have allowed researchers to utilize pre-trained models that have already been trained on a large dataset of facial expressions. This type of model can be quickly adapted to a particular application and can be used to achieve accurate results. Overall, recent advancements in emotion detection have enabled the development of innovative applications that can accurately interpret a person's emotional state from an image. By leveraging advanced computer vision algorithms, 3D facial modeling, and transfer learning, researchers are able to create more accurate models for emotion detection. In a single image to detect the face, four methods based on feature, template, knowledge, and appearance are used. Hybrid techniques, however, are also used for detection of emotion. A top-down approach is used in knowledge-based methods. In this case, the face is located using human-coded rules such as skin color, facial features and template matching.

Researchers in [1] used various classifications to identify human emotions. According to the study, there are six basic emotions known as universal emotions, which include fear, delight, grief, surprise, anger, and contempt. Humans experience these emotions around the world. These universal sentiments can always be divided into two categories evidently, positive or negative. More feelings, such as amusement, pride, satisfaction, shame, excitement, and embarrassment are included and discussed later in [2]. Realtime emotion recognition and crowd emotion prediction are an important tool for understanding the thoughts and feelings of people. It has a wide range of applications in fields such as marketing, customer service, education, health, and entertainment. In marketing, real-time emotion recognition can help companies better understand customer sentiment and tailor their product offerings accordingly. Companies can use this data to better target potential customers and drive sales. In customer service, real-time emotion recognition can help to improve customer satisfaction and reduce customer churn. By understanding customer sentiment in real-time, customer service agents can quickly address customer concerns and resolve issues more effectively. In education, real-time emotion recognition can help teachers better understand student engagement and adapt their teaching style accordingly. For example, teachers can quickly adjust their lesson plans if they detect that students are becoming bored or disinterested. Realtime emotion recognition and crowd emotion prediction is also important for the health sector. By detecting signs of depression and anxiety in real-time, healthcare providers can better address mental health issues and provide more effective treatments for their patients. Overall, real-time emotion recognition and crowd emotion prediction is an important tool for understanding people's thoughts and feelings. It can be used to improve marketing, customer service, education, health, and entertainment. In the digital age, where images are everywhere, emotion in images has become an increasingly important topic. Understanding the emotions in images has become a complicated task, with occlusions making it even more difficult. Occlusions are portions of an image that are purposely or accidentally blocked from view. Neural networks are now being used to interpret facial emotions in photographs with fragmentary occlusions. This not only helps to identify the emotions in the image, but also to better understand how the presence of occlusions can affect the emotion interpretation of images. With such advancements, computer vision has the potential to play a major role in interpreting emotion in images with occlusions. Neural networks are a powerful tool for machine learning, and they can be used to interpret facial emotions in photos. By recognizing patterns and features in the images, neural networks can be trained to identify facial expressions, even when there are partial occlusions. Neural networks are capable of taking in a lot of data and analyzing it to find patterns, which makes them well-suited for analyzing facial emotions. Additionally, neural networks have the ability to recognize subtle nuances in facial expressions, such as slight changes in emotions. This can be incredibly useful for recognizing emotions in photographs that have been partially obscured. Through this, neural networks have the potential to revolutionize the way we interpret facial emotions and can help us better understand the emotional context of a photograph. Multimodal emotion recognition is an exciting field of research that has the potential to revolutionize how we interpret facial emotions from photos. By utilizing neural networks and other machine learning algorithms, we are able to analyze photographs with fragmentary occlusions and accurately interpret the emotions present. Through the combination of multiple data sources, such as facial expressions, body language, and speech, multimodal emotion recognition has the potential to be more accurate than traditional approaches to emotion recognition. This technology can be applied to numerous areas, such as healthcare, customer service, and entertainment. In healthcare, it can be used to diagnose depression and anxiety by tracking changes in facial emotions over time. In customer service, it can be used to interpret customer feedback and improve customer service experiences. Finally, in entertainment it can be used to create more immersive and interactive experiences. Overall, multimodal emotion recognition is an exciting and rapidly evolving area of research. Its potential applications are vast and it holds great promise for the future.

2. Related works

Recent research in the field of Neural Networks has focused on the ability to detect facial emotions in photographs with occlusions. Recent studies have shown that using a Convolutional Neural Network (CNN) to detect facial emotions can be an effective approach. CNNs have been used to detect facial expressions, including happy, sad, angry, and neutral emotions. Previous studies [28-33] demonstrated that convolutional neural network (CNN) architecture is appropriate for the problem, and therefore architectures of this type are utilized in all studies in this work because it is relatively simple. Furthermore, these networks can be trained to detect occluded facial expressions, such as those in photographs with obstructions. The use of a CNN architecture has been shown to improve the accuracy of facial emotion detection when compared to traditional methods. In addition to the use of CNNs, researchers have also used other types of Neural Networks to detect facial emotions. Recurrent Neural Networks (RNNs) have been used to process sequence data, such as video streams, and to detect emotions. Additionally, Generative Adversarial Networks (GANs) have been used to generate realistic facial images and detect emotions from those images. Finally, a variety of other methods have been employed to detect facial emotions. For example, researchers have used a combination of facial feature extraction and a Support Vector Machine (SVM) to classify emotions. Similarly, rule-based approaches have been used to detect facial emotions. Although these methods are not as accurate as Neural Networks, they can be useful in certain cases. Overall, the use of Neural Networks to detect facial emotions in photographs with occlusions is an active area of research. Although researchers proposed to recognize a person's face in real-time videos using a cascade classifier with a Haar-like features and local binary pattern histogram [3] to determine emotions no significant work has been done. Convolution neural networks (CNN) are commonly used in recent academic research on emotion recognition [4, 5]. CNN has shown promise in the areas of classification, feature extraction, and face detection. This method automatically extracts and classifies a characteristic, eliminating the need for manual methods. CNN's four fundamental components are convolution layers, activation functions, subsampling, and dense layers (fully connected layer). However, a pre-trained deep learning based CNN model incorrectly identified several occlusion-based instances of perplexed face images. The study [6] presents a CNN-based FER model with three convolutional layers and five filter sizes. The dropout layer was used as the regularization layer by the authors. In 3 minutes, the proposed model achieves an emotion recognition accuracy of 96%. Two convolutional layers are employed in [7], the first with five filter sizes and the second with seven. To reduce the size, the max-pooling layer has a 2x2 kernel, whereas there are 256 hidden neurons in the dense layer. It has a learning rate of 0.01 and was trained over 2000 epochs. The study produced promising results despite ignoring occlusions and lighting variations. The issue of occlusions in face in the perspective of facial recognition was studied way back in [8], where the authors proposed that the amount of degradation in recognition performance varies based on the location of the facial occlusion. Several methods have been developed to make the identification algorithm more resistant to occlusions. Leonardis and Bischof [9] proposed early work in this direction, demonstrating to process occlusions in framework of eigenface [10]. Their primary idea was to use a robust hypothesis-and-test paradigm to extract eigenspace coefficients from subsets of image points rather than computing the coefficients by superimposing the records onto the Eigen images.

