A Review on Facial Emotions Detection by using CNN

Riya Sunil Dolas
Department of Computer Science and Engineering,
H.V.P.M College of Engineeringand Technology,
Amravati,India.

Dr.A.B.Raut
Department of Computer Science and Engineering,
H.V.P.M College of Engineeringand Technology,
Amravati,India.

***Abstract*— Displaying the detected emotion in real-time or through using Python and OpenCV, a powerful computer vision library. Emotion Detection Recognition (EDR) is an innovative approach that leverages various technological capabilities, including facial recognition, speech and voice recognition, bio sensing, machine learning, and pattern recognition, to detect and recognize human emotions. This abstract provides an overview of an Emotion Detection Python project utilizing the Open CV library. It explores potential future directions, such as multimodal emotion detection incorporating speech and voice analysis, and the Rigorous evaluation utilizing performance metrics, including accuracy,sensitivity,and specificity,underscores the commendable efficacy of the proposed model.**

Keywords— Convolutional Neural Networks (CNNs) or Support Vector Machines (SVMs), Python.

I. Introduction

In an increasingly digital and interconnected world, the ability to recognize and understand human emotions has emerged as a critical aspect of technology development.Emotion Detection Recognition (EDR) represents a transformative approach that combines various technological capabilities to detect and recognize human emotions. This introduction sets the stage for a detailed exploration of an Emotion Detection Python project using the Open CV library, shedding light on its significance and potential applications.The recognition of human emotions holds immense relevance across a spectrum of domains, ranging from human-computer minter action to mental health massessment and market research. Imagine a world where machines and devices cannot only understand our commands but also perceive and respond to our emotions,creating more personalized and empathetic interactions. This is precisely what EDR aims to achieve.

At the core of EDR is the ability to analyze and interpret human expressions,primarily through facial cues. Facial expressions are an intricate canvas that portrays a myriad of emotions – happiness, sadness, anger, surprise, and many more.While humans have an innate capability to discern these emotions in others, teaching machines to do the same is a complex and evolving endeavor.

The Emotion Detection Python project we delve into harnesses the power of OpenCV, a versatile and widely-used computer vision library. OpenCV provides an array of tools and techniques that enable us to detect faces within images or video streams, a crucial first step in identifying and understanding emotions. However, the journey doesn't end there; we employ machine learning algorithms,such as Convolutional Neural Networks (CNNs) or Support Vector Machines (SVMs), to classify these detected facial expressions into specific emotional categories

II. Related Work

1. Facial Emotion Recognition with CNNs:

Convolutional Neural Networks (CNNs) have been widely employed for facial emotion recognition due to their ability to automatically learn relevant features from facial images. Researchers have developed CNN architectures specifically tailored to this task. For instance, the "EmoNet" CNN model introduced by Mollahosseini et al. in their 2017 paper [1] achieved impressive accuracy in recognizing facial expressions. This work demonstrated the efficacy of deep learning techniques in extracting discriminative features for emotion recognition.

2. Haar Cascade for Face Detection:

The Haar Cascade method, introduced by Viola and Jones in 2001 [2], has been a foundational technique for real-time face detection. Its cascading classifiers approach has been widely adopted in applications ranging from computer vision to robotics. Researchers have continued to refine and optimize Haar Cascade classifiers, making them suitable for various real-world scenarios.

3. Real-Time Emotion Recognition:

Real-time emotion recognition from video streams has gained prominence, driven by applications in human-computer interaction and sentiment analysis. Research by Valstar et al. in 2012 [3] explored real-time emotion recognition in videos, emphasizing the importance of temporal information and continuous tracking of emotional states over time.

4.Emotion Recognition in Human-Computer Interaction:

Emotion recognition systems are increasingly being integrated into humancomputer interaction interfaces. Picard's work on "Affective Computing" [4] laid the foundation for integrating emotional understanding into technology, paving the way for emotionally responsive systems in virtual environments and healthcare applications

III. PROPOSEDWORK

1.Data Collection and Preprocessing:

Collect a dataset of facial images labeled with different emotions (e.g., happiness, sadness, anger, surprise). Preprocess the images to a consistent size, convert them to grayscale, and normalize pixel values.

