International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 06th Feb 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=ISSUE-02

DOI: 10.48047/IJIEMR/V12/ISSUE 02/28

Title **Classification of Facial Emotion Expressions using DeepLearning Approaches**

Volume 12, Issue 02, Pages: 175-179

Paper Authors
Jinka Chandra Kiran, Abbaraju Sai Sathwik, BeebiNaseeba, Nagendra Panini Challa

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Classification of Facial Emotion Expressions using Deep-Learning Approaches

[1]**Jinka Chandra Kiran**, B-Tech (CSE), VIT-AP University, India
chandrakiran.20bci7045@vitap.ac.in
[2]**Abbaraju Sai Sathwik**, B-Tech (CSE) , VIT-AP University, India
sathwik.20bce7392@vitap.ac.in
[3] **BeebiNaseeba**
School of Computer Science and Engineering (SCOPE), VIT-AP University, Amravati, India
beebi.naseeba@vitap.ac.in
[4] **Nagendra Panini Challa**
School of Computer Science and Engineering (SCOPE), VIT-AP University, Amravati, India
nagendra.challa@vitap.ac.in

Abstract
In this work emotion recognition and human face expression extraction are both done using machine learning and deep learning models. The many different body language tactics, including as eye stirring, idiomatic expressions, and body motions, are significant when applying for the interaction between computers and humans. Identification of human emotions is crucial for criminal investigation, detection, prevention and many other domains. In this work VGG-19, ResNet-50, ResNet-101, (CNN's) to identify contempt, sadness, anger, fear, disgust, happy and surprise. An overall accuracy of 98% is achieved through VGG-19 which is best suitable for facial recognition.

Index Terms–Convolution Neural Network, ResNet, VGG- 19, Efficient Net, Deep Learning.

Introduction

Mood-indicating facial expressions, such as those used to show happiness, surprise, disdain, sorrow, fear, disgust, and fury. A person's goals in a social setting may be inferred from their facial expressions, which can also reflect inner feelings. In people's social relationships, they are extremely important. Background surroundings in which faces are seen give crucial contextual information for comprehending facial emotions. Ostensive motions, like the brow flick, which denote the desire to communicate, are particularly significant among facial expressions. These movements suggest two things: first, that the sender may be trusted; and, second, that the recipient should take note of any further signals.

Facial emotion identification research is unquestionably a popular area in academic circles and business. The technology for recognizing facial expressions is frequently employed in daily life today, including in the identification of driver tiredness. According to the existing research, maximum of them employed the transfer learning technique with well-established pre-trained Deep-Learning methods for expression detection analysis. Since there are just a few publications on the subject of facial emotion detection. In this area, their plan has produced good outcomes. But it has three significant drawbacks. First, these models only perform adequately for binary classification. Second, they don't identify the top pre-trained DL techniques or the most effective combinations of them in order to get the highest performance; instead, For transfer learning, they just consider the VGG-19 DL model.

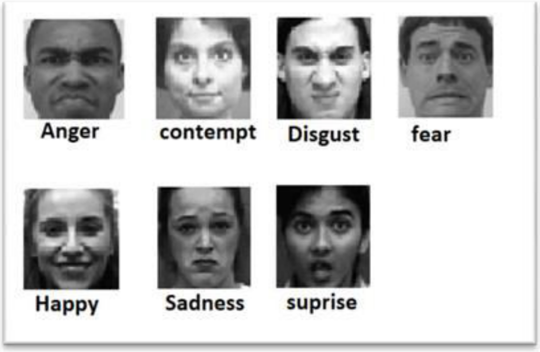

Fig1:Sample images from Dataset

Third, their models aren't clear enough to understand. First, in order to get beyond the aforementioned restrictions, we improve the 13 DL models that have already been trained using the same technique.

Literature Survey

DL approaches are used to categorize face emotion detection in various recent research reports. The bulk of the writers used deep learning models like VGG16, Inceptionnet-v2, EfficientNet b0, and CNN. And 65.3% for AleNet, 66.4 for VGG, 61.9 for RESNET, 70% for CNN, 85% for CFR- SVM, and 92 for CFR-CURL-SVM.