Ayvaci et al. [11] considered the classification of occlusion area as a variational optimization process by minimization the convex approximation under the Lambertian reflection and stationary lighting hypothesis. Georgescu et al. [12] researched on a CNN method that is based on transfer learning for training occlusion based teacher-student. With deep learning's remarkable success, recent studies focused on using techniques of deep learning to recognize face expressions on pictures with occlusions and occlusion-less recognition measures, classification and hand-crafted extractions of features. Cheng et al. utilized Gabor filter for extracting features in 2014, and a multi-layered deep network was later used for training the network [13]. Toser et al. [14] To cope with head pose variance utilized CNN and tested their technique for faces with occlusion images by using a 21x 21 pixel patch in the cheek and lip region. Ashwin et al. [15] used pyramid HOG as well as LBP features to present a complex approach for face expression classification. From patches, from the face which change dramatically hybrid features are taken when the expression changes. Using the CK+ dataset, Using SVM the experimental results show a 94 percent emotion recognition. CNN model was used by the authors of [16] to extract feature from depth data. The model is divided into two layers: The first layer's the kernel size is 5 and feature map is 6, so max pooling has been applied. Eventually, second layer is premised on 6 feature maps with a kernel size as 5, maximum pooling of 2, and thereafter 12 feature maps with Softmax. The proposed method is an illumination variation that achieves an accuracy of 87.98% over 1000 epochs. [17] The article presents a hybrid of two emotion recognition models to obtain static spatial features in the multi-signal convolutional model, whereas to obtain dynamic temporal features and combine them, the part-based hierarchical recurrent neural network is utilized. The PHRNN model possesses 12 layers, whereas the MSCNN model possesses 6. The pre-training and fine-tuning approach has demonstrated remarkable success in NLP [18-20] and computer vision [20-25]. Eventually, [26] used speckle noise to add to datasets like CREMA-D, RAVDESS, and TESS to implement CNN models for emotion detection. [27] when using mel-spectrograms and spectrograms, compared the performance of CNN models. Two distinct architectures were used for this: the well-known ResNet-18 as well as a CNN. The majority of the experiments demonstrated that accuracy can be greatly enhanced by utilizing metric mel-spectrograms as a feature extraction method. However, one study demonstrated that models trained on spectrograms surpassed those trained on mel-spectrograms only by a small margin of course. As a result, as the data processing technique it is usually preferable to use mel-spectrograms.

3. Methodology

The datasets used are The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) is an important dataset for emotion recognition. The RAVDESS contains 24 professional actors, reciting two lexically-matched statements in a neutral North American accent. Each actor speaks both lines in a calm and emotional manner, with 7 different emotions: neutral, calm, happy, sad, angry, fearful and surprised. The dataset includes 1440 audio files in total; 60 files for each emotion. Each audio file is 1.5-2 seconds long and has a sampling rate of 44.1 kHz. It also includes 24 video files that are synchronized with the audio files. The RAVDESS is widely used in many studies on emotion recognition. It has been used to assess the performance of different machine learning algorithms, such as Support Vector Machines, Neural Networks, and Gaussian Mixture Models. Researchers have also used the RAVDESS to compare different speech recognition systems and to study the impact of emotional expression on speech recognition. The RAVDESS is a valuable resource for researchers in the field of emotion recognition. It provides a large and diverse dataset that can be used to develop and evaluate emotion recognition systems. The Affective Computing Group at the University of Missouri developed the CREMA-D dataset, which stands for Crowd-sourced Emotional Multimodal Actors Dataset. This dataset contains over 7,000 clips of actors expressing various emotions and reactions. The dataset is used for emotion recognition tasks, such as facial expression recognition, audio emotion recognition, and speech emotion recognition. It also provides a large-scale dataset for training and testing multimodal emotion recognition systems. CREMA-D contains audio and video clips of actors performing a wide variety of emotions, such as anger, disgust, happiness, sadness, and surprise. Each clip is annotated with human-labeled labels such as valence, arousal, and emotion. As the name suggests, the dataset was created from the crowdsourcing of actors from Emotionally Valenced Actor (EVA) project, which is a part of the PURE-Passion project. CREMA-D is an ideal dataset for research related to emotion recognition, multimodal emotion recognition, and affective computing. It is available for free and can be used for any educational or research purpose. The Fer2013 dataset is one of the most popular datasets for emotion recognition. It was created by researchers at the Technical University of Madrid in 2013, and it consists of 350,000 facial images of over 10,000 different subjects. The images in the dataset are labelled as one of seven expressions: angry, disgust, fear, happy, sad, surprise, or neutral. Each image is also labelled with one of four intensities, which allows the dataset to be used for both emotion recognition and emotion intensity estimation. The Fer2013 dataset is particularly useful because it is large enough to be used for training deep learning models, yet it still allows for quick prototyping and model development. Additionally, the images in the dataset are realistic, as they were collected from online videos and are not artificially generated. Overall, the Fer2013 dataset is a great resource for emotion recognition and emotion intensity estimation tasks. It is one of the most popular datasets in the field, and it is a valuable tool for researchers. The datasets were resized to 48 X 48 format. The resized images were grey scaled as it aids in the simplification of algorithms and the elimination of complexities associated with computational requirements. Later an algorithm was implemented that inserts occlusions on the images. The altered dataset is used to evaluate the model as per the requirement.