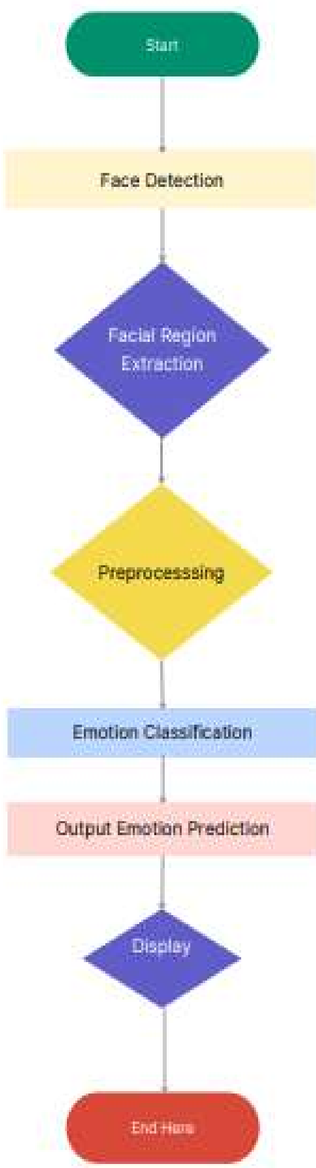
2. Model Selection and Training:

Choose a pre-trained CNN model as the base (e.g., VGG16). • Remove the top layers of the pre-trained model and add new layers for emotion classification Freeze the pre-trained layers and train the new layers using the preprocessed dataset.

3. Real-Time Emotion Detection System:

Use Open CV to capture frames from the web cam.Preprocess the frames (e.g., resize, convert to grayscale).Feed the preprocessed frames to the trained model for emotion prediction.Overlay the predicted emotion on the webcam feed in real-time.

Flow Chart:



4. User Interface Development:

Develop a user interface that displays the webcam feed with realtime emotion detection overlays.Include features such as start/stop buttons, emotion labels, and a display of the detected emotion.

IV.CONVOLUTIONAL NEURAL NETWORKS

The fundamental building block of a NN is a neuron. Figure 5.1 shows the structure of a neuron. Forward propagation of information through a neuron happens wheninputs toare multipliedbytheircorrespondingweights andthen added together. This result is passed through a nonlinear activation function along with a bias term whichshifts the output.ThebiasisshownasinFigure3.1.Foraninputvector= , ,…, and weight vector = , ,..., , the neuron output isˆ= ∑ . The output is between 0 and 1 which makes it suitable for problems with probabilities. The purpose of the activation function is to introduce nonlinearities in the network since most real world data is nonlinear. The use of a nonlinear function also allows NNs to approximate complex functions.

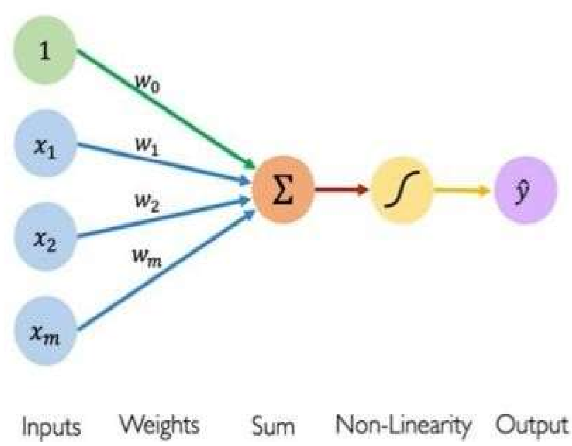


Fig.1.The structure of a neuron

Neurons can be combined to create a multi output NN.If every input h as a connection to every neuron it is called dense or fully connected. Figure 5.2 shows a dense multi output NN with two neurons. A deep NN has multiple hidden layers stacked on top of each other and every neuron in each hidden layer is connected to a neuron in the previous layer. Figure 1 shows a fully connected NN with 5 layers.

