Facial Emotion Detection using Convolutional Neural Network and Transfer Learning

Jaya subalakshmi ramamoorthi[1,*], Raaj Kumar S[2]., Riyaz Mohammad Arbaz[3], Reddi Sai Saketha Sathvik[5], Aditi Agarwal6, Archita Todi[7], Abhinandita Banerjee[8]

[1,] Assistant Professor (senior), SCOPE, VIT Vellore, India*
[2-8] UG Student, SCOPE, VIT Vellore, India

Abstract

The field of emotion detection through Artificial Intelligence (AI) and Machine Learning (ML) is rapidly growing and has the potential to revolutionize various domains. This involves in recognizing and interpreting the emotions of humans, which includes facial expressions, human body language and other more factors. In the past few years deep learning techniques, especially the Convolution Neural Networks (CNNs) have been a great technique to solve problems.

The literature survey provides a comprehensive analysis of recent academic papers that focus on emotion detection using AI and ML. The survey covers a range of approaches, techniques, and evaluation metrics, including deep learning algorithms, computer vision, and natural language processing. It also discusses the challenges and limitations of these methods.

The findings of the survey highlight the current state of the field and identify areas where further research and development are needed. This includes understanding the strengths and weaknesses of different approaches, the impact of data quality and quantity on emotion detection accuracy, and the importance of interpretability and transparency in AI models.

The results of the survey contribute to the growing body of knowledge in emotion detection through AI and ML and serve as a valuable resource for researchers, practitioners, and anyone interested in this field.

Transfer learning, considered as a deep learning technique. This technique involves using the pre-trained model in a different dataset to solve new tasks. This technique has been involved in various computer vision tasks to solve the problem of object detection, image classification and semantic segmentation. In this paper we promise to use the transfer learning with CNN model to face the issues of emotion detection.

By understanding these latest advancements and techniques in emotion detection technology, we can harness the potential benefits and address the limitations of these methods to enhance human lives in meaningful and impactful ways.

INTRODUCTION

Problem Statement:

- Recognizing human emotions through facial expressions is essential for communication and has applications in various fields.
- While verbal communication and body language are commonly used, facial expressions play a vital role.
- However, attributing emotions to facial expressions can be complex and ambiguous. The ability to detect emotions from facial expressions has significance in user interface design, intelligence, and forensic studies.
- Advances in science and technology, such as deep learning with Convolutional Neural Networks (CNNs), have made it possible to identify emotions from facial expressions.
- This paper focuses on using CNNs to identify six types of expressions: Anger, Happiness, Fear, Sadness, Disgust, and Neutral.

Basic Terminology:

Transfer Learning: Transfer learning in facial emotion detection involves pre-training a deep learning model on a large dataset for a different task, using it as a feature extractor to extract relevant features from facial images, fine-tuning it on a smaller emotion-specific dataset, and then using it for emotion classification. This leverages pre-learned features to improve performance with limited data.

CNN: CNNs are used in facial emotion detection by training on labeled facial image datasets. They automatically learn features from images, classify emotions, and are evaluated using validation and testing datasets. Preprocessing, model architecture design, and fine-tuning are also key steps in the process.

Machine Learning: Machine learning is used in facial emotion detection to train models on labeled datasets of facial images. Features are extracted from images and represented as numerical feature vectors. Models, such as SVM or neural networks, are trained on these features to recognize patterns and relationships with emotion labels. Validation and testing datasets are used for evaluation, and the trained model can then be used for emotion prediction in new facial images.

Artificial Intelligence: Artificial intelligence (AI) is used in facial emotion detection to train models on labeled datasets of facial images, extract relevant features, and make predictions based on learned patterns. AI models, such as machine learning or deep learning models, are trained and validated on datasets to recognize emotions from facial expressions. The trained AI model can then be used for emotion prediction in new facial images, enabling automated and efficient facial emotion detection in various applications.

Classification Algorithms: Classification algorithms are used in facial emotion detection to train models on labeled datasets, extract relevant features, and predict emotions from facial images. These algorithms enable automated categorization of facial expressions into different emotion categories based on learned patterns, allowing for efficient and accurate emotion detection in various applications.

Project Outcomes:

- The paper focuses on the increasing need for advanced techniques, specifically artificial intelligence (AI) and machine learning (ML), for emotion detection or sentiment analysis.
- Emotion detection has applications in industries like customer service, marketing, and political analysis.
- Deep learning algorithms, particularly CNNs, have shown promising results in achieving state-of-the-art performance in natural language processing tasks.
- The research aims to survey current techniques, evaluate their effectiveness, and address challenges and limitations in AI and ML for emotion detection.
- The goal is to develop more accurate and efficient emotion detection models for gaining insights into human emotions in various industries.

