A SMART STORY TELLING MODEL WITH EMOTION-BASED ENUNCIATION

M.Pallavi[1], P.H.S.S.Meghana[2], S.J.Manasa[3], P.Chandrakala[4], Y.Gayatri [5]
Assistant Professor[1,] Student[1,2,3,4,5 ,], [1,2,3,4,5] Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India

ABSTRACT
The study of emotion is significant in many disciplines, including biomedical engineering, psychology,neurology,healthandeducation.Thiscanbejudgedbydetectingtheemotionsofthepeoplefromtheirfacialexpressions. By analyzing the children mood, the narrative may be shown on the screen and delivered by avoice.consequently,itcanhelpchildrenwhoareexposedto readingearlytendtoperform betteracademically in various subjects, including language arts and mathematics. Listening to stories encourageschildren to sit still and focus, improving their attention span. Exposure to a variety of words, sentencestructures, and vocabulary enhances language skills. Children learn to articulate their thoughts and expressthemselves more effectively. Deep learning techniques are one of the best techniques for analyzing theimageswell.Amongthedeeplearningtechniquesconvolutional neuralnetworks(CNN)areusedfordetecting the student emotions and python libraries to deliver a voice. In this system the student emotion isdetectedasanyofthesevenfacialexpressionssuchasAngry,Disgust,Fear, Happy, Sad,SurpriseandNeutral.Storiesprovideanenjoyableescapefromreality,allowingindividualstoimmersethemselvesindifferent worlds, characters, and plots.

INTRODUCTION
A Smart Storyteller is a technology that uses deep learning to create stories that adapt to the emotionsand behavioral responses of the people. This technology can predict future emotional responses andgeneratestorytellingthatkeepsthe useron adynamiccurve ofinteractions.The technological components that the system uses voice delivery systems , deep learning algorithmsandcamerasto observeandunderstand facialexpressions.Thegathereddataisthenprocessedbythesystem,allowingittomakeassumptionsorpredictions about the user's emotional state. The software can be applied for learning and entertainment purposesaswell astherapeuticprograms.Itiscapableofmakingassumptionsbasedonpast andreal-timecollecteddata,influencingboththestorylineandtheemotionalresponsesoftheuser.We use a computer language called Python and deep learning techniques to explore how theseemotion-filledstoriescanmakekidsbetterattalking,payingattention,andlearningingeneral.

LITERATURESURVEY
1. Guillaume-Benjamin-Amand Duchenne de Boulogne was a French neurologist in the 19th century, who wasinterested in Physiognomy and wanted to understand how human face muscles work to produce facialexpressions,ashebelievedthattheseweredirectlylinkedtoahuman'ssoul.Todothis,heusedelectricprobesto trigger muscle contractions, and then took pictures, using newly developed camera technology, of

hissubjects' faces showing the distorted expressions he was able to create. In 1862, he published his research andthe photographs of the triggered facial expressions in the book "The Mechanism of Human Physiognomy "[4]. An example from his publication can be seen in Fig. 1, showing photographs of his subjects displaying adifferent expression oneachside oftheirfaces.

2. Later,in1872,CharlesDarwinusedthisworkasanimportantresource[5]forhisbookcalled"TheExpressionof EmotioninManandAnimals"inwhichhefocusedonthegeneticsofbehavior.However,inrecentDuchennede Boulogne'sbookhasbeenrediscoveredbyphotographersasunquestionably,EkmanisthemostIn 2016, Pramerdorfer and Kampel obtained state-of-theart, which is 75.2% accuracy on the FER2013, usingConvolutional Neural Networks (CNNs) [6]. The authors used an ensemble of CNNs using VGG, Inception,andResNetwithdepthsof10,16,and33,withparametersof1.8m,1.6m,and5.3m,respectively.Theauthorsusedthefaceimagesasgiveninthedataset,andforilluminationcorrection,theyusedhistogramequalization.They performed horizontal mirroring for training data augmentation and randomly cropped images to the sizeof 48 x 48 pixels. They also trained the architecture for up to 300 epochs and used stochastic gradient descenttooptimizethecross-entropyloss,withamomentumvalue0.9.Theotherparameterswerefixed,likelearningratewith 0.1,batch size with 128,and weightdecaywith 0.0001influentialresearchersinthefieldofemotionalexpressionofthiscentury,asdiscussedintheintroduction.