The authors achieved 96% accuracy in the below Literature survey using the VGG Fine-Tune DL algorithm, the highest accuracy among the following surveys.

TABLE 1

S. No	Paper Title	Computational Models	Accuracy
1.	Deep learning algorithm for Facial Expression Classification[1]	AlexNet, VGG, ResNet	65.3%, 66.4%, 61.9%
2.	Ensemble Algorithm of Convolution Neural Networks for Enhancing Facial Expression Recognition[2]	VGG16, Inceptionnet-v2, EfficientNetb0, CNN	66.6%, 67.7%, 67.4%, 70.5%
3.	Leveraging the Deep Learning Paradigm for Continuous Affect Estimation from Facial Expressions[3]	CFR-SVM, Zero-bias CNN*, VGG, CFR-ELM, VGG Fine-Tune, FN2EN, CFR-CURL-SVM, AURF, AUDN, CFR-CURL-ELM	85.1%, 85:6%, 88:7%, 89.0%, 89:9%, 96:8%, 90.2%, 92:1%, 93.7%, 93:7%

Proposed Work

The proposed work part includes the information about the process flow in the face expression detection. Mainly the first thing will be the dataset important for the detection. So, collecting the dataset followed by the data pre-processing technique to make the dataset incorrupt to receive better and accurate results followed by application of various classification algorithms followed by individual algorithm performance analysis and checking the accuracy for which algorithm it will be more and lastly doing the comparative analysis of all the algorithms with their accuracies.

The proposed method is described below for collecting data, screening experts, and augmenting the data. Additionally, we show how we develop our databases.

Data collection:

981 photos were obtained from the kaggle database, which is the world's largest community offering a data platform for machine learners to exploit. And there are 7 folds of Facial emotion expressions.

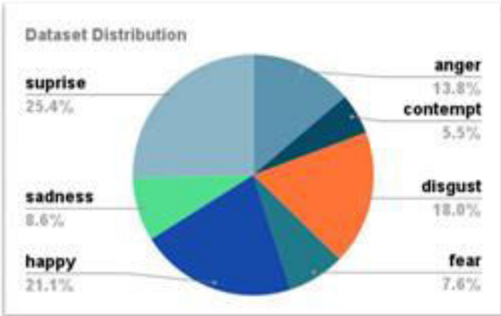

Fig.2:Original dataset distribution

Transfer learning

Transfer learning is the process by which a neural network that has been taught to do one task may be trained to perform another. This type of application is useful for problems with a little dataset. This type of application works well since the first layers model tries to identify properties like edges that are shared by both datasets. Because most real-life scenarios do not require complicated models to be trained using millions of labeled data points, this is particularly helpful in the field of data science.

In this study, we will freeze the layers

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

from the previously trained model while adding some new layers to it. Furthermore, by adding new layers, old features will be transformed into new dataset predictions (face emotion expression dataset)

A well-known collection of actual photographs, IMAGENET has 1000 classes, making it more powerful. We will make use of models that have already been trained using it. For this work, we choose three pre-trained DL models. VGG 19, ResNET101, ResNET50, and CNN are the models. We'll add the identical new layer to every pre-trained model.

Model Building

VGG

The Convolutional Neural Network (CNN) architecture is deep and has several layers. Convolutional neural networks' primary characteristics are the foundation of VGGNets (CNN). Input, output, and hidden layers make up convolutional neural networks. Small convolutional filters make up this network. All hidden layers' activation functions are provided by ReLU. All hidden layers' activation functions are provided by ReLU. Because ReLU enables speedier learning and lowers the frequency of gradient issues that vanish, it is more successful when used on computers.VGG19 in the study accepts input tensor sizes of 224, 244 with 3 RGB channels. These input photos are sent into the VGG network, which updates the weights of the newly added layers. Finally, the model generates a vector with 7 values which consists of face emotion expression probabilities.. The softmax function is used to ensure that these probabilities total up to 1.

RESNET

The exploding gradients issue that happened in VGG-16 and 19 is resolved using resnets. The creation of the ResNet model was made possible by the introduction of Residual blocks, which made it considerably simpler to train incredibly deep networks. We apply in this network, there is a technique called skip connections. The skip connection skips a few layers in between and connects layer activations to succeeding levels. Regularization will skip any layer that reduces architectural performance, which is an advantage of incorporating this sort of skip link. As a consequence, vanishing/exploding gradient problems may be avoided while training an exceedingly deep neural network.

