TECHNIUM

Romanian Journal of Applied Sciences and Technology
ISSN: 2668-778X

vol. 16 / 2023

The 7th International Conference on Science Technology

organized by
Faculty of Social Science and
Law Universitas Negeri Manado and
Consortium of International Conference
on Science and Technology

The Innovation Breakthrough in Digital and Disruptive Era

PLUS COMMUNICATION

International Communication & PR

Technium: Romanian Journal of
Applied Sciences and Technology
Vol. 16, 213-218, October, 2023
ISSN: 2668-778X
www.techniumscience.com

Developing Human Hand Gesture Data

Muhammad Fuad[1*], *Sri* Wahyuni[1], *Nuniek* Fahriani[2], *Ilham* Nurwahyudi[1], *Mochammad* Ilham Darmawan[1], and *Fahmi* Maulana[1]

[1]Dept. of Mechatronics Engineering, Faculty of Engineering, Universitas Trunojoyo, Bangkalan, Indonesia
[2]Universitas Muhammadiyah, Surabaya, Indonesia

Abstract. Gesture is a movement originating from one of the body's limbs, generally using the hand. The movement is usually used to interact with each other among human beings. In this study, gestures are considered as a way for humans to be able to communicate with a system so that it can understand human body language. The first step to be able to create a system that can recognize body gestures is to create a dataset of the gestures to be used. The process for creating a dataset in this study includes taking pictures from a webcam, then pre-processing the images so that they can display the characteristics of each image. The dataset can then be used to model gesture classification using Convolutional Neural Network (CNN). The collection of gestures in this dataset is carried out on the entire upper body to create data variations of male and female also with various human body shapes. This study has collected 450 images of gesture data and has prepared these data with ration of 80% for training and 20% for testing.

* Corresponding author: fuad@trunojoyo.ac.id

1 Introduction

A mobile robot is a robot that can move from one initial location to certain target position using some actuators. Generally, the actuators used are motors which are controlled by means of a microcontroller. There are various tools that can be used for robot control facilities. Starting from using a joystick that has a specific message from each button, progressing to control technology that allows users to interact with mobile robots without having to come into direct contact with the controller. Research related to controlling the motion of a mobile robot using speech recognition has previously been carried out [1]. This research discussed applications that can be used to control the movement of a mobile robot using voice commands in real-time. The application was created by utilizing Google's speech Application Programming Interface (API) which can convert voice into text before sending it to the mobile robot as a motion controller. Another study on the design of mobile robot navigation using speech recognition from the Android smartphone being 10-15 cm [2]. Based on the research that has been done, an idea emerged to develop motion control of a mobile robot with gesture recognition technology using robot vision.

Robot vision has been developed to control the longitudinal linear velocity and body rotation control of omni-directional mobile robot [3]. Distance estimation based on depth image and color detection based on RGB image were proposed to identify the environment landmarks. The landmarks that were represented by a movement of part of body such as hand or head for idea expression is known as gesture. The control of a robot manipulator to imitate the motion of operator’s hand was investigated in [4] by exploring Kinect sensor. Skeleton image was used to control SCORBOT-ER 9 Pro by proposing gesture recognition based on bone measurement and estimation of angle between bones. Law of cosines was utilized to calculate the angle of each robot’s joints.

Progressive probabilistic hough transform with law of cosines, quadrant principle, and voting mechanism based on the length of perpendicular line and region grouping was proposed in [5] to extract angle information of robot orientation from line segments from single-frame of fish eye image.

RGB-D data from Kinect sensor for estimation of translational motion was studied in [6]. Sensor was placed in the lateral direction of robot movements. Pairs of RGB frame point features were matched and synchronized with depth image to get the distance information by applying law of cosines.

The Convolutional Neural Network (CNN) was used to recognize gesture of palm and finger of human hand by using monocular camera [7]. Hand Gesture Recognition System based on contour finding was investigated in [8] to detect one to five finger gestures. Skin color of the face-based hand gesture detection was explored in [9] to move a mobile robot by utilizing the hand movement.

