Comparative Analysis and Implementation of Facial Expression Recognition Algorithms

Devanshu Gupta , Karus Manisha , Anushka Kalwale

Prof Ramesh K Bhukya

IIIT Allahabad

ABSTRACT—The field of automatic emotion recognition based on facial expressions is an intriguing area of research that has found applications in various domains such as safety, health, and human-machine interfaces. Researchers in this field aim to develop techniques that can effectively interpret and code facial expressions, extracting relevant features to enhance computer-based prediction. Deep learning has achieved notable success, prompting exploration of different architectures to improve performance. This paper aims to conduct a comprehensive study on recent advancements in automatic facial emotion recognition (FER) through deep learning. We highlight the contributions made by these studies, including the utilized architectures and databases, and present a comparative analysis of the proposed methods and their results. The objective of this paper is to provide valuable insights to researchers, offering a review of recent works and guiding potential improvements in this field.

I. **Introduction**:

FER stands for Facial Expression Recognition, which refers to the ability of a computer system to recognize human emotions from facial expressions.FER is a challenging research area[1] that has gained significant attention over the past few decades due to its potential applications in various fields, such as psychology, healthcare, security, entertainment, and human-computer interaction.FER has the potential to improve the accuracy of emotion recognition in various applications, such as mental health diagnosis, human-computer interaction, and security[2]. For example, in mental health diagnosis, FER can help diagnose depression and anxiety by analyzing facial expressions. In human-computer interaction, FER can enable more natural and intuitive communication between humans and computers.FER is a challenging problem that requires the development of advanced algorithms and models,[3] which can contribute to the development of the field of artificial intelligence and machine learning.

Based on research findings, the spoken message can be broken down into three components: the verbal component (7% contribution), the vocal part (38% contribution), and the speaker's facial expressions (55% contribution). This indicates that facial expressions play a crucial role in human communication, surpassing the importance of other factors [4]. Analyzing these facial expressions is instrumental in understanding an individual's emotions and is often linked to physiological changes. For instance, anger is associated with a higher heart rate compared to happiness, while sadness can lead to decreased skin resistance, indicating heightened stress levels [5]

A. Motivation

In recent times, there has been a growing interest in evaluating the effectiveness of traditional machine learning and deep learning (DL) techniques used in facial emotion recognition (FER). Several surveys have been conducted to investigate various aspects of FER, including specific algorithm types and application areas. The primary objective of this research is to gain insights into the underlying design principles of the most prominent FER approaches. By comprehensively examining each method, including their architectures, advantages, and limitations, we aim to understand how and why these strategies perform in different scenarios. This deeper understanding will facilitate the handling of new challenges and the development of improved algorithms. By leveraging the strengths and addressing the weaknesses identified in these approaches, novel FER designs can be created to enhance performance while mitigating potential drawbacks.

B. Challenges in the FER

Facial Expression recognition is a technique that humans do on a regular basis and effortlessly , but it is not yet readily accomplished by computers, despite new approaches demonstrating higher accuracies.In some cases (frontal face, controlled environment, and high-resolution photos), the accuracy is more than 95%. Many works in the literature do not use a consistent evaluation methodology (e.g., no subject overlap in training and testing) and thus present a misleading high-accuracy, but do not represent the majority of real-world face expression recognition problems. Low accuracy, on the other hand, has been documented in databases with uncontrolled environments and in cross-database comparisons. To overcome these constraints, various research projects have attempted to make computers achieve the same accuracy as humans.This problem continues to be a challenge for computers because it is difficult to separate the feature space of expressions, i.e. facial features from one subject in two different expressions may be very close in the feature space, whereas facial features from two subjects with the same expression may be very far apart. Furthermore, some expressions, for example, "sad" and "fear" are, in some cases, very similar.

Fig. 1. Three different subjects with the happy expression. As it can be seen, the images vary a lot from each other not only in the way that the subjects show their expression but also in light, brightness, position and background.

Fig. 1 shows three subjects with a happy expression. As observed in the picture, the photographs significantly differ from one another not just in how the participants display their emotion but also in lightness, posture, backdrop, and lighting. The uncontrolled training-testing situations (training photos might be substantially different from testing images in terms of ambient circumstances and subject ethnicity) are another issue with facial emotion recognition that is illustrated by this figure.

In recent years, machine learning algorithms based on convolutional neural networks have seen considerable success in the field of computer vision. It is widely employed in fields such as visual object recognition , Natural Language Processing , autonomous cars , and so on. It's also a potential strategy for FER research. Convolutional neural networks, unlike previous approaches, can accomplish tasks end-to-end by training both feature extraction and classification phases concurrently.

FER may be separated into four basic phases, as in a standard machine learning-based method: 1) a face detection step that automatically detects the face region in the input photos;
2) intensity normalization; 3) face data extraction and representation, which automatically collects and represents information about observed facial expressions; and 4) expression identification, which classifies the extracted features into suitable expressions.

C. Contributions

The paper provides a comparative analysis of different facial expression recognition algorithms, including traditional hand-crafted feature-based approaches and deep learning-based approaches. This analysis helps in understanding the strengths and weaknesses of each approach and provides insights into the state-of-the-art in facial expression recognition.The proposed model achieves significant improvements over previous models on multiple datasets, including FER-2013, CK+ and AFFECTNET. This performance improvement demonstrates the effectiveness of the proposed attentional convolutional network.

II. LITERATURE REVIEW

In the past few decades, there has been significant progress in facial expression recognition, thanks to the development of various approaches. Notably, the emergence of deep learning methods, such as Convolutional Neural Networks (CNNs), has played a crucial role in this advancement [6,7,8,9,10]. Deep learning approaches have become more feasible due to the availability of large datasets for training and advancements in GPU technology. The abundance of data enables the training of networks with deep architectures, while GPU technology provides cost-effective and high-performance numerical computations required for the training process. For further insights into facial expression recognition research, surveys can be found in references [11,12].

Addressing facial expression recognition in uncontrolled environments, where factors like non-frontal faces, partially overlapped images, and spontaneous expressions pose challenges, has been the focus of recent approaches [8,13,14]. Overcoming these difficulties remains a significant and ongoing problem in the field of facial expression recognition.

Liu et al. [6] introduced a novel approach called the boosted deep belief network (BDBN) for facial expression recognition. The BDBN consists of a set of classifiers referred to as weak classifiers, with each one responsible for classifying a specific expression. Their approach iteratively performs three learning stages (feature learning, feature selection, and classifier construction) within a unified framework. They conducted experiments using two publicly available databases of static images, namely Cohn-Kanade [15] and JAFFE [16], achieving accuracies of 96.7% and 91.8% respectively. In addition, they conducted experiments in less controlled scenarios using a cross-database configuration, training with CK+ and testing on JAFFE, achieving an accuracy of 68.0%. Preprocessing was performed on all images based on given eye coordinates, involving alignment and cropping. The training and testing utilized a one-versus-all classification strategy, employing a binary classifier for each expression. The network training took approximately 8 days, and the recognition was calculated based on the weak classifiers. Their method employed six or seven classifiers depending on the number of expressions to be recognized, with each classifier taking 30 ms to recognize an expression, resulting in a total recognition time of about 0.21 s. The reported recognition time was based on a 6-core 2.4 GHz PC.

Song et al. [7] developed a facial expression recognition system that utilizes a deep Convolutional Neural Network (CNN) and is designed to run on a smartphone. The proposed network comprises five layers and 65,000 neurons. The authors addressed the issue of overfitting that arises when training data is limited and the network size is substantial. They applied data augmentation techniques to increase the amount of training data and employed dropouts during network training [17]. The experiments were conducted using the CK+ dataset [14] along with three additional datasets created by the authors. For the CK+ dataset, the images were initially cropped to focus on facial regions that exhibit changes caused by expressions. The authors performed a 10-fold cross-validation; however, they did not specify whether there was any overlap among subjects in the training and test sets. Hence, it can be assumed that there was some overlap. They achieved an accuracy of 99.2% on the CK+ database when recognizing five expressions (anger, happy, sad, surprise, and neutral).