3. Zhang et al. [7] used a Siamese Network to introduce a method for understanding social relation behaviorsfrom images and achieved a test accuracy of 75.1% on the challenging Kaggle facial expression dataset. Theauthorsusedmultipledatasets,withvariouslabels,toincreasethetrainingdata;theyalsointroducedafeatureextractionmethodandpatch-basedregistration,aswellasworkingonfeatureintegrationviaearlyfusion.Kimet al. [8] proposed an ensemble of CNNs and demonstrated that during training and testing it is advantageousto use both registered and unregistered forms of given face images. The authors achieved a test accuracy of73.73% on the FER2013 dataset. They also conducted Intraface for a conventional 2-D alignment, which ispublicly available for landmark detector, and performed illumination normalization. To avoid the registrationerror,theyperformedregistrationselectively,basedontheresultsoffaciallandmarkdetection.

EXISTINGSYSTEM

The existing system is detecting the emotions based on facial expressions typically employ deep learningmodels, particularly convolutional neural networks (CNNs), to analyze facial features and patterns indicativeof different emotions. The process involves the extraction of facial landmarks, such as eye and mouthmovements,andthemappingofthesefeaturestospecificemotional categorieslikehappiness,sadness,anger,surprise,fear,disgust,andneutrality.Trainingdatasetsarecrucialformodel development,yettheyposechallengesrelatedtopotential biasesandculturalvariations.Theimplementationoffacialemotionrecognition often involves preprocessing steps, such as face detection and alignment, to enhance the accuracyofemotionclassification.

Disadvantages

1. Everyday facial expressions are often ambiguous and can convey a range of emotions simultaneouslyornone atall, and are notused in ourdailylife.
2. Facialemotionrecognitioncanbesensitivetoreal-worldconditions,suchaschangesinlighting, environmentalfactors,orocclusions(partialfacevisibility),impactingthesystem'saccuracy.
3. Collectingandprocessingfacialdataforemotionrecognitionraisesignificantprivacyconcerns,particularlywhenimplementedin publicspaces orwithoutexplicit user consent.
4. Implementingrobustcybersecuritymeasurestoprotectcollecteddatafrombreachesorunauthorizedaccessis anethicalresponsibilityto preventpotentialharmormisuse.

PROPOSEDSCHEME

A Smart Storyteller is a system that employs deep learning to generate stories that adapt to people's emotionsand behavioral reactions. Current technologies leverage deep learning techniques, with convolutional neuralnetworks(CNN)emergingasapopularchoiceforimageanalysistasks,includingemotiondetectionandstory teller. These systems aim to enhance the educational experience by tailoring content based on theemotionalstatesofusers,promotingengagementandpersonalizedlearning.Additionally,voicedelivery systems, employing libraries such as those in the Python programming language, have been integrated toprovideamultimodal learningenvironment wherenarrativesarenotonlydisplayedbutalsospokenaloud.

ADVANTAGES

1. Storytellingissoeffectivebecause storiescreateconnectionsbetweenpeople,andbetweenpeopleand ideas.
2. Storytellinghelpsteammembersgettoknowandunderstandeachotherbetter.Thisfostersanenvironment ofsharing,mutuality, and trust.
3. Storiesmaketheabstract concreteandproviderisk-freeavenuestoprocess andintegratechange.