Occlusion is applied to images read from datasets. The cascade classifier is utilized in the process to identify faces in images. as well as the eyes from photographs. As input, an image is utilized to occlude the face as fragments. Occlusions are applied to images. The images that have been occluded are saved in the occluded repository. Following that, all of the images in the occluded repository are read and split in an 80:20 ratio. FERCNN is used to train the model. The label is extracted from the Altered/Hybrid training and testing data. The feature is extracted from the Altered/Hybrid training and testing data. Later, CNN is initialized, and the model is compiled. Plot has been trained to generate validation accuracy/loss values. The confusion matrix is generated using the predicted and actual values. The columns depict the predicted instances for each label, while the rows depict the actual number of instances for each label. Printing precision and recall, as well as other metrics. A function is defined to load the image and convert it from a PIL image instance to a NumPy array. The shape of an input array that is passed to it is expanded. When a new axis is inserted, it appears in the axis position of the resulting expanded array shape. In NumPy, axis is defined for arrays with more than one dimension. Later, the Expression is predicted for an image received as a parameter. Test Data is used to evaluate original images and performance. Similarly, the model is trained on the original CREMAD and RAVDESS datasets and evaluated on the altered datasets. The three outcomes are compared, and graphs are generated later.

a) ALGORITHM

```

Step 1:  $l = \text{len}(\text{Input Image folder})$ 
Step 2 : if ( $l > 0$ ) go to Step 3 else go to Step 10
Step 3: for ( $i = 0 ; l > 0 ; i = i + 1, l = l - 1$ )
    // Read Images and Apply Occlusion
     $\text{img}_i = \text{Random}(\text{Select}(\text{image from Input Folder}))$ 
    Resize of ( $\text{img}_i, 256, 256$ )
     $\text{face}_i = \text{face.Cascade Classifier}(\text{img}_i)$ 
    //Identifies faces from images
 $\text{eyes}_i = \text{eyes.Cascade Classifier}(\text{face}_i)$ 
    //Identifies eyes from face images
     $\text{obs}_i = \text{occluded image equal to the size of}(\text{eyes}_i)$ 
    // Image to occlude on eyes

```

```

    Occ_imgi = facei + obsi
    // Occluded image is placed on the Facial image
    //Occluded images are stored in the distinct Folder
        Occluded Folder = Occluder Folder + Occ_imgi)
Input Folder = Remove( imgi)
    // Removes the used input image
Step 4: l1 = len ( Occluded Folder)
Step 5: Read all the Images of Occluded Folder
Trocc,Tsocc = Random(Split (Occluded Folder(80,20))
// Splits data in 80: 20 Ratio
    FERCNN = FERCNN(Trocc)
    //Train the FERCNN Model with Occluded Images
    Evaluate(FERCNN(Tsocc))
    //Evaluate the Performance with Test Data)
Step 6: Read all the Images of Input Folder
Tr,Ts = Random(Split (Input Folder(80,20))
    // Splits data in 80: 20 Ratio
    FERCNN = FERCNN(Tr)
    //Train the FERCNN Model with Original Images
    Evaluate(FERCNN(Ts))
    //Evaluate the Performance with Test Data)
Step 7: Read all the Images of Input Folder & Occluded Folder
Tr,Ts = Random(Split (Input Folder(80,20))
    // Splits data in 80: 20 Ratio
    FERCNN = FERCNN(Tr)
    //Train the FERCNN Model with Original Images
    Evaluate(FERCNN(Ts))
    //Evaluate the Performance with Test Data)
Step 8: Read all the Images of Input Folder & Occluded Folder
Tr,Ts = Random(Split (Input Folder(80,20))
    // Splits data in 80: 20 Ratio
    FERCNN = FERCNN(Trocc)
    //Train the FERCNN Model with Original Images
    Evaluate(FERCNN(Ts))
    //Evaluate the Performance with Test Data)
Step 9: Compare the Performances of Steps 4, 5, 6
Step 10: End

```

4. Numerical results and discussion

In TABLE I. the results obtained can be observed. FER2013 OCCLUDED Dataset has obtained an Precision of 0.94 and the Highest PRECISION is obtained by RAVDESS 0.99. Similarly the F1-SCORE is also highest for the RAVDESS Occluded dataset 0.98. The F1-Score is slightly less in case of FER2013 Occluded dataset with 0.96 score. Yet the recall is 0.98 for FER2013 and RAVDESS Occluded dataset. But for the CREMAD Occluded dataset the recall is quite low with 0.86. In case of Fig. 1, It is the Count of Train images and test images wrt FER2013 dataset and similarly Fig. 9 and 12 are the count of Train and test images wrt CREMAD and RAVDESS Occluded datasets. The model accuracy and loss obtained by the FER2013 model are depicted in Fig. 2 and 3. The Heatmap obtained is displayed in Fig. 4. It is a 2-dimensional matrix figure that visualises numerical information in the manner of cells. Each heatmap cell is coloured, and the tones of colour symbolise some kind of correlation between the value and the dataframe. The heat maps generated from the CREMAD and RAVDESS Occluded dataset are depicted in Fig.8 and 12.

Table 1: PRECISION, RECALL, F1-SCORE, AND ACCURACY OF CNN MODELS TRAINED WITH DISTINCT DATASETS

<i>DATASET</i>	<i>PRECISION</i>	<i>RECALL</i>	<i>F1-SCORE</i>	<i>SUPPORT</i>	<i>ACCURACY</i>
FER2013 OCCLUDED	0.94	0.98	0.96	831	0.9137
CREMAD OCCLUDED	0.82	0.86	0.84	356	0.8735
RAVDESS OCCLUDED	0.99	0.98	0.98	336	0.9852

