ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2, 2024

HAND GESTURE RECOGNITION USING CNN

[1]MR.MAHESH HUNDEKAR,[2]PAILU UDAY KIRAN,[3]MADISHETTY URVI,[4]K RAJ KUMAR,[5]SHAIK RAFI

[1]Assistant Professor, Department of CSE-AI&ML, Malla Reddy College of Engineering,secunderabad Hyderabad

[2,3,4,5]UG Students,Department of CSE-AI&ML, Malla Reddy College of Engineering,secunderabad Hyderabad

ABSTRACT

Automatic human gesture recognition from camera images is an interesting topic for developing intelligent vision systems. In this paper, we propose a convolution neural network (CNN) method to recognize hand gestures of human task activities from a camera image. To achieve the robustness performance, the skin model and the calibration of hand position and orientation are applied to obtain the training and testing data for the CNN. Since the light condition seriously affects the skin color, we adopt a Gaussian Mixture model (GMM) to train the skin model which is used to robustly filter out non-skin colors of an image. The calibration of hand position and orientation aims at translating and rotating the hand image to a neutral pose. Then the calibrated images are used to train the CNN. In our experiment, we provided a validation of the proposed method on recognizing human gestures which shows robust results with various hand positions and orientations and light conditions. Our experimental evaluation of seven subjects performing seven hand gestures with average recognition accuracies around 95.96% shows the feasibility and reliability of the proposed method.

I.INTRODUCTION

Hand gesture recognition has emerged as an important area of research with applications in various fields, including human-computer interaction, sign language recognition, and augmented reality. By interpreting hand movements and gestures, computers can understand and respond to user commands, enabling intuitive and natural interaction with digital devices and systems.

The "Hand Gesture Recognition Using Convolutional Neural Networks (CNN)" project aims to develop a robust and

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2, 2024

efficient system for recognizing hand gestures from image data. CNNs, a type of deep learning architecture, have demonstrated remarkable capabilities in image recognition tasks, making them well-suited for this project.

Objectives

The primary objective of this project is to design and implement a CNN-based model capable of accurately classifying hand gestures captured from image or video streams. By training the model on a diverse dataset of annotated hand gesture images, we aim to create a system that can recognize a wide range of gestures with high accuracy and robustness.

Significance

The significance of this project lies in its potential to enhance human-computer interaction and enable new applications in various domains. By enabling computers to understand and interpret hand gestures, we can create more intuitive and immersive user interfaces for devices ranging from smartphones and tablets to virtual reality headsets and smart appliances.

Additionally, hand gesture recognition has applications in accessibility technology, allowing individuals with disabilities to interact with digital devices and systems more effectively. It also has potential applications in healthcare, robotics, and automotive industries, among others.

In this introduction, we have outlined the motivation, objectives, and significance of the "Hand Gesture Recognition Using CNN" project. Subsequent sections will delve into the methodology, dataset collection, model development, and expected outcomes, providing a comprehensive overview of our approach and contributions to the field of computer vision and human-computer interaction.

II.EXISTING SYSTEM

This work is a CNN-based human hand gesture recognition system. CNN is a research branch of neural networks. Using a CNN to learn human gestures, there is no need to develop complicated algorithms to extract image features and learn them. Through the convolution and sub-sampling layers of a CNN, invariant features are allowed with little dislocation. To reduce the effect of various hand poses of a hand gesture type on the recognition accuracies, the principal axis of the hand is found to calibrate the image in this work. Calibrated images are advantageous to a CNN to learn and recognize correctly.

III.PROPOSED SYSTEM

From the camera image input, the hand is extracted by skin color segmentation. The skin model is trained by a Gaussian Mixture model to classify skin color and non-skin color. After that, the calibration of hand position and orientation is used to translate and rotate the hand image to a neutral pose. The calibrated image is fed to the CNN to train or test the network. For continuous hand motion, the post-processing is used to filter out the noises of the results from the CNN. Each block of the flowchart is described as follows.

IV.LITERATURE REVIEW

Dynamic Hand Gesture Recognition, Subash Chandra Bose Jaganathan; Kesavan R; Thevaprakash P; Krishna Basak; Shinjan Verma; Anisha Mital, Gestures were most likely utilised by our ancestors to communicate. Armstrong once stated that he believes movements using the hands were the earliest form of complex human communication. The beginning stage of human computer Interactions is a gesture recognition system. Here, we have designed a Dynamic Hand Gesture Recognition (HGR) System using a Neural Network, which can recognize the gesture using the Computer or Laptop's Web Camera and do the corresponding tasks. The model has been divided into mainly three modules. Firstly, Module 1 is about building a CNN model which we use to predict the gestures. Our second module is about predicting the gesture through live video feed. The final module is about assigning a specified task for a particular predicted gesture. For the predicted gesture, we use PyAutoGUI module to control devices like mouse and keyboard. Finally, the system is made to do some desired tasks in response to the gestures and obtains 99.84% of accuracy.