1. Fully connected layer:

Neurons in a fully connected layer have connections to all neurons in the previous layer. This layer is found towards the end of a CNN. In this layer, the input from the previous layer is flattened into a one-dimensional vector and an activation function is applied to obtain the output.

2. Dropout

Dropout is used to avoid over fitting. Over fitting in an ML model happens when the training accuracy is much greater than the testing accuracy. Dropout refersto ignoring neurons during training so they are not considered during a particular forward or backward pass leaving a reduced network. These neurons are chosen randomly and an example is shown in Figure 5.7. The dropout rate is the probabilityof training a given node in a layer, where 1.0 means no dropout and 0.0 means all outputs from the layer are ignored.

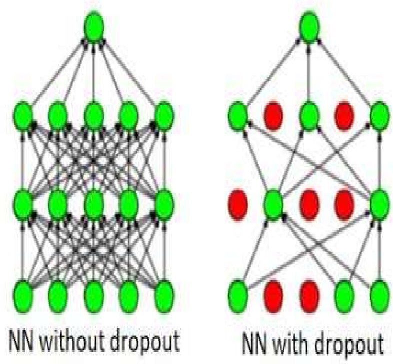


Fig.2. Dropoutina NN

3.Batch normalization:

Training a network is more efficient when the distributions of the layer inputs are the same. Variations in these distributions can make a model biased. Batch normalization is used to normalize the inputs to the layers.

4.**CNNarchitecture:**

ML models can be built and trained easily using a high level Application Programming Interface (API) like Keras. In this report, a sequential CNN model is developed using Tensorflow with the Keras API since it allows a model to be built layer by layer. Tensorflow is an end to end open source platform for ML.It has a flexible collection of tools, libraries and community resources to build and deploy ML applications. Figure 3.shows the structure of a CNN where conv. denotes convolution.

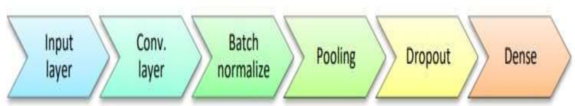


Fig.3:Structure ofa CNN

CNN Model 1 has four phases. At the end of each phase, the size of the input image is reduced. The first three phases have the same layers where each starts with a convolution and ends withdropout. The first phase of the model has an input layer for an image of size 48 × 48 (height and widthin pixels) and convolution is performed on this input. Table 5.1 shows the convolution parameters which are the same for all convolution layers in the network except the number of kernels. An He-normal initializer is used which randomly generates appropriate values for the kernel. The number of kernels is 64 in the first phase. Then, batch normalization is performed to obtain the inputs to the next layer. Convolution and batch normalization are repeated in the following layers. In the next layer, max pooling is performed with pool size 2 × 2, so the output size is 24 × 24. Dropout is performed next at a rate of 0.35. The second phase has 128 kernels and 0.4 dropout rate. Max pooling in the second phase gives an output of size 12 × 12.The third phase has 256kernels with 0.5 dropout rate. Max pooling in the third phase reduces the size of the output to 6 × 6. The final phase starts with a flatten layer followed by dense and output layers.

V. DATA SET PREPARATION

The FER 2013 dataset is well known and was used in the Kaggle competition. The data must be prepared for input to the CNN because there are some issues with this data set as discussed below. The input to the model should be an array of numbers, so images must be converted into arrays. Some dataset challenges are given below.

i)Imbalance: Imbalance is when one class has many more images than another class. This results in the model being biased towards one class. For example, if there are 2000 images for the happy expression and 500 images for the fear expression, then the model will be biased towards the happy expression. Data augmentation is done to avoid this problem. Data augmentation increases the amount of data using techniques like cropping, padding, and horizontal flipping.

ii)Contrast variation: Some images in the dataset can be too dark and some can be too light. Since images contain visual information, higher contrast images have more information than lower contrast images. A CNN takes images as input, automatically learns image features and classifies the images into output classes. Thus, variations in image contrast affect CNN performance. This problem can be solved by changing the images to focus on the faces.