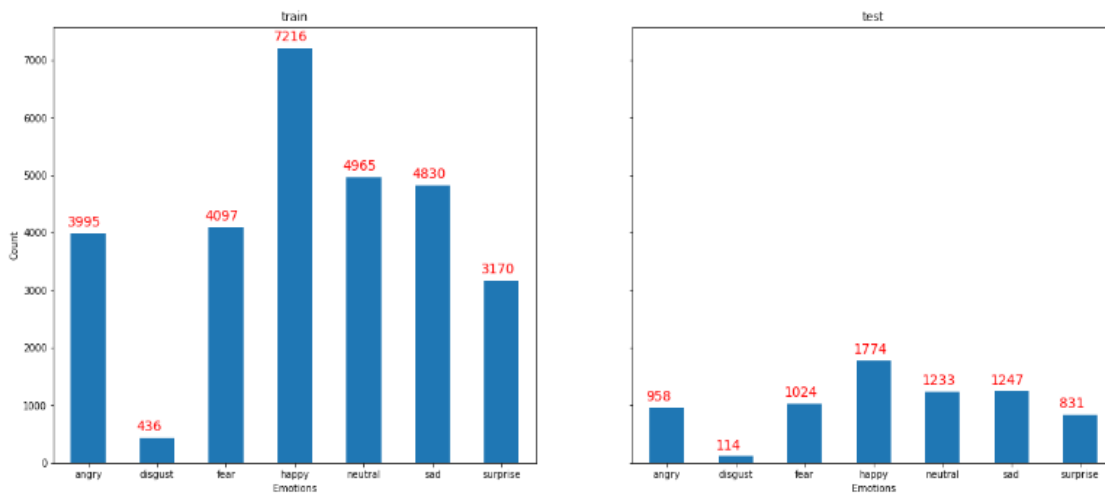


Figure 1. Count of Train images and test images wrt FER2013 dataset

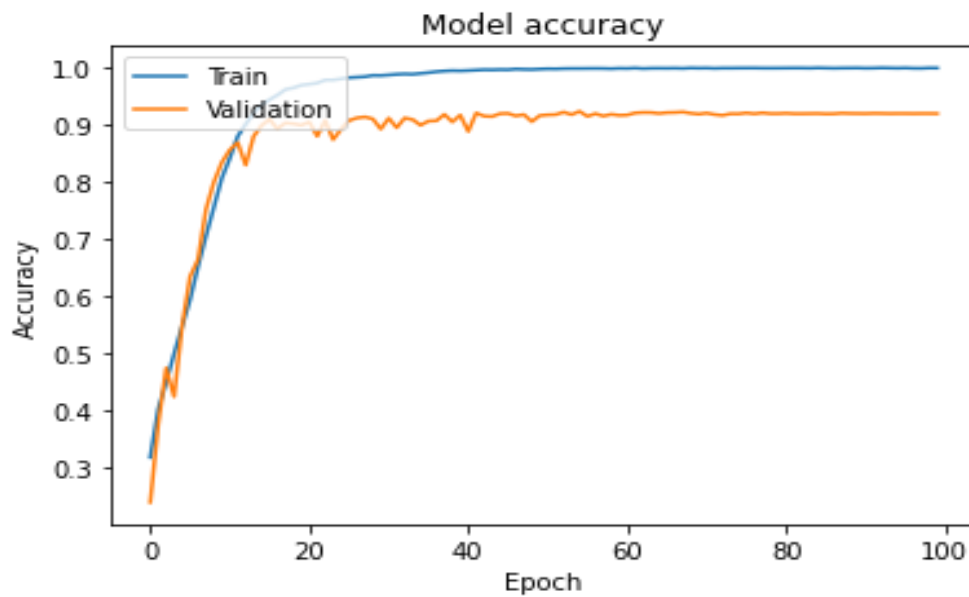


Figure 2. Model Accuracy obtained by evaluating with FER2013 Occluded dataset

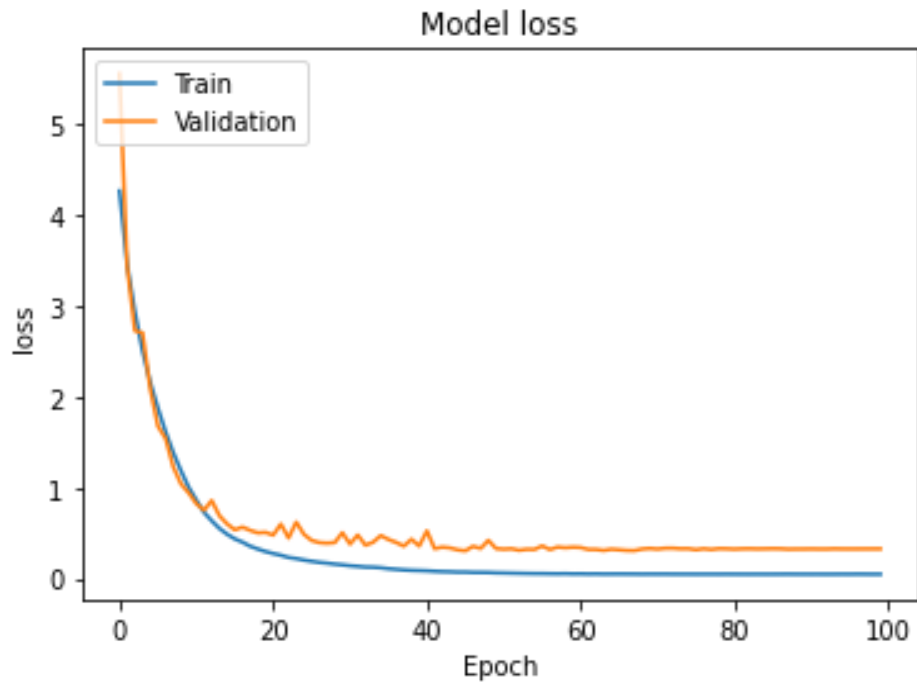


Figure 3. Model Loss obtained by evaluating with FER2013 Occluded dataset

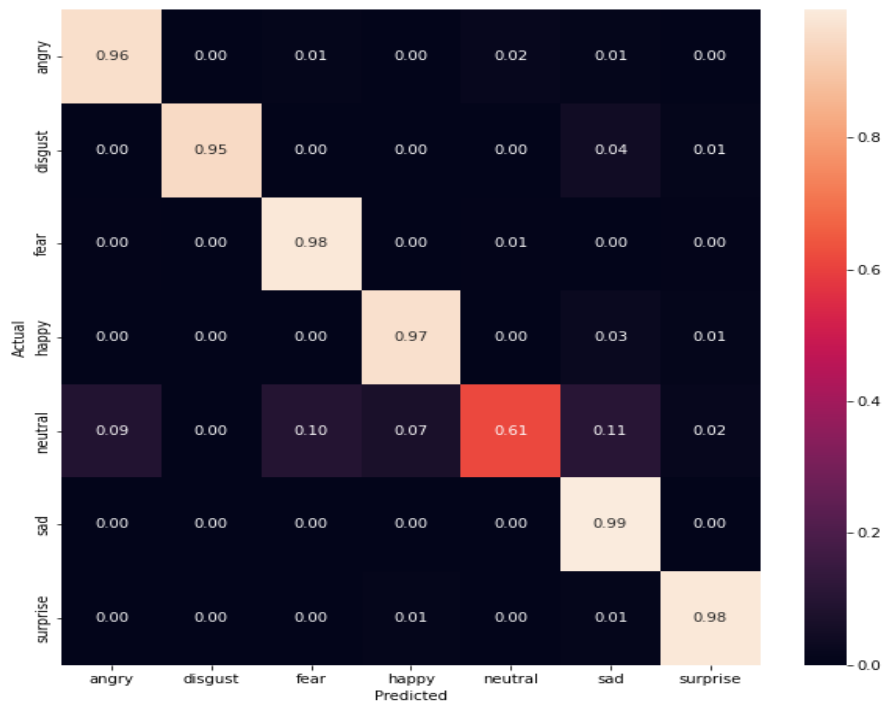


Figure 4. Heatmap of Model evaluated with FER2013 Occluded dataset.