Burkert et al. [10] introduced a method for facial expression recognition that is based on Convolutional Neural Networks (CNNs). The authors emphasize that their approach is independent of any hand-crafted feature extraction techniques and instead utilizes the raw image as input. Their network architecture consists of four main parts. The first part handles the automatic data preprocessing, while the remaining parts are responsible for the feature extraction process. The extracted features are then classified into the corresponding expression using a fully connected layer at the end of the network. The proposed architecture comprises a total of 15 layers, including 7 convolutional layers, 5 pooling layers, 2 concatenation layers, and 1 normalization layer. To evaluate their method, they used the CK+ database and the MMI database, achieving accuracies of 99.6% and 98.63% respectively. However, it is worth noting that their experimental methodology did not explicitly ensure that the subjects used in training were entirely separate from those used in testing.

Shan et al. [18] conducted a study using Local Binary Patterns (LBP) as a feature extractor for facial expression recognition. They compared and combined various machine learning techniques, including template matching, Support Vector Machine (SVM), linear discriminant analysis, and linear programming. The authors also investigated the impact of image resolution on accuracy and found that methods based on geometric features struggle with low-resolution images, whereas methods based on appearance, such as Gabor wavelets and LBP, are less sensitive to image resolution. The highest accuracy achieved in their work was 95.1% using SVM and LBP on the CK+ database. To evaluate the proposed system in a less controlled scenario, they performed a cross-database validation by training with CK+ and testing with JAFFE, achieving an accuracy of 41.3%. The images were initially cropped using the eye positions. For their experimental setup, they employed a 10-fold cross-validation scheme without any subject overlap.

III. **METHODOLOGY**

We have conducted research on facial expression recognition using three different datasets: FER2013, CK+, and AffectNet. We have also implemented and compared the results of three different algorithms on these datasets to achieve more accurate results.By using multiple datasets, we have evaluated the performance of our algorithm on different data distributions and identified its strengths and weaknesses. By using multiple algorithms, we have compared the effectiveness of different approaches and identified the best performing algorithm.Comparing the results of different algorithms can also provide insights into the strengths and weaknesses of each approach and help guide future research directions. For example, if one algorithm performs significantly better than the others on all three datasets, it may indicate that this approach is particularly effective for facial expression recognition.

A. DATASETS

a. FER2013

The FER2013 (Facial Expression Recognition 2013) dataset contains images along with categories describing the emotion of the person in it. This dataset was prepared by Pierre-Luc Carrier and Aaron Courville. The dataset contains 48×48 pixel grayscale images with 7 different emotions. The dataset contains 28709 examples in the training set, 3589 examples in the public testing set, and 3589 examples in the private test set. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression in to one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).Train.csv contains two columns, "emotion" and "pixels". The "emotion" column contains a numeric code ranging from 0 to 6, inclusive, for the emotion that is present in the image. The "pixels" column contains a string surrounded in quotes for each image. The contents of this string are space-separated pixel values in row major order. test.csv contains only the "pixels" column and your task is to predict the emotion column.

The FER2013 dataset has been widely used in the field of computer vision and machine learning as a benchmark for facial expression recognition algorithms. The dataset has also been used to develop and evaluate deep learning models for facial expression recognition, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs). One of the challenges with the FER2013 dataset is the class imbalance, where some classes such as happiness and neutral have a much larger number of images compared to other classes like disgust and fear. This can make it difficult for models to accurately predict the less frequent classes.

REFERENCE

b. CK+

CK+ dataset is a widely used dataset in the field of facial expression recognition. It is a collection of facial expression images with corresponding emotion labels, created by the Kanade Research Group at Carnegie Mellon University. The dataset consists of 593 image sequences of 123 subjects, with each sequence containing a series of images depicting the same subject with different facial expressions.

The images in the CK+ dataset were captured under controlled laboratory conditions, with neutral expressions, posed expressions, and peak expressions for five basic emotions: anger, fear, happiness, sadness, and surprise. Each image sequence is labeled with a corresponding emotion category, and the images are manually annotated with key facial landmarks to facilitate feature extraction.

The CK+ dataset is a valuable resource for developing and evaluating facial expression recognition algorithms. It is often used as a benchmark dataset for evaluating the performance of algorithms on recognizing basic emotions, and has been used in many research papers on facial expression recognition.

One of the strengths of the CK+ dataset is its controlled conditions, which ensure consistency in the data and enable researchers to compare the performance of different algorithms under the same conditions. However, a potential limitation is that the dataset may not fully capture the complexity and variability of facial expressions in real-world settings.

REFERENCE

c. AFFECTNET

AffectNet is a large-scale dataset for facial expression recognition, created by the AffecTech consortium in collaboration with the University of Trento and the University of Geneva. It is a diverse dataset that contains over one million images with corresponding emotion labels, making it one of the largest publicly available datasets for facial expression recognition.

The AffectNet dataset contains images of real-world facial expressions, captured from a variety of sources including internet searches, social media, and webcams. The images cover a wide range of emotions, from basic emotions such as happiness, anger, and sadness, to more complex emotions such as confusion, fear, and surprise.

In addition to emotion labels, the AffectNet dataset also provides information about the gender, age, and ethnicity of the individuals in the images, as well as image quality and face landmarks annotations. This additional information can be used to develop more accurate and robust facial expression recognition algorithms.

REFERENCE

One of the strengths of the AffectNet dataset is its diversity, which makes it more representative of real-world facial expressions than some other datasets that use controlled conditions. This diversity also enables researchers to evaluate the performance of algorithms on more complex emotions and to address issues related to demographic bias in facial expression recognition.

B. ALGORITHMS

a. CNN

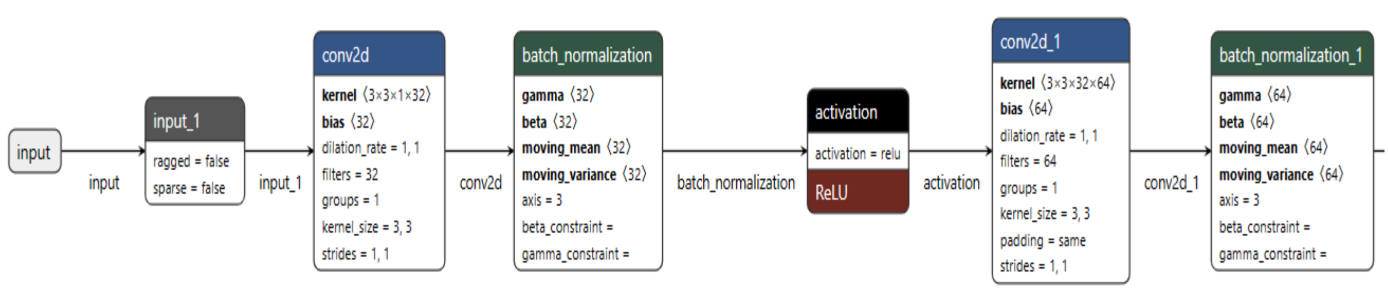
CNN, short for Convolutional Neural Network, is a type of deep learning algorithm commonly used for image and video recognition, processing, and analysis. CNNs are particularly well-suited for tasks that involve identifying patterns and features within visual data.

