**Research Review**

By Jacob Kennedy and Connor Maynes

1. INTRODUCTION

The entire purpose is feature extraction and classification of elbow flexing, specific finger movements, and wrist movement. The methods of independent component analysis ( ICA ) and fast fourier transforms ( FFT ) were used for feature extraction, while convolutional neural networks ( CNN ) and support vector machines ( SVM ) were used for classification. The goal of this research was to produce an algorithm capable of classifying these seven actions at an accuracy of 90-95% and to do so in real-time.

1. LITERATURE SURVEY
   1. GENERAL STATEMENTS

Rapid classification of sEMG signals could allow more accurate and functional prosthetics that require the minimum amount of training to start and to maintain the same level of accuracy despite displacement of the prosthetic on the body or other changes in the subject. Currently, the path to the aforementioned goal is likely through Blind Source Separation techniques that will create a set of data for robust and complex machine learning algorithms capable of handling such real-world scenarios to quickly ascertain specific movements based on their generated classifications. The design of machine learning systems is challenging and often progress is only made through trial and error of many different designs, but with each iteration the progress of the entire scientific community steps forward.

A CNN or SVM for classifying 4 to 12 actions should have an accuracy of about 90-95% [10]. Recent work with CNNs combining the techniques of using a 300 ms timeframe, pruning, and a hyperbolic tangent function for the action function in the neural network have proven highly effective at increasing classification accuracy [11-15]. SVMs have been shown to outperform multi-layer perceptrons [18]. SVMS using methods of multiple kernel learning ( MKL ), the one-against-one method, and linear kernels have been shown to produce accuracies ranging from 90-95% and to outperform ANNs and linear discriminant analysis ( LDA ) [21-23, 27]. For the aforementioned reasons, a combination of a CNN and SVM should yield similar classification accuracy.

* 1. FEATURE EXTRACTION - ICA & FFT

Fast Fourier Transform is an analytically method that extrapolates frequency components from a time domain-based signal. FFT by itself does not change the information that is trying to be sent but coupled with filtering allows for random noise or unnecessary information that is in a different frequency band to be ignored. FFT is an extremely useful tool especially when dealing with biopotentials since the biopotentials are clustered and hard to differentiate in the time domain [29-31].

ICA or Independent Component Analysis stems from a type of processing called BSS or Blind Source Separation [32]. ICA extracts statistically independent sources called independent components, or ICs from the mixed signals given [33]. There are a number of different ICA algorithms and they can be broken up into two general methods, minimizing mutual information, or the maximization of non-Gaussianity [32][34]. Both of these general methods work in terms of separating a source signal from a mixed signal, and both try to generate statistically independent matrices known as unmixing matrices [32]. This unmixing matrix is the product of the ICA and is a multi-dimensional matrix that consists of the statistically independent source signals and features that then can be used in post processing [35].

* 1. CLASSIFICATION - CNN & SVM

Pattern recognition has shown promising results in recent years for classifying hand movements from accelerometer data; however, these methods do not provide robust natural control [1-5]. A recent study used a CNN consisting of an input layer, four convolutional layers, four subsampling layers, and two fully-connected layers, both with a trained and untrained classifier. The results indicated better accuracy from the CNN versus the SVM and that that network design is somewhat user-adaptive [6]. A user-adaptive design reduces training time for each user by generalizing part of the training results, so they can be used across users.

The accuracy of any surface electromyography ( sEMG ) classification system is heavily dependent on the number of actions being classified with accuracy levels of 90-95% for 4 to 12 actions and 60-70% for 50 or more actions [10]. Before the data can be passed through the preprocessing and feature extraction sections of a classification system, an appropriate window size for the time-series data must be selected. Numerous time-windows have been attempted, often ranging from 100 to 125 ms or 280 to 300 ms [13-15]. The first time-window range of 100 to 125 ms results in lower classification latency, but often at the cost of accuracy, while the inverse is true of the second range. An additional method of affecting the efficiency of a CNN is through pruning. A pruning approach, whereby the network is trained, unimportant connections are pruned, and then the network is retrained, has proven to dramatically reduce the number of parameters in the network by as much as a factor of ten [11]. The activation function is also important in the efficiency of a CNN. The hyperbolic tangent function converges faster than the sigmoid or the the rectified linear unit ( ReLU ) and thus reduces training time [12].