2 Related works

In previous research, various approaches and techniques have been used to build a gesture dataset according to needs. In the field of human gesture classification, various datasets for gesture recognition have been produced which are differentiated based on size, number of classes, movements, and sensors used to capture data. This section describes the overview of several studies in creating dataset for classifying human gestures. The HaGRID dataset was designed for the Hand Gesture Recognition (HGR) system [10]. The dataset contained 552,992 data samples which were divided into 18 gesture classes. The dataset could be used in video conferencing services, the automotive sector, and services for people with speech and hearing impairments. Large-scale datasets were collected based on the real world through video frames for dynamic gesture recognition [11]. The created dataset includes 148,092 video clips with a ratio of 8:1:1 for training, validation and testing. The approach taken to test this research is by using a 3D Convolutional Neural Network (3DCNN). approximately the same size as the main text (10 point). Try to ensure that lines are no thinner than 0.25 point.

Dataset contained of actual and augmented images was worked in [12] for astronaut gesture recognition. Two gestures dataset comprised of static and dynamic was established in [13] to facilitate the learning process and aid the research work on human-robot interaction. Dataset of eight channel of surface electromyogram in non-ideal conditions was presented in [14]. Dataset based on wrist-worn with various subjects was explored in [15] to investigate about human-home appliances interaction in different work area.

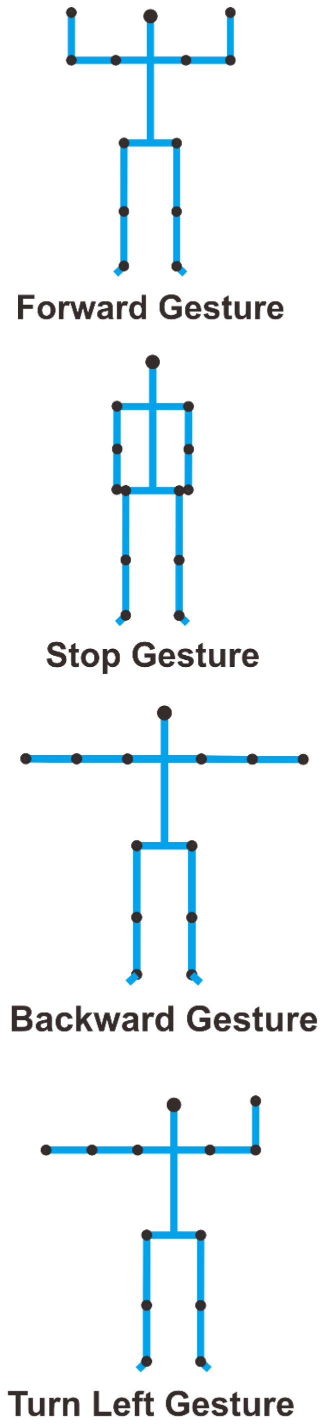
This study proposes CNN to recognize five human hand gestures. Gestures read by the webcam are used to control the movement of the mobile robot to move "forward", "back", "right", "left", "stop". By using a web camera for gesture detection, it is hoped that the interactions that are built between humans and robots will become easier and feel natural. Each gesture is stored in a different class in the dataset. Additional images for classes where there is no movement or empty images also added to this dataset.

3 The proposed method

The dataset consists of 450 image data with 5 motions and a "no motion" class. The design of gesture to control mobile robot is shown in Figure 1. The data recording process begins by using a webcam to take pictures of body gestures. Then, the captured images are pre-processed before being used as a dataset. Pre-processing also aims to improve image quality, reduce noise and highlight important features of each image. In this study, several stages were carried out for data pre-processing starting from color segmentation to reducing the size of the image to match the input data model. The complete dataset development process is described in Figure 2.

Technium: Romanian Journal of Applied Sciences and Technology
Vol. 16, 213-218, October, 2023
ISSN: 2668-778X
www.techniumscience.com

The RBG image that captured by the webcam is converted to HSV. The color segmentation is then executed by utilizing blue color backdrop. The color segmentation process is shown in Figure 3. The background color search is done manually using the HSV trackbar. Determination of the range of minimum and maximum values for hue, saturation, and value refers to the brightness level of the light at the time of testing.