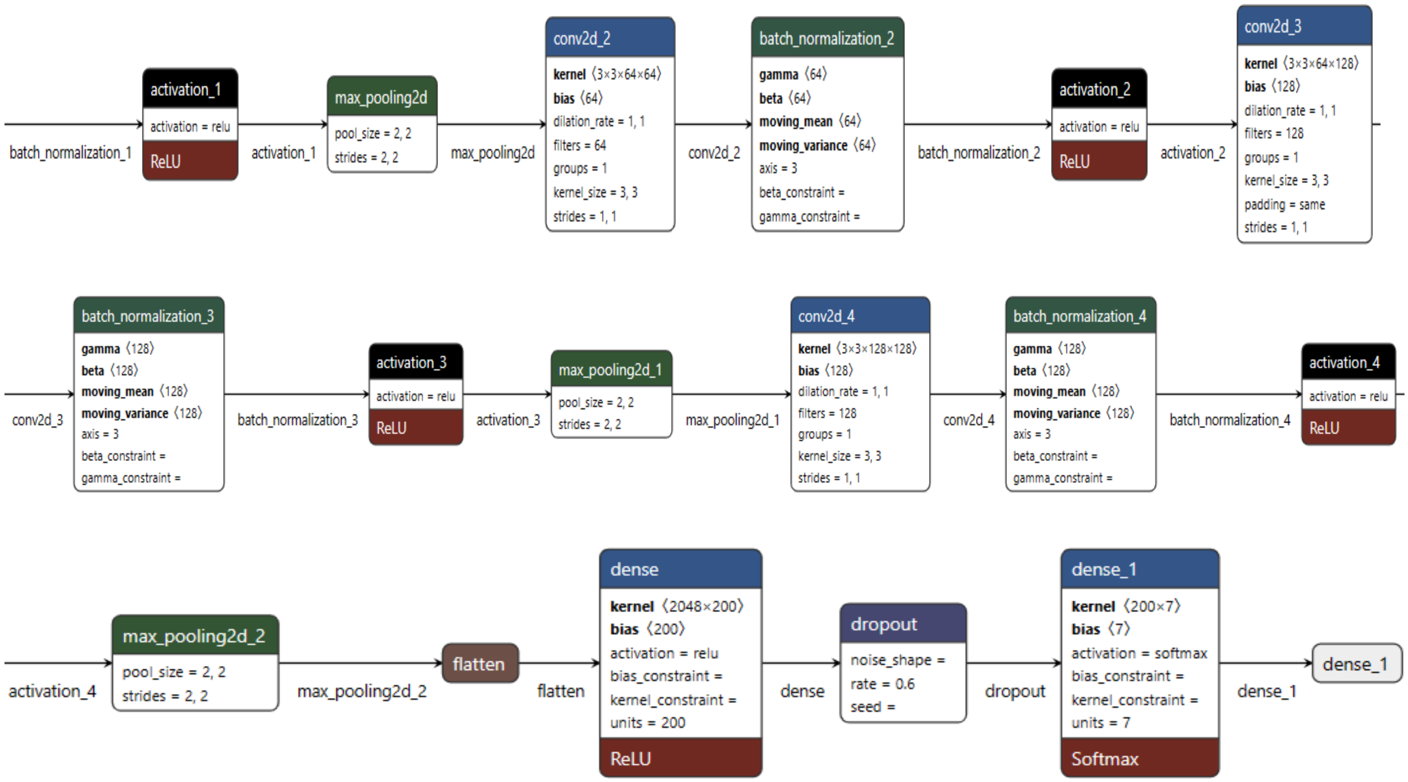
When using the CNN algorithm on an image dataset, the input data is typically a set of images that have been preprocessed to a standard size and format. The images are then fed into the network as input, and the algorithm learns to recognize patterns and features in the data through a series of convolutional, pooling, and fully connected layers.

The CNN architecture for facial expression recognition typically consists of multiple convolutional and pooling layers, followed by fully connected layers for classification. The convolutional layers apply a set of learnable filters to the input images, creating a set of feature maps that represent different features or patterns within the images. The pooling layers then downsample the feature maps to reduce the spatial size of the data and extract the most important features.

The output of the final convolutional or pooling layer is then flattened and fed into a

ARCHITECTURE OF CNN





set of fully connected layers, which perform the final classification of the image based on the learned features. The number of neurons in the output layer corresponds to the number of possible facial expressions in the dataset. During training, the CNN algorithm adjusts the weights of the filters and neurons to minimize the difference between the predicted and actual facial expressions in the training data. Once the model is trained, it can be used to predict the facial expression of new, unseen images.

b. LSTM

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) that has been used for facial expression recognition (FER). LSTMs are particularly useful for modeling sequences of data, such as a sequence of images representing facial expressions over time.

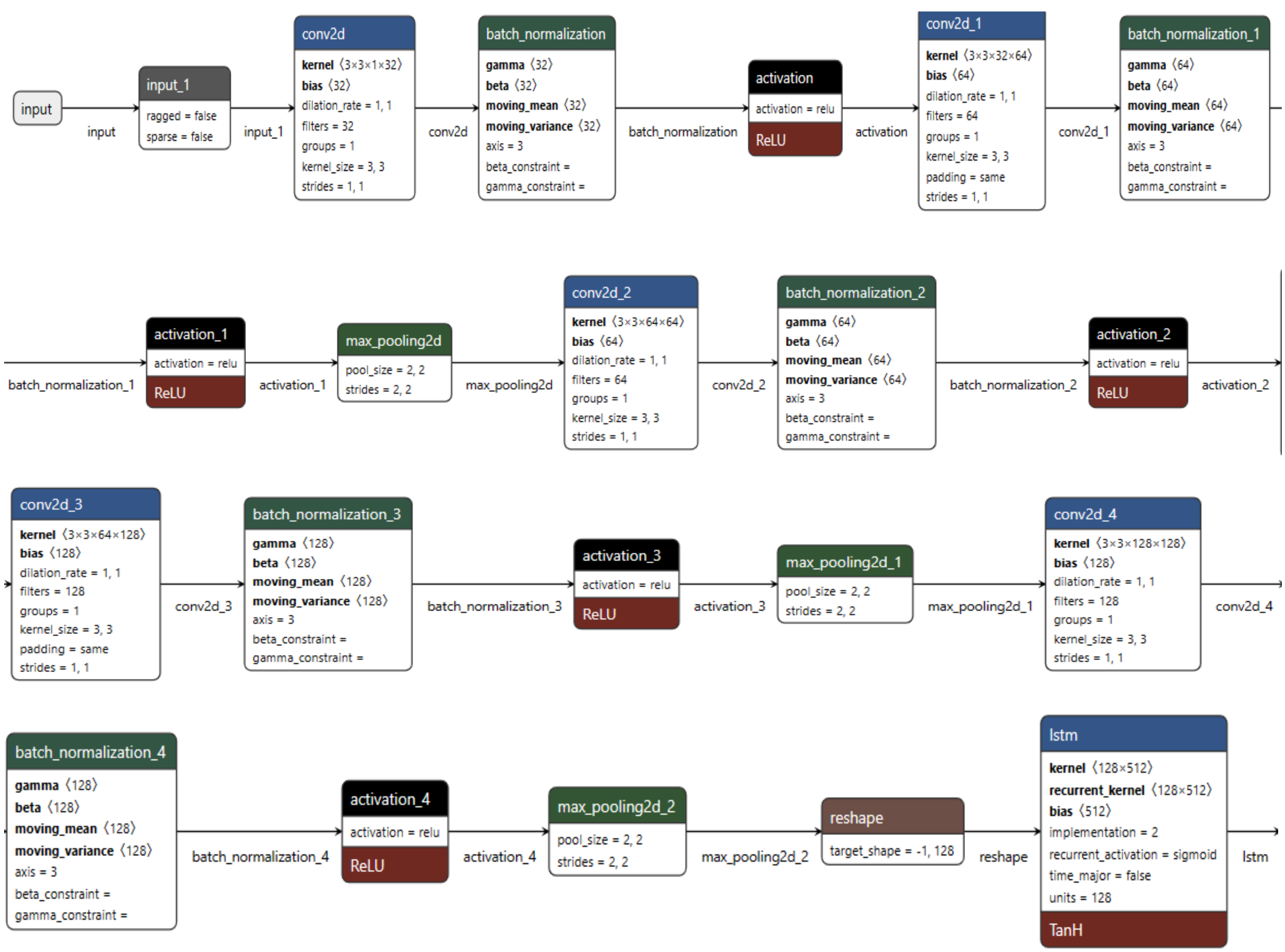
In an LSTM network for FER, the input sequence consists of a series of facial expression images, and the output is a prediction of the emotion category for each image. The LSTM network contains multiple memory cells that can remember information over long periods of time, making it well-suited for modeling sequences with complex temporal dependencies.