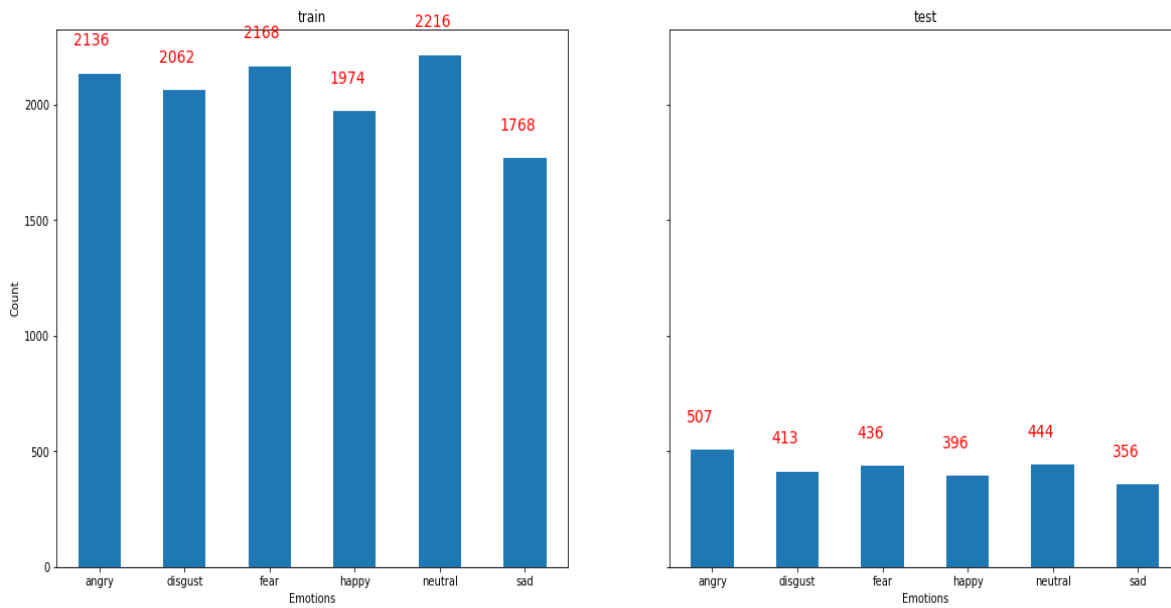


Figure 5. Count of Train images and test images wrt CREMAD Occluded dataset

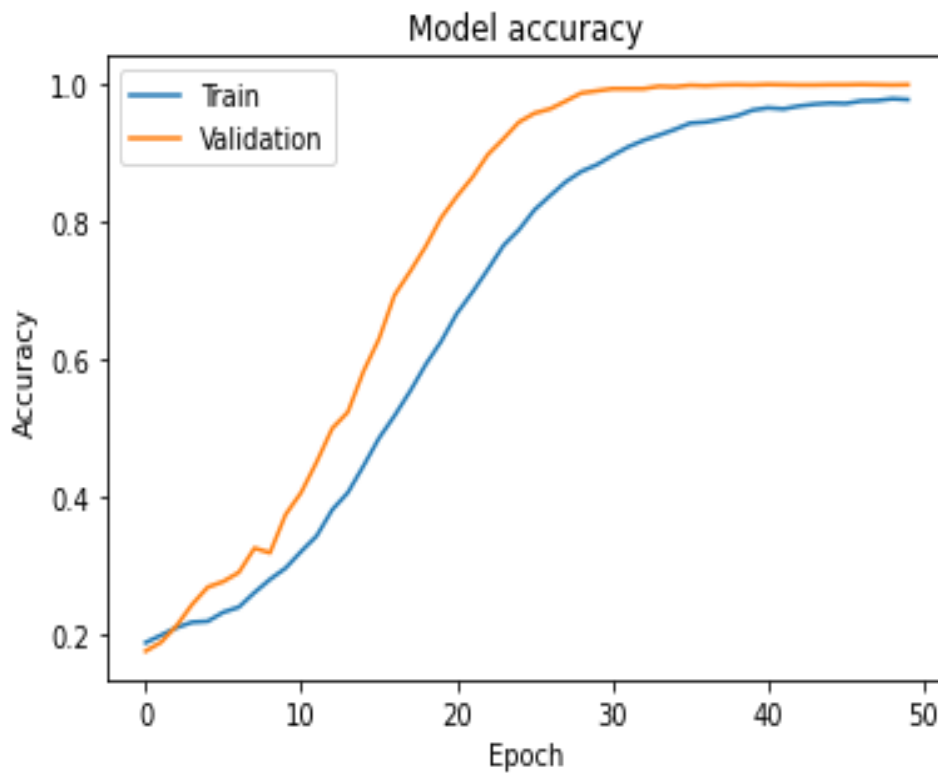


Figure 6. Model Accuracy obtained by evaluating with CREMAD Occluded dataset.

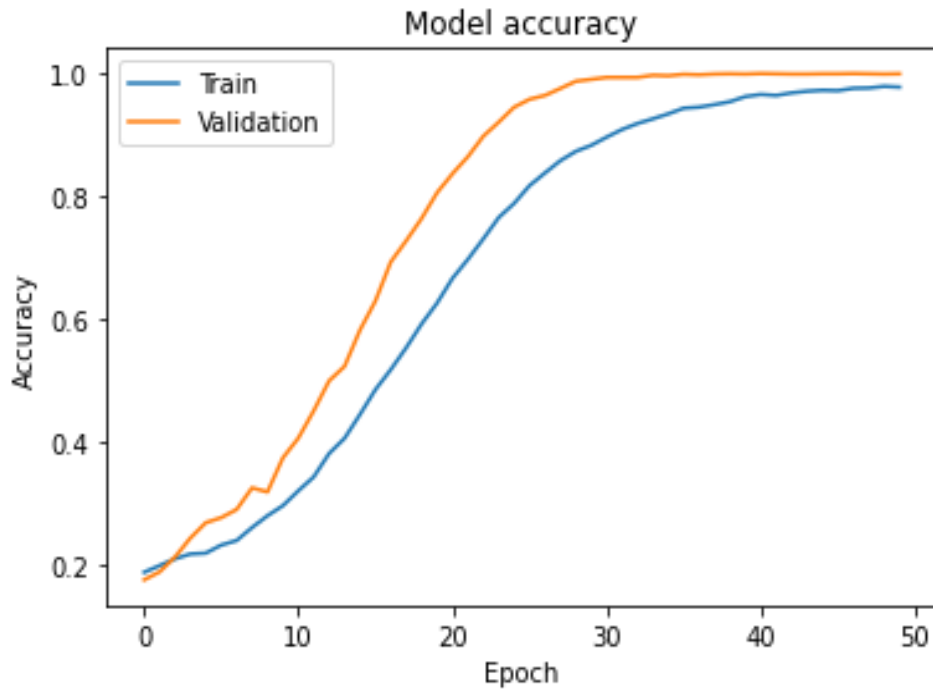


Figure 7. Model Loss obtained by evaluating with CREMAD Occluded dataset.