Multi-layer perceptrons ( MLPs ) have been compared against the performance of SVMs and shown to be outperformed by SVMs [18]. Known problems relating to the use of SVMs are the memory requirements of the computer used to train them and the time-cost of training them [19-20]. A combination of multiple kernel learning ( MKL ) and SVM has been attempted and resulted in accuracy above 95% [21]. A linear kernel tends to produce higher accuracy over other kernels [22]. A SVM has been shown to provide better classification accuracy compared to ANNs and linear discriminant analysis ( LDA ) [23]. Another known problem with SVMs is that classification performance degrades over time by as much as 14.6%, when the SVM is not recurrently trained [24]. Time-frequency domain features, when combined with an SVM, provide superior accuracy over SVMs fed with time domain features [25]. An approach called incremental online SVM, whereby only the samples closest to the current boundary are used to train the classifier, resulted in a dramatic improvement in accuracy over traditional SVM [26]. The one-against-one method has proven effective when working with high-dimensional vector spaces and SVMs [27]. In the one-against-one method, a binary classifier is created for each pair of classes using SVM. All of the binary classifiers are executed and then the class with the highest number of votes is selected.

CNNs and SVMs have consistently shown higher classification accuracy over many alternatives independently. The combination of CNNs and SVMs has rarely, if ever, been used for sEMG signals, but has proven highly accurate in image processing and should provide comparable results of 90-95% for the seven actions that are being classified [28]. The problems of an SVMs accuracy dropping over time should be mitigated by the use of a CNN which has smaller drops in accuracy over time.

1. PREPROCESSING & LEARNING
   1. PREPROCESSING TECHNIQUES

FFT is a very simple technique to instantiate, MATLAB already has prebuilt functions that will automatically cycle through the gathered data and present the data’s frequency spectrum plot. The data gathered and the signals within the data become far more apparent in the frequency spectrum since the signals, though still very intertwined, can exist in different frequencies. This creates an opportunity to manually look at the data and create filters based on keeping desired frequency components, and filtering out undesired components, such as 60 Hz noise, and DC.

With mutual information and noise taken out, the data that exists is now cleaned and ready to be analyzed by use of ICA. As stated previously, there are many different ICA algorithms, each with given weaknesses and advantages, but for this specific project, and the fact that this project is desired to be as instantaneous as possible in terms of processing, this data will be processed by MATLAB’s FastICA package. This package requires certain information, such as number of independent components, and a matrix consisting of [n samples, dimensional data]. The much more general ICA package requires more information such as period of independent components and signal to noise ratio, which may generate slightly more accurate unmixing matrix, but would take longer. Once the unmixing matrix has been created pre-processing is complete.

* 1. LEARNING METHODS

A CNN is a class of ANN, which consists of convolution, pooling, fully-connected ANN(s), and normalization layers. Convolution can be thought of as a filter - a matrix of numbers - sliding around an input matrix, stopping after each slide to perform a dot product and then a sum of those products. After the filter has slid across the entire input matrix, the feature map that contains all of the sums of the dot products is complete. The number of feature maps generated is called the depth of the convolution. The feature map represents extracted features as a result of the filtering. The quantity of cells the filter slides over with each move is called the stride. The larger the stride, the smaller the feature map. One additional concept in convolution is called zero-padding, where zeroes of some width coat the outside of the input matrix. Zero-padding helps with feature extraction on the bordering elements. Convolution using zero-padding is called wide convolution, while convolution not using it is called narrow convolution. Convolution is often followed by running every datapoint through a rectified linear unit ( ReLU ) function, which outputs 0 if the input is less than 0 and outputs the value otherwise.

The next step in convolution is pooling, or subsampling, which reduces the feature map while retaining the most important information that the convolution process extracted. There are many different types of pooling, such as max, average, sum, and others. The pooling operation consists of defining a spatial neighborhood - a relatively small sliding window over the feature maps - and then breaking the feature maps up into a bunch of these spatial neighborhoods and performing some calculation, such as max, sum, or average, on the data points in each neighborhood. Pooling reduces the feature space and thus decreases the likelihood of overfitting the model to the training data. The pattern of convolution, ReLU, and pooling may be repeated multiple times with the output of one feeding into the input of the next. All three of the aforementioned steps are not required and can be mixed and matched to suit the situation.