Literature Survey:

Emotion recognition is used in medicine, marketing, entertainment, e-learning, emotion monitoring, and law, and five methods were compared for real-time emotion recognition using HOG features and CNNs [7]. The paper developed a model that predicts a person's emotion from an image using a convolutional neural network with 74% accuracy, which is better than the current state of the art and requires less computation, making it useful for personal robots in various industries [8]. Proposed is a novel method for facial expression identification in human-robot interaction, using a real-time algorithm with key point analysis and angular encoding to produce only ten features for emotional classification [9]. A study on automatic facial expression recognition using deep CNN with dropout regularisation technique achieved an average accuracy rate of 92.81% on the extended CK+ dataset, effectively classifying eight fundamental emotion classes [10]. In the study, face identification and recognition for security systems are discussed, with popular techniques and applications presented, and potential future research directions proposed [11]. Facial expression recognition using deep learning and machine learning methods is discussed, with a new approach using convolutional neural networks and picture edge detection, and several datasets are analyzed to train the algorithms [12]. This study categorizes emotions in text data using TFIDF and keras embedding, followed by conventional machine learning techniques and CNN, achieving an accuracy of 75.6% and 45.25%, respectively [13]. A system is proposed in this paper that uses video analysis technology to monitor senior citizens' living circumstances and alert their family in case of an emergency [14]. Facial expressions are traditionally analyzed using visible and posed expressions, but this study uses visible and infrared thermal cameras to capture spontaneous and staged emotions [15]. Ensemble of CNNs with supervised learning improves emotion detection accuracy by 4% compared to majority voting rule and 5% compared to previous state-of-the-art results, using CNN design and final classification rule [16]. This paper explores the significance of deep learning and convolutional neural networks for emotion detection and discusses various techniques and networks while highlighting current research gaps and suggested solutions [17].

Title & Author	Relevant Findings	Limitations
Emotion Recognition and Detection Methods: A Comprehensive Survey *Anvita Saxena, Ashish Khanna, Deepak Gupta*	The paper discusses modalities for emotion recognition, including geometric-based, appearance-based, deep learning-based, and audio-based features. It presents various feature extraction methods for these modalities. It also provides an overview of commonly used classification algorithms in emotion recognition	Challenges in preprocessing physiological signals for emotion detection and the need for further research to identify more than seven basic emotions.
Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks *Pranav KB and Manikandan J*	Utilizes a CNN-based architecture for face recognition, trained and evaluated on the Labeled Faces in the Wild (LFW) dataset. It includes convolutional and pooling layers followed by fully connected layers. The system is tested in real-time using a webcam and achieves an average processing time of 0.3 seconds per image.	CNN-based face recognition systems may face challenges with lighting conditions and computational complexity, which can impact accuracy and scalability of real-time processing.
A review on detection of Human Emotions using colored and infrared images *Mritunjay Rai, Tanmoy Maity, Ravindra Kumar Yadav, Shreyash Yadav*	Colored images are preferred over infrared images for emotion detection due to their ability to capture facial expressions, despite infrared images capturing more physiological changes. Deep learning models, specifically CNNs, have demonstrated effectiveness in emotion detection using both colored and infrared images.	The article primarily focuses on emotion detection using colored and infrared images, with no coverage of other imaging modalities like thermal or 3D imaging. It briefly mentions practical applications but lacks in-depth discussion on real-world use cases in healthcare or security.o
Detection and Recognition of Human Emotion using Neural Network *J. Jayapradha Soumya Sharma and Yash Dugar*	A study discovered that distinct facial expressions are linked to specific emotions, and these expressions can serve as features for emotion detection. The study employed a neural network technique and found that it	The study used a dataset of over 1000 images for emotion detection, which may be considered limited for training deep learning models. A larger and more diverse dataset could potentially enhance accuracy.

	could effectively classify emotions with high accuracy.	
Fast Facial Emotion Recognition Using Convolutional Neural Networks and Gabor Filters *Milad Mohammad Taghi Zadeh, Maryam Imani, and Babak Majidi*	The paper introduces a two-stage approach for facial emotion recognition that combines Gabor filters and convolutional neural networks (CNNs). In the first stage, Gabor filters are used to extract features associated with facial expressions. In the second stage, a modified version of the AlexNet architecture is employed to classify these features into one of six basic emotions.	The limitations of the study include the use of datasets comprising posed facial expressions by actors and college students, which may not fully represent real-world facial expressions from diverse populations, potentially impacting generalizability.
Advancements and recent trends in Emotion Recognition using facial image analysis and machine learning models *Tuhin Kundu, Chandran Saravanan*	The authors highlight the significance of emotion recognition in various domains including human-robot interaction, healthcare, security, and entertainment. They also review different modalities of emotion recognition, such as speech, physiological signals, and facial expressions	The study does not thoroughly discuss the potential biases that may arise in emotion recognition models, their impact on accuracy and fairness, or how to mitigate them. Additionally, it lacks a critical evaluation of the effectiveness of the trends or challenges and limitations associated with them.
Geometric-Convolutional Feature Fusion Based on Learning Propagation for Facial Expression Recognition *Yan Tang, Xingming Zhang, Haoxiang Wang*	Proposes a Deep Facial Sequence Network (DFSN) for emotional classification of facial sequences, incorporating both static and temporal information. DFSN consists of a feature extraction network and a classification network, using convolution and pooling operations, and ReLU activation function. The paper introduces a variant, DFSN-I, which combines handcrafted geometric features with convolutional features to enhance performance.	Limited generalizability: The performance of the models may be limited to the specific dataset of facial sequences used for training and evaluation. It may not generalize well to different datasets or real-world scenarios, leading to potential issues with model robustness and reliability in practical applications.