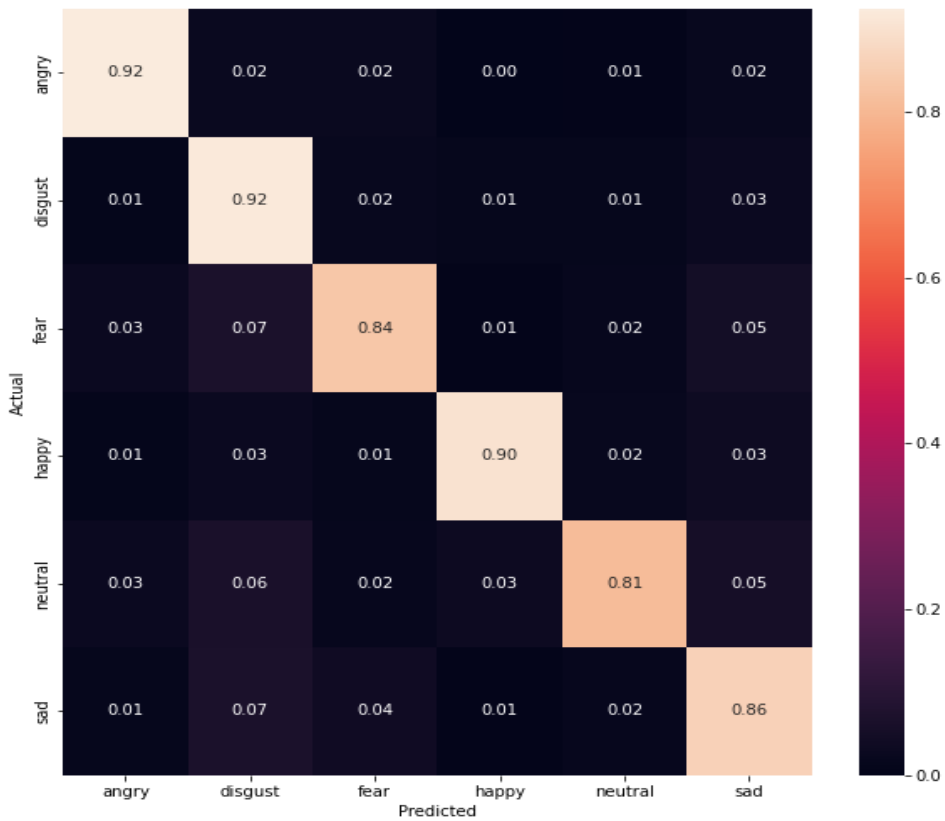


Figure 8. Heatmap of Model evaluated with CREMAD Occluded dataset RAVDESS Occluded

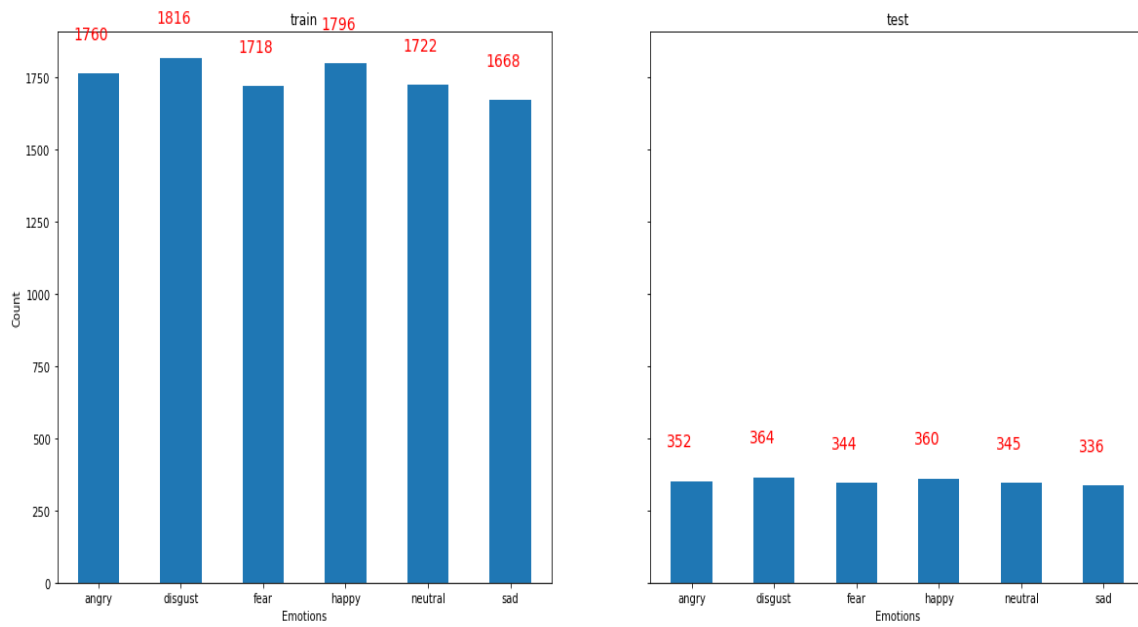


Figure 9. A Count of Train images and test images wrt RAVDESS Occluded dataset.

