# Real- time visual and EMG signals recognition to control dexterous prosthetic hand based on deep learning and machine learning

**Noof T. Mahmood[1], Mahmuod H. Al-Muifraje[2], Sameer K. Salih[3]**

[1]Department of Electrical Engineering, University of Technology, Iraq
[2] Department of Computer Techniques Engineering, Al- Rasheed University Collage, Iraq
[3] Ministry of Sciences and Technology, Iraq

## ABSTRACT

The revolution in prosthetic hands allows the evolution of a new generation of prostheses that increase artificial intelligence to control an adept hand. A suitable gripping and grasping action for different shapes of the objects is currently a challenging task of prosthetic hand design. The most artificial hands are based on electromyography signals. A novel approach has been proposed in this work using deep learning classification method for assorting items into seven gripping patterns based on EMG and image recognition. Hence, this approach conducting two scenarios;

The first scenario is recording the EMG signals for five healthy participants for the basic hand movement (cylindrical, tip, spherical, lateral, palmar, and hook). Then three time-domain (standard deviation, mean absolute value, and the principal component analysis) are used to extract the EMG signal features. After that, the SVM is used to find the proper classes and achieve an accuracy that reaches 89%.

The second scenario is collecting the 723 RGB images for 24 items and sorting them into seven classes, i.e., cylindrical, tip, spherical, lateral, palmar, hook, and full hand. The GoogLeNet algorithm is used for training based on 144 layers; these layers include the convolutional layers, ReLU activation layers, max-pooling layers, drop-out layers, and a softmax layer. The GoogLeNet achieves high training accuracy reaches 99%. Finally, the system is tested, and the experiments showed that the proposed visual hand based on the myoelectric control method (Vision-EMG) could significantly give recognition accuracy reaches 95%.

| **Keywords**: | EMG, RGB, GoogLeNet, SVM, ReLU. |
| --- | --- |

*Corresponding Author:*

Noof T. Mahmood
Department of Electrical Engineering
University of Techonlogy, Iraq
E-mail: noofthabit@gmail.com

## 1. Introduction

Physical therapy plays a significant role in the rehabilitation stage for the people who suffer from an amputee. Thus, many efforts are made to elucidate the leverage of medical/ clinical and human-machine interface (HMI) applications [1]. Therefore, robotic has the effectiveness to increase the independence of the individuals living lifestyle with their disabilities. This work aims to improve life quality by empowering people to achieve a wide range of daily responsibilities within few minutes. The most widely used are robotic hands and arms because the robotic hands would have the ability to achieve the primary skills, like grasping and transferring objects from one place to another, similar to what non-amputees individuals do [2].

At present, designing a prosthetic hand-based myoelectric control plays a dominating role in the re-habitation side. This requires recording for surface electromyography (EMG) signals; this technique follows up electrical activity associated with skeletal muscles [3]. The muscles contain some motor unit action potentials (MUAPs) to record grasping and clipping actions [4]. The combination of muscle fiber action potentials from all the muscle fibers of a single motor unit is called the MUAP [5].

During clipping or grasping; the tactical information, electromyography information and visual information are involved, where the visual part occupies 83% of all information [6]. The visual feedback helps the people compare the size of the hand and the target object. In that way, adjusting occurs when the grasping gestures and the aperture size match the target object. For this reason, the visual information is achievable to obtain the necessary control parameters for the artificial prosthetic hand [7]. Also, the technology of combining more than one information source will provide an efficient way for the control of smart prostheses [8], [9]. For increasing human-machine interaction, the visual control system is used to fill the gaps of designing prostheses based on EMG signals. The artificial vision system based on image recognition can obtain the features of the object, for example, the shape and the size of the required item [10].

After collecting the information, EMG or visual information, the artificial intelligence (AI) is used to extract the main feature and classify it using the Machine Learning algorithms (ML) or Deep Learning (DL) algorithms [11].

ML is a learning algorithm to train random data in an intelligent way [12]. The learning algorithms are classified into supervised and unsupervised. Supervised learning occurs when a group of known data and the algorithm is trained for a specific function. Meanwhile, unsupervised learning occurs when the inputs are given, and the algorithm learns to find features or patterns to produce the output [13].

