Dimensionality Reduction in Hand Gesture Classification

Hara Kumar
harak@kth.se

Vikash Dodani
dodani@kth.se

Balint Kovacs
balintk@kth.se

***Abstract*—An apparatus containing five sensors, one for each finger on the human hand, records position data for various gestures. Using this data, a k-NN classifier is built to distinguish different gestures. The nearest neighbors are calculated using both euclidean and DTW. The experiment compares the accuracy of the models with and without dimensionality reduction. The results show a small drop in accuracy, but careful selection of parameters can mitigate some these losses. On the positive, dimensionality reduction saw a significant improvement in classification speed.**

I. Introduction

This paper explores whether dimensionality reduction can be applied on data coming from a hand gesture apparatus to increase the accuracy of identifying hand gestures using sensors located on the hand. For this paper, an apparatus containing a sensor on each finger was used. Each of these sensors record the pitch, yaw, and roll of the finger, giving a rough idea of how the finger is positioned in 3D space. Please refer to Figure 1. Existing research shows that gloves embedded with sensors is a viable way to record hand gestures [8]. Existing research also shows that this data can be used to identify gestures for the purposes of sign language translation, among other potential uses [9].

The data coming from such an apparatus, however, exists in high dimensional space. A single moment in time is 15 dimensions (3 positions for 5 fingers) and a single gesture can take place over a period of time. Often reducing the number of dimensions in high dimensional problems like this one results in an improvement to the accuracy of the classifier [1] [2] [3] [6]. This paper will compare two dimensionality reductions techniques: Principle Component Analysis and Partial Least Squares. The accuracy of models using these dimensional reductions will be compared to a control that uses no dimensionality reduction. This data then will be used by a k-NN classifier with euclidian distance, and an other k-NN classifier combined with Dynamic Time Warping.

II. Method

A. Hand Gesture Data

A sample of 37 gestures were gathered. The sample contains 13 handshakes, 13 waves (as in waving someone hello), and 11 Italian "what?" gestures. The number of features in a sample is 1005 after applying all preprocessing. The data was divided into test and training samples using 10-fold cross validation. However, due to the small sample size, this results in test sets containing only 3-4 samples. Thus using dimensionality reduction on the test set separately results in poor results. To mitigate this, the algorithm defines the lower dimensional space using the training data, and applies these dimensions both on the training data set and the test data set.

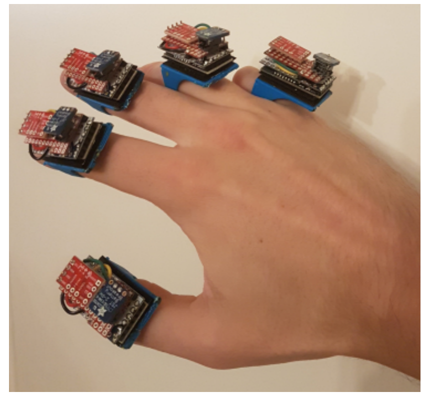

Fig. 1. Sensors used to collect gesture data.

The sensors were programmed to send data at 100ms intervals, but in reality the data did not arrive consistently. In some data sets, the starting position of a sensor was record nearly 500ms after another sensor on different finger started sending data. Furthermore, the interval of data ranged well below and above the 100ms setting. The data had to be processed so that each sample has 15 dimensions of data at 100 second intervals, starting and ending at the same time.

B. Algorithms

The classification algorithm used for this paper is k-NN. The k-NN algorithum was custom built, and did not use existing python packages. K values of 1, 3, and 5 were used in an un-weighted majority-vote classifier. Various types of weighted voting systems were researched, but ultimately an un-weighted voting system was settled on to isolate the effect of dimensionality reduction [4] [5] [7] [10]. To compare the accuracy of the control, PCA, PLS, and DTW models, the average Brier score and accuracy are calculated for each trial using 10-fold cross validation. Existing Python packages were used for these models.

III. Results

The results of the classifiers built without applying any dimensionality reduction method can be seen in Figure 3.

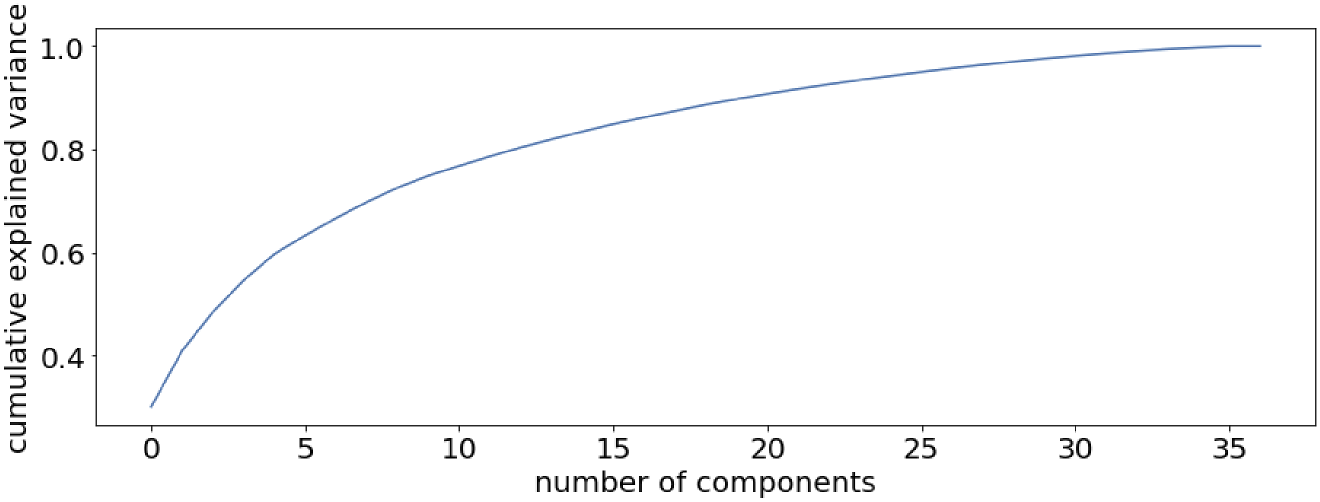

Fig. 2. Plotting the cumulative variance in context of the number of components.

