Fuzzy Systems and Soft Computing
ISSN : 1819-4362

REAL TIME FACIAL EXPRESSION RECOGNITION

[1]T. RAVI KIRAN KUMAR, [2]G. VIKAS REDDY ,[3]G. VISHNU,[4]B. PRASHANTH,[5]TALHA HUSSAIN

[1]ASSISTANT PROFESSOR, [2,3,4&5] UG STUDENTS

DEPARTMENT OF CSE, MNR COLLEGE OF ENGG. & TECHNOLOGY, MNR NAGAR, FASALWADIGUDA, SANGA REDDY-502294

ABSTRACT

We offer Real-time Facial Expression Recognition in Python with CNN for detecting facial emotions in real-time and in bulk to obtain a higher classification result in this project. We create a real-time vision system to see if our model is effective. The objective of emotion classification is accomplished by this system, which uses CNN Model Architecture. Facial expression recognition computer technology may extract emotional information from a person's expression in order to determine the person's condition and purpose. A model of a convolutional neural network is proposed in this article (CNN). This model is used to recognize facial expressions. The article starts by constructing a CNN model and learning the local features of the eyes, brows, and lips. Finally, the model's output result is chosen and fused to get the final recognition result. The system will compile the model and use the fit function to apply it. There will be 32 batches in all. The average validation accuracy was 90.00%, and the average training accuracy was 90.00%.

1.INTRODUCTION

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the back propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

2.LITERATURE SURVEY

1) Constants across cultures in the face and emotion AUTHORS: P. Ekman and W. V. Friese

Investigated the question of whether any facial expressions of emotion are universal. Recent studies showing that members of literate cultures associated the same emotion concepts with the same facial behaviors could not demonstrate that at least some facial expressions of emotion are universal; the cultures compared had all been exposed to some of the same mass media presentations of facial expression, and these may have taught the people in each culture to recognize the unique facial expressions of other cultures. To show that members of a preliterate culture who had minimal exposure to literate cultures would associate the same emotion concepts with the same facial behaviors as do members of Western and Eastern literate cultures, data were gathered in New Guinea by telling 342 Ss a story, showing them a set of 3 faces, and asking them to select the face which showed the emotion appropriate to the story. Ss were members of the Fore linguistic-cultural group, which up until 12 yr. ago was an isolated, Neolithic, material culture. Results provide evidence in support of the hypothesis

2) Challenges in representation learning: A report on three machine learning contests AUTHORS: I. J. Goodfellow et al.,

The ICML 2013 Workshop on Challenges in Representation Learning focused on three challenges: the black box learning challenge, the facial expression recognition challenge, and the multimodal learning challenge. We describe the datasets created for these challenges and summarize the results of the competitions. We provide suggestions for organizers of future challenges and some comments on what kind of knowledge can be gained from machine learning competitions.

3) Face recognition: A convolutional neural-network approach AUTHORS: S. Lawrence, C. L. Giles, A. Chung Tsoi, and A. D. Back

We present a hybrid neural-network for human face recognition which compares favourably with other methods. The system combines local image sampling, a self-organizing map (SOM) neural network, and a convolutional neural network. The SOM provides a quantization of the image samples into a topological space where inputs that are nearby in the original space are also nearby in the output space, thereby providing dimensionality reduction and invariance to minor changes in the image sample, and the convolutional neural network provides partial invariance to translation, rotation, scale, and deformation. The convolutional network extracts successively larger features in a hierarchical set of layers. We present results using the Karhunen-Loeve transform in place of the SOM, and a multilayer perceptron (MLP) in place of the convolutional network for comparison. We use a database of 400 images of 40 individuals which contains quite a high degree of variability in expression, pose, and facial details. We analyze the computational complexity and discuss how new classes could be added to the trained recognizer.

4) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,

AUTHORS: H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao,
D. Mollura, and R. M. Summers

Remarkable progress has been made in image recognition, primarily due to the availability of large-scale annotated datasets and deep convolutional neural networks (CNNs). CNNs enable learning data-driven, highly representative, hierarchical image features from sufficient training data. However, obtaining datasets as comprehensively annotated as ImageNet in the medical imaging domain remains a challenge. There are currently three major techniques that successfully employ CNNs to medical image classification: training the CNN from scratch, using off-the-shelf pre-trained CNN features, and conducting unsupervised CNN pre-training with supervised fine-tuning. Another effective method is transfer learning, i.e., fine-tuning CNN models pre-trained from natural image dataset to medical image tasks. In this paper, we exploit three important, but previously understudied factors of employing deep convolutional neural networks to computer-aided detection problems. We first explore and evaluate different CNN architectures. The studied models contain 5 thousand to 160 million parameters, and vary in numbers of layers. We then evaluate the influence of dataset scale and spatial image context on performance. Finally, we examine when and why transfer learning from pre-trained ImageNet (via fine-tuning) can be useful. We study two specific computer-aided detection (CADe) problems, namely thoraco-abdominal lymph node (LN) detection and interstitial lung disease (ILD) classification. We achieve the state-of-the-art performance on the mediastinal LN detection, and report the first five-fold cross-validation classification results on predicting axial CT slices with ILD categories. Our extensive empirical evaluation, CNN model analysis and valuable insights can be extended to the design of high performance CAD systems for other medical imaging tasks.