Meanwhile, DL is the trending algorithm due to the algorithm structure that contains a deep network (multiple hidden layers) that learns different features with multiple levels [14]. DL problem is summarized in the hierarchy of concept, where each concept is constructed on top of the others. Therefore, the lower layers are considered as the primary representation of the problem [15]. Besides, the consecutive layers can be learned via sub-models that are organized in layers stacked. The main problem in the ML is the feature extraction; this problem has been solved in DL due to DL's capability to learn the useful features by itself. Utilized and unsupervised DL models have been grown so fast due to their achievement in solving complex problems [16].

The novelty of this work is satisfied by sorting the RGB images of the objects into the six basic hand movements using deep learning (GoogLeNet) to improve the operability of the prosthetic hand by designing a pattern recognition for a prosthetic hand by using the Myo armband, also supporting the prosthetic hand by the visual recognition to distinguish the objects. The Myo armband is used to collect the EMG signal; then, the signal is classified into the six basic hand movements (cylindrical, tip, spherical, lateral, palmar, and hook) using machine learning (SVM).

## 2. Methodology

This research works on a novel idea of designing a prosthetic hand for an amputee's person based on an electromyography (EMG) sensor and camera. Therefore, this work is divided into two parts; the EMG signals part and the image recognition part.

In the EMG part, the first step is collecting the EMG signals from the elbow muscles for six basic hand movements, i.e., cylindrical, spherical, tip, palmar, hook, and lateral. The EMG data are collected from more than six participants of different gender and ages by using the Myo armband sensor. Then, simplify the raw collecting EMG signals by extracting time-domain features to prepare the extracted signals for the classifications stage. The final step will be choosing suitable classifiers from the Machine Learning (ML) algorithm (Support Vector Machine) groups to classify the EMG signals for different movements.

In The image recognition part, the image data are collected for many different shapes to prepare the right environment for the recognition step. Then, process all the collected images by making all the images in the same resolution and size with an empty background. Sequentially, the features are extracted from the images then the images are classified using Deep Learning (DL) algorithm (GoogLeNet) into cylindrical, spherical, tip, palmar, hook, and lateral. The final step is adding a camera in the palm of the prosthetic hand. The camera is used to recognize the shapes of the objects that appear in front of the camera. The camera is working online in real-time while moving the hand and objects to test system ability and accuracy to recognize 24 items.

### 2.1. Data collection and processing

Designing a prosthetic hand, as proposed in this work, will be depended on collecting both EMG signals from the elbow and image data for different items. For this reason, the data are divided into two parts; EMG data and image data.

- Recording six basic hand movements, i.e., cylindrical, spherical, hook, palmar, tip, and lateral) for two males and three females. Eight sensors of the Myo armband are corresponding to eight inputs in the time domain, and the amplitude of the signal represents the voltage of the required muscles; and each time series consists of 3000 samples, the eight sensors cover the elbow. As a notice, fig.1 shows the raw EMG signals for basic hand movements; the amplitude of the EMG signals represents muscle contraction for high voltage and muscle relaxation for low voltage. In this part, the obtained EMG signals have low amplitude due to the low tension on muscles responsible for finger movements.
- All the world items can be held by the human hand based on six hand movements, i.e., cylindrical, spherical, hook, palmar, tip, and lateral. This research captures 723 pictures for 24 items from different capturing corners, background, and distance from the camera. These items are classified into seven categories according to how the prosthetic hand will hold the items, i.e., basic hand movements. Table.1 clarifies the categories containing the items and the number of pictures captured for each item.