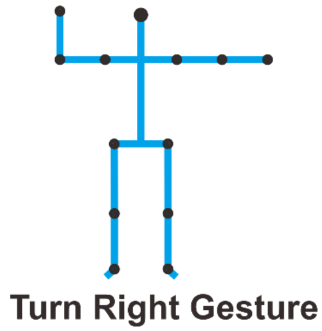


Fig. 1. Gesture design to represent forward, stop, backward, turn left, and turn right.

So that the range of values for HSV can be different if testing is done for different times. Then thresholding is done to get a black and white image. The process aims to distinguish between background and foreground. The entire image pre-processing process is used to lighten the computational load while reducing the number of datasets used. Then resize the image size from 640 x 480 down to 100 x100 pixels. The purpose of resizing is to adjust the image to match the size expected by the model. The results of the image data after pre-processing is depicted in Figure 3.

After the image data has been obtained, then the labelling process is carried out. The correct labels allow the model to learn and identify patterns or relationships between input and output features appropriately. In other words, data labelling allows the model to learn the rules and correlations present in the data. Image data that has been pre-processed will be captured and then stored in each class folder. The data labelling process uses the Python programming language is presented in Figure 4.

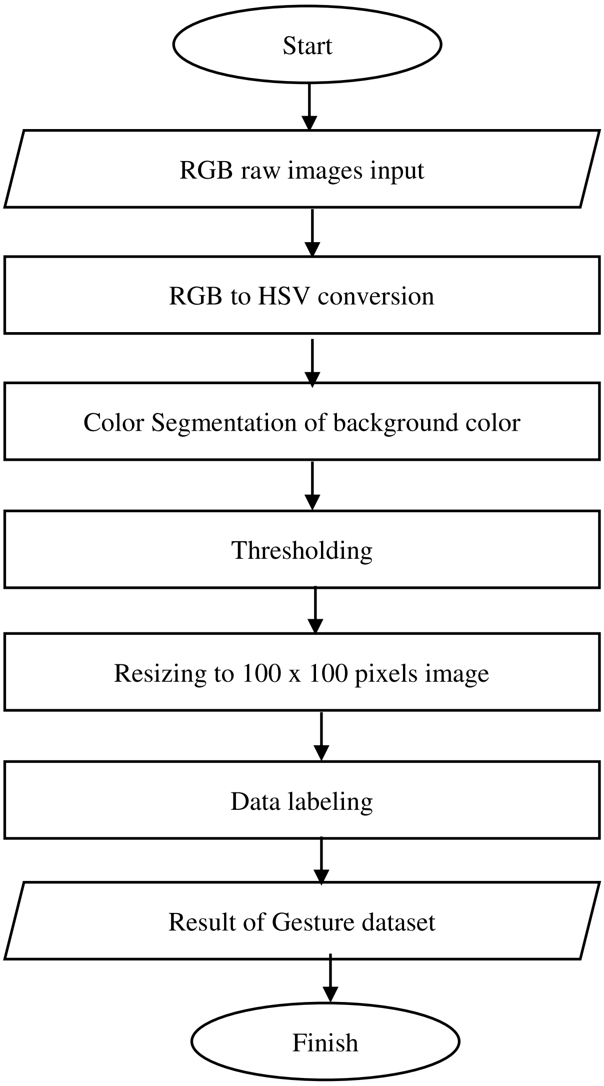
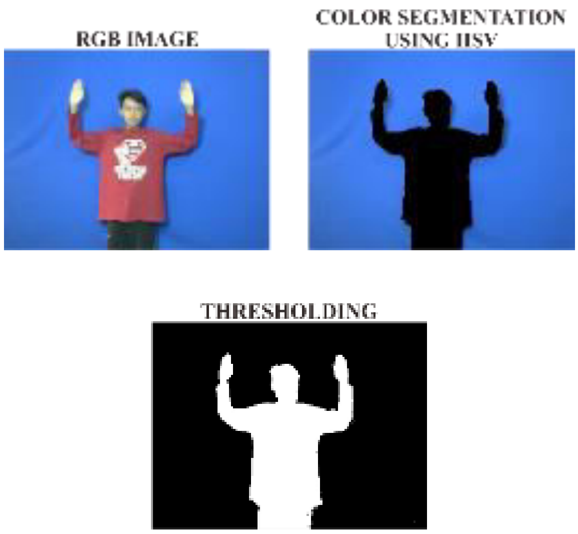


Fig. 2. The flow of dataset development.