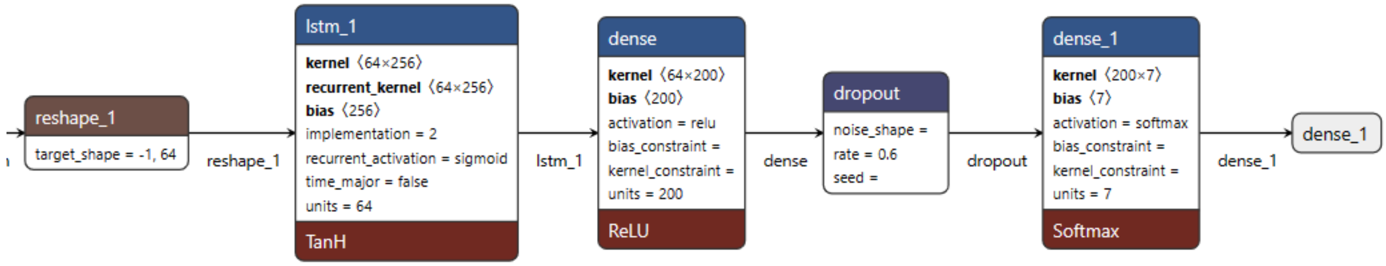
In an LSTM network for FER, the input images are typically preprocessed to extract features, such as facial landmarks or deep features from a convolutional neural network (CNN). These features are then fed into the LSTM network, which processes the sequence of features and produces an output prediction for each image.

One of the advantages of using LSTM for FER is its ability to capture temporal dependencies between facial expressions. For example, an LSTM network can learn to recognize a sequence of subtle changes in facial expression that lead up to a particular emotion.

LSTM networks have been shown to achieve high accuracy on several FER datasets, including the CK+ dataset and the FER2013 dataset.

ARCHITECTURE OF LSTM





c. BILSTM

Bidirectional Long Short-Term Memory (BiLSTM) is an extension of LSTM that has been used for facial expression recognition (FER) with promising results. BiLSTM networks can capture both forward and backward dependencies in a sequence, which makes them well-suited for modeling temporal dynamics in facial expressions.

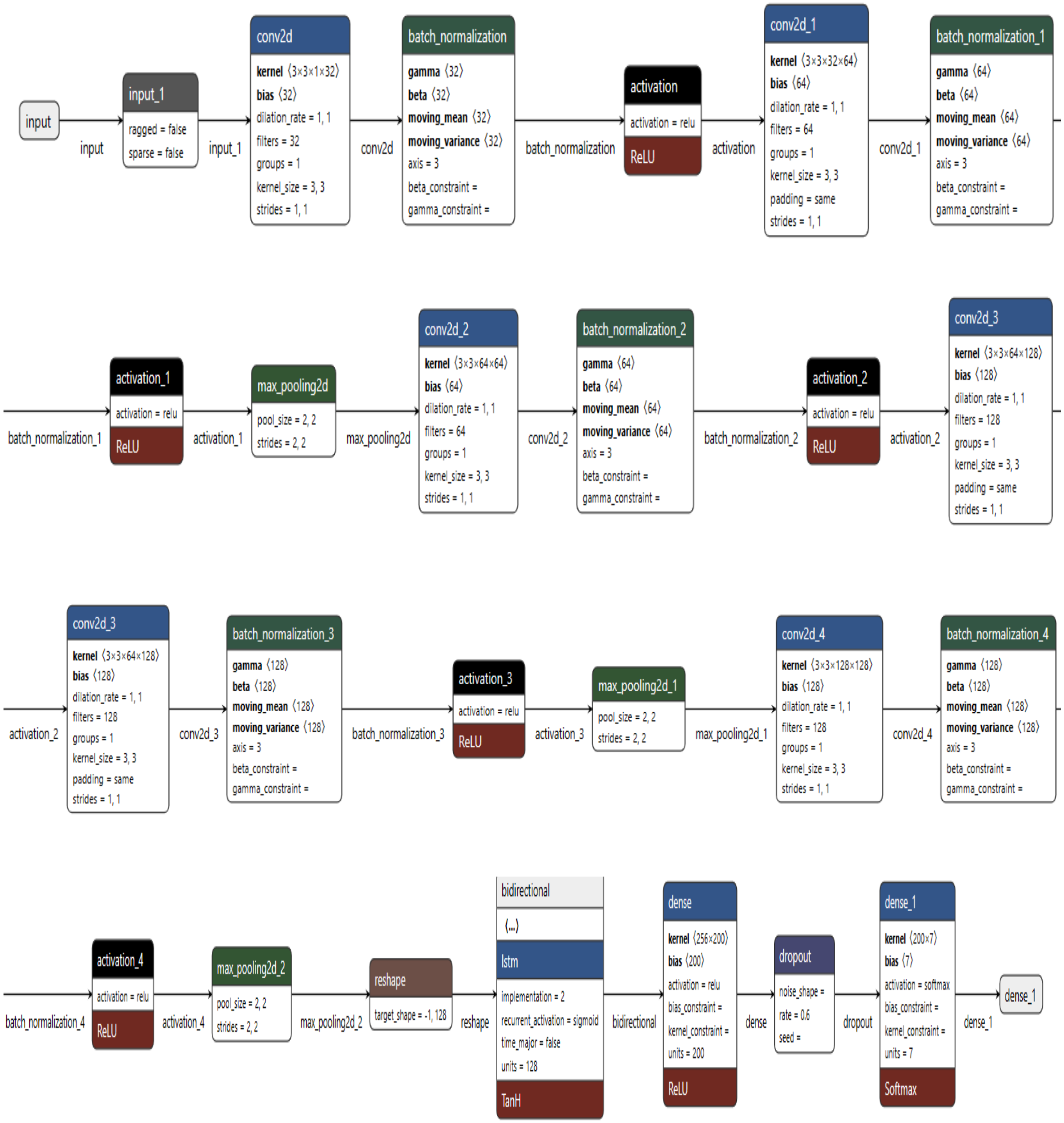
In a BiLSTM network for FER, the input sequence consists of a series of facial expression images, and the output is a prediction of the emotion category for each image. The BiLSTM network contains two LSTM layers, one that processes the input sequence in the forward direction, and another that processes the sequence in the backward direction. The outputs of both layers are then concatenated and used to make the final prediction.

One of the advantages of using BiLSTM for FER is its ability to capture both short-term and long-term temporal dependencies in facial expressions. This can lead to better recognition of subtle changes in facial expression that occur over time.

BiLSTM networks have been shown to achieve high accuracy on several FER datasets, including the CK+ dataset and the FER2013 dataset. They have also been used in conjunction with other techniques, such as attention mechanisms and data augmentation, to further improve accuracy.