For the data processing part, this work is made simple processing for the collected data. When using the Myo Armband sensor, the level of the noise is considered very low. Therefore, no need to use filtering in the EMG signals for noise reduction. In the image data, all the image is resized to be 224*224 pixels, and the background is deleted. The processing step is important for preparing the data for feature extraction, either using machine learning or deep learning.
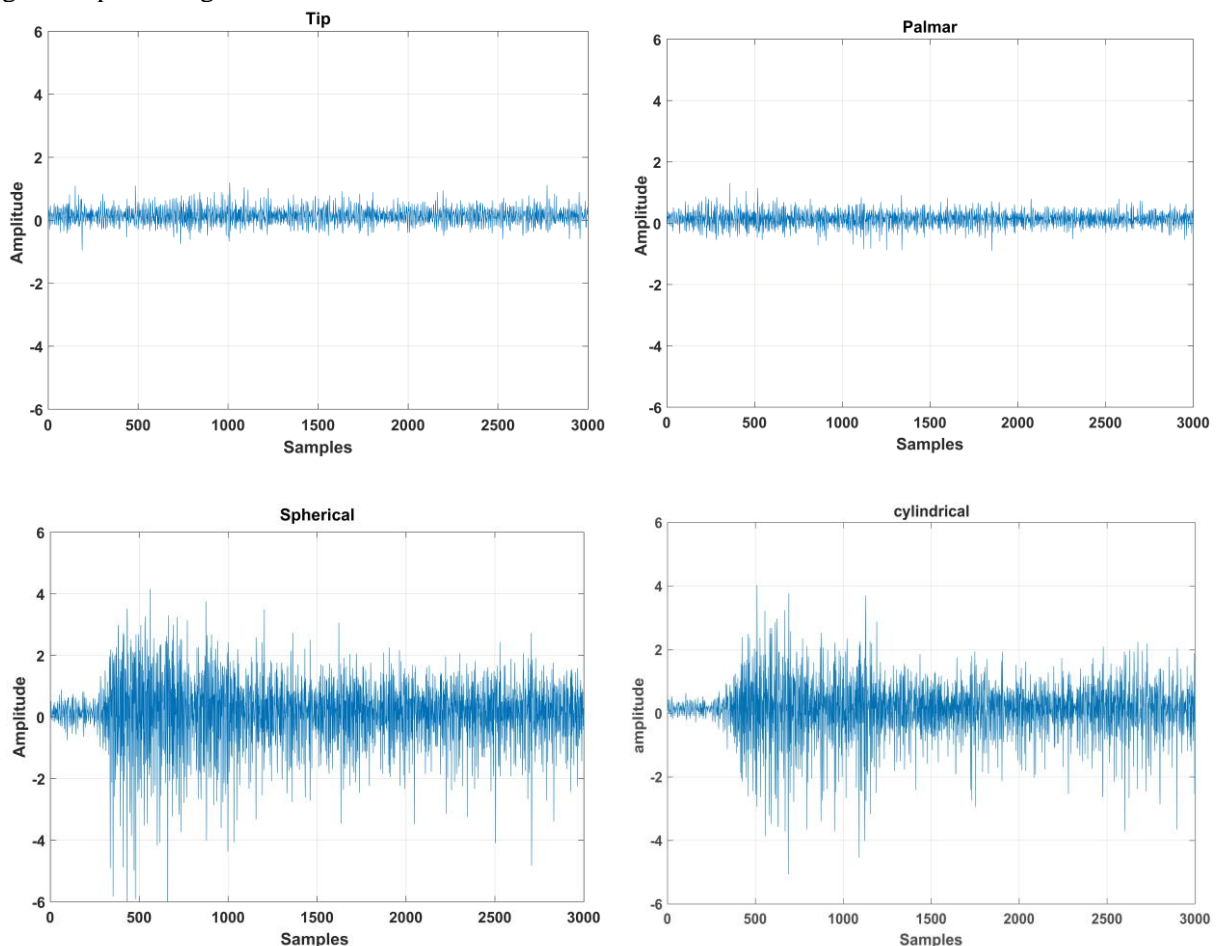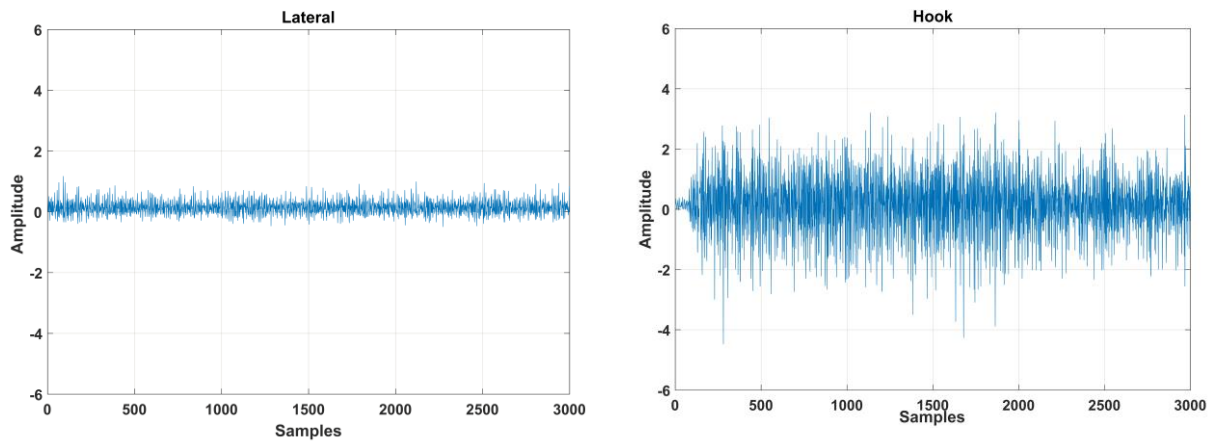
Figure 1. The raw EMG signals for basic hand movements