Finally, the CNN ends with one or more fully-connected layers or ANN’s. A ANN consists of layers of nodes, with the nodes of one layer connected to all nodes of the following layer. A bias node is often included in each layer in order to avoid a node’s weights getting stuck at zero once and then remaining dead for the rest of training. Each node takes the sum of its inputs and passes it through a nonlinear function called the activation function. The function must be non-linear, so that during the process of backpropagation, whereby the network learns, the network will converge to a classification of the input. The most common types of activation function are sigmoid, which normalizes the input to a range between zero and one, hyperbolic tangent, which normalizes the input to a range between -1 and 1, and ReLU. Each node has a weight assigned to it that varies as the network learns to either increase or decrease the strength of a signal at a connection. Each node also has a threshold that will either enable or disable the node from sending its output. Learning, or gradient descent through backpropagation, amounts to varying the node weights and the output thresholds based on feedback generated by comparing what the network predicted to what it actually was. The error between the predicted and actual values is calculated at the output nodes and then propagated back through the network, applying the weights of the connections between the nodes along the way and summing the backwards-propagated error for each node. Once the errors have been propagated, the weights are recalculated in the network, using the combination of the original weight, a learning coefficient that affects how quickly the system converges, the error for each node, and the derivative of the activation function of the node evaluated at its original input value. All of the filter values, parameters ( depth, stride, and zero-passing ), and weights are taught the appropriate values during the learning process. The input layer does nothing more than pass the data along to the next layer. The output layer often uses either a softmax function or a support vector machine ( SVM ) as its activation function. The softmax function is used because it will produce a set of output probabilities whose sum equals one. The output node with the highest value is the most probable answer and should be selected [16].

A support vector machine ( SVM ) takes in labeled training data and outputs an optimal hyperplane capable of classifying new inputs. More complex data forms may require transformation of the data in order to separate them, by projecting the points to a higher dimension. By tuning the regularization and gamma parameters of the SVM, a more accurate classifier can be produced.

1. EXPECTED RESULTS

This goal of this research is to classify 7 different gestures. Based on the literature survey, an accuracy of 90-95% is expected when using a CNNs or SVMs. A functional design for real-time classification of sEMG signals using FFT, ICA, CNN, and SVM is expected. A latency of at most 300 ms is expected during testing of the classification algorithm based on the results of previous work.

1. APPROACH
2. RESULTS
3. DISCUSSION
4. REFERENCES

[1] Došen, S., Cipriani, C., Kostić, M., Controzzi, M., Carrozza, M. C., and Popović, D. B. (2010). Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation. J. Neuroeng. Rehabil. 7:42. doi: 10.1186/1743-0003-7-42

[2] Markovic, M., Dosen, S., Cipriani, C., Popovic, D., and Farina, D. (2014). Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. J. Neural Eng. 11:46001. doi: 10.1088/1741-2560/11/4/046001

[3] Atzori, M., Gijsberts, A., Müller, H., and Caputo, B. (2014c). “Classification of hand movements in amputated subjects by sEMG and accelerometers,” in Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (Chicago, IL), 3545–3549.

[4] Gijsberts, A., Atzori, M., Castellini, C., Muller, H., and Caputo, B. (2014). The movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification. IEEE Trans. Neural Syst. Rehabil. Eng. 22, 735–744. doi: 10.1109/tnsre.2014.2303394

[5] Krasoulis, A., Vijayakumar, S., and Nazarpour, K. (2015). “Evaluation of regression methods for the continuous decoding of finger movement from surface EMG and accelerometry,” in Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference (Montpellier), 631–634.

[6] Park, K., and Lee, S. (2016). “Movement intention decoding based on deep learning for multiuser myoelectric interfaces,” in 2016 4th International Winter Conference on Brain-Computer Interface (BCI), 1–2.

[7] Fukuda, O., Tsuji, T., Kaneko, M., Otsuka, A., and Tsuji, O. F. T. (2003). A human-assisting manipulator teleoperated by EMG signals and arm motions. IEEE Trans. Robot. Autom. 19, 210–222. doi: 10.1109/tra.2003.808873

[8] Li, G., Schultz, A. E., and Kuiken, T. A. (2010). Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses. IEEE Trans. Neural Syst. Rehabil. Eng. 18, 185–192. doi: 10.1109/TNSRE.2009.2039619

[9] Tenore, F. V. G., Ramos, A., Fahmy, A., Acharya, S., Etienne-Cummings, R., and Thakor, N. V. (2009). Decoding of individuated finger movements using surface electromyography. IEEE Trans. Biomed. Eng. 56, 1427–1434. doi: 10.1109/TBME.2008.2005485

[10] Atzori, M., Cognolato, M., & Müller, H. (2016). Deep learning with convolutional neural networks applied to electromyography data: a resource for the classification of movements for prosthetic hands. Frontiers in neurorobotics, 10, 9.

[11] Han, S., Pool, J., Tran, J., & Dally, W. (2015). Learning both weights and connections for efficient neural network. In Advances in neural information processing systems (pp. 1135-1143).

[12] Allard, U. C., Nougarou, F., Fall, C. L., Giguère, P., Gosselin, C., Laviolette, F., & Gosselin, B. (2016, October). A convolutional neural network for robotic arm guidance using sEMG based frequency-features. In Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on(pp. 2464-2470). IEEE.

[13] B. Hudgins, P. Parker, R. Scott, "A new strategy for multifunction myoelectric control", IEEE Transaction on Biomedical Engineering, vol. 40, no. 1, pp. 82-94, 1993.