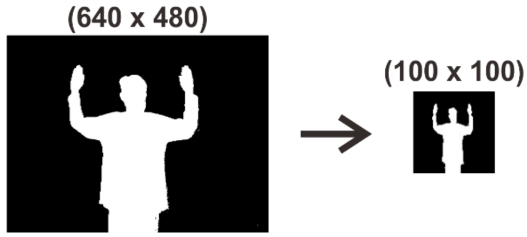


Fig. 3. Pre-processing data involves color conversion, segmentation, thresholding, and resizing.

```
if interrupt & 0xFF == ord('1') :

    cv2.imwrite(
        dirs   +   "0/"   +   "Maju"   +
str(i)      + ".jpg",  frame)

    print("data  gesture  Maju  ke-
",i)

    i+=1


if interrupt & 0xFF == ord("2") :

    cv2.imwrite(dirs + "1/" +
        "Mundur" + str(j) +
        ".jpg",frame)

    print("data gesture Mundur ke-
        ",j)

    j+=1
```

Fig. 4. Data labelling for collection of images.

The recorded gestures total 450 image data which are divided into 380 train data and 70 test data. The process of creating the gesture dataset involves several research subjects with differences in gender and body posture so that the model becomes more robust and accurate in recognizing body movements from various users. The image data in this dataset displays wide variations in gender differences, and body posture. Variance of gender in dataset is described in Figure 5. The variations in the posture of the dataset are shown in Figure 6.

Fig. 5. Gender variance in dataset.

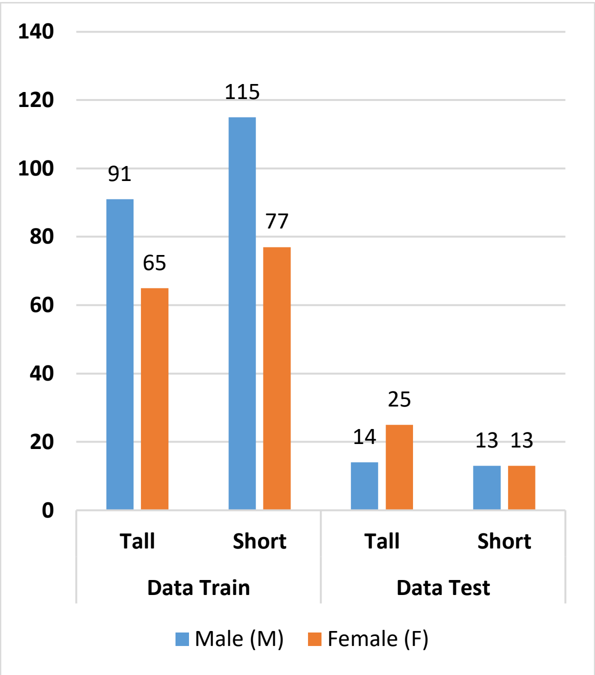


Fig. 6. Posture variance in dataset.

4 Environment setting and gesture recording step

The body gesture dataset is used as input to train a CNN model that can recognize body movements with high accuracy. Before recording body movements, the optimal recording position and angle is planned. Ensuring that there is sufficient light and a clean background will help produce clear footage of body movement that is not distracted by external factors. Based on Figure 7, determined that the distance between the webcam and the operator is 250 cm with a webcam height of 100 cm and an angle of 90°. Using a blue background to simplify the color segmentation process. Based on the picture, the operator (red box) stands in front of the webcam and lighting (yellow box) with a distance of 250 cm. The function of lighting is to help distribute light evenly so as to facilitate the process of color segmentation as seen on computer (blue box). The image dataset is saved by pressing numbers 1-6 on the laptop keyboard. The dataset will be automatically collected according to the class.