SystemBlockDiagram

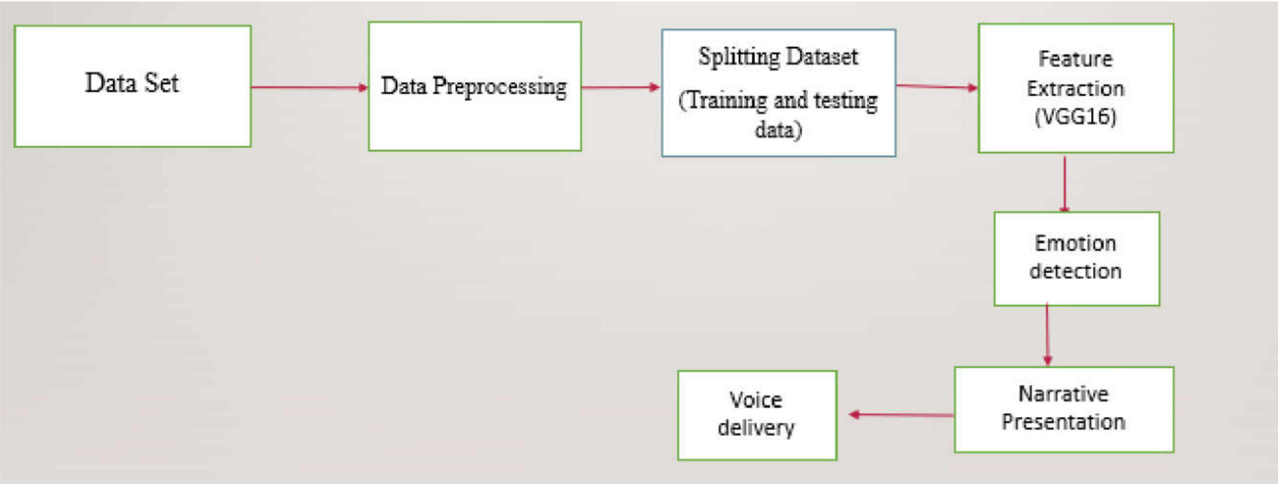

Fig1: SystemArchitecture

RESULTS

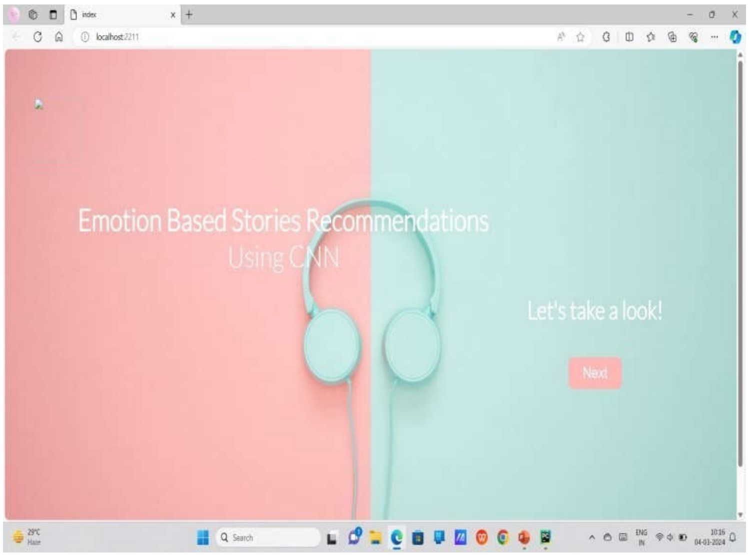

Fig:2 A Login Page

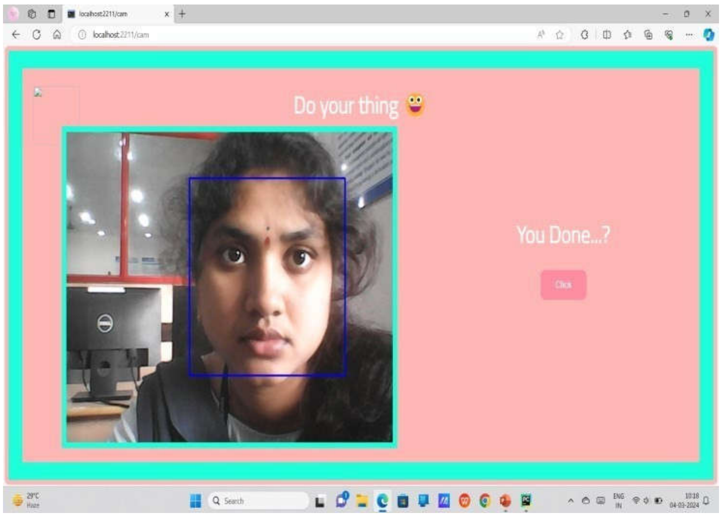

Fig:3Face DetectionPage captures emotion by facial expression

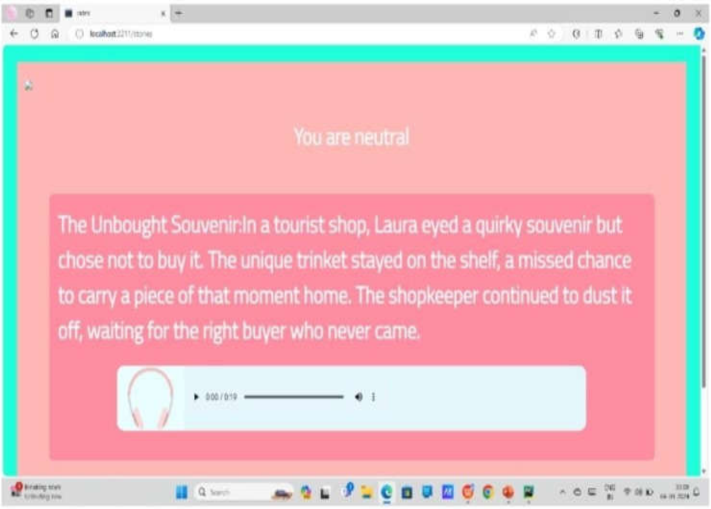

Fig:4OutputPage that displays story based on facial expression

CONCLUSION

The paper explores methodologies to detect emotions from children's fairytale sentences and express themappropriately in the text-to-speech conversion. To further enhance this, a Question-Answering system isimplemented, which caters to any doubts or queries the listener has regarding the story. A valid dataset withtext,emotion,and audiohasbeencreated according tothe paper'sscope totestandtrainthemodelaccurately.Thepredictionofemotionsfrom each sentenceisaccuratetomoveto thenextstep, convertingthetextintoaudiostripswithemotionsattached.Theaudiostripsfromeachsentencearemergedtoprovid eacontiguous and seamless experience to the user. The Question and Answering module implemented providesaninteractive experience tothe user.

REFERENCES

[1] Application:FacialExpressionRecognition.In:MachineLearninginComputerVision.ComputationalImaging andVision,vol29.Springer,Dordrecht(2005).

[2] https://en.wikipedia.org/wiki/Paul_Ekman,Jan2020.