RESNET 5O AND RESNET 101 in the study accept input tensor sizes of 224, 244 with 3 RGB channels. These input photos are sent into the ResNet network, by doing so, a new layer is added, and its weights are updated s. Finally, the model generates a vector with 7 values which consists of face emotion expression probabilities. The softmax function is used to ensure that these probabilities total up to 1.

Performance Evaluation

With the use of the most popular statistical techniques, including accuracy, precision, recall and F1-score, the total experimental result is measured.The dataset for this study is rather small. The possibility of misclassification thus exists. For instance, the patient might not be identified if true positive (TP) face emotion expression data points are categorized as other class. In order to choose the best model from all deployed machine learning and deep learning models, we must carefully review statistical measures.

Accuracy: The number of predictions made by a model thatare true across all cases is its accuracy

$$Accuracy = \frac{Tn + Tp}{Fn + Fp + Tn + Tp} \quad (1)$$

Precision: An evaluation of precision will be based on proportion of accurately predicted events to all expected positive outcomes.

$$Precision = \frac{Tp}{Fp + Tp} \quad (2)$$

Recall: An algorithm's recall refers to how many relevant results it correctly recognizes out of all possible results.

$$Recall = \frac{Tp}{Tn + Fp} \quad (3)$$

F1- score: This statistic is used to assess the model's recall and accuracy performance. The F1 score is widely used in NLP, and it may be adjusted to priorities precision over recall or the other way around.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

$$\text{F1} - \text{score} = 2x\,\frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Results and Discussion

DL models were used in this study. In deep learning, we used vgg19, resent50, resnet101. After applying deep learning methods to the face expression dataset, we achieved the accuracy values mentioned below. We got 97% with CNN, 98% with VGG-19, 95% with RESNET–50 and 97% accuracy with RESNET-101. Among all DL techniques, VGG19 has the highest accuracy of 98%.

TABLE 2 Accuracy comparisons

Methods	Accuracy
CNN	97
VGG - 19	**98**
RESNET - 50	95

CNN

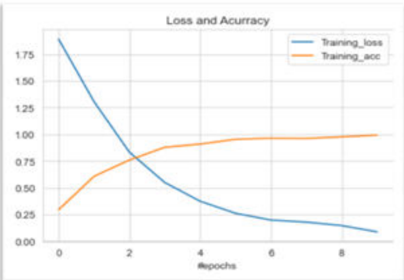

Fig. 3: Train vs validation accuracy and loss graph followed by confusion matrix for CNN

VGG – 19

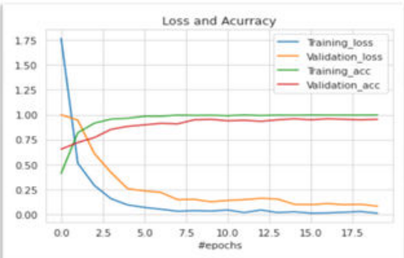

Fig. 4: Train vs validation accuracy and loss graph followed by confusion matrix for VGG19

RESNET – 50

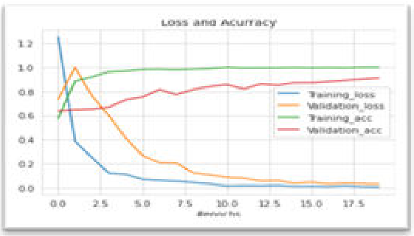

Fig. 5: Train vs validation accuracy and loss graph followed by confusion matrix for resent 50

RESNET – 101

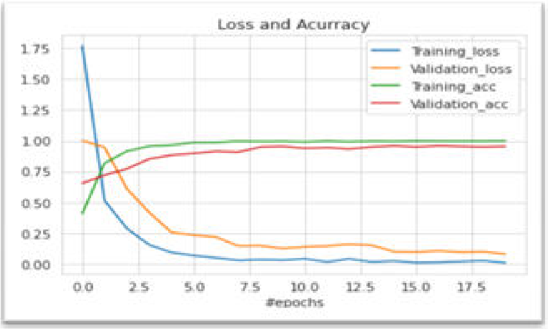

Fig. 6: Train vs validation accuracy and loss graph followed by confusion matrix for resent 101

