Advances in Informatics, Management and Technology in Healthcare
J. Mantas et al. (Eds.)

doi:10.3233/SHTI220694

Comparison of Data Classification Results for Leap Motion Recovery Gestures

Stelian NICOLA[a,1], Oana-Sorina CHIRILA[a] and Lacramioara STOICU-TIVADAR[a]
[a] *Department of Automation and Applied Informatics*
Politehnica University Timisoara, Romania

Abstract. Static and dynamic gestures are frequently used in activities supporting learning, recovery healthcare, engineering, and 3D games to increase the interactivity between man and machine. The gestures are detected via hardware devices and data is processed using different software methods. This paper presents the manner of detection and interpretation of two gestures, a hand rotation gesture and a palm closing and opening gesture, using the Leap Motion device. These two dynamic gestures are very often used in hand recovery exercises. For comparing the two gestures we use data classification methods, Support Vector Machine (SVM) and Multilayer Perceptron (MLP). The data for the gesture classification were 80% training data and 20% testing data. The metrics for comparison are precision, recall, F1-score, and the total number of testing cases (support). The SVM classifier gives an accuracy of 99.4% and the MLP classifier a 96.2%. We built two confusion matrices for better visualizing the results.

Keywords. Multimodal interaction, Machine learning, Linear-SVM, MLP, Leap Motion, Recovery

1. Introduction

Currently, gestures are frequently used either to detect an action, a state or a word or to control user interfaces in domains as engineering, IT, recovery healthcare [1], music [2] and education. They support the users' in controlling various hardware or software interfaces without the need for classic input devices such as the mouse and the keyboard. as the alternatives to these, are: Leap Motion (LM), Microsoft Kinect, Microsoft HoloLens, distance sensors or video cameras. The environment where the gestures are rendered for further visualization and interpretation use technologies based on Virtual Reality (VR) [3], Augmented Reality (AR) or Mixed Reality (MR) [4].

Different LM gesture classification methods are addressed in existing literature. Thus, the predefined LM gestures as circle, swipe and Key Tap were detected in video games. They play and important role in human-computer interaction. Predefined gestures reported an accuracy of 96.7% [5]. Genetic algorithms (GA) processed data to detect gestures used in sign language. Different models were applied: Support Vector Machine (SVM), Random Forest (RF) and Naive Bayes (NB) wherein an accuracy of 74% was achieved for gesture recognition [6]. As results from the examples above, proper gesture detection and classification for the LM device is a challenge for the specialists in the

[1] Corresponding Author, Stelian NICOLA, Politehnica University Timisoara, P-ta Victoriei no. 2, Timisoara, Romania; E-mail: stelian.nicola@aut.upt.ro.

field. Given that the most frequently used LM gestures are dynamic, incorrect detections thereof occur. This paper will present and compare 2 LM gesture classification modalities, using SVM classification algorithms and algorithms based on Multi-layer Perceptron classifier (MLP) neural networks. It will contemplate the manner of classification of two LM gestures: fist closing and hand rotation. These gestures are the ones of interest from the clinician for patient's hand recovery exercises.

2. Methods and tools

Two virtual 3D models of the hands were used; they copy the real movement of the users' hands and enable the visualization of the hands in a virtual environment. For the visual exposure of the LM gestures built for hand recovery, the 3D application editor Unity version 2020 was used. C# scripts were written to detect the gestures and collect the data describing the gestures.

For classification of the two gestures, 29 data were collected: position within the 3D space of the fingers (15 distinct data: the position of 5 fingers on the x, y, z axes), the distances between the fingers and the palm (5), the direction vector for the thumb (3 directions for the x, y, z axes), the distance between the thumb and the other fingers (4), the hand rotation angle to a fixed point situated at the (0, 0, 0) point within the 3D virtual space and the gesture class (palm closing and opening gesture and hand rotation gesture). Since the gestures are used for hand recovery, the two gestures were divided into 3 levels each, that are relevant for the clinical assessment. The first gesture has 3 value ranges that define each level, and so does the second gesture. For the first gesture the value range is given by the semiperimeter of the geometric shapes with points on the thumb, the palm, and the middle finger. If the semiperimeter formed by the 3 points is larger than 150 units of the Unity virtual space, then we have a level 1 (palm completely open), if the semiperimeter is within the [150;100) value range, then we have level 2, and if it is smaller than 100, we have level 3, i.e. a closed fist. For the hand rotation gesture, we took into consideration the angle formed by the coordinates of the thumb, a point situated in the origin (coordinates 0, 0, 0) and the y axis within the virtual 3D space. Therefore, for level 1 of the rotation gesture, the angle must be smaller than 75 degrees, for level 2 we have the [75;100] range, while for level 3, the angle must be larger than 100 degrees.

