International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 20, Issue 7 (July, 2024), PP. 188-201

Facial Emotion Recognition

Zriyanka[1], Siya[2]and Muskan Meena[3]

Abstract
Facial expressions are among the most common ways that people communicate.Inthispaper,weimplementnumerousdeeplearningmodelsto go in-depth for recognizing facial expressions (FER). Facial expressionrecognition is crucial to people's daily lives and jobs. Facial expression-basedautomaticemotion recognition is anintriguing study area thathasbeenpresentedandusedinnumberoffields,including safety,health,andhuman-machineinteractions.Researchers in this disciplineare interested in creating methods to decipher, encode, and extract thesecharacteristicsfromfacialexpressionsordertoimprovecomputerprediction.Duetoitsbuilt-infeatureextractionprocess from images,DeepNeuralNetworks,particularlytheConvolutionalNeuralNetwork(CNN),areemployedextensivelyinFER. Onlyafew layers have beenusedinnumberofinitiativesthathavebeenpublishedon CNN toaddress FER issues. Standard shallow CNNs, on the other hand, havealimitedabilitytoextract features thatcan extract emotional informationfromhigh-resolutionphotosusingsimple learning algorithms. Thefact that most current approaches only takeintoaccount frontal photosand overlook profile views from different angles out of convenience is anoteworthyflaw,astheseviewsarecrucialforaworkableFERsystem.

Date of Submission: 03-07-2024 Date of Acceptance: 16-07-2024

---

1 Introduction

Applicationsthatimprovehumanlifeandworkaredevelopedto create information technology. Artificial intelligence technology, or simply"Artificial intelligence," is the current trend in thedevelopment of modern information technology (AI). The intro-duction of new forms of interaction, such as the use of buttons,screens, touch screens, voice commands, security control systems,human-computer interfaces, and other applications, changed theway modern technological developments were made. The firststep of human recognition face is also one of these new forms ofinteraction. The study's citation, "musa2017analisis," suggests anEEG signal filtering procedure that modifies data on human emotional features utilizing wavelet packet decomposition (WPD) andempirical mode decomposition (EMD) methodologies. The usageof a face expression recognition system is one of her methods forfiguringoutuserreactions.[1]

As a kind of nonverbal communication, facial expressions are theoutcomeoffacialmovementsorexpressionsthatdisplaytheposition of the human facial muscles. They are crucial for conveyingone's emotions as a form of sentiment, intention, or desire. It is atool.Additionally,thereareotherpeople'sopinions. [2]
Humansareconditionedtoreadthefeelingsofothers;infact,attheageofjust14months,infantscanalreadydiscernbetweenhappy and sad emotions. But can technology access human emotions more quickly and efficiently than humans? We developed adeep-learningneuralnetworkthatenablescomputerstodeducedetailsabout humanemotional states in order to react.To putitanother way, we offer them the ability to see what we see. Simpleclassesforhumanfacialexpressionsencompasshappy,unhappy,surprised,scared,indignant,disgusted,andimpartial.Uniqueagenciesoffacialmusclegroupsarebroughtaboutwhenoneexperiencesfacialemotion.Thosechangingtinybutcomplicatedindicatorsinone'sfacialexpressionsregularlydisplayawealthofstatisticsapproximatelyhowoneisfeeling.onecanwithoutdifficultyandinexpensivelyexaminetheoutcomesthatcontentmaterialandserviceshaveonaudiencesandcustomersbyusing

faceemotionrecognition.Shopsmightmakeuseoftheseindicators,asaninstance,togaugecustomerinterest.Havingextraknow-how approximately the emotional situation of patients whilstreceiving treatment can assist healthcare carriers to serve patientsbetter. So one can continuously supply preferred cloth, and amusementmanufacturerscansongtargetmarketparticipationinthecourseofactivities.
The initiative point of this project was to find a data set thatcan be used to work with. FER 2013 data set was chosen for thepurposeoflearningandexperimenting.FERstandsforFacialemotion recognition data set

which includes 48 x 48pixel gray scaleimages of faces from seven different classes given in lablesfrom 0to 6 (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise,6=Neutral). The data set is distributed into testing and trainingsets. Hetrainingsetconsistsof28,709examplesandthepublictestset consists of 3,589 examples. Work began with the exploratorydata analysis which was done in order to identify any error, get betterunderstandingofdata,detectoutliers,andtounderstanddataset variables and the relationship among them.Then proceeded itwith the feature extraction process where SIFT (Scale-InvariantFeatureTransform)[3]isanalgorithmincomputervisiontodetectand describe local features of an image. After this process, a fewof ML models were taken into account like K-nearest neighbour,decision tree, naive bayes model, logistic regression, ans supportvector machine and DL models Like CNN.[4] The objective of thisresearch is to determine which emotions a person's face transmitsfrom a gray-scale image of their face. Using real-time facial featuresreputation with synthetic intelligence within side the device canenhance accuracy. This allows direct reputation of facial expressions. The accuracy for each emotion will serve as our evaluationmetric, and a confusion matrix will be included to show certainemotionsaremoreaccuratelyidentifiedthanothers.