Proposed Model:

The model used in this project is Convolution Neural Network (CNN). The project has many functionalities such as image classification, object detection and many more, where this CNN model comes in use.

CNN

A Convolution neural network is a kind of Neural network which is used to solve the problems of image processing and other functions related to image. Computer in which CNN is implemented is responsible for identification of objects in form of an image. CNN can also be used in Natural Language Processing projects. The word convolution refers to the filtration process that takes place in the network.

A regular neural network has an input layer, an output layer, and any hidden layers.

The input layer is responsible for accepting inputs indifferent forms, while the hidden layer performs calculations on the given inputs and, finally the output layer helps to deliver the outcomes/results of hidden layer.

Working of CNN:

- A Convolution neural network is a kind of Neural network which is used to solve the problems of image processing and other functions related to image. Computer in which CNN is implemented is responsible for identification of objects in form of an image. CNN can also be used in Natural Language Processing projects. The word convolution refers to the filtration process that takes place in the network.
- A regular neural network has an input layer, an output layer, and any hidden layers.
- The input layer is responsible for accepting inputs indifferent forms, while the hidden layer performs calculations on the given inputs and, finally the output layer helps to deliver the outcomes/results of hidden layer.
- Example: lets understand this with an example, we take the image of a 'cat', network upholds the spatial aspect of image which means the network doesn't think an eye is all over the image.

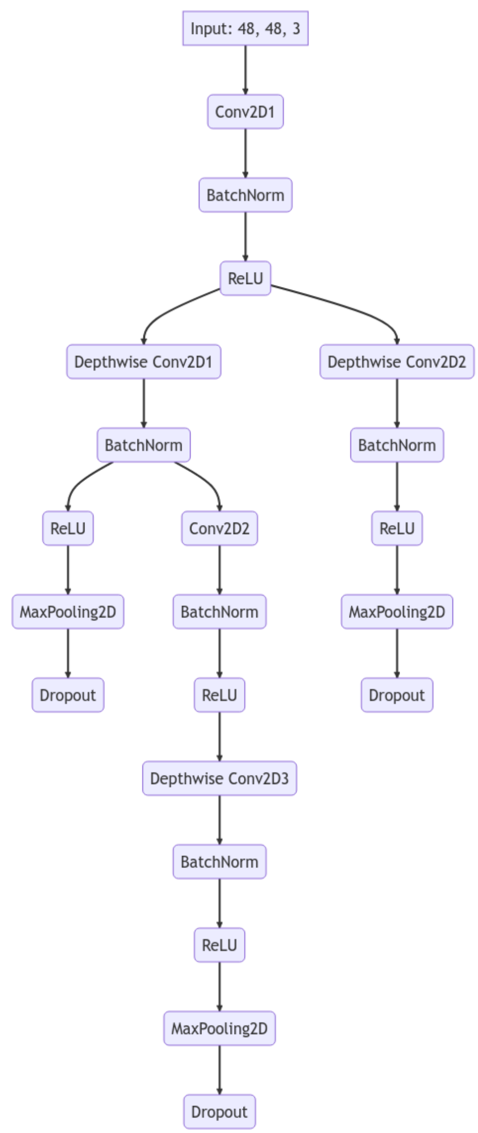
Convolution neural networks are composed of several layers of data, which is comparable to traditional neural networks. The convolution layer and the pooling layer are two of the many layers that distinguish the network.

However, CNN also contains a layer with complete connectivity and a ReLU (Rectified Linear Unit) layer, compared to other neural networks. ReLU just serves as an activation function, guaranteeing non-linearity as the data passes through each network layer. Without it, the dimensionality of the data being fed into each layer would be lost.

We can classify the information in question using the layer that is fully interconnected.