5) Facial expression recognition based on complexity perception classification algorithm

AUTHORS: T. Chang, G. Wen, Y. Hu, and J. Ma

Facial expression recognition (FER) has always been a challenging issue in computer vision. The different expressions of emotion and uncontrolled environmental factors lead to inconsistencies in the complexity of FER and variability of between expression categories, which is often overlooked in most facial expression recognition systems. In order to solve this problem effectively, we presented a simple and efficient CNN model to extract facial features, and proposed a complexity perception classification (CPC) algorithm for FER. The CPC algorithm divided the dataset into an easy classification sample subspace and a complex classification sample subspace by evaluating the complexity of facial features that are suitable for classification. The experimental results of our proposed algorithm on Fer2013 and CK-plus datasets demonstrated the algorithm's effectiveness and superiority over other state-of-the-art approaches.

3. SYSTEM ANALYSIS

3.1. EXISTINGSYSTEM:

- The current facial expression identification methods are mostly separated into two categories: traditional manual approaches and network models based on deep learning. Although the traditional approach is frequently utilised, its practical applicability are severely constrained. Learning how to employ strong supervision methods to describe the emotional aspects of large sample data is usually the first step in using deep learning to categorise facial expressions.

- For the formalisation of the Facial Channel neural network for Facial Expression Recognition, Barros et al. suggested a network model based on the topological structure of VGG-16 (FER).
- Koujan et al. proposed a CNN that recognized human emotions from a single face image.
- Xiao et al. combined the Region of Interest (ROI) and K-Nearest Neighbor algorithm for facial expression recognition and solved the problem of the poor generalization ability of deep neural networks in the case of small data.
- Liu et al. proposed a deep learning method based on the geometric model of the facial region for facial expression recognition.
- Zhao et al. proposed a lightweight expression detection model that can solve the delay problem under natural conditions.
- Abate et al. proposed a neural network model for face attributes recognition based on transfer learning to group faces according to common facial features.

DISADVANTAGES OF EXISTING SYSTEM:

- Too many parameters
- Slowing down the training speed
- It is easy to over fitting problem
- When the network is deeper, it means that the parameter space is larger, and the optimization problem becomes more difficult. Therefore, simply increasing the depth of the network causes more training errors.

3.2. PROPOSEDSYSTEM:

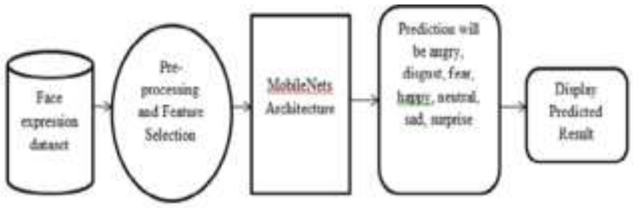
- We propose and design a convolution neural networks framework for identifying facial emotions in real-time and in large batches in this study.
- The categorization model is based on data from kaggle, and this dataset contains all forms of expressions. The dataset is also preprocessed before being used to develop the model. We can get all of the information in the dataset with the use of preprocessing. It assisted us in determining the quality of data and, on the other hand, preventing data redundancy. Preprocessing the data set improves both of our models, which is significant in our research.
- Following the training of our CNN Model Architecture, it was discovered that the model was successfully trained and provided individual training accuracy. Furthermore, the epochs were expanded to a specific limit, and it was discovered that the accuracy was improving as well as the production.

ADVANTAGES OF PROPOSED SYSTEM:

- The proposed system model will have a high anti-disturbance capability as well as a high recognition rate.
- In the final experimental test, we got good findings.
- We remove the interference components of the various faces in the image, considerably improving the effect of emotion recognition.
- Our proposed model has a 90% accuracy rate, which is the highest among existing system models.
- We found that the proposed method outperforms existing state-of-the-art methods in terms of accuracy, and we statistically examined it.

4.IMPLEMENTATION

4.1 SYSTEM ARCHITECTURE



4.2. MODULE'S:

Dataset:

In the first module, we developed the system to get the input dataset for the training and testing purpose. We given the data set in model folder

The dataset Link: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data.