2 RelatedWork

Artificialintelligenceandpsychologicalhumanemotionperceptionare two distinct fields of research that are crucial to automaticemotionrecognition(AI).Aperson'semotionalstatecanbeascertainedusingtheverbalandnon-verbalinformationgatheredbythevarioussensors,includingchangesinfacialexpression,

speech tones, and physiological reactions. According to Mehrabian'sresearchfrom1967,55%ofemotionalinformationisvisual,38% is vocal, and 7% is linguistic. The majority of scholars are particularlyinterestedinthismodalitysincechangesinthefaceduringcommunication are the first indicators of the emotional state that isbeingsent.Inordertogetabetterclassification,itisachallenginganddelicatetasktoextractfeaturesfromonefacet oanother.

Ekman and Freisen, who were some of the first researchers tobecome interested in facial expression [5], developed the FACS(FacialActionCodingSystem)in1978.Thehumanfacewasdivided into 46 AUs, each of which was connected to one or morefacialmuscles.

Compared to certain other modalities of statistics created by,the automated FER is the one that researchers have researched themost.philippetal[6],howeveritisadifficultundertakingbecauseeveryone expresses emotion differently. Some of the challenges andissuesthatoneshouldnotdisregardinthissectorincludethevariation in head postures, luminosity, age, species, and backdrop, aswellastheproblemofocclusiongeneratedbysunglasses,scarves,skindiseases,etc.

Geometrical and texture features like the Gabor wave, localbinary patterns, facial action units, and development practitionerspatterns are used in traditional methods for extracting facialfeatures. Deep learning has lately shown to be a very successfulandeffectiveapproachasaresultoftheoutcomesprovidedbyits architectures, such as the convolutional neural network CNNand the recurrent neural RNN, that enable the automatic featuresareextractedandcategorization.Thisiswhatledresearcherstobeginapplyingthismethodtorecognisehum anemotions.Researchers have made several attempts to create deep neural net-works. network topologies, which in this regard yield quite pleasingresults.

Deep features for automated face expression and emotion detection produce state-of-the-art outcomes, according to recent studiesby Li and Deng , who also won the Emotion Recognition in theWildChallenge[7](EmotiW)andtheFacialExpressionRecognition and Analysis Challenge (FERA) . Therefore, in this work, deepneuralnetworksareusedtodirectlylearnthefeaturesneededto

describe the collected information rather than creating them fromscratch.

The primary methods utilized to recognize emotions at the timeare physiological signal detection, scale approach, and lab simulation.Eachofthesemethodshasthedrawbackofrelyingondataonpeople'semotionsthatarebeingmonitoredinreal-time.Atechniqueforilluminationaugmentationwithadaptiveattenuationquantificationwasputforthby Boubennaand Leetoreducethechallenginglightinginfluenceonfacialphotography[8].KangandYoondevelopedamulti-structure[9],variable-parameterexpressions processes can be classified that fixes the vanishing gradientproblemcausedbyhavingtoomanynetworklayerswhilepreservingthesequentialfeaturesoffacialexpressions.

Fanetal.extractedthesequentialfeaturesfromthefacialexpressionimages[10]afterperformingblock-basedpreprocess-ing on them. They then used the emotional index to measure thecorrelations between the various facial expressions and emotions,leading to continuous description of distinct facial expressions andeffectiveemotionrecognition.

Themicro-facialexpressionswererecognisedbyBalouchianand Foroosh [11] using an end-to-end deep neural network. Theyemployedthefocuslossfunctiontominimisetheimbalancebetweenboththevariousclassesofmicro-expressiondataandchangedtheparametersofthepretrainedmodelusingtransferlearningtomakeupforthelimitedsizeof thesampleset.

3 DatasetAndDataAnalysis

3.1 SourceoftheDataset

1. TheFER2013(FacialExpressionRecognition 2013) data set has beentaken from the research paper of Standford University"Facial ExpressionRecognitionwithDeepLearning"[12].
2. Pierre-Luc Carrier and Aaron Courville presented the 2013 Facial Expression Recognition data set (FER-2013) at the International Conference onMachineLearning(ICML)in2013.[13]
3. Each face in this data-set has been classified according to various moodcategories. The FER2013 data-set is not balanced, though, as it includespictures of seven different facial expressions, including angry (4,953), dis-gusted (547), fear (5,121), happy (8,989), sad (6,077), surprise (4,002), andneutral(6,198)

Table1:AttributeTable

Attributes	Labels	TestImages	TrainingImages	TotalImages
Angry	0	958	3995	4953
Disgust	1	111	436	547
Fear	2	1024	4097	5121
Happy	3	1774	7215	8989
Sad	4	1274	4830	6077
Surprise	5	831	3171	4002
Neutral	6	1233	4965	6198
Total	7	7178	28709	35887

3.2 SizeoftheDataset

TheFER2013(FacialExpressionRecognition2013)datasetincludes pictures of people and categories that describe their emotions.The48x48pixelgrayscaleimagesinthedatasetdepictsevenvarious emotions, including rage, disgust, fear, happiness, sadness,surprise, and neutrality. There are 28709 training examples, 3589examples in the public testing set, and 3589 examples in the privatetestingsetinthedataset.(referTable1).