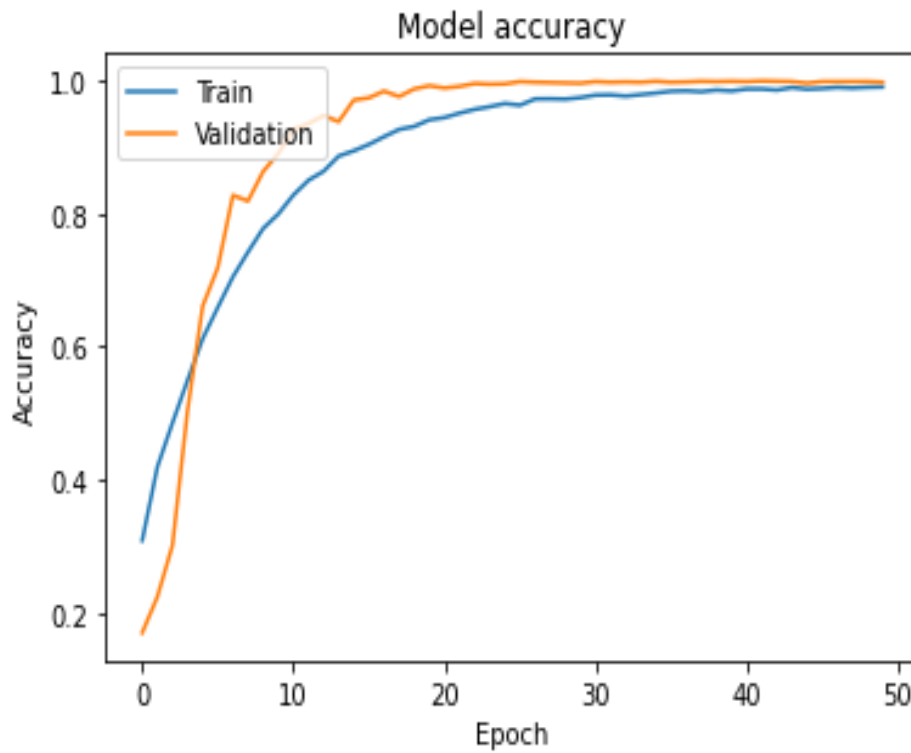


Figure 10. Model Accuracy obtained by evaluating with RAVDESS Occluded dataset

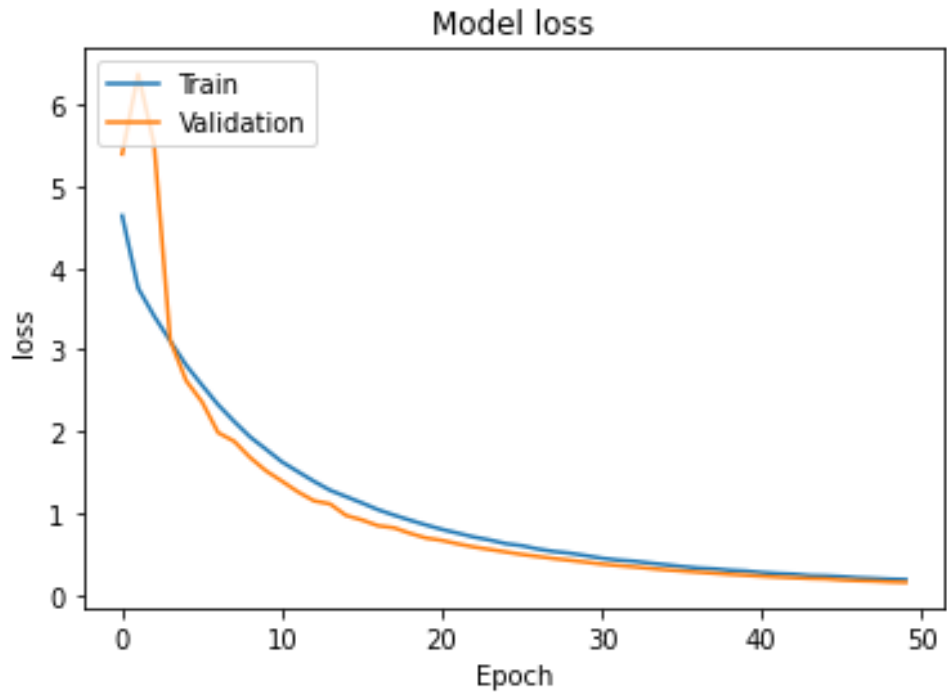


Figure 11. Model Loss obtained by evaluating with RAVDESS Occluded dataset

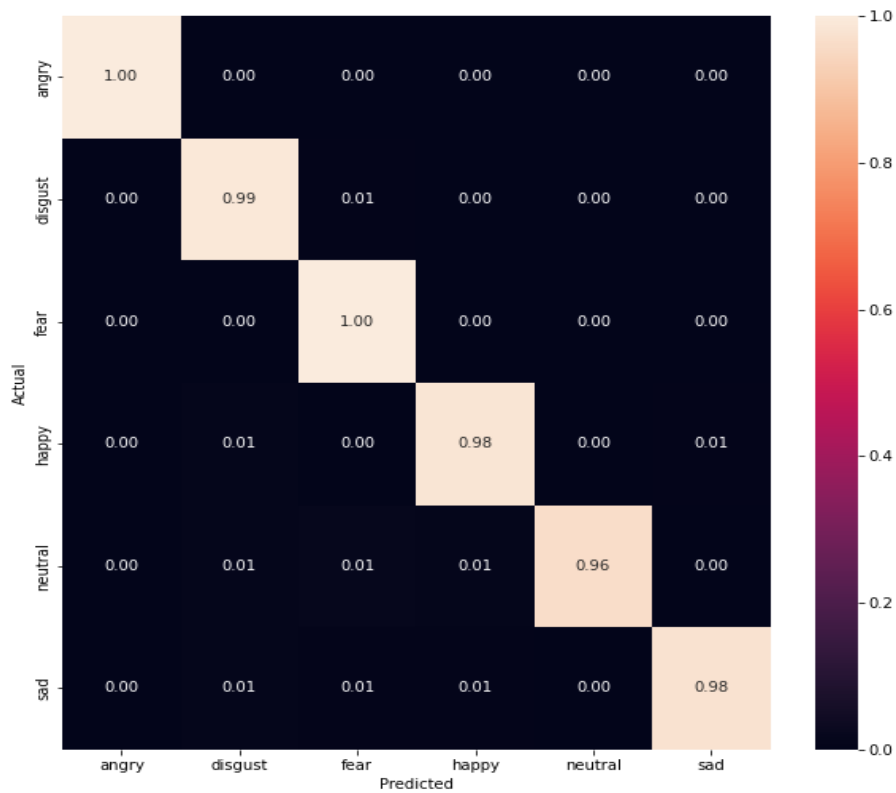


Figure 12. Heatmap of Model evaluated with RAVDESS Occluded dataset

5. Conclusion

In conclusion, neural networks have been shown to be a reliable tool for detecting emotions in images with occlusions. With further development and optimization, they can become even more accurate at recognizing emotional expressions and recognizing the context of an image. The potential applications of using neural networks to detect emotions in images with occlusions are vast, and their use can help to provide a better understanding of emotions and the complexities of facial expressions. Emotion detection using the facial expressions in images has become a critical tool in understanding human behavior. It offers insight into how people are feeling, making it easier to gauge how they might respond to certain situations or events. By using facial expressions and other cues to interpret emotions, organizations can better understand their customers and build stronger relationships with them. With the help of machine learning techniques, emotion detection has become a powerful and useful tool for businesses and professionals.