Table 1. The collected images data

| Items | No. of Images | Items | No. of Images | Items | No. of Images | Items | No. of Images |
|---|---|---|---|---|---|---|---|
| Cylindrical | | | Tip | | Hook | | Full Hand |
| Cans | 28 | Pins | 40 | Mug | 70 | | |
| Water bottles | 31 | Coins | 7 | Door's hand | 55 | Soft things | 40 |
| Liquid bottles | 55 | Nails | 42 | | | | |
| Lateral | | | Palmar | | Spherical | | |
| Book | 40 | Spon | 17 | Balls | | 46 | |
| Key | 20 | Knife | 33 | Egg | | 30 | |
| Flash memory | 25 | Fork | 20 | Orang | | 20 | |
| Card | 28 | Pen | 21 | Apple | | 10 | |
| Comb | 25 | Glasses | 20 | Balls | | 46 | |

## 2.2. Feature extraction

After the data processing step, for construction pattern recognition (PR) system with considerable scope to transform the raw data into a feature vector or another suitable representation, to avoid time-consuming and extensive data. The feature vectors are given the ability to the learning system, i.e., classifiers, to detect the correct pattern [15]. Therefore, transforming the raw data to feature vectors requires time-domain analysis (TDA) or frequency domain analysis (FDA).

In this work, for extracting features for EMG signals, the TDA is used instead of FDA because TDA is proper to use with EMG signals [17]. Thus, three TDA are used to analyze EMG signals in offline mode using Matlab 2019a. The recoded signals are then segmented in a window size of 200 ms and an increment of 150 ms. In this work, the mean absolute value (MAV) and the Standard deviation (SD) are used to extract the features from the raw EMG signals and simplify them by using the following mathematical models [18];

$$MAV = \frac{1}{N}\sum_{n=1}^{N} |x_n| \qquad (1)$$

$$SD = \sqrt{1/N \sum_{n=1}^{N}(x_n - MAV)^2} \qquad (2)$$

Where; x: input EMG signal, and N: the length of the input signals.

After that, the Principal Component Analysis (PCA) is used for dimension reduction of the EMG signals because of the reduction in time and space complexities. The dimensions of the new components will be uncorrelated and orthogonal to each other. The reduction happens when PCA selects the maximum variance and forms new directions.

## 2.3. Learning algorithms

After that, the Principal Component Analysis (PCA) is used for dimension reduction of the EMG signals because of the reduction in time and space complexities. after extracting the features from the raw data, the learning algorithm is chosen to classify the data according to the required categories. In this work, two learning algorithms are used; Machine Learning and Deep Learning.

The difference between these two algorithms is the ML deals with the features extraction first, then uses the extracted feature to fed it to the classifiers. In comparison, deep learning takes the features extraction and the classifiers as one package to deals with.

### 2.3.1. Support vector machine (SVM)

ML contains many algorithms that are based on fixed logic; one of these algorithms is Support vector machines (SVM) that is used in this work. SVM is a powerful supervised machine learning technique that is used for regression or classification issues. The SVM finds the optimum separating hyperplane in data classification by building hyperplane or hyperplane groups with infinite dimensionality space [19].

The SVM can deal with linearly separable data and not linearly separable. This work deals with not linearly separable because it maps the data to a higher dimension and uses the radial basis function (RBF) to obtain a high classification rate. Therefore, the RBF is considered as the kernel function to generate non-linear classifiers [20]. Also, SVM uses a subgroup of training (support vectors), and diverse Kernel functions can be particular for choosing the support vectors. The classifier used in medical diagnosis, also SVM, is useful for low memory [21, 22]. The cubic SVM is a type of SVM, and the equation of cubic SVM can be given as (3):e dimensions of the new components will be uncorrelated and orthogonal to each other. The reduction happens when PCA selects the maximum variance and forms new directions.

$$k(x_i, x_j) = (x_i^T, x_j)^3 \quad (3)$$

The cubic SVM is preferred because the short time required form training in this work training time ranges from 3 sec to 12 sec.