[3] DeepFER:ASurveyhttps://arxiv.org/pdf/1804.08348.pdf,Jan2020.

[4] Duchenne, G.-B. (1862), Mécanisme de la physionomie humaine, ou analyse électro-physiologique de sesdifférentsmodesdel'expression.Paris:Archivesgénéralesdemédecine,P.Asselin;vol.1,p.29-47,152-174.[5]Darwin,CharlesRobert(1872),Theexpressionoftheemotionsinmanandanimals.London:JohnMurray London.

[6] C.PramerdorferandM.Kampel,“Facialexpressionrecognitionusingconvolutionalneuralnetworks:Stateoftheart,”arXiv preprintarXiv:1612.02903,2016.

[7] Z. Zhang, P. Luo, C.-C. Loy, and X. Tang, “Learning Social Relation Traits from Face Images,” inProc.IEEE Int.ConferenceonComputerVision (ICCV),2015, pp.3631–3639.

[8] B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, “Fusing aligned and non-aligned faceinformationforautomaticaffectrecognitioninthewild:Adeeplearningapproach,”inProceedingsoftheIEEEConferenceonComputerVisionandPatternRecognitionWorkshops,2016,pp.48–57.

[9] Raghuvanshi, A., & Choksi, V. (2016). Facial Expression Recognition with Convolutional NeuralNetworks.

[10] "Challenges in Representation Learning: A report on three machine learning contests." I Goodfellow,D Erhan,PLCarrier,ACourville,MMirza,BHamner,WCukierski,YTang, DHLee,YZhou,CRamaiah,FFeng,RLi,XWang,DAthanasakis,JShawe-Taylor,MMilakov,JPark,RIonescu,MPopescu,CGrozea,JBergstra,J Xie,LRomaszko,BXu, ZChuang,andY.Bengio.arXiv2013.