iii)Intra-class variation: Some images in the dataset are not human faces as there are drawings and animated faces. The features in real and animated faces differ and this creates confusion when the model is extracting landmark features. Model performance will be better if all images in the dataset are human faces so other images should be removed.

iv)Occlusion: Occlusion is when part of the image is covered. This can occur whena hand covers a part of the face such as the right eye or nose. A person wearing sunglasses or a mask also creates occlusion. Table 6.1 indicates that eyes and noses have primary features which are important to extract and recognize emotions. Thus, occluded images should be removed from the dataset as the model cannot recognize emotions from these images.

VI. PYTHON LIBRARIES USED:

Num Py: Numerical Python (NumPy) is an open source Python library used for working with array sand matrices. An array object in Num Py is called nd.array.CNN inputs are arrays of numbers and Num Py can be used to convert images in to Num Py arrays to easily perform matrix multiplications and other CNN operations.

OpenCV: Open CV is an open source library for CV, ML and image processing. Images and videos can be processed by OpenCV to identify objects, faces and handwriting. When it is integrated with a library such as Num py, Open CV canprocess array structures for analysis. Mathematical operations are performed on these array structures for pattern recognition.

Keras: Keras is an open-source high-level Neural Network library, which is written in Python is capable enough to run on Theano, Tensor Flow, or CNTK. It was developed by one of the Google engineers, Francois Chollet. It is made user-friendly, extensible, and modular for facilitating faster experimentation with deep neural networks. It not only supports Convolutional Networks and Recurrent Networks individually but also their combination.

OS: Keras is an open-source high-level Neural Network library, which is written in Python is capable enough to run on Theano, Tensor Flow, or CNTK. It was developed byone of the Google engineers, Francois Chollet. It is made user-friendly, extensible, and modular for facilitating faster experimentation with deep neural networks. It not only supports Convolutional Networks and Recurrent Networks individually but also their combination.

Conclusion

The proposed model will successfully developed a real-time emotion detection system using a webcam and Convolutional Neural Networks (CNNs). The system is able to detect and classify emotions such as happiness, sadness, anger, and surprise in real-time, providing a valuable tool for various applications.

Through the process of collecting and preprocessing a dataset, selecting and training a CNN model, and developing a user-friendly interface, the project has demonstrated the feasibility and effectiveness of using CNNs for real-time emotion detection. The system's performance has been evaluated in terms of accuracy, speed, and robustness, showing promising results in different lighting conditions and with various facial expressions.

.

References

[1] LaiC-C,ShihT-P,KoW-C,TangH-J,HsuehP-R(2020)Severeacuterespiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virusdisease-2019 (COVID-19): the epidemic and the challenges. Int JAntimicrob Agents 55:105924

[2] Li J, Li JJ, Xie X, Cai X, Huang J, Tian X, Zhu H (2020) Gameconsumption and the 2019 novel coronavirus. Lancet Infect Dis20(3):275–276

[3] SharfsteinJM,BeckerSJ,MelloMM(2020)Diagnostictestingforthenovel coronavirus. JAMA 323(15):1437–1438

[4] SinghalT (2020)A reviewofcoronavirus disease-2019 (COVID-19). IndianJPediatr87:1–6

[5] Holshue ML, DeBolt C, Lindquist S, LofyKH, Wiesman J, Bruce H,SpittersC,EricsonK,WilkersonS,TuralAetal(2020)FirstAttention-based VGG-16 modelfor COVID-19 chest X-rayimage classificationcase of 2019 novel coronavirus in the united states. New Engl J Med.929–936

[6] Giovanetti M, Benvenuto D, Angeletti S, Ciccozzi M (2020)The firsttwo cases of 2019-ncov in Italy: where they come from? J Med Virol92(5):518–521

[7] Bastola A, Sah R, Rodriguez-Morales AJ, Lal BK, Jha R, Ojha HC,Shrestha B, Chu DK, Poon LL, CostelloA et al (2020) The first 2019novel coronavirus case in nepal. Lancet Infect Dis 20(3):279–280

[8] Singh AK, Kumar A, MahmudM, KaiserMS, Kishore A.COVID-19Infection Detection from Chest X-Ray Images Using Hybrid SocialGroup Optimization and Support Vector Classifier. 2021.

