Feature Extraction and Prediction for Hand Hygiene Gestures with KNN Algorithm

Rashmi Bakshi
School of Electrical and Electronic Engineering
City Campus, TU Dublin

Abstract

This work focuses upon the analysis of hand gestures involved in the process of hand washing. There are six standard hand hygiene gestures for washing hands as provided by World Health Organisation—hand hygiene guidelines. In this paper, hand features such as contours of hands, centroid of the hands and extreme hand points along the largest contour are extracted with the use of computer vision library, OpenCV. These hand features are extracted for each data frame in a hand hygiene video. A robust hand hygiene dataset of video recordings was created in the project. A subset of this dataset is used in this work. Extracted hand features are further grouped into classes based on KNN algorithm with cross fold validation technique for the classification and prediction of the unlabelled data. A mean accuracy score of >95% is achieved and proves that KNN algorithm with an appropriate input value of K=5 is efficient for classification. A complete dataset with six distinct hand hygiene classes will be used with KNN classifier as a future work.

Keywords: Hand Hygiene, Hand Washing, KNN algorithm, Supervised Learning, Contours, Feature Extraction

1. Introduction

The current ongoing coronavirus pandemic has highlighted the importance of hand hygiene practices in our daily lives, with governments and healthcare authorities around the world promoting good hand hygiene practices. More than 1 million cases of hospital-acquired infections (HAIs) are reported in Europe annually [1]. Hand hygiene compliance may reduce the risk of cross-transmission of microbes such as MRSA, bacteria commonly known for causing the spread of HAIs thereby reducing the number of infections as well as health-care expenditures [2]. There are well-documented and structured guidelines for washing hands as per World Health Organization (WHO) [3]. In advance of developing, a real time hand-hygiene gesture detection system, a preliminary analysis for feature extraction associated with each hand gesture is required. In this paper, features for hand hygiene gestures—“Rub hands palm to palm”, “Hands with fingers interlocked” and “Hands with fingers interlaced” are extracted with the use of basic computer vision, OpenCV library. The features are then passed to a KNN classifier for hand hygiene gesture classification and prediction.

2. Hand Hygiene Dataset

Thirty volunteers were asked to perform the various hand washing gestures as per WHO guidelines. For recording the hand gestures, an aluminium rig was placed over the laboratory's sink and a digital camera was mounted over. The camera was mounted in such a way that only the hand movements of the participants were recorded to ensure the user’s privacy. The participants were informed about the study and were asked to complete an information sheet along with the consent for their inclusion in the study. WHO hand hygiene stages were demonstrated before capturing the video recording. The participants originated from diverse ethnic backgrounds thereby providing skin colour variation. A robust dataset of hand hygiene gestures was created where skin extraction algorithm was applied to extract hands for users with different skin color. The video length for the hand washing activity was recorded for 25-30 seconds. Every hand-washing step was followed by a pause where in the user was instructed to move their hands away from the camera. Video format for this data set is MP4 file with a size of range 40-60 MB and a frame rate of 29.84 frames/s. All of the six hand washing movements were recorded in one video for each participant. In this paper, a subset of the data set is utilised with an aim for the use of a complete dataset in future studies.
3. Edge Detection in Computer Vision

A contour is a closed curve of points or line segments that represents the boundaries of an object in an image, having the same colour and intensity [4]. Finding contours can be applied to determine the shape of an object (arc length; number of vertices), the number of objects in an image (number of contours) and measure the size of the objects [4].

Contour features: Contour features are the attributes of a contour that can be derived from the contour. They are used as feature extraction and classification of images. The common contour features are:

1. Image Moments: Image moments are statistical properties of a section of an image. Image moments can be used to extract useful information from the contour such as the centroid, area, etc. The moments are used as features for shape recognition.
2. Centroid: The centroid is defined as a coordinate (cx, cy) and is derived from the image moments.
3. Contour area: The contour area is the image area outlined by the contour.
4. Contour perimeter: It is also called the ‘arc length’; it is the length of the contour in pixels.
5. Convex hull: The Convex hull of a shape or a group of points is a tight fitting convex boundary around the points or the shape of the object, which is tracked. They are the minimum enclosing polygon of all points of the input shape. With a given set of X points in the Euclidean space, the convex hull is the smallest possible convex set containing these X points [5,6]

Contour tracking is widely used in the field of computer vision. The idea behind contour tracking is to traverse the border of a region completely and detect the edge points. Xie et al [7] used the contour detection method to determine the number of copper cores in the wire. Poda et al [8] found the perimeter of a contour to detect a specific shape such as a pentagon, and the initial character ‘P’ is saved and transmitted to an Arduino for the movement of a mechanical arm. Bochkarev et al [9] compares object characteristics such as area, perimeter and compactness of the contour of regular shapes to that of irregular shapes using Computer vision-open source library-OpenCV.