3.3 Exploratorydataanalysis

Observationsmade:

1. AnalysismadefromFig.1:
   (a) TypeofthevariableisMulticlass.
   (b) Happinessisthemostfrequentlydetectedemotion,andtherearealsothemostimagesofitinthedataset.
   (c) Disgustisleastidentifieddata.
2. Theplotfortrainingsetwhichshowstheimagescontaininghappyemotionsaremaximum(morethan7000)andtheimagesofdisgustemotionsareveryless(lessthan1000)(Fig.2)
3. The plot for test set which shows the images containing happy emotions aremaximum (nearly 1750) and the images of disgust emotions are very less(lessthan250)(Fig.3)

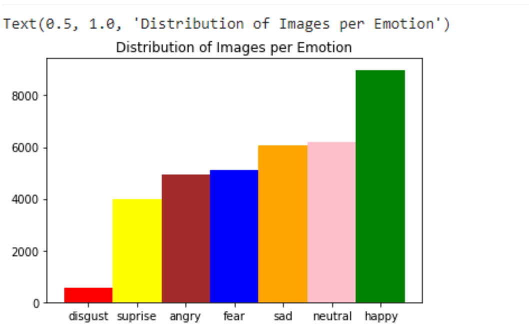

Fig.1:Thebargraphdisplaysdatainstancesforeachclass.

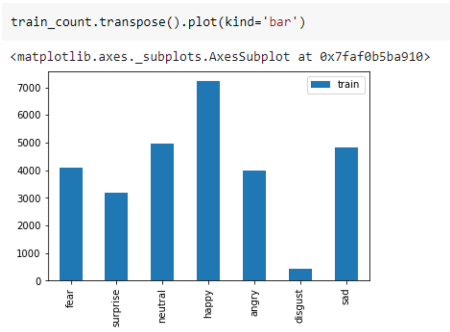

Fig.2: PlotfortrainingsetforFER(2013)Dataset

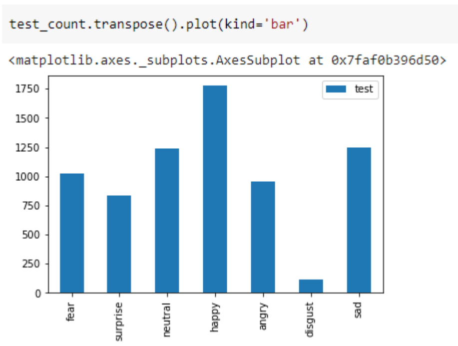

Fig.3: PlotfortestsetinFER(2013)Dataset

3.4 FeatureExtraction

Thedimensionalityreductionmethod,whichdividesandcondensesastartingsetofrawdataintosmaller,easier-to-managegroupings,

includesfeatureextraction.Asaresult,processingwillbesimpler.Features are elements or patterns that assist identify an object inan image. For instance, a square has four corners and four edges,which are known as the square's characteristics and aid in humanrecognition of the shape. Features include things like ridges, corners,edges,andpointsofinterest.

Observationsmade:

1. Edge detection is a method of image processing that locates the edges ofobjectsinpictures.Itoperatesbylookingforchangesin brightness. Infields including image processing, machine learning, and machine vision,edge detection is utilized for image segmentation and data extraction. TheimagesareshowninFig.4andFig.5

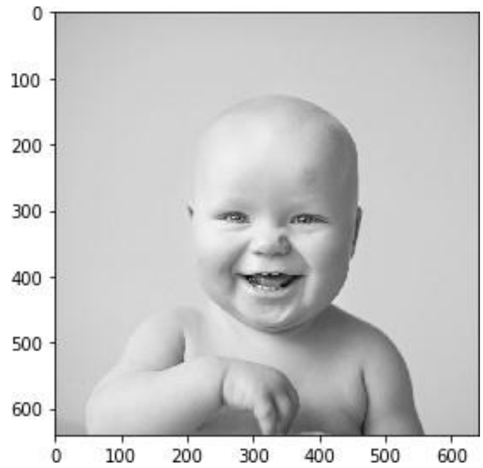

Fig.4:Grayscaleimageof'happy'emotion

Fig.5:Grayscaleimageof'happy'emotionafterapplyingedgealgorithm

2. SIFT: In image classification problems, the scale-invariant feature transform(SIFT) is a popular feature extraction technique. Local features in an image,also referred to as"key points" in the image, can be found using SIFT. Asthey are scale- and rotation-invariant, these fundamental concepts can beappliedtoanumberofcomputervisiontasks,includingpicturematching,objectdetection,andscenedetecti on.

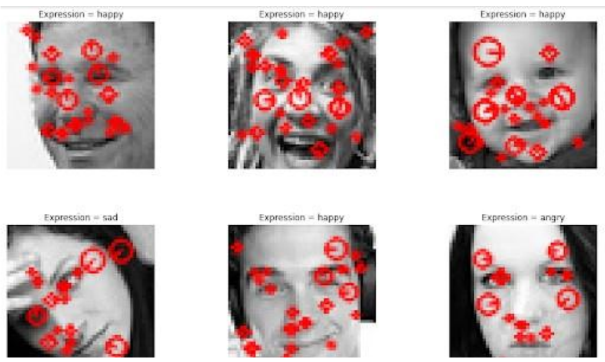

Fig.6:ImagesshowingkeypointsafterapplyingSIFTalgorithm

4 Experiments Design

4.1 MLMODELS

Modelsofmachinelearningarecomputeralgorithmsthathavebeentrainedtoidentifytrendsinnewdataandpredictconsequences. These a mathematical function that gets input from a modelandrequestsaremadeintheformofinputdata,andthedataisprocessedtoprovideestimationsbeforeprovidinganooutput.Thesearethefirstmodelsorasetofdataisusedtotrain,afterwhichanalgorithmisprovidedtoindividualssotheycanreviewtheinformation,lookforpatterns,andunderstandbasedonthedata. These simulations can be used to predict the unforeseen afterbeingtrainedonadataset.

