SIGN LANGUAGE RECOGNITION

K.SaiCharitha[1], A.Maneesha[2], Sd.karishma[3], J.Ramya[4], [1,2,3,4]IV B.Tech, Department of Computer Science and Engineerig, Vignan's Nirula Institute of Technology & Science for Women, Peda Palakaluru, Guntur-522009, Andhra Pradesh, India.

Mr.K.V.S.S.RamaKrishna (Asst. Professor), Computer Science and Engineering Department, VNITSW, Guntur, Andhra Pradesh, India.

ABSTRACT

Sign Language Recognition (SLR) targets on interpreting the sign language into text or speech, so as tofacilitate the communication between deaf-mute people and ordinary people. This task has broad socialimpact, but is still very challenging due to the complexity and large variations in hand actions. Existingmethods for SLR use hand-crafted features to describe sign language motion and build classification modelsbased on those features. However, it is difficult to design reliable features to adapt to the large variations ofhand gestures. To approach this problem, we propose a novelconvolutional neural network (CNN) whichextracts discriminative spatial-temporal features from raw video stream automatically without any priorknowledge,avoidingdesigningfeatures.Toboosttheperformance,multi-channelsofvideostreams,including color information, depth clue, and body joint positions, are used as input to theCNN in order tointegrate color, depth and trajectory information. We validate the proposed model on a real dataset collectedwith Microsoft Kinect and demonstrate its effectiveness over the traditional approaches based on hand-craftedfeatures.

INTRODUCTION

Signlanguage,asoneof the mostwidely usedcommunicationmeansforhearing-impairedpeople, is expressed by variations of hand-shapes,body movement, and even facial expression. Sinceitisdifficulttocollaborativelyexploittheinformationfromhand-shapesandbodymovement trajectory, sign language recognition isstillaverychallengingtask.Thispaperproposesaneffectiverecognitionmodeltotranslatesignlanguage into text or speech in order to help thehearing impaired communicate with normal peoplethroughsign language.

Technically speaking, the main challenge of signlanguage recognition lies in developing descriptorsto express hand-shapes and motion trajectory. Inparticular,hand-shapedescriptioninvolvestracking hand regions in video stream, segmentinghand-shape images from complex background ineachframeandgesturesrecognitionproblems.Motiontrajectoryisalsorelatedtotrackingofthe key points and curve matching. Although lots ofresearch works have been conducted on these twoissues for now, it is still hard to obtain satisfyingresultforSLRduetothevariationandocclusionof hands and body joints. Besides, it is a nontrivialissuetointegratethehand-shapefeaturesandtrajectoryfeaturestogether.Toaddressthesedifficulties,wedevelopaCNNstonaturallyintegrate hand-shapes,trajectoryofactionandfacial expression. Instead of using commonly usedcolor images as input to networks like [1, 2], wetake color images, depth images and body skeletonimagessimultaneouslyasinputwhichareallprovided byMicrosoft Kinect.

Kinect is a motion sensor which can provide colorstreamanddepthstream.WiththepublicWindowsSDK, the body jointlocationscan beobtained in real-time as shown in Fig.1. Therefore,we choose Kinect as capture device to record signwordsdataset.Thechangeofcoloranddepthin pixel level are useful information to discriminatedifferent sign actions. And the variation of bodyjointsintimedimensioncandepictthetrajectoryofsignactions.Usingmultipletypesofvisualsourcesas input leadsCNNs paying attentiontothe change not only in color, but also in depth andtrajectory.Itisworthmentioningthatwecanavoid the difficulty of tracking hands, segmentinghands from background and designing descriptorsforhandsbecauseCNNshavethecapability tolearn features automatically from raw data withoutanypriorknowledge[3].

CNNshavebeenappliedinvideostreamclassificationrecentlyyears.Apotentialconcernof CNNs is time consuming. It costs several weeksor months to train a CNNs with million-scale inmillion videos. Fortunately, it is still possible toachievereal-timeefficiency,withthehelpofCUDAforparallelprocessing.WeproposetoapplyCNNstoextractspatialandtemporalfeaturesfromvideostreamforSignLanguageRecognition (SLR). Existing methods for SLR usehand-craftedfeaturestodescribesignlanguagemotionandbuildclassificationmodelbasedonthesefeatures.Incontrast,CNNscancapturemotioninformationfromrawvideodataautomatically,avoidingdesigningfeatures.Wedevelop a CNNs taking multiple types of data asinput. This architecture integrates color, depth andtrajectory information by performing convolutionandsubsamplingonadjacentvideoframes.Experimental results demonstrate that 3D CNNscansignificantlyoutperformGaussianmixturemodel with Hidden Markov model (GMM-HMM)baselinesonsomesignwordsrecordedbyourselves.