One potential limitation of BiLSTM networks is their computational complexity, which can make them slower to train and more resource-intensive than simpler models. However, recent advancements in hardware and software have made it easier to train and deploy BiLSTM networks for FER

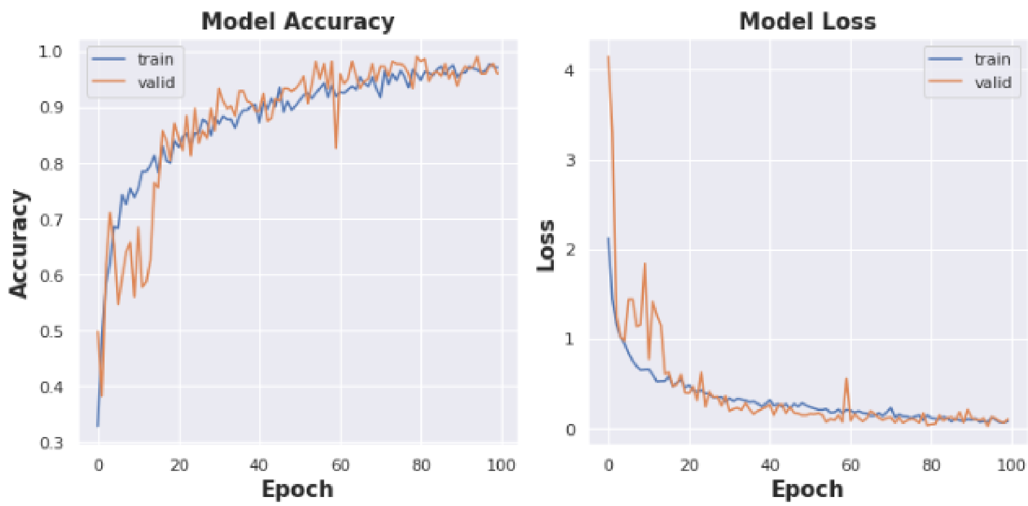
ARCHITECTURE Of BILSTM

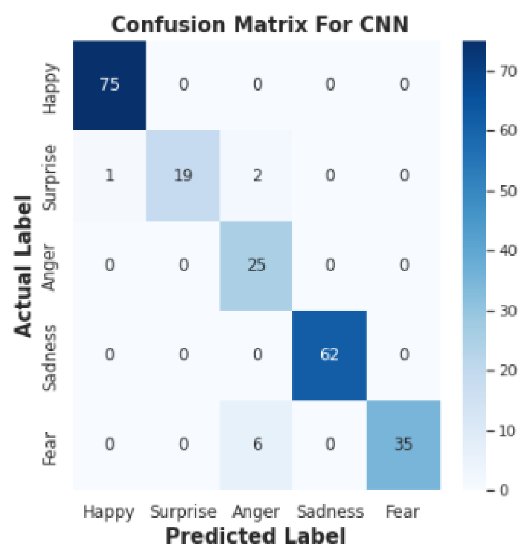


C. Output of Algorithms on different dataset

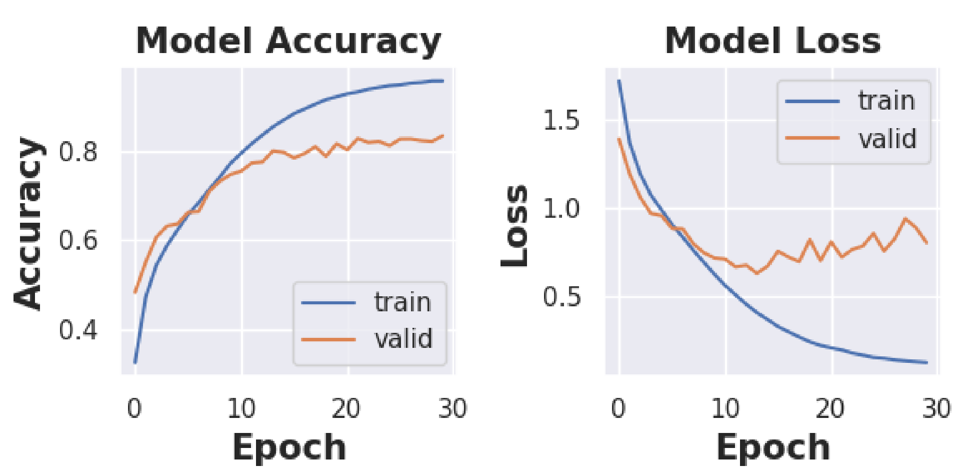
a. CNN

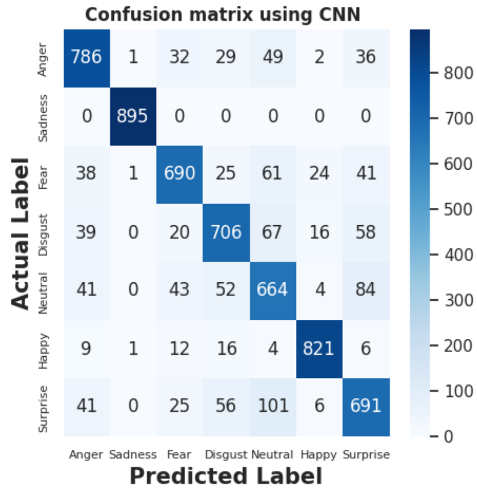
i. CK+



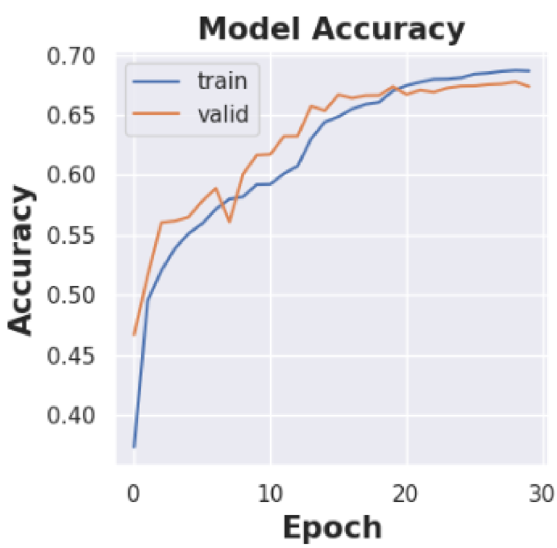


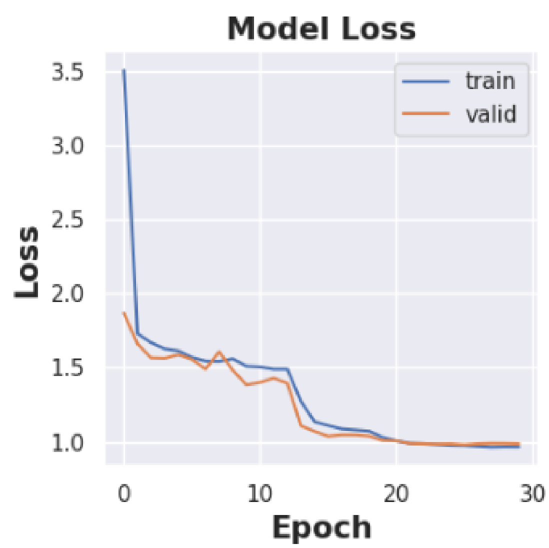
ii. FER-2013

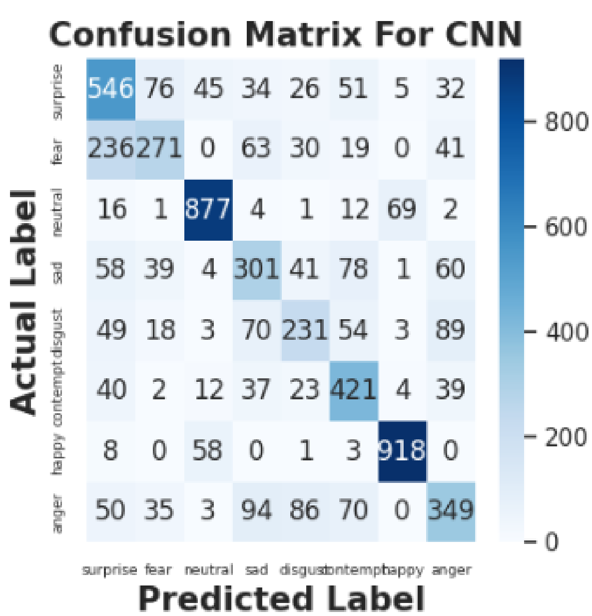




iii. AFFECTNET

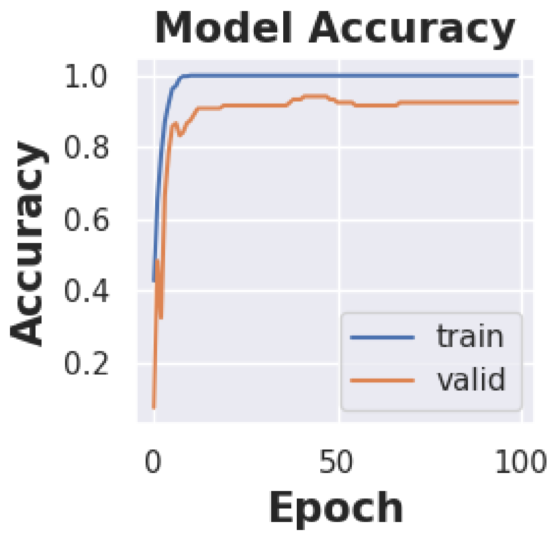


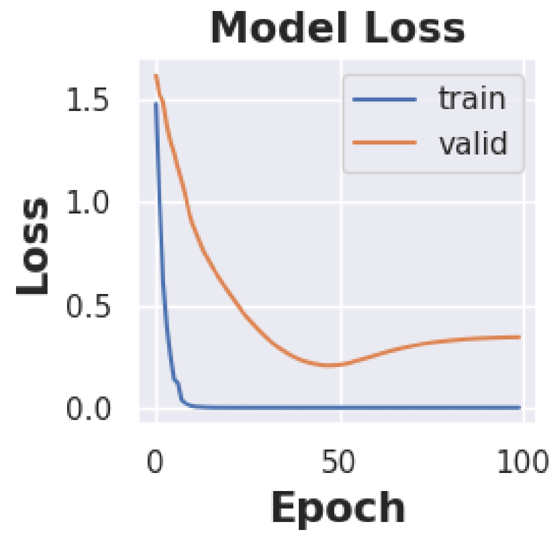


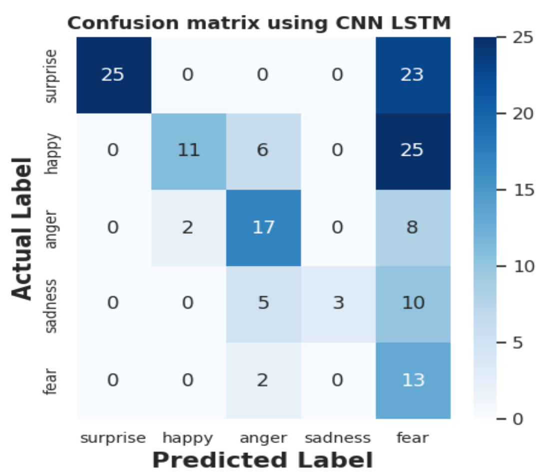


b. CNN-LSTM

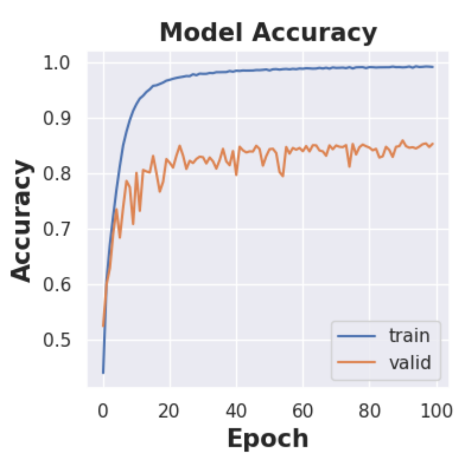
i. CK+

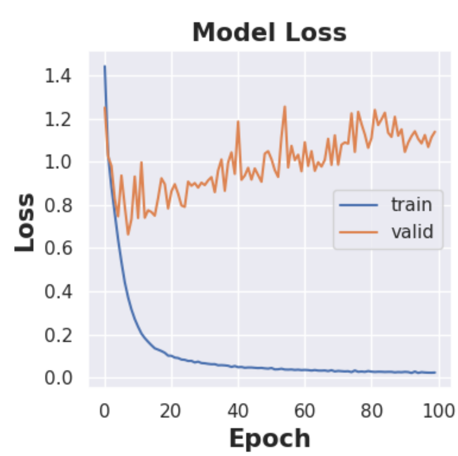


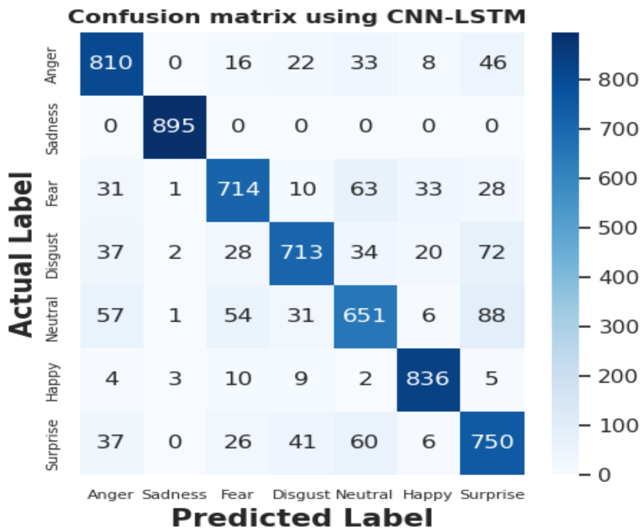




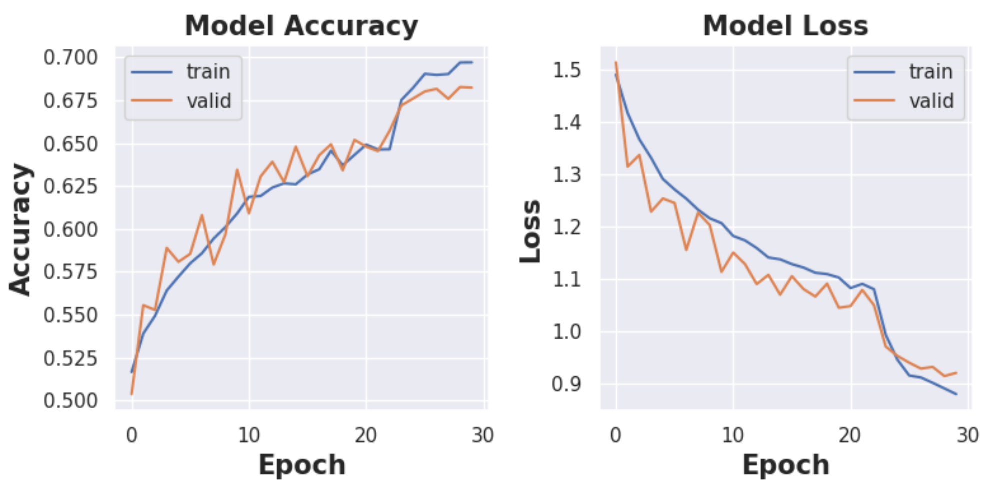
ii. FER-2013

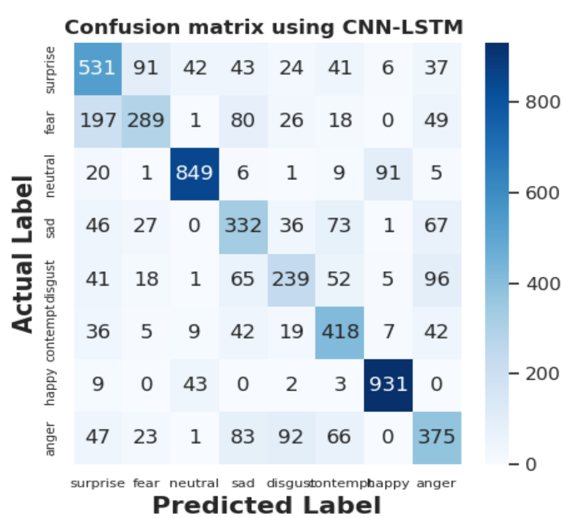






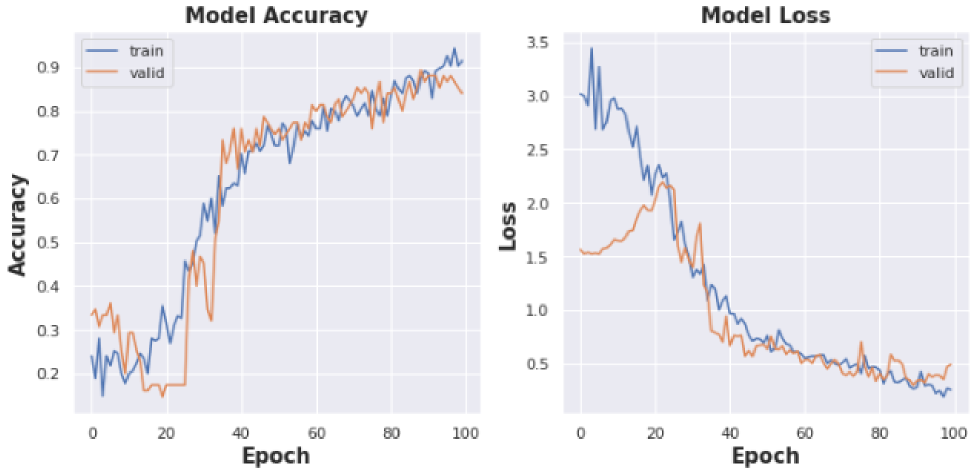
iii. AFFECTNET

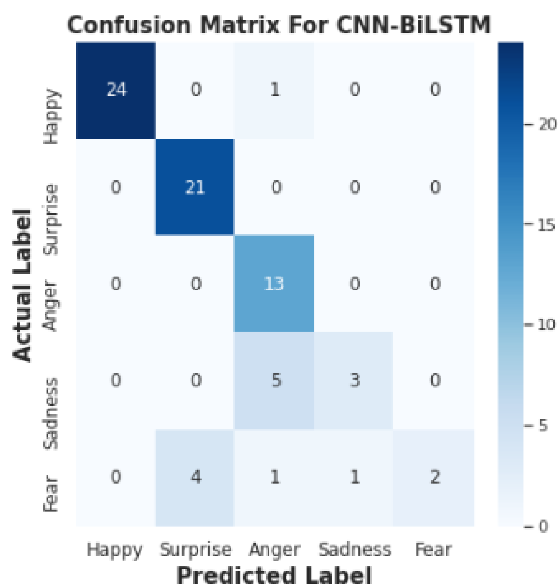




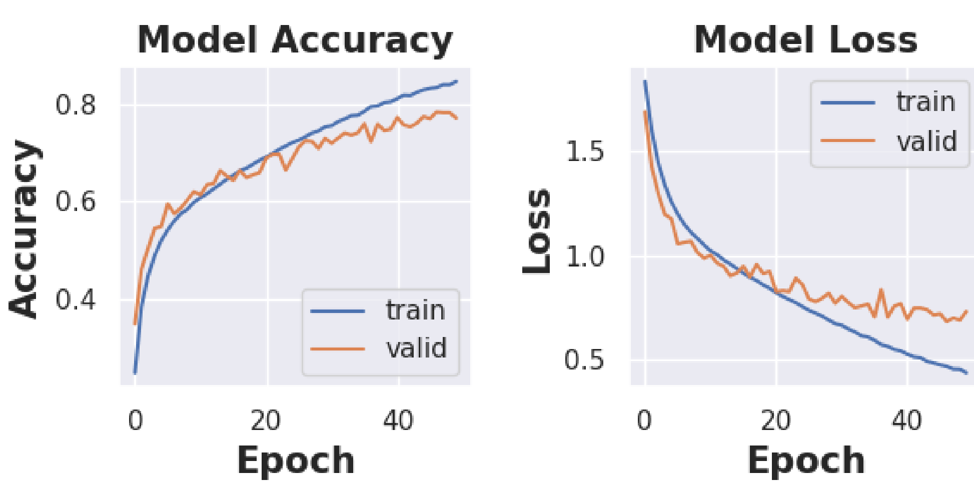
c. CNN-BILSTM

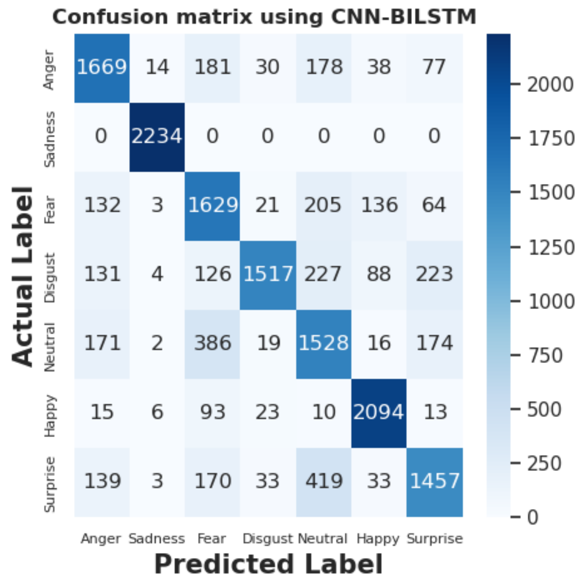
i. CK+



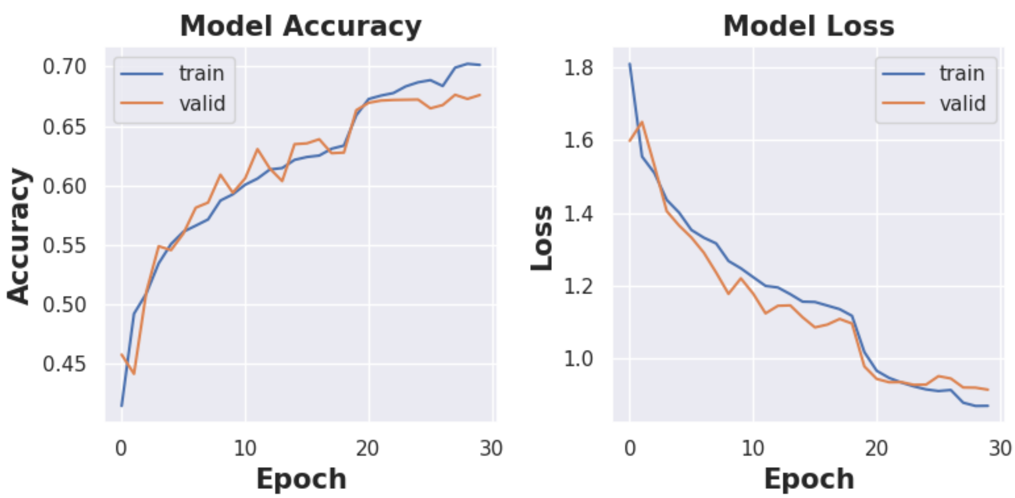


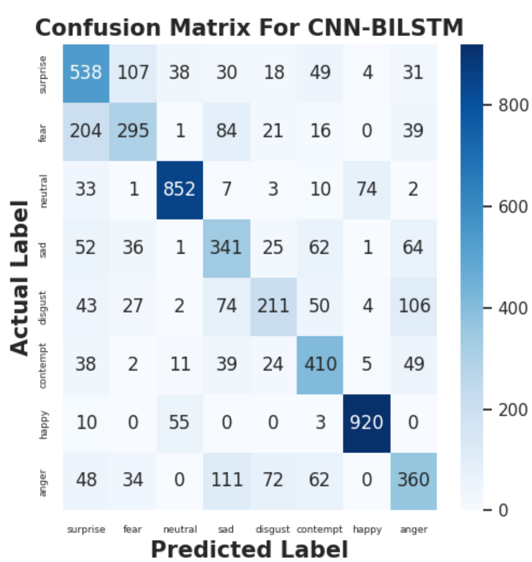
ii. FER-2013





iii. AFFECTNET





III. RESULTS

DATASET	CK+		FER2013		AFFECTNET	
METHOD	Training accuracy	Testing accuracy	Training accuracy	Testing accuracy	Training accuracy	Testing accuracy
CNN	97.14%	96%	95.82%	83.47%	68.67%	67.38%
CNN-LSTM	100.00%	92.50%	99.13%	85.31%	69.71%	68.24%
CNN-BILSTM	91.43%	84.00%	98.97%	83.68%	70.13%	67.60%

The table displays the performance of three different deep learning models: CNN, CNN-LSTM, and CNN-BILSTM, on three different facial expression recognition datasets: CK+, FER2013, and AFFECTNET. The performance of each model is measured in terms of accuracy and reported as percentages in the table.

For the CK+ dataset:

- CNN method achieved a training accuracy of 97.14% and a testing accuracy of 96%.