The scripts for collecting the data were applied in the applications described in [7] and [8]. We have collected approximately 10k data for each of the 29 types of data describing the gestures. Python scripts were written for the gesture classification, using libraries as: *sklearn*, *pandas*, *numpy*, and *matplotlib*.

In order to consolidate the results, we followed a comparison using two models for the gesture classification: linear-SVM and MLP. Linear-SVM is used for data that may be classified into 2 classes, which means that the data may be divided linearly. Although we have 6 classes, this model may be used because such classes describe 3 levels of the same hand rotation gesture, respectively the closing and opening of the palm. The Multi-layer Perceptron (MLP) is a supervised learning algorithm, where an $f(\cdot) : R^m \rightarrow R^o$ function is learned by training on a dataset with an *m* number of dimensions for the input data and an *o* number of dimensions for the output data [9]. For MLP we use 100 hidden layers, 1000 iterations, and the alpha regularization (L2 regularization) parameter that was set to 1.

3. Results

The SVM algorithm had the best results for the data classification with a gesture classification accuracy of 99.4%. The data for the gesture classification were divided in 80% training data and 20% testing data. The confusion matrix (Figure 1.a.) shows how gestures were classified. Out of a total of 788 cases representing gesture 1 (the palm closing and opening gesture), 784 gestures were classified correctly. Level 2 of the same gesture was detected with an accuracy of 100% (288 cases), while level 3 was detected correctly in 527 cases and in 2 cases a confusion was made with level 2 of the same gesture. The hand rotation gesture (gesture 2) on level 1 was classified correctly in 174 cases, while it was only once mistaken for level 1 of gesture 1. Such level of gesture 2 has the best results for classification based on the number of cases and the correct classification thereof. For level 2 of gesture 2, 3 cases were detected incorrectly out of a total of 33 cases, while for level 3 of gesture 2, a case was detected incorrectly out of a total of 56 cases.

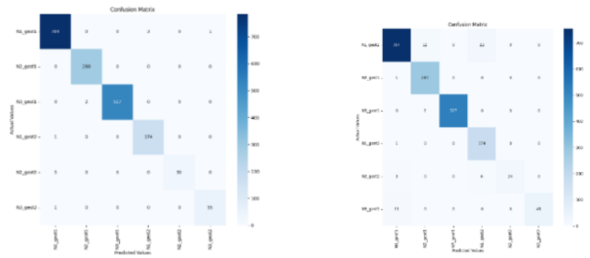

Figure 1. a. Confusion Matrix for linear-SVM model **Figure 1. a.** Confusion Matrix for MLP model

Using the second classification model, MLP, an accuracy of 96.2% was achieved for the gesture classification. The data were divided the same way as the data in the first model, the Linear-SVM, 80% training data and 20% testing data. In the confusion matrix (Figure 1. b) the level 1 for the gesture 1 was classified correctly in 754 cases, while in 34 cases it was mistaken. In 12 cases a confusion was made with level 2 of the same gesture and in 24 cases a confusion was made with level 2 of gesture 2. The best results were achieved for the classification of gesture 1 on level 2. Only one case out of 288 cases was classified incorrectly. On level 3 for the gesture 1, 2 cases out of a total of 529 cases were classified incorrectly. For the second gesture, on level 1 we have only one case incorrectly classified out of 175 cases. On levels 2 and 3 of the same gesture, we have the lowest performance. Hence out of a total of 89 cases, 20 cases were incorrectly classified, as being another gesture or on another level of gesture 2.

Table 1 presents the values for the following metrics: precision, recall, F1-score, and the total number of testing cases (support).