LITERATURESURVEY

Technically speaking, the main challenge of signlanguage recognition lies in developing descriptorsto express hand-shapes and motion trajectory. Inparticular,hand-shapedescriptioninvolvestracking hand regions in video stream, segmentinghand-shape images from complex background ineachframeandgesturesrecognitionproblems. Motion trajectory is also related to tracking of thekey points and curve matching. Although lots ofresearch works have been conducted on these twoissues for now, it is still hard to obtain satisfyingresultforSLRduetothevariationandocclusionof hands and body joints. Besides, it is a nontrivialissuetointegratethehand-shapefeaturesandtrajectoryfeaturestogether.Toaddressthesedifficulties,wedevelopaCNNstonaturallyintegrate hand-shapes,trajectoryofactionandfacial expression. Instead of using commonly usedcolor images as input to networks like [1, 2], wetake color images, depth images and body skeletonimagessimultaneouslyasinputwhichareallprovided byMicrosoft Kinect.

Kinect is a motion sensor which can provide colorstreamanddepthstream.WiththepublicWindowsSDK, the body jointlocationscan beobtained in real-time as shown in Fig.1. Therefore,we choose Kinect as capture device to record signwords dataset. The change of color and depth inpixel level are useful information to discriminatedifferent sign actions. And the variation of bodyjointsintimedimensioncandepictthetrajectoryofsignactions.Usingmultipletypesofvisualsourcesas input leadsCNNs paying attentiontothe change not only in color, but also in depth andtrajectory.Itisworthmentioningthatwecanavoid the difficulty of tracking hands, segmentinghands from background and designing descriptorsforhandsbecauseCNNshavethecapabilitytolearn features automatically from raw data withoutanypriorknowledge[3].

CNNshavebeenappliedinvideostreamclassificationrecentlyyears.Apotentialconcernof CNNs is time consuming. It costs several weeksor months to train a CNNs with million-scale inmillion videos. Fortunately, it is still possible toachievereal-timeefficiency,withthehelpofCUDAforparallelprocessing.WeproposetoapplyCNNstoextractspatialandtemp oralfeaturesfromvideostreamforSignLanguageRecognition (SLR). Existing methods for SLR usehand-craftedfeaturestodescribesignlanguagemotionandbuildclassificationmodelbasedonthesefeatures.Incontrast, CNNscancapture motioninformationfromrawvideodataautomatically,avoidingdesigningfeatures.Wedevelop a CNNs taking multiple types of data asinput. This architecture integrates color, depth andtrajectory information by performing convolutionandsubsamplingonadjacentvideoframes.Experimental results demonstrate that 3D CNNscansignificantlyoutperformGaussianmixturemodel with Hidden Markov model (GMM-HMM)baselinesonsomesignwordsrecordedbyourselves.

PROPOSEDMETHODOLOGY

Toapproachthisproblem,weproposeanovelconvolutionalneuralnetwork(CNN)whichextractsdiscriminativespatial-temporalfeaturesfrom raw video stream automatically without anyprior knowledge, avoiding designing features. Toboosttheperformance,multi-channelsofvideostreams, including color information, depth clue,and body joint positions, are used as input to theCNNinordertointegratecolor,depthandtrajectory information. We validate the proposedmodel on a real dataset collected with MicrosoftKinect and demonstrate its effectiveness over thetraditionalapproachesbasedonhand-craftedfeatures

LIBRARIES USED

Tensorflow

To pursue research, Tensorflow, an interface forexpressing machine learning algorithms, is used toimplement ML systems into fabrication across avarietyofcomputerscienceareas,includingsentiment analysis, voice recognition, geographicinformationextraction,computervision,textsummarization,information retrieval,computational drug discovery, and flaw detection.Tensor flow is used at the backend of the proposedmodel'sSequentialCNNarchitecture(whichconsists of numerous layers). It's also used in dataprocessingto restructure thedata(picture).

Keras

Kerasprovideessentialreflectionsandbuildingunits for thedesign and transfer of machine leaningarrangementsatahighiterationrate.Tensorflow'sscalabilityandcross-platformfeaturesare fullyutilized.Kerasprimarydatastructuresarelayers and models. Keras is utilized to implementall of the layers in the CNN model. It aids in thecompilation of the overall model, as well as theconversion of the class vector to the binary classmatrixin dataprocessing.

Algorithm

Convolutional Neural Network(ConvNet/CNN)isadeeplearningalgorithm which can take in an input image, assignimportance(learnableweightsandbiases)tovarious aspects/objects in the image and be able todifferentiateonefromtheother.Thepre-processing required in a ConvNet is much lower ascomparedto otherclassification algorithms.