- CNN-LSTM method achieved a training accuracy of 100% and a testing accuracy of 92.5%.
- CNN-BILSTM method achieved a training accuracy of 91.43% and a testing accuracy of 84%.

For the FER2013 dataset:

- CNN method achieved a training accuracy of 95.82% and a testing accuracy of 83.47%.
- CNN-LSTM method achieved a training accuracy of 99.13% and a testing accuracy of 85.31%.
- CNN-BILSTM method achieved a training accuracy of 98.97% and a testing accuracy of 83.68%.

For the AFFECTNET dataset:

- CNN method achieved a training accuracy of 68.67% and a testing accuracy of 67.38%.
- CNN-LSTM method achieved a training accuracy of 69.71% and a testing accuracy of 68.24%.
- CNN-BILSTM method achieved a training accuracy of 70.13% and a testing accuracy of 67.60%.

From the table, we can make the following observations:

1. Across all three datasets, the CNN-LSTM model performs the best, achieving the highest accuracy scores among the three models.

2. The CNN model is the second-best performing model on all three datasets.
3. The CNN-BILSTM model has the lowest accuracy scores on all three datasets.
4. AFFECTNET appears to be the most challenging dataset, with the lowest accuracy scores among the three datasets for all three models.
5. CK+ is the easiest dataset, with the highest accuracy scores among the three datasets for all three models.
6. It's true that CK+ is a relatively small dataset with only 5 emotions (anger, disgust, fear, happiness, and sadness), which can make it easier for deep learning models to achieve higher accuracy scores compared to larger datasets with more emotions or more variations in facial expressions.
7. AffectNet is a relatively challenging dataset due to its large size and the high variability in the facial expressions and lighting conditions of the subjects.This can make it more difficult to extract relevant features from the images and to train accurate models that can generalize well to new subjects or expressions.

IV. CONCLUSION