Importing the necessary libraries:

https://doi.org/10.36893/FSSC.2024.V19I2.037

We will be using Python language for this. First we will import the necessary libraries such as keras for building the main model, sklearn for splitting the training and test data, PIL for converting the images into array of numbers and other libraries such as pandas, numpy, matplotlib and tensorflow.

Retrieving the images:

We will retrieve the images and their labels. Then resize the images to (224,224) as all images should have same size for recognition. Then convert the images into numpy array.

Splitting the dataset:

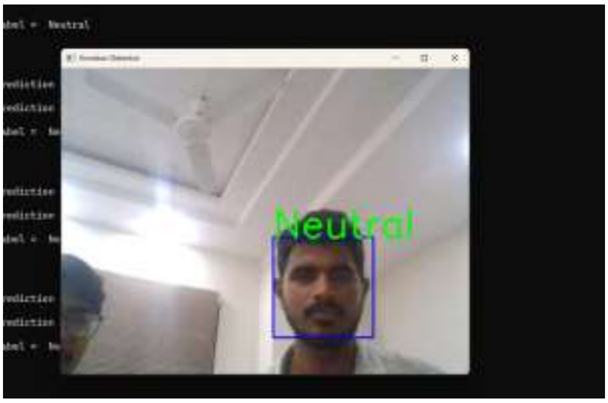
Split the dataset into train and test. 80% train data and 20% test data.

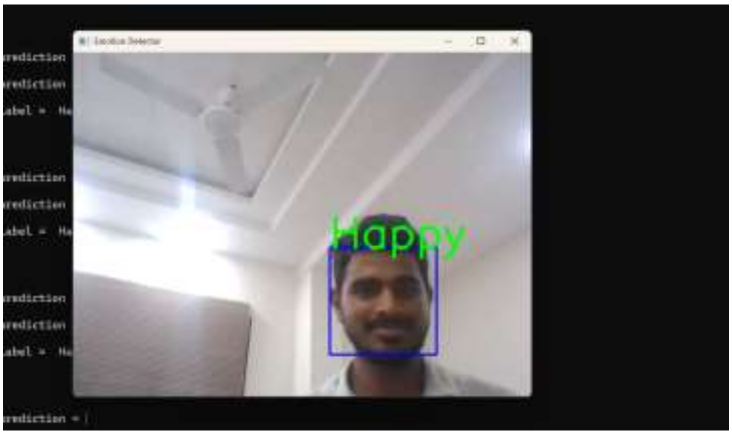
MobileNet | CNN model

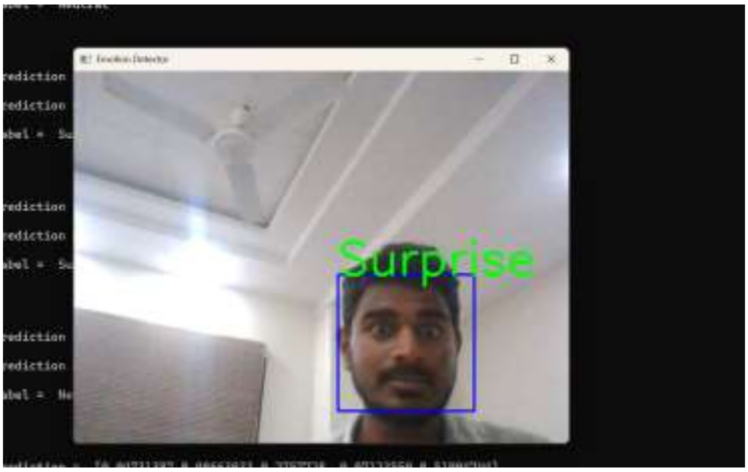
Architecture:

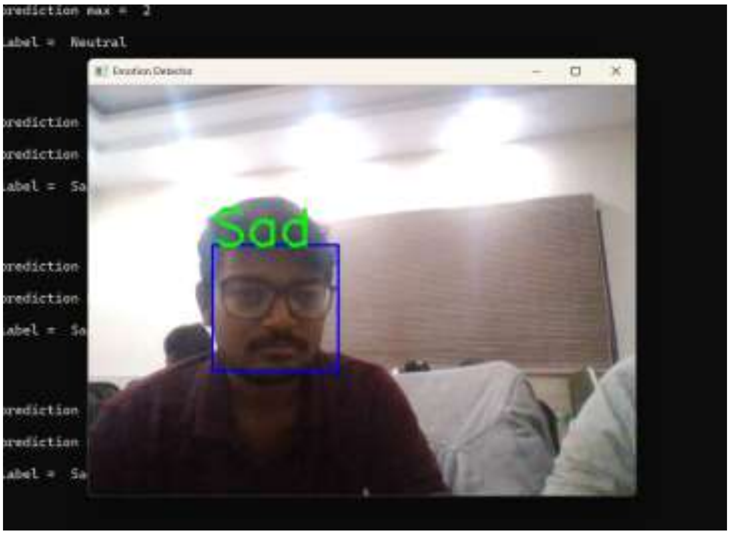
We shall be using Mobilenet as it is lightweight in its architecture. It uses depthwise separable convolutions which basically means it performs a single convolution on each colour channel rather than combining all three and flattening it. This has the effect of filtering the input channels. Or as the authors of the paper explain clearly: “ For MobileNets the depthwise convolution applies a single filter to each input channel. The pointwise convolution then applies a 1×1 convolution to combine the outputs the depthwise convolution. A standard convolution both filters and combines inputs into a new set of outputs in one step. The depthwise separable convolution splits this into two layers, a separate layer for filtering and a separate layer for combining. This factorization has the effect of drastically reducing computation and model size. ”