ARCHITECTURE

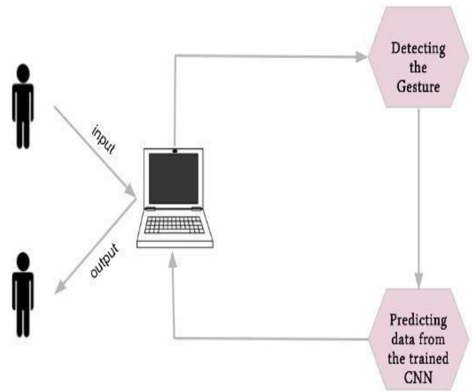

Fig: flow diagram for sign language recognition.

RESULT

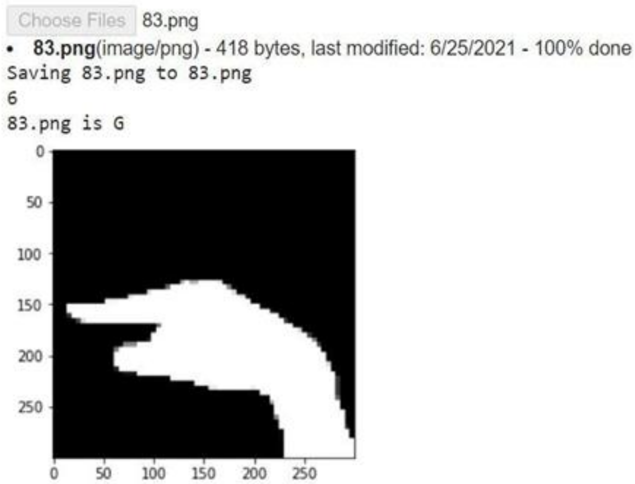

Fig: Output for given hand gesture.

CONCLUSION

We developed aCNN model for sign languagerecognition. Our model learns and extracts bothspatialandtemporalfeaturesby performing3Dconvolutions.Thedevelopeddeeparchitectureextractsmultipletypesofinformationfromadjacentinputframesandthenperformsconvolution and subsampling separately. The finalfeature representation combines information fromallchannels.Weusemultilayerperceptronclassifier to classify these feature representations.Forcomparison,weevaluatebothCNNandGMM-HMMonthesamedataset.Theexperimental results demonstrate the effectivenessofthe proposed method.

REFERENCES

[1] Alex Krizhevsky, Ilya Sutskever, and GeoffreyEHinton,“Imagenetclassificationwithdeepconvolutionalneuralnetworks,”inAdvancesinneural information processing systems, 2012, pp.1097–1105.

[2] AndrejKarpathy,GeorgeToderici,SankethShetty, Thomas Leung, Rahul Sukthankar, and LiFei-Fei,“Large-scalevideoclassificationwithconvolutionalneuralnetworks,”inCVPR,2014.

[3] Yann LeCun, Leon Bottou, YoshuaBengio, andPatrick´Haffner,"Gradient-basedlearningapplied to document recognition," Proceedings ofthe IEEE,vol. 86, no.11,pp. 2278–2324, 1998.

[4] HueihanJhuang, Thomas Serre, Lior Wolf, andTomaso Poggio, "A biologically inspired systemfor action recognition," in Computer Vision, 2007.ICCV 2007. IEEE 11th International Conferenceon. Ieee, 2007, pp. 1–8. [5] Shuiwang Ji, Wei Xu,Ming Yang, and Kai Yu, "3D convolutional neuralnetworksforhumanactionrecognition,"IEEETPAMI,vol. 35, no. 1, pp. 221–231, 2013.

[6] KirstiGrobelandMarcellAssan,"Isolatedsignlanguagerecognitionusinghiddenmarkovmodels," in Systems, Man, and Cybernetics, 1997.Computational Cybernetics and Simulation., 1997IEEEInternationalConferenceon.IEEE,1997,

vol. 1, pp. 162–167.

[7] ThadStarner,JoshuaWeaver,andAlexPentland,“Realtimeamericansignlanguagerecognitionusingdeskandwearablecomputerbased video,” IEEE TPAMI, vol. 20, no. 12, pp.1371–1375, 1998.

[8] ChristianVoglerandDimitrisMetaxas,“Parallel hidden markov models for american signlanguage recognition,” in Computer Vision, 1999.TheProceedingsoftheSeventhIEEEInternational Conference on. IEEE, 1999, vol. 1,pp. 116–122.