In the domain of machine learning, categorization refers to the process of choosing the type or class of an item from a limited set ofalternatives. The outcome of classification is always a categoricalvariable. Identifying whether an email is spam or not is an example of a typical binary classification task. Several key models forclassificationproblemswillnowbelisted.

1. k-Nearestneighbouralgorithm.
2. Decisiontree
3. LogisticRegression
4. Supportvectormachine.
5. NaiveBayesModel.

4.1.1 k-NEARESTNEIGHBOUR(KNN)Model

The k-Nearest Neighbors method for pattern recognition algorithm(also known as k-NN) is a non-parametric technique that utilizedfor regression and classification. The output of k-NN classificationis a class. member. Objects are categorised by a majority of theirthe object being a member of the class most similar to its typicalof its k closest neighbours (k is a positive number) k1 (integer).The item is just assigned to if k = 1 the category of that particularclosestneighbour.[14]

```
print(confusion_matrix(y_val,y_pred))
```

```
[[147   9  46 118  47  18  82]
 [  6  20   5   8   6   5   6]
 [ 69   9 133 115  65  39  66]
 [ 82   8  84 413  92  47 169]
 [104  15  77 160 139  21 137]
 [ 38   6  60  63  22 173  53]
 [ 56  11  76 161  78  35 190]]
```

Fig.7:confusionmatrixforKNN

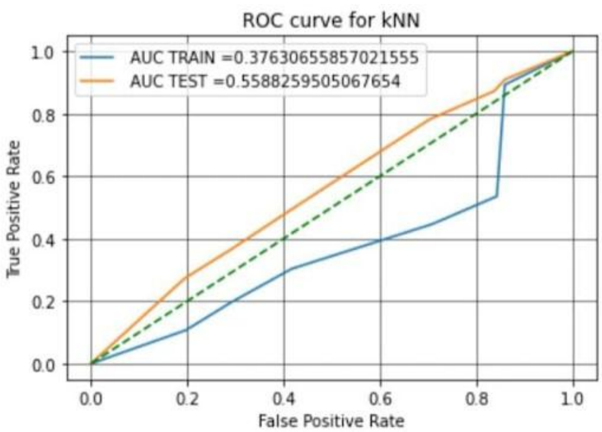

Fig.8:ROCcurveforKNNModel

4.1.2 DecisionTreeModel

Decisiontreesandotherguidedmachinelearningtechniquesrequireongoingdatasegmentationbasedonacertainparameter. The two elements that can be utilized to explain the tree aredecisionnodesandleaves.[15]

```
Confusion matrix

[[ 37   0   0 369  57  28   0]
 [  4   0   0  43   5   3   0]
 [ 39   0   0 394  53  42   0]
 [ 35   0   0 750  71  23   0]
 [ 34   0   0 442  90  28   0]
 [ 22   0   0 272  43  79   0]
 [ 17   0   0 508  72  29   0]]
```

Fig.9:confusionmatrixforDecisionTreeModel

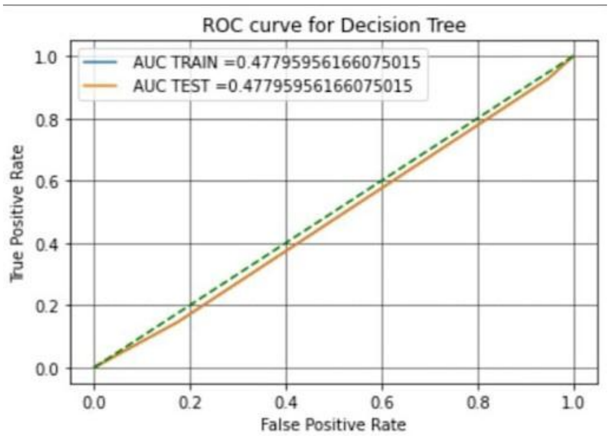

Fig.10:ROCcurveforDecisionTreeModel

4.1.3 LogisticRegressionModel

Using prior observations from a data set, a statistical analysis technique called logistic regression predicts a binary outcome, includingsuchyesorno.Thetechniqueoflogisticregressionhasriseninsignificanceinthefieldofmachinelearning.Itenablesmachinelearning algorithms to categorize incoming input based on previousdata. By enabling data sets to be processed for analysis by puttingthem into precisely defined buckets throughout the extract, trans-form, and load process, logistic regression can also be used in datapreparationoperations.[16]

```
[[ 40   1  78 145   3 197  27]
 [  5   0  12  17   0  17   4]
 [  9   1  98 114   4 272  30]
 [ 24   0  49 536   5 230  35]
 [ 32   1  93 164   8 232  64]
 [  2   0  30  56   2 317   9]
 [ 20   0  80 157   8 264  97]]
```