Table 1. Metrics of the two models: Linear-SVM and MLP

Model	Gesture	Precision	Recall	F1-score	Support
	N1_gest1	0.99	0.99	0.99	788
Linear-SVM	N2_gest1	0.99	1.00	1.00	288
	N3_gest1	1.00	1.00	1.00	529
	N1_gest2	0.98	0.99	0.99	175
	N2_gest2	1.00	0.91	0.95	33
	N3_gest2	0.98	0.98	0.98	56
MLP	N1_gest1	0.98	0.96	0.97	788
	N2_gest1	0.95	1.00	0.97	288
	N3_gest1	1.00	1.00	1.00	529

N1_gest2	0.86	0.99	0.92	175
N2_gest2	1.00	0.73	0.84	33
N3_gest2	1.00	0.80	0.89	56

4. Conclusions

This paper presents a comparison when using two data classification methods of the LM gestures. It resulted that the Linear-SVM model performs better as compared to the MLP model and has a higher classification performance than the latter. The results show that the Linear-SVM model has a gesture classification accuracy of 99.4%, while the MLP model has an accuracy of 96.2%. These two models are effective as compared to the literature [10], [11] and they may be used in gesture classification.

Since we have few cases where the gestures were incorrectly classified on both models, they may be taken into consideration as optimal models for the classification of these dynamic gestures. Moreover, the models have high performances because the dataset is quite large, and the manner of detection of the gestures depends on all the input data described in chapter 2. Compared to the results presented in the literature in the previous chapters, we perform better at classifying dynamic recovery gestures. As future work we aim at using other models such as: Logistic Regression, Linear Discriminant Analysis or Random Forest to have an even better comparison among the different classification models and to select the best ones for the best results in clinical practice.

References

[1] Kavian M, Nadian-Ghomsheh A. "Monitoring Wrist and Fingers Range of Motion using Leap Motion Camera for Physical Rehabilitation", International Conference on Machine Vision and Image Processing (MVIP), 2020, p. 1-6.
[2] Hariadi RR, Kuswardayan I. "Design and implementation of Virtual Indonesian Musical Instrument (VIMi) application using Leap Motion Controller", International Conference on Information & Communication Technology and Systems (ICTS), 2016, p. 43-48.
[3] Caro L, Villablanca E, Peralta B. "A Virtual Reality Application using Hand Motion Tracking with Leap Motion Sensor", 39th International Conference of the Chilean Computer Science Society (SCCC), 2020, p. 1-6.
[4] Dhanasree KS, Nisha KK, Jayakrishnan R. "Hospital Emergency Room Training Using Virtual Reality and Leap Motion Sensor", Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, p. 924-928.
[5] Dzikri A, Kurniawan DE. "Hand Gesture Recognition for Game 3D Object Using The Leap Motion Controller with Backpropagation Method", International Conference on Applied Engineering (ICAE), 2018, p. 1-5.
[6] Rakesh S, Kovács G, Mokayed H, Saini R, Pal U. "Static Palm Sign Gesture Recognition with Leap Motion and Genetic Algorithm", Swedish Artificial Intelligence Society Workshop, 2021, p. 1-5.
[7] Nicola S, Chirila OS, Stoicu-Tivadar L. "Gesture Classification for a Hand Controller Device Using Neural Networks", Studies in Health Technology and Informatics, 2020, Vol. 270, p. 756-760.
[8] Nicola S, Chirila OS, Stoicu-Tivadar L. "Enhancing Precision in Gesture Detection for Hand Recovery After Injury Using Leap Motion and Machine Learning", Studies in Health Technology and Informatics, 2020, Vol. 262, p. 320-323.
[9] Huang KY, Shen LC, Jiun-Der You. "Proof of hidden node number in MLP and experiments on well log data inversion", International Joint Conference on Neural Networks (IJCNN), 2015, p 1-6.
[10] Naglot D, Kulkarni M. "Real time sign language recognition using the leap motion controller", International Conference on Inventive Computation Technologies (ICICT), 2016, p. 1-5
[11] Li W, Hsieh C, Lin L, Chu W. "Hand gesture recognition for post-stroke rehabilitation using leap motion", International Conference on Applied System Innovation (ICASI), 2017, p. 386-388