ISSN: 2278-4632
Vol-11 Issue-01 2021

SIGN LANGUAGE RECOGNITION

K.SaiCharitha[1], A.Maneesha[2], Sd.karishma[3],J.Ramya[4]. Computer Science and Engineering Department, VNITSW, Guntur, AndhraPradesh, India. Project under the guidance of: MR.K.V.S.S. Rama Krishna (Asst.Professor),[1,2,3,4]IVB.Tech, Department of Computer Science and Engineerig, Vignan's Nirula Institute of Technology & Science for Women, PedaPalakaluru, Guntur-522009, 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-crafted features.

INTRODUCTION

Sign language, as one of the most widely used communication means for hearing-impaired people, is expressed by variations of hand-shapes, body movement, and even facial expression. Since it is difficult to collaboratively exploit the information from hand-shapes and body movement trajectory, sign language recognition is still a very challenging gask. This paper proposes an effective recognition model to translate sign language into text or speech in order to help thehearing impaired communicate with normal people through sign language.

Technically speaking, the main challenge of signlanguage recognition lies in developing descriptorsto express hand-shapes and motion trajectory. In particular, hand-shape description involves tracking hand regions in video stream, segmentinghand-shape images from complex background in each frame and gestures recognition problems. Motion trajectory is also related to tracking of the key points and curve matching. Although lots ofresearch works have been conducted on these twoissues for now, it is still hard to obtain satisfying result for SLR due to the variation and occlusion of hands and body joints. Besides, it is a nontrivial issue to integrate the hand-shape features and trajectory features together. To address these difficulties, we develop CNNs to naturally integrate hand-shapes, trajectory of action and facial expression. Instead of using commonly usedcolor images as input to networks like [1, 2], wetake color images, depth images and body skeleton images simultaneously as input which are all provided by Microsoft Kinect.

Kinect is a motion sensor which can provide color stream and depth stream. With the public Windows SDK, the body joint locations can be obtained in real-time as shown in Fig.1. Therefore, we choose Kinect as capture device to record sign words dataset. The change of color and depth in pixel level are useful information to discriminate different sign actions. And the variation of body joints in time dimension and epict the trajectory of sign actions. Using multiple types of visual sources as input leads CNNs paying attention to the change not only in color, but also in depth and trajectory. It is worth mentioning that we can avoid the difficulty of tracking hands, segmenting hands from background and designing descriptors for hands because CNNs have the capability to learn features automatically from raw data without any prior knowledge [3].

ISSN: 2278-4632
Vol-11 Issue-01 2021

CNNs have been applied in video stream classification recently years. A potential concern of CNNs is time consuming. It costs several weeksor months to train a CNNs with million-scale inmillion videos. Fortunately, it is still possible to achieve real-time efficiency, with the help of CUDA for parallel processing. We propose to apply CNNs to extract spatial and temporal features from video stream for Sign Language Recognition (SLR). Existing methods for SLR use hand-crafted features to describe sign language motion and build classification model based on these features. In contrast, CNNs can capture motion information from raw video data automatically, avoiding designing features. We develop CNNs taking multiple types of data asinput. This architecture integrates color, depth and trajectory information by performing convolution and sub sampling on adjacent video frames. Experimental results demonstrate that 3D CNNs can significantly outperform Gaussian mixture model with Hidden Markov model (GMM-HMM) baselines on some sign words recorded by ourselves.

LITERATURE SURVEY

Technically speaking, the main challenge of signlanguage recognition lies in developing descriptorsto express hand-shapes and motion trajectory. In particular, hand-shape description involves tracking hand regions in video stream, segmentinghand-shape images from complex background in each frame and gestures recognition problems. 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 satisfying result for SLR due to the variation and occlusion of hands and body joints. Besides, it is a non trivial issue to integrate the hand-shape features and trajectory features together. To address these difficulties, we develop CNNs to naturally integrate hand-shapes, trajectory of action and facial expression. Instead of using commonly usedcolor images as input to networks like [1, 2], wetake color images, depth images and body skeleton images simultaneously as input which are all provided by Microsoft 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 is useful information to discriminatedifferent sign actions. And the variation of body joints in time dimension and epict the trajectory of sign actions. Using multiple types of visual sources as input leads CNNs paying attention to the change not only in color, but also in depth and trajectory. It is worth mentioning that we can avoid the difficulty of tracking hands, segmenting hands from background and designing descriptors for hands because CNNs have the capability to learn features automatically from raw data without any prior knowledge [3].

CNNs have been applied in video stream classification recently years. A potential concern of CNNs is time consuming. It costs several weeksor months to train CNNs with million-scale inmillion videos. Fortunately, it is still possible to achieve real-time efficiency, with the help of CUDA for parallel processing. We propose to apply CNNs to extract spatial and temporal features from video stream for Sign Language Recognition (SLR). Existing methods for SLR use hand-crafted features to describe sign language motion and build classification model based on these features. In contrast, CNNs can capture

Motion information from raw video data automatically, avoiding designing features. We develop CNNs taking multiple types of data asinput. This architecture integrates color, depth andtrajectory information by performing convolution and sub sampling on adjacent video frames. Experimental results demonstrate that 3D CNNs can significantly outperform Gaussian mixture model with Hidden Markov model (GMM-HMM) base lines on some sign words recorded by ourselves.