Fig.11:confusionmatrixforlinearregression

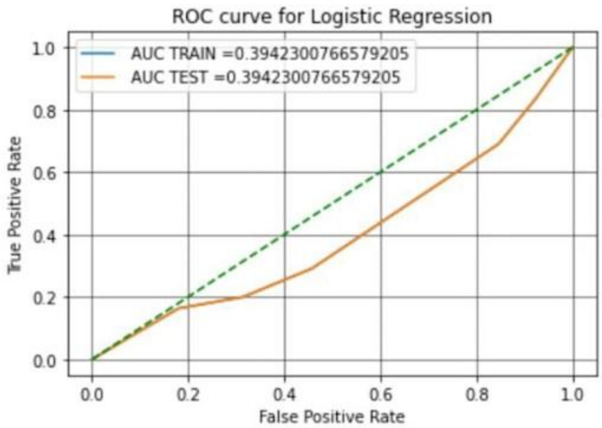

Fig.12:ROCcurveforLogisticRegressionModel

4.1.4 SupportVectorMachine(SVM)Model

Support Vector Machine is a supervised computer vision approachthat can be applied to classification or regression problems. How-ever,classification-relatedissuesaretheapplicationthatismostfrequentlyused.TheSVMalgorithmconvertseachdatapointintoa point inside an n-dimensional dimension, where n represents thenumberoffeatures,soeachfeature'svalueisassignedtoaspecificplace.Then,classificationiscarriedoutbyidentifyingthehyper-planethatcorrectlydividesthetwogroups.[17]

```
        Confusion Matrix
[[109   0  46 142  96  20  78]
 [  7   3   3  24   6   2  10]
 [ 39   0 122 131 111  50  75]
 [ 25   0  30 636 101  18  69]
 [ 35   0  50 136 241  11 121]
 [ 13   0  40  68  31 219  45]
 [ 28   0  25 182 101  15 275]]
```

Fig.13:confusionmatrixforSupportVectorMachine

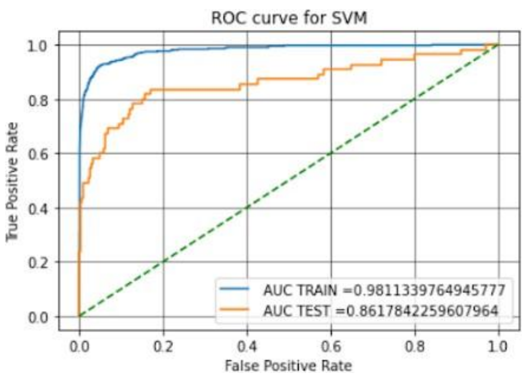

Fig.14:ROCcurveforSVM

4.1.5 NaiveBayesModel

TheNaiveBayesclassificationalgorithmisaprobabilisticclassifier.Probability models with strong independence assumptions serveas its cornerstones. The independence factors frequently have noimpactonreality.Asaresult,theyareseenasbeingnaive.[18]

```
         Confusion Matrix
[[ 94  58  94  70  24  63  88]
 [  8  14   9  12   0   6   6]
 [ 73  59 101  75  19 109  92]
 [111  71 133 254  36 103 171]
 [ 93  77  99  97  46  79 103]
 [ 36  31  80  51  16 145  57]
 [ 51  76  95 110  24 103 167]]
```

Fig.15:confusionmatrixforNaiveBayesModel

Classification Report

	precision	recall	f1-score	support
0	0.20	0.19	0.20	491
1	0.04	0.25	0.06	55
2	0.17	0.19	0.18	528
3	0.38	0.29	0.33	879
4	0.28	0.08	0.12	594
5	0.24	0.35	0.28	416
6	0.24	0.27	0.25	626
accuracy			0.23	3589
macro avg	0.22	0.23	0.20	3589
weighted avg	0.26	0.23	0.23	3589

Fig.16:ClassificationReportofNaiveBayesModel

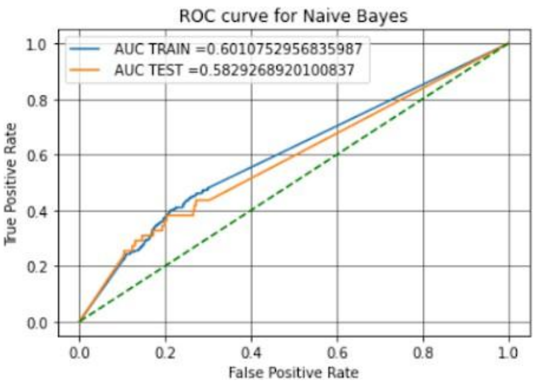

Fig.17:ROCcurveforNaiveBayesModel

4.2 DLMODEL

4.2.1 ConvolutionalNeuralNetwork(CNN)

1. In deep learning, convolutional neural networks (CNN/ConvNet) are indeedafamilyofdeepneuralnetworksthat areoftenusedtoanalysevisualinput.
2. CNN is a specific kind of deep learning network design that is used forprocessingpixelinputandperformingimagerecognition.
3. Classification of images and segmentation, detection, video analysis, naturallanguageprocessing,andspeechrecognitionareafewofCNN’sintriguingapplicationareas.DeepCNN’sgreatlearningcapacityisexplainedbyits extensive usage of extracting features phases, which may automaticallydiscoverrepresentationsfromdata.[19]
4. Evaluation:
   (a) TestLoss:0.89
   (b) TestAccuracy:0.67
5. Lossplotandaccuracyplotisgiveninfig.18.
6. Confusionmatrixisgiveninfig.19.