2. Human hand gesture recognition using a convolution neural network,Hsien-I Lin; Ming-Hsiang Hsu; Wei-Kai Chen,Automatic human gesture recognition from camera images is an interesting topic for developing intelligent vision systems. In this paper, we propose a convolution neural network (CNN) method to recognize hand gestures of human task activities from a camera image. To achieve the

robustness performance, the skin model and the calibration of hand position and orientation are applied to obtain the training and testing data for the CNN. Since the light condition seriously affects the skin color, we adopt a Gaussian Mixture model (GMM) to train the skin model which is used to robustly filter out non-skin colors of an image. The calibration of hand position and orientation aims at translating and rotating the hand image to a neutral pose. Then the calibrated images are used to train the CNN. In our experiment, we provided a validation of the proposed method on recognizing human gestures which shows robust results with various hand positions and orientations and light conditions. Our experimental evaluation of seven subjects performing seven hand gestures with average recognition accuracies around 95.96% shows the feasibility and reliability of the proposed method.

V.IMPLEMENTATION

To implement the "Hand Gesture Recognition Using CNN" project, we begin by collecting a diverse dataset of hand gesture images or video clips, ensuring it covers various gestures relevant to the project's objectives. Following this, we preprocess the dataset, standardizing and enhancing image quality through resizing, normalization, and augmentation techniques. Next, we select a suitable CNN architecture for the task, such as VGG, ResNet, or custom-designed networks, and train it on the preprocessed dataset using techniques like SGD, Adam optimization, or transfer learning. Throughout training, we fine-tune hyperparameters to optimize model performance. Once trained, we evaluate the model on validation data, adjusting as necessary to improve accuracy and generalization. After validation, we test the final model on unseen testing data to assess real-world performance, iterating on improvements as needed. The deployed model is then integrated into a user-friendly interface or application, with APIs for inference and a GUI for input processing. Continuous monitoring, maintenance, and optimization ensure the system's reliability, scalability, and efficiency over time. By following this implementation method, we can develop a robust and efficient system for hand gesture recognition, enabling intuitive human-computer interaction and various practical applications.

VI.CONCLUSION

In this paper, we developed a CNN-based human hand gesture recognition system. The salient feature of the system is that there is no need to build a model for every gesture using hand features such as fingertips and contours. To have robust performance, we applied a GMM to learn the skin model and segment the hand area for recognition. Also, the calibration of the hand pose was used to rotate and shift the hand on the image to a neutral pose. Then, a CNN was trained to learn seven gesture types in this paper. In the experiments, we conducted 4-fold cross-validation on the system where 600 and 200 images from a subject were used to train and test, respectively and the results showed that the average recognition rates of the seven gesture types were around 99%. To test the proposed method on multiple subjects, we trained and tested the hand images of the seven gesture types from seven subjects. The average recognition rate was 95.96%. The proposed system also had the satisfactory results on the transitive gestures in a continuous motion using the proposed rules. In the future, a high-level semantic analysis will be applied to the current system to enhance the recognition capability for complex human tasks.

VII.REFERENCES

1. Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. In CVPR (Vol. 1, No. 2).

2. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

3. Ciftci, U., Ercil, A., & Littlewort, G. (2010). Real-time hand gesture recognition using Haar-like features and support vector machines. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops (pp. 68-75). IEEE.

4. Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th international conference on Machine learning (pp. 160-167).

5. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical

image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). IEEE.

6. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning (Vol. 1). MIT press Cambridge.

7. Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing (pp. 6645-6649). IEEE.

8 .Gu, Y., Wang, G., Xiong, Y., & Yu, G. (2017). Recent advances in convolutional neural networks. Pattern Recognition, 77, 354-377.

9.Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

10.Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

11.LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

12.Li, L., Su, H., Peng, Z., Zhu, Y., & Li, J. (2017). Hand gesture recognition using a depth camera for human-computer interaction. Journal of Visual Communication and Image Representation, 43, 279-289.

13.Li, X., & Kwok, J. T. (2017). Hand gesture recognition using leap motion controller. IEEE Transactions on Human-Machine Systems, 47(4), 513-523.

14.Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).

15. Liu, Z., Li, X., Luo, P., Loy, C. C., & Tang, X. (2016). Semantic image segmentation via deep parsing network. In Proceedings of the IEEE international conference on computer vision (pp. 1377-1385).

16. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE transactions on

pattern analysis and machine intelligence, 39(6), 1137-1149.

17. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). ImageNet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.

18. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

19. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-First AAAI Conference on Artificial Intelligence.

20. Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6848-6856).