[14] T. Farrell, R. Weir, "The optimal controller delay for myoelectric prostheses", IEEE Transaction on Neural systems and rehabilitation engineering, vol. 15, no. 1, pp. 111-118, 2007.

[15] B. Peerdeman, D. Boere, H. Witteveen, R. H. in ‘t Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, S. Misra, "Myoelectric forearm prostheses: State of the art from a user-centered perspective", Journal of Rehabilitation Research & Development, vol. 48, no. 6, pp. 719-738, 2011.

[16] Yao, X. (1999). Evolving artificial neural networks. Proceedings of the IEEE, 87(9), 1423-1447.

[17] “Convolutional Neural Network.” Unsupervised Feature Learning and Deep Learning Tutorial, ufldl.stanford.edu/tutorial/supervised/ConvolutionalNeuralNetwork/.

[18] N.F. Guler, S. Koçer, Classification of EMG signals using PCA and FFT, J. Neurosci. Methods, Vol. 29, Iss. 3, 2005, 241-250

[19] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov., Vol. 2, 1998, 121-167

[20] G. Horvath, J. Suykens, S. Basu, Advances in Learning Theory: Methods Models And Applications, 2003, IOS Press, Amsterdam

[21] Chen, X., & Wang, Z. J. (2013). Pattern recognition of number gestures based on a wireless surface EMG system. Biomedical Signal Processing and Control, 8(2), 184-192. 10.1016/j.bspc.2012.08.005

[22] Stango, A., Negro, F., & Farina, D. (2015). Spatial correlation of high density EMG signals provides features robust to electrode number and shift in pattern recognition for myocontrol.IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23(2), 189-198. 10.1109/TNSRE.2014.2366752

[23] M.A. Oskoei, H. Hu  
Support vector machine-based classification scheme for myoelectric control applied to upper limb  
IEEE Trans. Biomed. Eng., 55 (2008), pp. 1956-1965

[24] P. Kaufmann, K. Englehart, M. PlatznerFluctuating EMG signals: investigating long-term effects of pattern matching algorithms

Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires(2010), pp. 6357-6360

[25] Kamali, T., Boostani, R., & Parsaei, H. (2014). A multi-classifier approach to MUAP classification for diagnosis of neuromuscular disorders. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22(1), 191-200. 10.1109/TNSRE.2013.2291322

[26] Xu, X., Zhang, Y., Luo, Y., & Chen, D. (2013). Robust bio-signal based control of an intelligent wheelchair. Robotics, 2(4), 187-197. 10.3390/robotics2040187

[27] C. Hsu, C. Lin, "A comparison of methods for multi-class support vector machines", IEEE Trans. Neural Netw., vol. 13, no. 2, pp. 415-425, 2002.

[28] Agarap, A. F. (2017). An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification. *arXiv preprint arXiv:1712.03541*.

[29] Tidwell, R., Akumalla, S., Karlaputi, S., Akl, R., Kavi, K., & Struble, D. (2013, July). Evaluating the feasibility of EMG and bend sensors for classifying hand gestures. In Proceedings of the International Conference on Multimedia and Human Computer Interaction (Vol. 63, pp. 1-8).

[30] Cannan J, Hu H. Automatic user identification by using forearm biometrics. IEEE; 2013.

[31] Oliver, A. S., Maheswari, N., & Samraj, A. (2014). Stable and critical gesture recognition in children and pregnant women by SVM classification with FFT features of signals from wearable attires. Research Journal of Applied Sciences, Engineering and Technology, 7(23), 4917-4926.

[32] Zhiqiang Zhang, Huihui Li, Danilo Mandic (2016). Blind Source Seperation and Artefact Cancellation for Single Channel Bioelectrical Signal. Wearable and Implantable Body Sensor Networks (BSN).

[33] Bogdan Mijovic, Maarten De Vos, Ivan Gligorijevic, Joachim Taelman, Sabine Van Huffel (2010). Source Separation From Single-Channel Recordings by Combining Empirical-Mode Decomposition and Independent Component Analysis. IEEE Transactions on Biomedical Engineering (Vol. 57, Issue 9, pp 2188-2196).

[34] José AG Mendes, Ricardo R. Robson, Sofiane Labidi, Allan Kardec Barros (2008). Subvocal Speech Recognition Based on EMG Signal Using Independent Component Analysis and Neural Network MLP. Image and Signal Processing, ISBN: 970-0-7695-3119-9.

[35] Marwa Farouk Ibrahim Ibrahim, Adel Ali Al-jumaily (2016). ICA based feature learning and feature selection. Electronic Devices, Systems and applications (ICEDSA). IEEE ISSN: 2159-2055