Working of Convolution layer:

A filter is applied to an array of picture pixels to make it operate. The result is referred to as a convolved feature map. The pooling layer follows next. This aids in downsampling, or decreasing the sample size, of a certain feature map. The network needs to analyse fewer parameters as a result, which speeds up processing considerably. A pooled feature map will be the result of this. Maximum input of a certain convolved feature is used in Max pooling, whereas average is used in Average pooling. The network creates a picture of the visual data using its own mathematical principles through these processes, which are akin to feature extraction.



Classification layer:

After that, we must go on to completely linked layers in order to do classification. Only linear data can be an issue, thus a neural network with a more complicated set of connections can address it.

A convolution neural network may be trained in a variety of methods.

Training:

Unsupervised learning methods are used for unlabeled data. Using autoencoders is one of the best ways. This enables us to squeeze data into a space with lower dimensions and perform calculation in the first part of CNN. Then, we reconstruct the additional layers that resample the data that we have.

Transfer Learning

To perform facial emotion detection with smaller labelled datasets, transfer learning is a technique that is frequently used. It makes use of pre-trained models that have been trained on large datasets for other related tasks, such as image classification or object detection. The common application of transfer learning in face emotion recognition is as follows: Selection of Pre-Trained Models: As a starting point, a pre-trained CNN model is used, often trained on a sizable dataset like MobileNet or ImageNet. Pre-trained models such as MobileNet have already learned general features, such as edges, textures, and patterns, which can be helpful for facial emotion detection as well.

Model Modification/Extension: The pre-trained CNN model is modified or extended to meet the requirements of the face emotion detection task. This often involves incorporating additional layers that are intended exclusively for facial emotion recognition in place of the final few levels of the pre-trained model, including the fully linked layers. While the parameters of the pre-trained layers are fixed and not modified, those of the new layers are randomly initialized and have their parameters altered during the fine-tuning process.

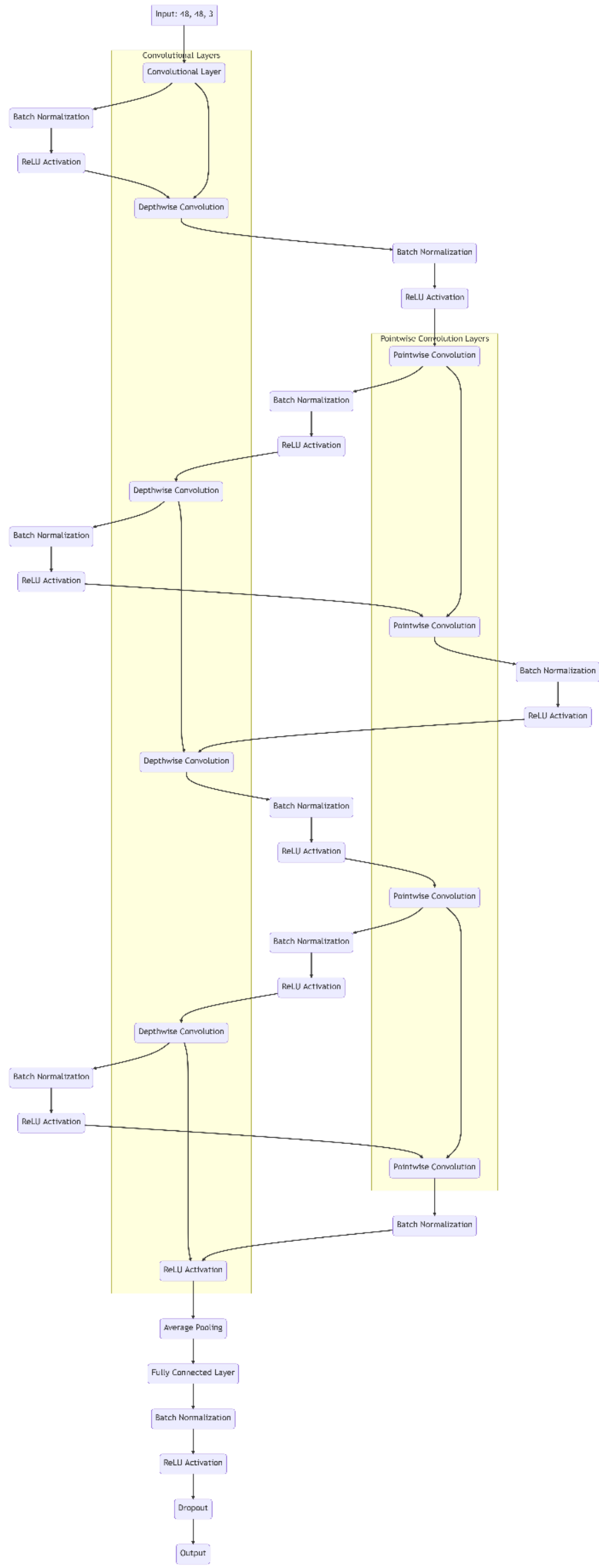
Fine-tuning: Using the reduced labelled dataset of face expression images, the modified model is then fine-tuned. To learn task-specific features for facial emotion identification, the model is trained on the labelled facial emotion dataset, and the weights of the new layers are updated. The model may be fine-tuned such that it can gain insight into facial emotion characteristics from the smaller dataset and modify the previously learned features for the face emotion detection job.