5.RESULT

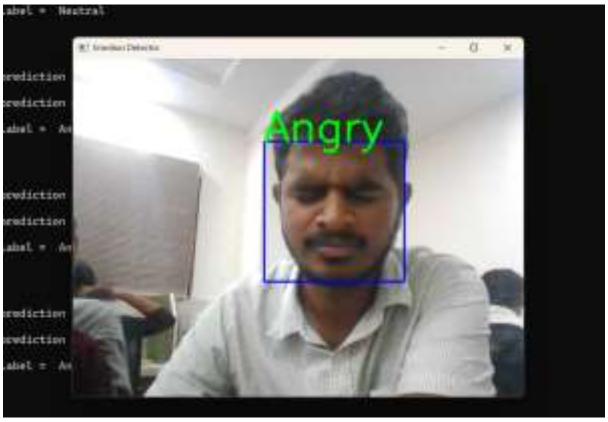








https://doi.org/10.36893/FSSC.2024.V19I2.037



CONCLUSION Our group proposes and constructs a lightweight convolutional neural network for face expression recognition in this study. By removing the fully connected layer from the convolutional layer, our network model minimises the number of parameters in the convolutional layer. Furthermore, our model has no discernible negative impact on detection and categorization. Our model achieves good detection results by detecting photographs outside of the dataset, demonstrating that the model developed in this study is suitable for facial expression multiclassification. In summary, we've developed a visual system that may be used to classify face expressions and decrease a huge number of parameters on devices with limited computing capability. After comparing our model to other current models, we found that ours is more accurate, and it has obtained good identification results in photos outside the dataset based on the experimental results.

FUTURE WORKS Although our model produced some findings, there may be a lot of noise in the facial expressions captured in real life, such as photographs with too bright or too dark illumination, blurred images, the majority of the face being blocked, and other circumstances that make recognition difficult. We must maintain our efforts in order to solve such a situation.

REFERENCES

1. P. Ekman and W. V. Friesen, ``Constants across cultures in the face and emotion.,'' J. Personality Social Psychol., vol. 17, no. 2, pp. 124-129, 1971.
2. I. J. Goodfellow et al., ``Challenges in representation learning: A report on three machine learning contests,'' Neural Netw., vol. 64, pp. 59-63, Apr. 2015.
3. S. Lawrence, C. L. Giles, A. Chung Tsoi, and A. D. Back, ``Face recognition: A convolutional neural-network approach,'' IEEE Trans. Neural Netw., vol. 8, no. 1, pp. 98113, Jan. 1997.
4. H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, ``Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,'' IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 12851298, May 2016.
5. T. Chang, G. Wen, Y. Hu, and J. Ma, ``Facial expression recognition based on complexity perception classication algorithm,'' 2018, arXiv:1803.00185. [Online]. Available: http://arxiv.org/abs/1803.00185
6. M.-I. Georgescu, R. T. Ionescu, and M. Popescu, ``Local learning with deep and handcrafted features for facial expression recognition,'' IEEE Access, vol. 7, pp. 6482764836, 2019.
7. C. Du and S. Gao, ``Image segmentation-based multi-focus image fusion through multi-scale convolutional neural network,'' IEEE Access, vol. 5, pp. 1575015761, 2017.
8. M. Z. Uddin, M. M. Hassan, A. Almogren, A. Alamri, M. Alrubaian, and G. Fortino, ``Facial expression recognition utilizing local direction-based robust features and deep belief network,'' IEEE Access, vol. 5, pp. 45254536, 2017.
9. M. Z. Uddin, W. Khaksar, and J. Torresen, ``Facial expression recognition using salient features and convolutional neural network,'' IEEE Access, vol. 5, pp. 2614626161, 2017.
10. D. Amodei et al., ``Deep speech 2: End-to-end speech recognition in english and mandarin,'' in Proc. Int. Conf. Mach. Learn., Jun. 2016, pp. 173182.

https://doi.org/10.36893/FSSC.2024.V19I2.037