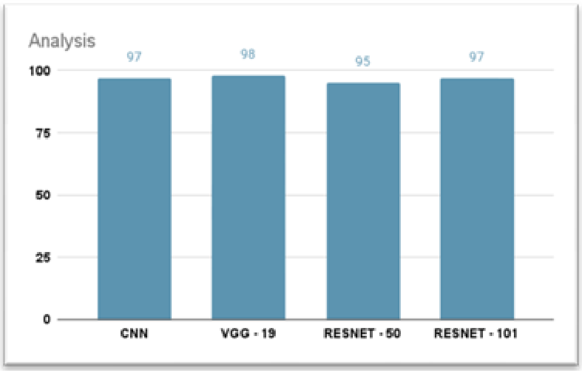

Fig. 7: Comparative analysis of accuracies

Conclusion and Future Work

On the face expression dataset, we examined 3 different pre- trained DL models using transfer learning. We got 98% accuracy with CNN, 98% VGG 19, 95% with RESNET 50 and 97% accuracy with RESNET 101. We identified the VGG19 algorithm as the best-performing DL model for predicting the face expression. The model performed best accuracy: 98% while detecting the face expression. Firstly, while the sample is very small, adding more data might increase performance even further. Second, ensemble models may be used to improve model accuracy.

References

[1]B. He, "Deep learning algorithm for Facial Expression Classification," 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications

(CVIDL & ICCEA), Changchun, China, 2022, pp. 386-391, doi: 10.1109/CVIDLICCEA56201.2022.9824661.
[2] G. -C. Lee, Z. -Y. Li and T. -W. Li, "Ensemble Algorithm of Convolution Neural Networks for Enhancing Facial Expression Recognition," 2022 IEEE 5th International Conference on Knowledge Innovation and Invention (ICKII), Hualien, Taiwan, 2022, pp. 111-115, doi: 10.1109/ICKII55100.2022.9983573.
[3] M. C. Oveneke, Y. Zhao, E. Pei, A. D. Berenguer, D. Jiang and H. Sahli, "Leveraging the Deep Learning Paradigm for Continuous Affect Estimation from Facial Expressions," in IEEE Transactions on Affective Computing, vol. 13, no. 1, pp. 426-439, 1 Jan.-March 2022, doi: 10.1109/TAFFC.2019.2944603.
[4] Roopa, S.N., Research on Fa.ce Expression Recognition. International Journal of Innovative Technology and Exploring Engineering (IJITEE),2019. 8(9s2).
[5] Bettadapura, V., Face Expression Recognition and Analysis: The State ofthe Art. 2012.
[6] Lei Xu, M.F.W.Z., Face Expression Recognition Based on Convolutional Neural Network. 2018.
[7] Căleanu, C., Face Expression Recognition: a Brief Overview of the Last Decade. 2013.
[8] Cotter, S., MobiExpress Net: A Deep Learning Network for Face Expression Recognition on Smart Phones. 2020.
[9] B., P.H., V.A. S. and D.V. K., Concise CNN model for face expressionrecognition. Intelligent Decision Technologies, 2021. 15(2).
[10] Wu, S., S. Zhong and Y. Liu, Deep residual learning for image steganalysis. 2018.
[11] Ruhua, L., et al., Facial Expression Recognition Based on Convolutional Neural Network. Journal of Physics: Conference Series, 2021.1757(1).
[12] Savchenko AV, Facial expression and attributes recognition based on multi-task learning of lightweight neural networks. 2021.
[13] Good fellow, I.J., et al., Challenges in Representation Learning: A report on three machine learning contests. 2013.
[14] Bin, L. and L. Dimas, Facial expression recognition via ResNet-50, International Journal of Cognitive Computing in Engineering,2021(preprint).
[15] Mazen, F.M.A., A.A. Nashat and R.A.A.A. Seoud, Real Time Face Expression Recognition along with Balanced FER2013 Dataset using Cycle GAN. International Journal of Advanced Computer Science and Applications (IJACSA), 2021. 12(6).