OpenCV offers cv2.findContours function, which can retrieve contours from the binary image and return the number of detected contours. In this work, RGB image is converted to YCbCr color space, creating a mask with widely accepted values for skin detection. Then, a largest contour for the skin pixels is retrieved and Image Moments are utilised to extract the centre of the mass. Argmin and argmax functions to slice a numpy array are used in order to extract the extreme points for each frame. Centroid and extreme hand points are saved in a csv file for each hand hygiene video recording for further processing and classification.

4. KNN Classification algorithm

The K-Nearest-Neighbors (KNN) is a nonparametric classification algorithm, i.e. it does not make any presumptions on the given dataset. It is known for the simplicity and effectiveness [10]. It is a supervised learning algorithm where a training dataset is provided and the data points are categorised into pre-defined classes. The training of the classifier is used to predict the class of unlabelled data. KNN has a high cost of classifying new instances as nearly all computation takes place at classification time rather than when the training examples are first encountered and therefore also known as lazy learning algorithm. It is used to classify the data based on closest or neighbouring training examples in a given region. For continuous data, Euclidian distance between the data points is calculated to determine the closest neighbours. The input value of ‘K’ is used to determine the number of neighbours and to build the boundaries for each class. It is usually combined with cross validation method, which is a resampling procedure and splits the data in K number of groups [10, 11]. In this work, K-fold cross validation is implemented in python as an adaptation from [12].

System requirements
The hardware used for all the experiments were Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz 2.19GHz with 4 GB RAM with 64-bit Windows Operating System.
5. Methodology

Hand Features such as centroid of the mass; extreme hand points (top, bottom, left and right) along with the largest contour are extracted for each frame in hand hygiene video recording. The extracted features are saved in hand_features.csv file. Figure 5.1 elaborates the steps involved in feature extraction.

Extracted features are passed to a KNN classifier that determines the nearest neighbours with K=5. The set of observations is randomly divided into K groups or folds of approximately an equal size. The first fold serves as a validation set and the method is fit on the remaining K-1 folds. The mean evaluation score is presented for each fold. Figure 5.2 presents the sequential workflow for training of the dataset; classification and prediction of the unlabelled data.

6. Results

This section provides the results for feature extraction and the classification of the unlabelled data. A sample of the extracted features for a given hand hygiene video recording is presented. Centroid, Extreme hand points-left, right, top and bottom- x, y coordinates are saved as a tuple in a csv file. The class name is appended in the csv file for each row of features.

Centroid: (112,175); Extreme left: (74,192); Extreme right: (153,149); Extreme top: (122,104); Extreme Bottom (74,239)

A sample of features saved in Hand_features.csv
Table 6.1 lists three hand hygiene classes corresponding to three hand hygiene videos that are utilised from a robust hand hygiene dataset, discussed in section 2. The integer number is assigned to each class as a class label. Number of rows for the extracted features for each video is presented.

| Class Name                | Class Label | # of Rows for Extracted Features |
|---------------------------|-------------|----------------------------------|
| Rub hands Palm to Palm    | 2           | 173                              |
| Fingers Interlocked       | 0           | 149                              |
| Fingers Interlaced        | 1           | 42                               |

Figure 6.1 is the result of image processing where the RGB pixels are converted to YCBCR color space for skin detection with minimum (0, 133, 77) and maximum (255, 179, 127) range used as a mask. The largest contour max (contour area) is displayed along with the contour features such as centroid and extreme hand points along the contour.

Class Predictions on the new unlabelled data that the classifier did not encounter before:

K Nearest Neighbors (KNN) algorithm classifies the data based on the similarity found by the Euclidian distance among the data points. Similar and closest data points are grouped together in one class. The success rate of the algorithm is dependent on the input value of K. An appropriate value of K with cross-fold validation can increase the accuracy of the classifier. The accuracy score for each fold is presented along with a mean accuracy metric is given in Table 6.2.