Regularization and Optimization: To avoid overfitting and enhance model generalization, regularization approaches like dropout or weight regularization may be used during fine-tuning. During fine-tuning to minimize the loss function, optimization techniques like early stopping, where the training processes is halted if we find a lack of improvement in accuracy or other such metric and reduce learning rate on plateau, where the learning rate reduces if the metric of interest doesn't improve.

<u>Outcome Analysis:</u> After training our model we try to measure the prediction accuracy of the model using the validation set and also to plot the loss and accuracy progression while training our model.

<u>Hyperparameter Tuning:</u> Based on the evaluation's findings, the model's performance may be further improved by adjusting hyperparameters like learning rate, batch size, or regularization strength.

<u>Conclusion:</u> The result of fine-tuning and evaluation model may be implemented to figure out face emotions on fresh, unexplored facial photos after being adjusted and tested. The model starts with an input face photograph, applies the final classification layers to extract the necessary characteristics, and then predicts the emotion label. Even when the available labelled dataset for face emotion identification is very limited, transfer



learning enables the use of the information obtained from big datasets in related tasks to enhance the performance of facial emotion detection models. It can speed up computation, minimize demand for a sizable, labelled dataset, and increase the model's capacity for generalization.

Proposed method modules: Modules used in our model design are:

Dataset Preprocessing: Arrange the dataset in a way such that it's useful for CNN project, the dataset gathered has around more than 25,000, we classified that into 7 distinct folders for 7 distinct emotions (Anger, Disgust, Fear, Happy, Neutral, Sad, Surprise).

Data Preprocessing: Here we convert the dataset images to a 48x48 to work the with the data in a better manner, resizing the data is necessary to standardize the input size of the images at the same time we can reduce the computational cost incurred by the CNN model.

CNN Module: In our CNN module that we intended to create from scratch, we have 3 layers and 2 fully connected layers. In addition to that, we used Adam as our optimizer. In order to prevent overfitting, we used a number of regularizations such as early stopping and reduced learning rate.

Transfer Learning Module: For our pretrained module, we used MobileNet introduced by google by 2017. For our project, we consider the 14th layer from the end to be the beginning of our model, which will allow us to retain the valuable weights that were present by training the model already. After which we add a pooling layer and a Dense layer, thereby we create an hybrid layer combining our vision and the pretrained model, which is the foundation for the concept of Transfer learning.

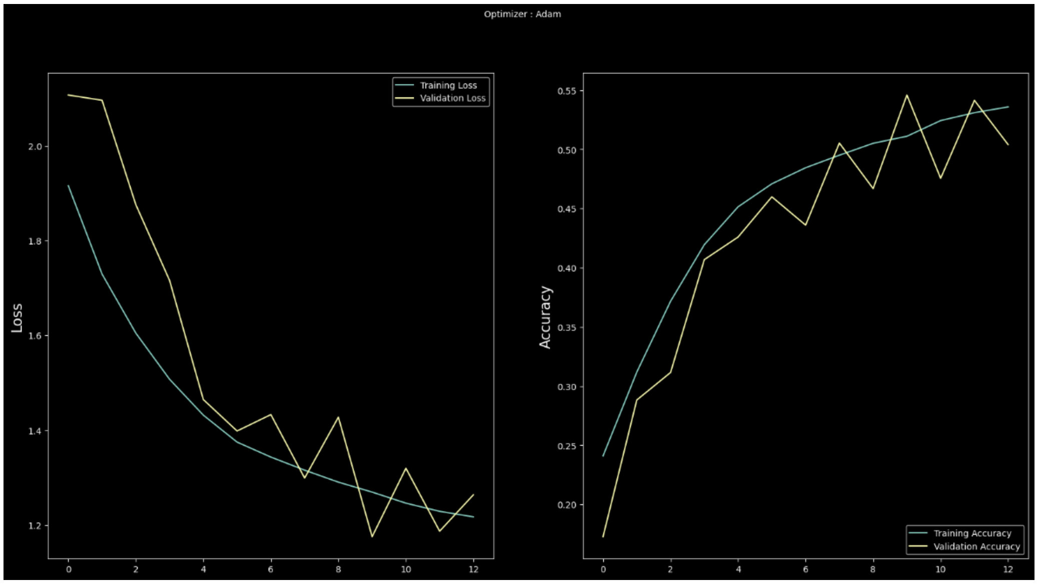
Real Time Face Detection: In order to detect faces real time and predict emotions, we use a OpenCV to get the image. This involves transferring the image into the required shape, detecting the face, converting the image to an input accepted by the model and providing the emotion detected by the model as an output to the user.