[9] Kanumalli, S.S., Chinta, A.,ChandraMurty, P.S.R. (2019). Isolation of wormhole attackers in IOV using WPWP packet. Revue d'IntelligenceArtificielle, Vol. 33, No. 1, pp. 9-13. https://doi.org/10.18280/ria.330102

[10] Narayana, Vejendla Lakshman, et al. "Secure Data Uploading and Accessing Sensitive Data Using Time Level Locked Encryption to Provide an Efficient Cloud Framework." Ingénierie des Systèmesd'Information 25.4 (2020).

[11] Kotamraju, Siva Kumar, et al. "Implementation patterns of secured internet of things environment using advanced blockchain technologies." Materials Today: Proceedings (2021).

[12] Krishna, Komanduri Venkata Sesha Sai Rama, et al. "Classification of Glaucoma Optical Coherence Tomography (OCT) Images Based on Blood Vessel Identification Using CNN and Firefly Optimization." Traitement du Signal 38.1 (2021).

[13] Satya Sandeep Kanumalli, Anuradha Ch and Patanala Sri Rama Chandra Murty, “Secure V2V Communication in IOV using IBE and PKI based Hybrid Approach” International Journal of Advanced Computer Science and Applications(IJACSA), 11(1), 2020. http://dx.doi.org/10.14569/IJACSA.2020.0110157

[14] CHALLA, RAMAIAH, et al. "Advanced Patient’s Medication Monitoring System with Ardunio UNO and NODEMCU." 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2020.

[15] Kanumalli, Satya Sandeep, Anuradha Ch, and Patanala Sri Rama Chandra Murty. "Advances in Modelling and Analysis B." Journal homepage: http://iieta. org/Journals/AMA/AMA_B 61.1 (2018): 5-8.

[16] Venkatramulu, S., et al. "Implementation of Grafana as open source visualization and query processing platform for data scientists and researchers." Materials Today: Proceedings (2021).

[17] Sandeep, Kanumalli Satya, Anuradha Chinta, and PatanalaMurty. "Isolation of Wormhole Attackers in IOV Using WPWP Packet." Rev. d'IntelligenceArtif. 33.1 (2019): 9-13.

[18] Gopi, ArepalliPeda, et al. "Classification of tweets data based on polarity using improved RBF kernel of SVM." International Journal of Information Technology (2020): 1-16.

[19] Narayana, Vejendla Lakshman, ArepalliPeda Gopi, and Kosaraju Chaitanya. "Avoiding Interoperability and Delay in Healthcare Monitoring System Using Block Chain Technology." Rev. d'IntelligenceArtif. 33.1 (2019): 45-48.

[20] Arepalli, Peda Gopi, et al. "Certified Node Frequency in Social Network Using Parallel Diffusion Methods." Ingénierie des Systèmesd'Information 24.1 (2019).

[21] Narayana, Vejendla Lakshman, ArepalliPeda Gopi, and R. S. M. Patibandla. "An Efficient Methodology for Avoiding Threats in Smart Homes with Low Power Consumption in IoT Environment Using Blockchain Technology." Blockchain Applications in IoT Ecosystem. Springer, Cham, 2021. 239-256.

[22] Kotamraju, Siva Kumar, et al. "Implementation patterns of secured internet of things environment using advanced blockchain technologies." Materials Today: Proceedings (2021).

[23] Bharathi, C. R., et al. "A Node Authentication Model in Wireless Sensor Networks With Locked Cluster Generation." Design Methodologies and Tools for 5G Network Development and Application. IGI Global, 2021. 236-250.

[24] Vejendla, Lakshman Narayana, Alapati Naresh, and Peda Gopi Arepalli. "Traffic Analysis Using IoT for Improving Secured Communication." Innovations in the Industrial Internet of Things (IIoT) and Smart Factory. IGI Global, 2021. 106-116.

[25] Narayana, Vejendla Lakshman, ArepalliPeda Gopi, and Kosaraju Chaitanya. "Avoiding Interoperability and Delay in Healthcare Monitoring System Using Block Chain Technology Avoiding Interoperability and Delay in Healthcare Monitoring System Using Block Chain Technology."

[26] Yamparala, Rajesh, and Balamurugan Perumal. "EFFICIENT MALICIOUS NODE IDENTIFICATION METHOD FOR IMPROVING PACKET DELIVERY RATE IN MOBILE AD HOC NETWORKS WITH SECURED ROUTE." Journal of Critical Reviews 7.7 (2020): 1011-1017.

[27] Devi, S. Pramela, et al. "Likelihood based Node Fitness Evaluation Method for Data Authentication in MANET." International Journal of Advanced Science and Technology 29.3 (2020): 5835-5842.

[28] Yamparala, Rajesh, and Balamurugan Perumal. "Secure Data Transmission with Effective Routing Method Using Group Key Management Techniques-A Survey Secure Data Transmission with Effective Routing Method Using Group Key Management Techniques-A