Fig. 7. Environment setting and gesture recording step.

5 Conclusion

This study presents a proposed approach to develop body gesture datasets. A webcam is used to capture pixels from the gesture image, then extract the image features to model the visual appearance of the gesture. This study proposed to use Convolutional Neural Network (CNN) for the dataset retrieval process. This method is implemented by using the Python programming language. This study has successfully

Technium: Romanian Journal of
Applied Sciences and Technology
Vol. 16, 213-218, October, 2023
ISSN: 2668-778X
www.techniumscience.com

obtained a gesture dataset which consists of 6 classes. Those classes are 'no movement'-class, 'forward'-class, 'backward'-class, 'stop'-class, 'right turn'-class, and 'left turn'-class. This gesture dataset can be used to control the motion of the mobile robot for future research. The dataset consists of 450 images of hand gesture data with a ratio of 80% for training and 20% for validation.

Author would like to acknowledge to the Institute for Research and Community Service (Lembaga Penelitian dan Pengabdian kepada Masyarakat, LPPM) University of Trunojoyo Madura for financial support with contract No. 5892/UN46.4.1/PT.01.03/2023 and also give the highest appreciation to the Department of Mechatronics Engineering Faculty of Engineering University of Trunojoyo Madura for providing laboratory and technical support in order to make this research to be realized.

References

1. I. P. A. Cendana, A. A. K. A. C. Wiranatha, and K. S. Wibawa, Merpati **5**, 2 (2017)
2. A. P. Asmoro, N. E. Khomariah, Repository UNTAG (2020)
3. M. Fuad, R.E. Abdul Kadir, D. Purwanto, JATIT, **79**, 3 (2015)
4. M. Fuad, *Skeleton based gesture to control manipulator*, in Proceedings of the 2015 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation, ICAMIMIA, 15-17 Oct 2015, Surabaya, Indonesia (2015)
5. M. Fuad, T. Agustinah, D. Purwanto, T. A. Sardjono, R. Dikairono, *Robot orientation estimation based on single-frame of fish eye image*, in Proceedings of the 2019 International Conference on Science and Technology, ICST, 17-18 Oct 2019, Surabaya, Indonesia (2019)
6. M. Fuad, *Translation motion estimation using Kinect*, in Proceedings of the 2019 International Conference on Science and Technology, ICST, 17-18 Oct 2019, Surabaya, Indonesia (2019)
7. P. Xu, arXiv CVPR, **1704**, 07296 (2017)
8. T. Jha, B. Singh, A. Sharma, S. Sharma, and R. Kumar, *Real Time Hand Gesture Recognition for Robotic Control*, in Proceedings of the 2nd Int. Conf. Green Comput. Internet Things, ICGCIoT, 16-18 Aug 2018, Bangalore, India (2018)
9. S. Ikegami, C. Premachandra, B. H. Sudantha, and S. Sumathipala, *A Study on Mobile Robot Control by Hand Gesture Detection*, in Proceedings of the 3rd Int. Conf. Inf. Technol. Res., ICITR, 5-7 Dec 2018, Moratuwa, Sri lanka (2018)
10. K. Alexander, M. Andrew, K. Karina, arXiv CVPR, **2206**, 08219 (2022)
11. J. Materzynska, G. Berger, I. Bax and R. Memisevic, *The Jester Dataset: A Large-Scale Video Dataset of Human Gestures*, in Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop, ICCVW, 27-28 Oct 2019, Seoul, Korea (South) (2019)
12. G. Lingyun, Z. Lin and W. Zhaokui, IEEE Access, **8**, (2020)
13. M. A. Kassab, M. Ahmed, A. Maher and B. Zhang, IEEE Access, **8**, (2020)
14. B. Zhu, D. Zhang, Y. Chu, Y. Gu and X. Zhao, IEEE TNSRE, **30**, (2022)
15. H. -Q. Nguyen, T.-H. Le, T. -K. Tran, H. -N. Tran, T. -H. Tran, T. -L. Le, H. Vu, C. Pham, T. P. Nguyen, and H. T. Nguyen, IEEE Access, **11**, (2023)