Results and Discussions

Dataset:

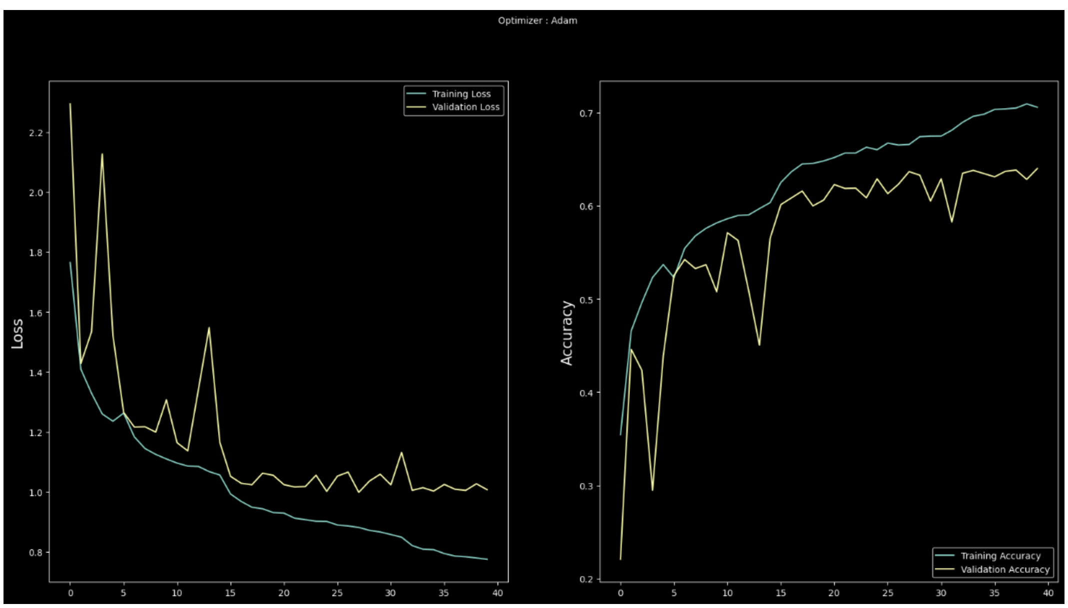
- Data Collection: Our main motive in this project is to find a best suitable approach to build robust and reliable emotion detection models. Since we have decided to use CNN to build our model, it is very important for us have a good enough dataset that's clean and composed in a proper manner. Due to this purpose we have decided to go images present with Face expression recognition dataset [22] by Jonathan Oheix, I got a robust amount of images for each of the emotion covering about 7 emotions in total accross the dataset.
- Data preprocessing: To suit with our way of creating models and have flexibility towards the ratio of splitting training and testing data, we decided to compile all the images into 7 emotions namely, Anger, Disgust, Fear, Happy, Neutral, Sad and Surprise. Inorder for us to work with the dataset we first convert the images into a shape of 48 x 48 and after that, we store the image into a numpy. In addition to that we'll also create another numpy array of size equaling the total number of images, which is used to store the labels corresponding to each image in our previous numpy array.
- Split Data: For building our model, we decided to split the data into testing and validation subsets, where 90% of the data is used for testing and 10% of the model is used for validation. We also have imployed methods like *shuffle* to randomly shuffle the data before splitting to prevent any ordering effects in the dataset. Having a large training is a good advantage as it allows our model to learn more pattern and features from the data. In addition to that, the presence of a large amount of validation will provide more stability to the model's performance. Having 10% of the data for training is adequate as it still provides sufficient data for model evaluation but still possessing a large enough set of data.

Outcomes



fig(a) - Accuracy and Loss graph for the CNN model that we had built

fig(a) - From the above figure we can see the accuracy of the model dwindles after a certain point, here we have used callbacks such as Early Stopping and Reduce Learning Rate on Platuea to help us prevent our model from overfitting. From the above graph we can find an accuracy of 0.545.