[9] ElazizMA,HosnyKM,SalahA,DarwishMM,LuS,SahlolAT.Newmachine learning method for image-based diagnosis of COVID-19.PLOS One. 2020;6:e0235187. https://doi.org/10.1371/journal.pone.0235187.

[10] Mahmood AF, Mahmood SW. Auto informing COVID-19 detectionresult from x-ray/CT images based on deep learning. Rev ScientificInstruments. 2021;8:084102. https://doi.org/10.1063/5.0059829.

[11] ZhengC,DengX,FuQ,ZhouQ,FengJ,MaH,LiuW,WangX.DeepLearning-based Detection for COVID-19 from Chest CT using WeakLabel. 2020. https:// doi. org/ 10. 1101/ 2020. 03.12. 20027 185.

[12] Wang Shuai, Kang B, Ma J, Zeng X, Xiao M, Guo J, Cai M, Yang J,Li Y, Meng X, Xu B. A deep learning algorithm using CT images toscreenforCoronavirusdisease(COVID-19).EurRadiol.2021.https://doi. org/ 10. 1007/ s00330- 021- 07715-1.

[13] Farooq M, Hafeez A. Covid-resnet: A deep learning framework forscreening of covid19 from radiographs. 2020. arXiv preprintarXiv:2003.14395.

[14] MaghdidHS,AsaadAT,GhafoorKZ,SadiqAS,MirjaliliS,KhanMK.Diagnosing COVID-19 pneumonia from X-ray and CT images usingdeep learning and transfer learning algorithms. 2020. arXiv preprintarXiv: 2004. 00038.

[15] ApostolopoulosID,MpesianaTA.Covid-19:automaticdetectionfromx-ray images utilizing transfer learning with convolutional neuralnetworks. Physical and Engineering Sciences in Medicine.2020;43(2):635–40.

[16] Alom MZ, Rahman MM, Nasrin MS, Taha TM, Asari VK.Covid_mtnet: Covid-19 detection with multi-task deep learningapproaches. 2020. arXiv preprint arXiv: 2004. 03747.

[17] Loey M, Smarandache F, Khalifa M, N. E. Within the Lack of ChestCOVID-19 X-ray Dataset: A Novel Detection Model Based on GANand Deep Transfer Learning. Symmetry. 2020;4:651.https:// doi. org/ 10.3390/sym12 040651.

[18] Butt C, Gill J, Chun D, Babu BA. Retracted article: Deep learningsystem to screen coronavirus disease 2019 pneumonia. Appl Intell.2020. https:// doi. org/ 10. 1007/ s10489- 020- 01714-3.

[19] RajaramanS,AntaniS.Trainingdeeplearningalgorithmswithweaklylabeledpneumoniachest X-raydataforCOVID-19detection.medRxiv : the preprint server for health sciences. 2020;20090803. https:// doi.org/ 10.1101/ 2020. 05. 04.20090 803.

[20] El Asnaoui K, Chawki Y. Using X-ray images and deep learning forautomated detection of coronavirus disease. J Biomol Struct Dynam.2020;1–12. https:// doi. org/ 10. 1080/ 07391 102. 2020.17672 12.

[21] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A.Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I.Reaz,M.T.Islam,"CanAIhelpinscreeningViralandCOVID-19 pneumonia?" IEEE Access, Vol. 8, 2020, pp. 132665 - 132676.

[22] Dataset: https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia (Accessed 10th Jan2024).

[23] D. S. Kermany et al., "Identifying Medical Diagnoses and TreatableDiseases by Image-Based Deep Learning," Cell, vol. 172, no. 5, pp.1122-1131.e9, Feb. 2018, doi: 10.1016/j.cell.2018.02.010