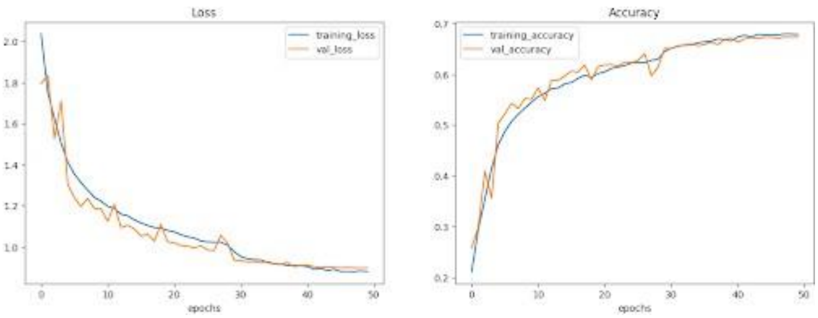

Fig.18:LossplotansaccuracyplotforCNNmodel

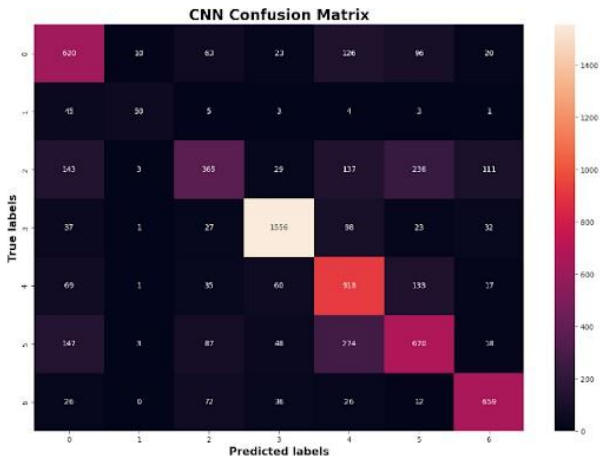

Fig.19:confusionmatrix

5FineTuning

Thetechniqueoffine-tuninginvolvesadjustingmodelparameterstobetterfitaspecificobservation.Applyingorutilizingtransferlearninginvolvesfine-tuning.Inparticular,fine-tuningistheprocess of altering or optimising a model that has already been trainedto carry out a specific task in order to carry out a second, relatedtask.

1. TestLoss:0.90
2. TestAccuracy:0.70

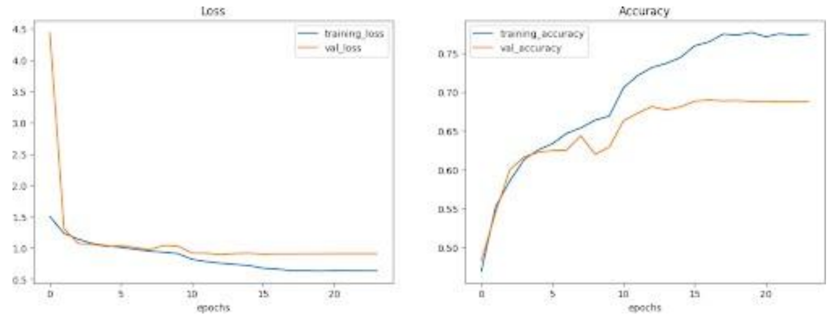

Fig.20:LossandAccuracyplotofResNet50V2model

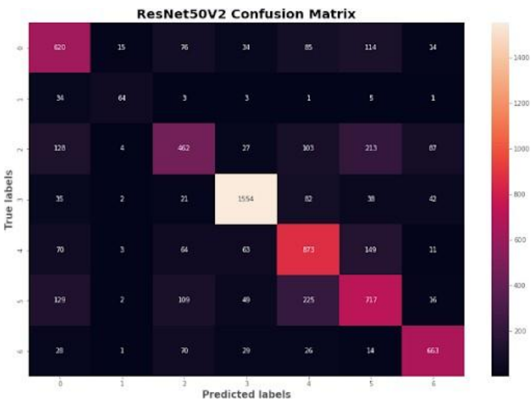

Fig.21:confusionmatrix

6 DataAugmentation

The process of creating new data points or providing significantlymodified versions of already existing data is known as data augmentation.Thisentailseitherutilizingmachinelearningmodelsto create additional data points in the subspace of the originaldataoraddingtothedatasetbymakingsmalladjustmentstothedata.Thisservesasregularisationandlowers over-fittingwhenmachine learning models are trained. This and oversampling indatagatheringarecloselyrelated.

7 TransferLearning

Transferlearningisindeedadeepvisionresearchproblemthatis concerned with storing knowledge obtained while resolving oneproblem and using it to solve another problem that is closelyrelated. For instance, while attempting to identify trucks, knowledgeobtainedthroughlearningtoidentifycarscanbeused.