Dynamic Time Warping performs slightly better, however, both classifiers are in the range of 90% accuracy. On the other hand, comparing the avarage time each classifier took to complete, we see a significant difference: dynamic time warping took on average more than 12 seconds to complete, while k-NN completed in 0.01 second on average. This clearly motivates the need to apply dimensionality reduction on the data.

		Accuracy	Brier score
euclidean	1	0.891892	0.216216
	3	0.891892	0.210210
	5	0.891892	0.216216
dtw	1	0.972973	0.054054
	3	0.945946	0.078078
	5	0.918919	0.108108

Fig. 3. Performance of classifiers without applying any dimensionality reduction.

The performance of the classifiers with data reduced by PCA can be seen in Figure 3. By plotting the cumulative explained variance in context to the number of components, we can see that by reducing to 20 components, 90% of the total variance is kept. This can be seen in Figure 2. Based on this, we projected our data into 20 components using PCA.

The drop of accuracy and increase of Brier score is less significant on the k-NN classifier, but it is high in the other case. However, the time the classifier using dynamic time warping took, is much less when PCA is applied on the data, it is under 0.5 seconds on average.

As can be seen in Figure 5, PLS leads to an even higher decrease in performance when compared to PCA. The performance regarding time is similar to PCA.

		Accuracy	Brier score
euclidean	1	0.891892	0.216216
	3	0.891892	0.210210
	5	0.891892	0.229189
dtw	1	0.837838	0.324324
	3	0.918919	0.210210
	5	0.783784	0.354595

Fig. 4. Performance of classifiers with PCA.

		Accuracy	Brier score
euclidean	1	0.837838	0.324324
	3	0.864865	0.210210
	5	0.864865	0.218378
dtw	1	0.783784	0.432432
	3	0.837838	0.342342
	5	0.783784	0.395676

Fig. 5. Performance of classifiers with PLS.

IV. Discussion

To conclude, dimensionality reduction leads to a significant decrease in the time the classifiers using dynamic time warping take to complete. This comes at a cost of accuracy. However, with careful selection of parameters, these costs can be mitigated, as can be seen in the case of 3 nearest neighbours. Nearest neighbour with euclidean distance function was not sensitive to the reduction of dimensions in the data.

Given that the number of samples was much less than the number of features, when applying dimensionality reduction algorithms, we have lost a lot of information. For further research, it would be optimal to have as many samples as features.An other improvement could be using a weight function for the k-NN classifier.

The next step in the development of the device is to make it to be able to scan a real time stream of positional data and find where the positional data indicates a hand gesture. To achieve this, we need fast classifiers with high accuracy. We will also have a lot more classes in the future. The time a classifier takes to make a prediction should be in the focus in the future of the project besides accuracy.

REFERENCES

[1] Archanah., Telgaonkar, and Deshmukh Sachin. Dimensionality Reduction and Classification through PCA and LDA. International Journal of Computer Applications, vol. 122, no. 17, 17 July 2015, pp. 48., doi:10.5120/21790-5104.

[2] Deegalla, Sampath, and Henrik Bostrm. Classification of Microarrays with KNN: Comparison of Dimensionality Reduction Methods. Intelligent Data Engineering and Automated Learning - IDEAL 2007 Lecture Notes in Computer Science, pp. 800809., doi:10.1007/978-3-540-77226-280.

[3] Deegalla, Sampath, and Henrik Bostrm. Improving Fusion of Dimensionality Reduction Methods for Nearest Neighbor Classification. 2009 International Conference on Machine Learning and Applications, 2009, doi:10.1109/icmla.2009.95.

[4] Gou, Jianping, et al. A Novel Weighted Voting for K-Nearest Neighbor Rule. Journal of Computers, vol. 6, no. 5, 3 May 2011, doi:10.4304/jcp.6.5.833-840.

[5] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, Kieran Greer. "KNN Model-Based Approach in Classification." Unpublished.

[6] Hany Yan, Hu Tianyu. "Unsupervised Dimensionality Reduction for High-Dimensional Data Classification." Machine Learning Research. Vol. 2, No. 4, 2017, pp. 125-132., doi:10.11648/j.mlr.20170204.13.

[7] Klaus Hechenbichler, Klaus Schliep. "Weighted k-Nearest-Neighbor Techniques and Ordinal Classification." Sonderforschungsbereich 386. 13 October 2004.

[8] Mehdi, S.a., and Y.n. Khan. Sign Language Recognition Using Sensor Gloves. Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., 1 Dec. 2002, doi:10.1109/iconip.2002.1201884.

[9] Radzi Ambar, Chan Kar Fai, Chew Chang Choon, Mohd Helmy Abd Wahab, Muhammad Mahadi Abdul Jamil and Ahmad Alabqari Ma'Radzi (2018) Preliminary Design Of A Dual-Sensor Based Sign Language Translator Device, Recent Advances On Soft Computing And Data Mining, Springer, Cham , 353, ISBN:978-3-319-72550-5

[10] Wei Zheng, HaiDong Wang, Lin Ma, RuoYi Wang. "An Improved k-Nearest Neighbor Classification Algorithm Using Shared Nearest Neighbor Similarity." Metallurgical and Mining Industry. No. 10, 2015, pp. 133-137.