Deep learning algorithms can achieve high accuracy in facial expression recognition tasks, but the complexity of the dataset can impact their performance. The accuracy of a deep learning model is influenced by various factors such as the size and quality of the dataset, the number and type of facial expressions, lighting and other environmental conditions, and the specific architecture and hyperparameters of the model.

Therefore, it's important to carefully choose the appropriate deep learning algorithm and architecture for a given dataset and task, taking into account the unique characteristics and challenges of the data. Different algorithms may be better suited for different types of facial expression data, depending on factors such as the number of emotions, the amount of data available, and the level of variability in the facial expressions.

Ultimately, achieving high accuracy in facial expression recognition requires careful data preparation, effective feature extraction techniques, and robust model architectures that can handle the challenges of the data. It's important to carefully evaluate and compare the performance of different models on specific datasets to determine the most effective approach for a given task.

V. REFERENCE

1. Khaireddin, Yousif, and Zhuofa Chen. "Facial emotion recognition: State of the art performance on FER2013." *arXiv preprint arXiv:2105.03588* (2021).
2. Karnati, Mohan, et al. "Understanding Deep Learning Techniques for Recognition of Human Emotions using Facial Expressions: A Comprehensive Survey." *IEEE*

Transactions on Instrumentation and Measurement (2023).
3. Nonis, Francesca, et al. "3D approaches and challenges in facial expression recognition algorithms—a literature review." *Applied Sciences* 9.18 (2019): 3904.
4. N. N. Khatri, Z. H. Shah, and S. A. Patel, “Facial expression recognition: A survey,” Int. J. Comput. Sci. Inf. Technol., vol. 5, no. 1, pp. 149–152, 2014.
5. D. A. Rosenbaum, Human Motor Control. New York, NY, USA: Academic, 2009.
6. P. Liu, S. Han, Z. Meng, Y. Tong, Facial expression recognition via a boosted deep belief network, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1805–1812.