Accuracy:0.30

F-1Score:0.30

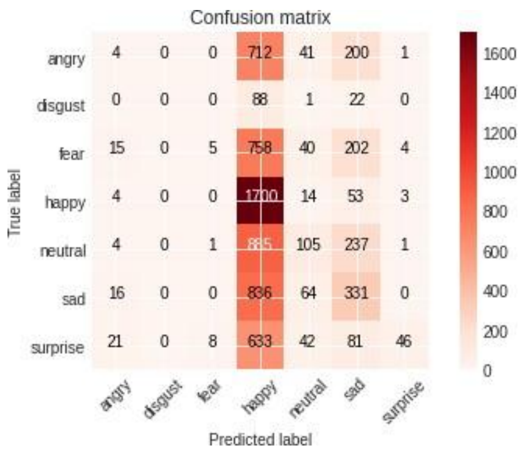

Fig.22:Confusionmatrix

```
Confusion matrix, without normalization
[[   4    0    0  712   41  200    1]
 [   0    0    0   88    1   22    0]
 [  15    0    5  758   40  202    4]
 [   4    0    0 1700   14   53    3]
 [   4    0    1  885  105  237    1]
 [  16    0    0  836   64  331    0]
 [  21    0    8  633   42   81   46]]
```

Fig.23:Confusionmatrix

8 EnsembleLearning

Ensemble methods use multiple learning algorithms to achieve betterpredictiveperformancethancanbeachievedwithindividuallearningalgorithmsalone.Theprocessofcreatingmultiplemodelssuchas B.theprocessofbuildingnumerousmodels,suchas B.Classifiersorexperts,thenstrategicallycombiningthemtoaddresscertainissuesinartificialintelligence.Ensemblelearningisprimarily useful for improving (classification, prediction, functionapproximation,etc.).[20]

8.1 HybridModel

Ensemblelearningisawidemetastrategytomachinelearningthattriestoenhancepredictiveperformancebycombiningtheresultsfromseveralmodels.Althoughyoucancreatean

apparentlyendlessamountofensemblestosolveyourpredictivemodellingissue.[21]

AccuracyofLogisticRegressionModelwas0.35.AccuracyofDecisionTreeModelwas0.27.
AccuracyofSupportVectorMachinewas0.41.Accuracyofk-NNModelwas0.30.
AccuracyofNaiveBayesModelwas0.21.AccuracyofEnsembledModelwas0.55.

9 Result

1. TheObservationsforMLmodelsaregiveninthetable2.
2. The Accuracy for DL and other operations are given in the table3.
3. Comparingresultsfrombaseresearchpaper:

(a) In the Fig. 24the training accuracy is reaching over 0.8 andtestingaccuracyismarkingnearly0.7,thisisthecaseofover-fitting.

(b) In the Fig. 25the training and testing accuracy are overlap-pingandbothreachingto0.7.

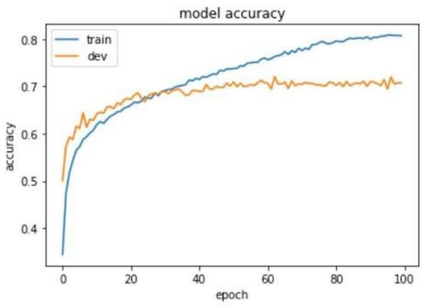

Fig.24:Accuracyplotfor“FacialExpressionRecognitionwithDeepLearning”:StanfordUniversity-CS230DeepLearning(2020)

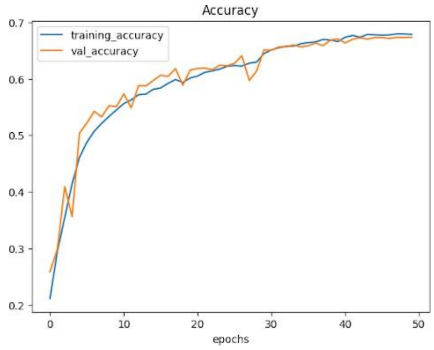

Fig.25:Accuracyplotofourmodel

Table2:MLModels

S.No.	Models	Accuracy	Recall	Precision	F-1Score
1	NaiveBayes	0.21	0.23	0.26	0.23
2	DecisionTree	0.27	0.27	0.17	0.17
3	LogisticRegression	0.35	0.37	0.33	0.34
4	KNN	0.30	0.34	0.34	0.34
5	SVM	0.41	0.45	0.45	0.43

Table3:Models

S.No.	Models	Accuracy
1	CNN	0.67
2	FineTuning	0.70
3	TransferLearning	0.30
4	EnsembleLearning	0.55

10 Limitations,ConclusionsandFutureWork

10.1 Conclusion

This project had two goals, one was to achieve higher accuracyand to apply models to the real world. We explored several modelsincluding k-NN, decision tree, naive Bayes, support vector machine,logistic regression, and CNN. We have also performed fine tuning,dataaugmentation,transferlearning,andensemblelearningonthe data set. After ensembling the several ML models results we haveachieved the accuracy to 0.55. The model is recognising the imagesand predicting their emotions on the basis of the training of themodel.

10.2 Limitations

The FER 2013 data set has only few range of emotions, as animprovementwewilltrytoimplementthesameprocesswithdataset having wide range of emotions. The images in the data set areonly grey scaled, we will also try to implement to work on coloredimages for the better result to come. As an improvement in thisproject, we will try to implement the same process on the real timevideos.

10.3 FutureWork

Tomakeourmodelmoreusableandaccessible,weintendtoofferitawebsiteinterface.Weaimtousefaciallandmarkrecog nitionandalignment,selectiveattentionCNNs,andretrainthenetworkby obstructing facial features unrelated to emotion recognition tofurtherenhancetheaccuracyofourmodels.