References

- [1] Ekman, P.; Friesen, W.V.; Ellsworth, P. *Emotion in the Human Face: Guidelines for Research and An Integration of Findings*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 11.
- [2] Dalgleish, T.; Power, M. *Handbook of Cognition and Emotion*; John Wiley & Sons: Hoboken, NJ, USA, 2000.
- [3] Hemamalini, Selvamani, and Visvam Devadoss Ambeth Kumar. (2022). Outlier Based Skimpy Regularization Fuzzy Clustering Algorithm for Diabetic Retinopathy Image Segmentation. *Symmetry*, 14(12), 2512
- [4] Mehendale, N. *Facial Emotion Recognition Using Convolutional Neural Networks (FERC)*; Springer: Berlin/Heidelberg, Germany, 2021; Volume 2, pp. 1–8. 18.
- [5] Akhand, M.; Roy, S.; Siddique, N.; Kamal, M.A.S.; Shimamura, T. Facial emotion recognition using transfer learning in the deep CNN. *Electronics* 2021, 10, 1036. [CrossRef]
- [6] Sathya Preiya, V., and V. D. Ambeth Kumar. (2023). Deep Learning-Based Classification and Feature Extraction for Predicting Pathogenesis of Foot Ulcers in Patients with Diabetes. *Diagnostics* 13(12), 1983.
- [7] Soleymani, M.; Asghari-Esfeden, S.; Fu, Y.; Pantic, M. Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Trans. Affect. Comput.* 2015, 7, 17–28. [CrossRef]
- [8] H. K. Ekenel and R. Stiefelhagen. Why is facial occlusion a challenging problem? In *Proc. International Conference on Advances in Biometrics (ICB)*, 2009.
- [9] Kumar, V.D.A., Sharmila, S., Kumar, A. et al. (2023). A novel solution for finding postpartum haemorrhage using fuzzy neural techniques. *Neural Comput & Applic.* 35(33), 23683–23696
- [10] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
- [11] Alper Ayvaci, Michalis Raptis, and Stefano Soatto. 2012. Sparse occlusion detection with optical flow. *International journal of computer vision* 97, 3 (2012), 322–338
- [12] Georgescu, M., Ionescu, R.T.: Teacher-student training and triplet loss for facial expression recognition under occlusion. *ArXiv*, abs/2008.01003(2020)
- [13] Cheng, Y., et al.: A deep structure for facial expression recognition under partial occlusion. In: *Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 211–214 (2014)
- [14] Tóser, Z., Jeni, L.A., Lőrincz, A., Cohn, J.F.: Deep learning for facial action unit detection under large head poses. In: Hua, G., Jégou, H. (eds.) *ECCV 2016*. LNCS, vol. 9915, pp. 359–371. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49409-8_29
- [15] Ashwin, T.; Jose, J.; Raghu, G.; Reddy, G.R.M. An e-learning system with multifacial emotion recognition using supervised machine learning. In *Proceedings of the 2015 IEEE Seventh International Conference On Technology for Education (T4E)*, Warangal, India, 10–13 December 2015; pp. 23–26.
- [16] Ghimire, D.; Lee, J. Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines. *Sensors* 2013, 13, 7714–7734.
- [17] Bost, R.; Popa, R.A.; Tu, S.; Goldwasser, S. Machine learning classification over encrypted data. *Cryptol. Eprint Arch.* 2014. Available online: <https://www.ndss-symposium.org/ndss2015/ndss-2015-programme/machine-learning-classification-overencrypted-data/>
- [18] Ambeth Kumar, V.D. Ramakrishnan, M. (2013). Temple and Maternity Ward Security using FPRS. *Journal of Electrical Engineering & Technology*, 8(3), 633-637.
- [19] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *ACL*, 2019.
- [20] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Proc. NeurIPS*, 2020.
- [21] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *Proc. ICML*, 2020a.

- [22] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In Proc. CVPR, 2020.
- [23] Jean-Bastien Grill, Florian Strub, Florent Altche, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap your own latent—a new approach to self-supervised learning. Proc. NeurIPS, 2020.
- [24] Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. In Proc. NeurIPS, 2020.
- [25] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollar, and Ross Girshick. Masked autoencoders are scalable vision learners. arXiv preprint arXiv:2111.06377, 2021.
- [26] Chitra, P., Indumathi, A., Rajasekaran, B., Babu, M.M. (2022). Emotion Recognition Using Speech Based Tess and Crema Algorithm. In: Suma, V., Baig, Z., Kolandapalayam Shanmugam, S., Lorenz, P. (eds) Inventive Systems and Control. Lecture Notes in Networks and Systems, vol 436. Springer, Singapore. https://doi.org/10.1007/978-981-19-1012-8_53.
- [27] Zielonka, Marta & Piastowski, Artur & Czyżewski, Andrzej & Nadachowski, Paweł & Operlejn, Maksymilian & Kaczor, Kamil. (2022). Recognition of Emotions in Speech Using Convolutional Neural Networks on Different Datasets. Electronics. 11. 3831. 10.3390/electronics11223831.
- [28] Zheng, W.; Yu, J.; Zou, Y. An experimental study of speech emotion recognition based on deep convolutional neural networks. In Proceedings of the 2015 International Conference on IEEE Affective Computing and Intelligent Interaction (ACII), Xi'an, China, 21–24 September 2015; pp. 827–831.
- [29] P. Sherubha, P Amudhavalli, SP Sasirekha, “Clone attack detection using random forest and multi-objective cuckoo search classification”, International Conference on Communication and Signal Processing (ICCSP), pp. 0450-0454, 2019.
- [30] S. Dinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji, “Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms”, Hindawi Journal of Healthcare Engineering, Volume 2022.
- [31] Sherubha, “Graph Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks”, Sādhanā(Springer), 45:212, <https://doi.org/10.1007/s12046-020-01451-w>
- [32] Piyush K. Pareek, Pixel Level Image Fusion in Moving objection Detection and Tracking with Machine Learning “,Fusion: Practice and Applications, Volume 2 , Issue 1 , PP: 42-60, 2020
- [33] Shivam Grover, Kshitij Sidana, Vanita Jain, “Egocentric Performance Capture: A Review”, Fusion: Practice and Applications, Volume 2, Issue 2 , PP: 64-73, 2020.
- [34] Abdel Nasser H. Zaied, Mahmoud Ismail and Salwa El- Sayed, A Survey on Meta-heuristic Algorithms for Global Optimization Problems, Journal of Intelligent Systems and Internet of Things, Volume 1 , Issue 1 , PP: 48-60, 2020
- [35] Mahmoud H. Alnamoly, Ahmed M. Alzohairy, Ibrahim M. El-Henawy, “A survey on gel images analysis software tools, Journal of Intelligent Systems and Internet of Things, Volume 1 , Issue 1 , PP: 40-47, 2021.