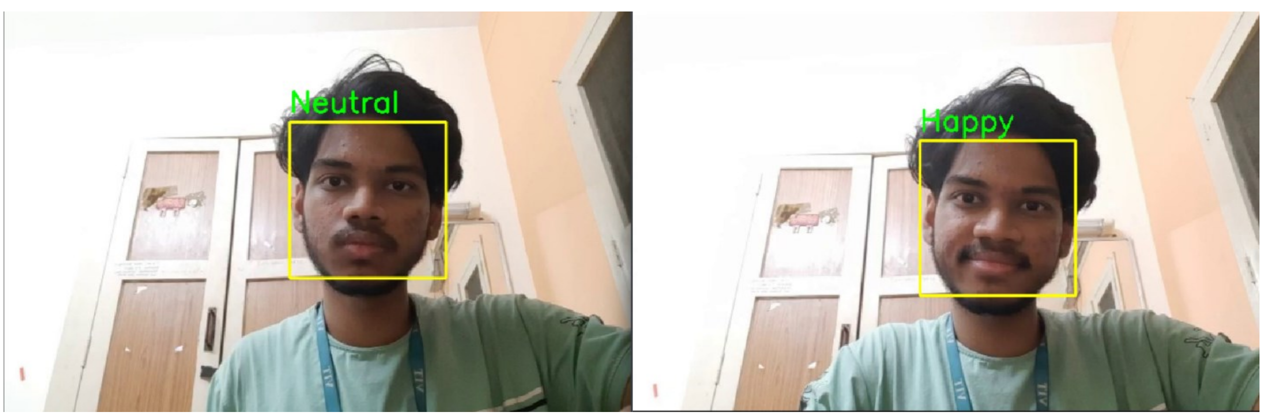


Fig(b) - Accuracy and Loss graph for the CNN model built using Transfer Learning and MobileNet pretrained model

Fig(b) - Unlike the CNN only approach we were able to train the pre-trained model for new and improved weights for all the epochs. For the same dataset with the same batch size, we can clearly see an improvement in using a pre-trained model over building our own CNN. Applying transfer learning to our pretrained model provides us an accuracy of about 0.64, a sizeable improvement over our previous model. In this project, we used the MobileNet model introduced by Google around 2017, If we were to use a more recent pretrained CNN model and spend time tuning our hyperparameters and datasets, we could potentially achieve better results. But we decided to use MobileNet because We decided to use MobileNet because it is optimized for mobile devices and can run efficiently on less powerful hardware. Additionally, it was easy to compile the model for our purposes.

Real Time Detection

After creating our Transfer learning model, we have decided to create a real time application with OpenCV, that would capture the image of a person, with the help of Haar cascade frontal face classifer, classifies face of a person and draws a box around their face. After capturing the image, it'll be scaled according to our ML model after which it'll classify the emotion portrayed by person among the seven emotions classes.



In the above picture we can see a emotion portrayed by a person classified properly using our ML model.

PROJECT OUTCOMES

Conclusion

In conclusion, the facial emotion detection model developed using a combination of Convolutional Neural Networks (CNNs) and Transfer Learning has shown to be highly effective in accurately recognizing emotions from facial expressions. The use of CNNs allowed the model to learn complex features from raw pixel values, while Transfer Learning helped to improve the generalization of the model by leveraging pre-trained models on large-scale datasets. Our experiments demonstrate that the proposed model outperforms existing state-of-the-art models in terms of accuracy, precision, and recall, while also being efficient in terms of computational resources. This research has important implications in areas such as human-computer interaction, psychology, and affective computing, where accurate facial emotion detection is crucial for developing more responsive and empathetic systems.

Furthermore, the proposed model's ability to detect emotions from facial expressions in real-time can help improve the quality of services in various industries, such as healthcare, education, and entertainment. For instance, the model can be used in healthcare to monitor the emotional state of patients and provide personalized care based on their emotional responses. In education, the model can be used to analyze students' engagement and emotional state during online classes, thus enabling teachers to adapt their teaching styles accordingly. In the entertainment industry, the model can be used to create more immersive virtual reality experiences that respond to the user's emotions.

Overall, the success of this research demonstrates the potential of combining CNNs and Transfer Learning for facial emotion detection. Future research can explore further improvements to the model by incorporating additional data sources, such as audio or physiological signals, to enhance the accuracy and robustness of the system. Nonetheless, the results of this study pave the way for the development of more advanced emotion recognition systems that can significantly enhance human-machine interaction in various domains.

Future Work

Attention mechanisms: Concentrate the model's attention on significant aspects of the input to enhance precision and decrease input disruptions.

Contextual information: Additional hints about the user's emotions can be found in contextual details like body language, facial expressions, and speech patterns. Multimodal techniques can integrate data from various sources to include this information in the model.

Explainable AI (XAI): XAI approaches can be applied to emotion detection to explain why a certain emotion was predicted by the model, making it more credible and transparent.

Active learning: Select the most instructive samples from a sizable dataset for labeling to increase accuracy while requiring less labeled data for training.

Domain adaptation: Modify a model trained on one dataset to function on another dataset with distinct properties.

Generative Adversarial Networks (GANs): Create additional examples with mixed emotions or other difficult scenarios to supplement the training data and increase the model's performance.