Additionally,wethinkpipelinemodels,whichfeedtypicallyincorrect emotion pairs (such neutral and depressed) to secondarynetworkshavinghigheraccuracylevelsbetweenthoseparticularemotions,havealotofroomforimprovement .Weintendtoincorporatecurrentpsychologicalresearch,particularlythearousal-valence emotional framework, as well as multi-label classification to more effectively handle images with numerous potentialemotion labels in order to further adapt their models to the realworld.

References

[1] Saputra,A.,Wirayuda,T.A.B.:Pengenalanekspresiwajahmenggunakanlocalbinarypattern(lbp).eProceedingsofEngineering**2**(3)(2015)

[2] Wijaya, I.G.P.S., Firdaus, A.A., Dwitama, A.P.J., Mustiari, M.: Penge-nalan ekspresi wajah menggunakan dct dan lda untuk aplikasi pemutarmusik (moodsic). Jurnal Teknologi Informasi dan Ilmu Komputer **5**(5),559–566(2018)

[3] Sushama, M., Rajinikanth, E.: Face recognition using drlbp and sift fea-ture extraction. In: 2018 International Conference on Communication andSignalProcessing(ICCSP),pp.994–999(2018).IEEE

[4] Khan, A.R.: Facial emotion recognition using conventional machine learn-inganddeeplearningmethods:Currentachievements,analysisandremainingchallenges.Information**13**(6)(2022)

[5] Ekman, P., Freisen, W.V., Ancoli, S.: Facial signs of emotional experience.Journalofpersonalityandsocialpsychology**39**(6),1125(1980)

[6] Zhao, S., Precioso, F., Cord, M., Philipp-Foliguet, S.: Actor retrieval sys-tem based on kernels on bags of bags. In: 2008 16th European SignalProcessingConference,pp.1–5(2008).IEEE

[7] Liu, Z., Li, S., Deng, W.: Boosting-poof: boosting part based one vs onefeature for facial expression recognition in the wild. In: 2017 12th IEEEInternationalConferenceonAutomaticFace&GestureRecognition(FG2017),pp.967–972(2017).IEEE

[8] RaiJain,P.,MKQuadri,S.,Lalit,M.:Recenttrendsin artificial intel-ligence for emotion detection using facial image analysis. In: 2021 Thir-teenth International Conference on Contemporary Computing (IC3-2021),pp.18–36(2021)

[9] Oh, J., Nam, S., Park, K.-S., Yoon, S., Kang, I., Jeong, J.: A study ofencapsulationstructurefortftreliabilityintopemissionoleddisplay.In:Proceedings of the International Display Workshops, vol. 2, pp. 910–913(2019).InternationalDisplayWorkshops

[10] Hasani, B., Mahoor, M.H.: Facial expression recognition using enhanceddeep3dconvolutionalneuralnetworks.In:ProceedingsoftheIEEECon-ference on Computer Vision and Pattern Recognition Workshops, pp.30–40(2017)

[11] Li, P., Zhou, Z., Liu, Q., Sun, X., Chen, F., Xue, W.: Machine learning-basedemotionalrecognitioninsurveillancevideoimagesinthecontextofsmartcitysafety.TraitementduSignal**38**(2)(2021)

[12] Khanzada, A., Bai, C., Celepcikay, F.T.: Facial expression recognitionwithdeeplearning.arXivpreprintarXiv:2004.11823(2020)

[13] Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Ham-ner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.-H., *et al.*: Challengesin representation learning: A report on three machine learning con-tests. In: International Conference on Neural Information Processing, pp.117–124(2013).Springer

[14] Yadav, K.S., Singha, J.: Facial expression recognition using modified viola-john’s algorithm and knn classifier. Multimedia Tools and Applications**79**(19),13089–13107(2020)

[15] Salmam, F.Z., Madani, A., Kissi, M.: Facial expression recognition usingdecision trees. In: 2016 13th International Conference on ComputerGraphics,ImagingandVisualization(CGiV),pp.125–130(2016).IEEE

[16] Lv, Y., Feng, Z., Xu, C.: Facial expression recognition via deep learn-ing. In: 2014 International Conference on Smart Computing, pp. 303–308(2014).IEEE

[17] Mazloom, M., Kasaei, S., Neissi, H.A.: Construction and application ofsvm model and wavelet-pca for face recognition. In: 2009 Second Inter-nationalConferenceonComputerandElectricalEngineering,vol.1,pp.391–398(2009).IEEE

[18] Azmin, S., Dhar, K.: Emotion detection from bangla text corpus usingnaive bayes classifier. In: 2019 4th International Conference on Electri-cal Information and Communication Technology (EICT), pp. 1–5 (2019).IEEE

[19] Akhand, M., Roy, S., Siddique, N., Kamal, M.A.S., Shimamura, T.: Facialemotion recognition using transfer learning in the deep cnn. Electronics**10**(9),1036(2021)

[20] Zhang, J., Yin, Z., Chen, P., Nichele, S.: Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review.InformationFusion**59**,103–126(2020)

[21] Alreshidi, A., Ullah, M.: Facial emotion recognition using hybrid features.In:Informatics,vol.7,p.6(2020).MDPI