PROPOSED METHODOLOGY

To approach this problem, we propose a novel convolutional neural network (CNN) which extracts

ISSN: 2278-4632
Vol-11 Issue-01 2021

discriminative spatial-temporal features from raw video stream automatically without any prior knowledge, avoiding designing features. To boost the performance, multi-channels of video streams, including color information, depth clue, and body joint positions, is used as input to the CNN in order to integrate color, depth and trajectory information. We validate the proposed model on a real dataset collected with Microsoft Kinect and demonstrate its effectiveness over the traditional approaches based on hand-crafted features

LIBRARIES USED

Tensorflow

To pursue research, Tensor flow, an interface for expressing machine learning algorithms, is used to implement ML systems into fabrication across a variety of computer science areas, including sentiment analysis, voice recognition, geographic information extraction, computer vision, text summarization, information retrieval, computational drug discovery, and flaw detection. Tensor flow is used at the backend of the proposed model's Sequential CNN architecture (which consists of numerous layers). It's also used in data processing to restructure the data (picture).

Keras

Keras provide essential reflections and building units for the design and transfer of machine leaning arrangements at a high iteration rate. Tens or flow's scalability and cross-platform features are fully utilized. Keras primary data structures are layers 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 class matrixin data processing.

Algorithm

A ConvolutionalNeural Network (Conv Net/CNN) is a deep learning algorithm which can take in an input image, assign importance (learn able weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a Conv Net is much lower as compared to other classification algorithms.

ARCHITECTURE

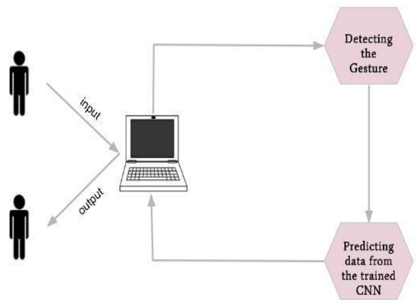
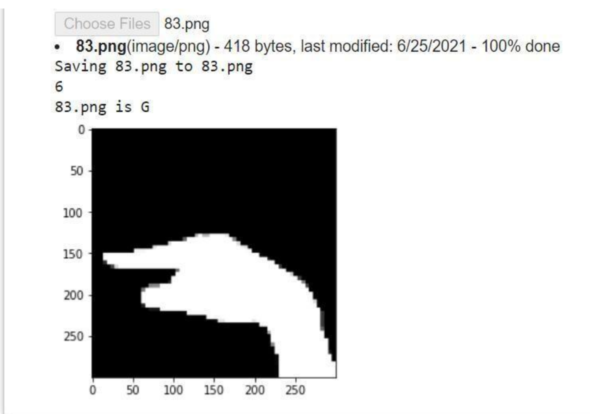


Fig: flow diagram for sign language recognition.

RESULT

Fig: Output for given hand gesture.



CONCLUSION

We developed a CNN model for sign language recognition. Our model learns and extracts both spatial and temporal features by performing 3D convolutions. The developed deep architecture extracts multiple types of information from adjacent input frames and then performs convolution and sub sampling separately. The final feature representation combines information from all channels. We use multilayer perception classifier to classify these feature representations. For comparison, we evaluate both CNN and GMM-HMM on the same data set. The experimental results demonstrate the effectivenessofthe proposed method.

REFERENCES

1. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "I magenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp.1097–1105.
2. Andrej Karpathy, George To derici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and LiFei-Fei,"Large-scalevideoclassificationwithconvolutionalneuralnetworks,"inCVPR,2014.
3. Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick´ Haffner," Gradient-based learning applied to document recognition," Proceedings ofthe IEEE,vol. 86, no.11,pp. 2278–2324, 1998.
4. Hueihan Jhuang, Thomas Serre, Lior Wolf, and Tomaso 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.
5. Kirsti Grobel and Marcell Assan," Isolated sign language recognition using hidden markov models," in Systems, Man, and Cybernetics, 1997.Computational Cybernetics and Simulation., 1997IEEEInternationalConferenceon.IEEE,1997,vol. 1, pp. 162–167.
6. ThadStarner,JoshuaWeaver,andAlexPentland,"Realtimeamericansignlanguagerecognitionusingdeskand wearablecomputerbased video," IEEE TPAMI, vol. 20, no. 12, pp.1371–1375, 1998.
7. ChristianVoglerandDimitrisMetaxas,"Parallel hidden markov models for american signlanguage recognition," in Computer Vision, 1999.TheProceedingsoftheSeventhIEEEInternational Conference on. IEEE, 1999, vol. 1,pp. 116–122.
8. 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
9. 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).
10. Kotamraju, Siva Kumar, et al. "Implementation patterns of secured internet of things environment using advanced blockchain technologies." Materials Today: Proceedings (2021).
11. 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).
12. 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
13. 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.

ISSN: 2278-4632
Vol-11 Issue-01 2021

14. 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.
15. Venkatramulu, S., et al. "Implementation of Grafana as open source visualization and query processing platform for data scientists and researchers." Materials Today: Proceedings (2021).
16. Sandeep, Kanumalli Satya, Anuradha Chinta, and PatanalaMurty. "Isolation of Wormhole Attackers in IOV Using WPWP Packet." Rev. d'IntelligenceArtif. 33.1 (2019): 9-13.
17. 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.
18. 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.
19. Arepalli, Peda Gopi, et al. "Certified Node Frequency in Social Network Using Parallel Diffusion Methods." Ingénierie des Systèmesd'Information 24.1 (2019).
20. 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.
21. Kotamraju, Siva Kumar, et al. "Implementation patterns of secured internet of things environment using advanced blockchain technologies." Materials Today: Proceedings (2021).
22. 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.
23. 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.
24. 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."
25. 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.
26. 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.
27. 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