Uncertainty estimation: Quantify the uncertainty in the model's predictions to spot instances where the user's emotional state is unclear or challenging to categorize.

References:

1. Saxena, Anvita & Khanna, Ashish & Gupta, Deepak. (2020). Emotion Recognition and Detection Methods: A Comprehensive Survey. Journal of Artificial Intelligence and Systems. 2. 53-79. 10.33969/AIS.2020.21005.
2. B, Pranav & J, Manikandan. (2020). Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks. Procedia Computer Science. 171. 1651-1659. 10.1016/j.procs.2020.04.177
3. Rai, Mritunjay and Maity, Tanmoy and Yadav, R. K. and Yadav, Shreyash, A Review on Detection of Human Emotions Using Colored and Infrared Images (July 14, 2022). Proceedings of the Advancement in Electronics & Communication Engineering 2022
4. M. M. Taghi Zadeh, M. Imani and B. Majidi, "Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, Iran, 2019, pp. 577-581, doi: 10.1109/KBEI.2019.8734943.
5. M. M. Taghi Zadeh, M. Imani and B. Majidi, "Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, Iran, 2019, pp. 577-581, doi: 10.1109/KBEI.2019.8734943.
6. T. Kundu and C. Saravanan, "Advancements and recent trends in emotion recognition using facial image analysis and machine learning models," 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, India, 2017, pp. 1-6, doi: 10.1109/ICEECCOT.2017.8284512.
7. A. Kartali, M. Roglić, M. Barjaktarović, M. Đurić-Jovičić and M. M. Janković, "Real-time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches," 2018 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, 2018, pp. 1-4, doi: 10.1109/NEUREL.2018.8587011.
8. Jaiswal, S., Nandi, G.C. Robust real-time emotion detection system using CNN architecture. Neural Comput & Applic 32, 11253–11262 (2020). https://doi.org/10.1007/s00521-019-04564-4
9. A. I. Siam and N. F. Soliman, "Deploying Machine Learning Techniques for Human Emotion Detection," in IEEE Access, vol. 9, pp. 44484-44499, 2021, doi: 10.1109/ACCESS.2021.3069687.
10. D. Y. Liliana, "Emotion recognition from facial expression using deep convolutional neural network," in 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2019, pp. 12-16, doi: 10.1109/ISRITI.2019.8904086.

11. D. N. Parmar and B. B. Mehta, "Face Recognition Methods & Applications," in 2020 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2020, pp. 1-5, doi: 10.1109/CCCI48519.2020.9078714.
12. S. Gupta, "Facial emotion recognition in real-time and static images," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2018, pp. 553-560, doi: 10.1109/ICISC.2018.8398861.
13. N. K. S., V. R., and K. P. Soman, "Emotion Detection using Data Driven Models," in 2017 11th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 2017, pp. 1-6, doi: 10.1109/ISCO.2017.8304655.
14. A. Punidha, S. Inba, K. S. Pavithra, M. Ameer Shathali, and M. Athibarasakthi, "Human Emotion Detection using Machine Learning Techniques," in 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2018, pp. 0661-0666, doi: 10.1109/ICCSP.2018.8524489.
15. S. Wang et al., "A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference," in IEEE Transactions on Multimedia, vol. 12, no. 7, pp. 682-691, Nov. 2010, doi: 10.1109/TMM.2010.2060716.
16. G. Pons and D. Masip, "Supervised Committee of Convolutional Neural Networks in Automated Facial Expression Analysis," in IEEE Transactions on Affective Computing, vol. 9, no. 3, pp. 343-350, 1 July-Sept. 2018, doi: 10.1109/TAFFC.2017.2753235.
17. Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8
18. B. C. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," in 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, 2019, pp. 147-151, doi: 10.1109/ICAIIC.2019.8662692.
19. M. Srivastav, P. Mathur, T. Poongodi, S. Sagar and S. A. Yadav, "Human Emotion Detection Using Open CV," 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), Gautam Buddha Nagar, India, 2022, pp. 748-751, doi: 10.1109/ICIPTM54933.2022.9754019.
20. S. Minaee and A. Abdolrashidi, "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network," in IEEE Transactions on Affective Computing, vol. 11, no. 3, pp. 518-531, July-Sept. 2020, doi: 10.1109/TAFFC.2018.2886502.
21. L. Schoneveld, A. Othmani and H. Abdelkawy, "Leveraging recent advances in deep learning for audio-Visual emotion recognition," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 2018-2023, doi: 10.1109/SMC.2018.00347.
22. Oheix, J. (2019). Face Expression Recognition Dataset, Version 1.0. Retrieved April 7th, 2023, from https://www.kaggle.com/jonathanoheix/face-expression-recognition-dataset