A data frame is extracted from different hand hygiene video recordings that the classifier has not encountered before. Features for the subsequent frames: centroid and the extreme points are extracted and passed as an input for the class prediction. Table 6.3 lists the unlabelled data frames with class predictions and it can be observed that KNN correctly classifies 4 out of 5 frames.

| Class Name                | Class Label | # of Rows for Extracted Features |
|---------------------------|-------------|----------------------------------|
| Rub hands Palm to Palm    | 2           | 173                              |
| Fingers Interlocked       | 0           | 149                              |
| Fingers Interlaced        | 1           | 42                               |

Table 6.2 Accuracy score for each fold passed to KNN algorithm

| Accuracy Score | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 |
|----------------|--------|--------|--------|--------|--------|
| (Correctly Predicted/Actual)*100 | 88.88  | 94.44  | 100.0  | 98.61  | 95.83  |

Mean accuracy : 95.556%
Table 6.3 Class Prediction for unlabelled data frame

| Unlabeled data | Predicted | Actual class |
|----------------|-----------|--------------|
| [203,150,48,234,95,221,76,190,48,239] | 0 | 0-Interlaced |
| [222,165,144,236,301,239,204,54,160,239] | 1 | 1-Interlocked |
| [185,107,116,236,177,205,141,168,116,239] | 2 | 2-P2P |
| [140,131,94,128,189,97,174,42,140,221] | 1 | 2-P2P |
| [180,102,122,118,231,108,162,54,148,138] | 1 | 1-Interlocked |

7. Future Work

In this paper, a preliminary classification of distinct hand hygiene videos were carried out with the use of a supervised learning algorithm, KNN. The hand features such as contours, centroid and extreme hand points were extracted for the videos and saved in a csv format. The saved features were further processed to compute the Euclidian distance to group the similar data points. The mean accuracy >95% was achieved with the cross fold validation technique. A complete hand hygiene dataset consists of six hand hygiene gestures captured in the video format for thirty participants imitating the health care setting. It is available and can be accessed online. The future work will incorporate all the six classes of hand washing and the use of KNN algorithm for the classification and prediction of the unlabelled data.

Supplementary Materials: The hand hygiene video recordings along with the python code; features-csv file for the following are available online:
https://tudublin-my.sharepoint.com:/f:/r/personal/d16126930_mytudublin_ie/Documents/Hand%20Hygiene%20Research/KNN Classification?csf=1&web=1&e=EFHrc9
The Complete robust hand hygiene dataset can be accessed at https://tudublin-my.sharepoint.com:/f:/r/personal/d16126930_mytudublin_ie/Documents/Hand%20Hygiene%20Research/Hand WashData?csf=1&web=1&e=mMwzfp.

References

[1] A. Cassini et al. Burden of Six Healthcare-Associated Infections on European Population Health: Estimating Incidence-Based Disability-Adjusted Life Years through a Population Prevalence-Based Modelling Study. PLoS Med., 2016, vol. 13, no. 10, pp. 1–16
[2] M. ŠLširak; A. Zvizdić.; M. Hukić, Methicillin-resistant staphylococcus aureus (MRSA) as a cause of nosocomial wound infections. Bosn. J. Basic Med. Sci., 2010, vol. 10, no. 1, pp. 32–37.
[3] WHO Guidelines on Hand Hygiene in Healthcare. Available online: URL, http://apps.who.int/iris/bitstream/handle/10665/44102/9789241597906_eng.pdf;jsessionid=5FCA460BF882961840326820865681E8?sequence=1 (accessed on 29-09-2021)
[4] Image processing with python. https://datacarpentry.org/image-processing/09-Contours/, accessed: 2018.
[5] Contour features-opencv documentation. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_contours/py_contour_features.py_contour_features.html accessed 2018.
[6] Advanced contour properties. "https://gurus.pyimagesearch.com/lesson-sample-advanced-contour-properties/#", accessed 2018.
[7] Guobo Xie and Wen Lu. Image Edge Detection Based On OpenCv. International Journal of Electronics and Electrical Engineering, 1(2):104–106, 2013.
[8] Xhensila Poda and Olti Qirici. Shape detection and classification using OpenCV and Arduino Uno. Proceedings of the 3rd International Conference RTA-CSIT, pages 128–136, 2018.
[9] Sergey O Bochkarev, Igor B Litus, and Natalia S Kravchenko. Irregular Objects. Shape Detection and Characteristic Sizes. Proceedings of the 1st International Workshop on Radio Electronics and Information Technologies, 1(2), 2017.

[10] K. Taunk, “A Brief Review of Nearest Neighbor Algorithm for Learning and Classification,” 2019 Int. Conf. Intell. Comput. Control Syst., no. Icics, pp. 1255–1260, 2019.

[11] G. Guo, H. Wang, D. A. Bell, and Y. Bi, “KNN Model-Based Approach in Classification KNN Model-Based Approach in Classification,” On the move to meaningful Internet Systems Intl conference, 2003.

[12] J. Brownlee. Develop k-Nearest Neighbors in Python From Scratch. URI: https://machinelearningmastery.com/tutorial-to-implement-k-nearest-neighbors-in-python-from-scratch/ (accessed on 30-12-2021)