7. I. Song, H.-J. Kim, P.B. Jeon, Deep learning for real-time robust facial expression recognition on a smartphone, in: International Conference on Consumer Electronics (ICCE), Institute of Electrical & Electronics Engineers (IEEE), Las Vegas, NV, USA, 2014.
8. M. Liu, S. Li, S. Shan, X. Chen, Au-inspired deep networks for facial expression feature learning, Neurocomputi
9. Y.-H. Byeon, K.-C. Kwak, Facial expression recognition using 3d convolutional neural network. International Journal of Advanced Computer Science and Applications(IJACSA), 5 (2014).
10. P. Burkert, F. Trier, M.Z. Afzal, A. Dengel, M. Liwicki, Dexpression: Deep Convolutional Neural Network for Expression Recognition, CoRR abs/1509.05371 (URL ⟨http://arxiv.org/abs/1509.05371⟩). [12] M. Liu, S. Li, S. Shan, X. Chen, Au-inspired
11. S.Z. Li, A.K. Jain, Handbook of Face Recognition, Springer Science & Business Media, Secaucus, NJ, USA, 2011.
12. C.-D. Caleanu, Face expression recognition: a brief overview of the last decade, in: 2013 IEEE 8th International Symposium on Applied Computational Intelligence and Informatics (SACI), 2013, pp. 157–161.
13. M.K.A.E. Meguid, M.D. Levine, Fully automated recognition of spontaneous facial expressions in videos using random forest classifiers, IEEE Trans. Affect. Comput. 5 (2) (2014) 141–154, http://dx.doi.org/10.1109/TAFFC.2014.2317711.
14. C. Turan, K. M. Lam, Region-based feature fusion for facial-expression recognition, in: 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 5966–5970 (http://dx.doi.org/10.1109/ICIP.2014.7026204).

15. P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, The extended Cohn–Kanade dataset (CKþ): a complete dataset for action unit and emotionspecified expression, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2010, pp. 94–101.
16. M. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, Coding facial expressions with gabor wavelets, in: Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998, 1998, pp. 200–205.
17. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
18. C. Shan, S. Gong, P.W. McOwan, Facial expression recognition based on local binary patterns: a comprehensive study, Image Vis. Comput. 27 (6) (2009) 803–816.