### 2.3.2. GoogLeNet

Convolutional Neural Network (CNN) is the trendiest applied supervised Deep Learning used in this work. The CNN has obtained the features from input at the higher layers then combine them for more complex features at the lower layers [23] [24].

GoogLeNet is a convolution architecture that has 22 layers that use an inception module to help in reduction parameters in the network [25]. Also, GoogLeNet is a concatenated layer of convolutions ($3 \times 3$ and $5 \times 5$ convolutions) and sub-layers for pooling process at different scales, the output of the pooling layer feeds to the filter banks to concatenate into a single output vector, then making this vector the input for the succeeding stage [26]. The sub-layers are connected in parallel as shown in Fig. 2, the GoogLeNet model has two convolutional layers, nine inception layers, four max-pooling layers, and a softmax layer [27].

## 3. Results

## 3.1. Evaluation methods

This paper works to assess the usability of the proposed pattern recognition system as a step toward designing a prosthetic hand. Therefore, the work is divided into two phases; The first one is applied to the collected EMG signals for the basic hand movements using three time-domain features and SVM as a classifier. The second one is applied to the collected image of 24 items to classify these items according to the hand movements using the DL algorithm, i.e., GoogLeNet. The performance of the classifiers model is described using the confusion

matrix. This matrix clarifies the relationship between the predictive and the actual events. Table. 2 shows the confusion matrix model.
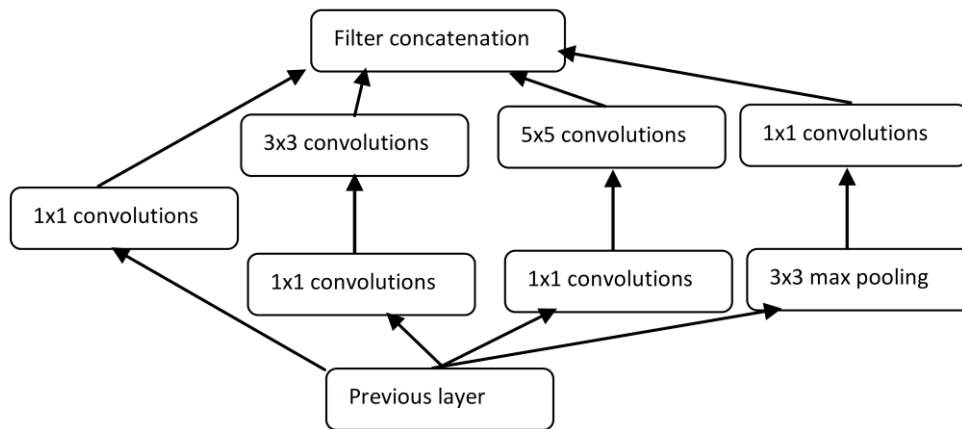


Figure 2. Inception Module [27].

Table 2. Confusion matrix model

| Confusion Matrix | Predicted No | Predicted Yes |
|---|---|---|
| Actual No | TN | FN |
| Actual Yes | FP | TP |

Where TP (True Positive) and TN (True Negative) mean the predicted value matches the actual value, FN (False Negative) means the predicted value was falsely predicted, and False Positive (FP) represents the actual value was falsely predicted. The accuracy of the confusion matrix can determine from the following equation [28]:

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \quad (4)$$

### 3.2. Training data

### 3.2.1. EMG Data

The first phase contains the training session for the basic hand movements. In this session, five participants (two males, three females) have been taken to collect the raw EMG signals by applying the MAV, SD, and PCA to extract the features. After that, the SVM and algorithm are used to classify six basic hand movements. Table. 3 shows the three different experiments; each experiment has different numbers of features. In the first experiment, only the MAV is used to extract the feature; for this reason, the training accuracy reaches 57.4%. While in the second one, both MAV and SD are applied together, so the training accuracy increases. In the last experiment, the MAV, SD, and PCA are used, and they achieve the highest training accuracy. As concluded, the increment in features with an acceptable limit affects increasing the accuracy percentage.

Table 3. ML experiments for basic hand movements.

| Basic hand movements | Features | Classifier | Time (sec) | Accuracy |
|---|---|---|---|---|
| Experiment 1 | MAV | | 11.91 | 57.4% |
| Experiment 2 | MAV and SD | SVM | 4.35 | 87% |
| Experiment 3 | MAV, SD, and PCA | | 6.39 | 89% |

Fig.3 shows the confusion matrices of SVM for the three experiments with different number of features and clarifies in details the accuracy percentage for each hand movements.
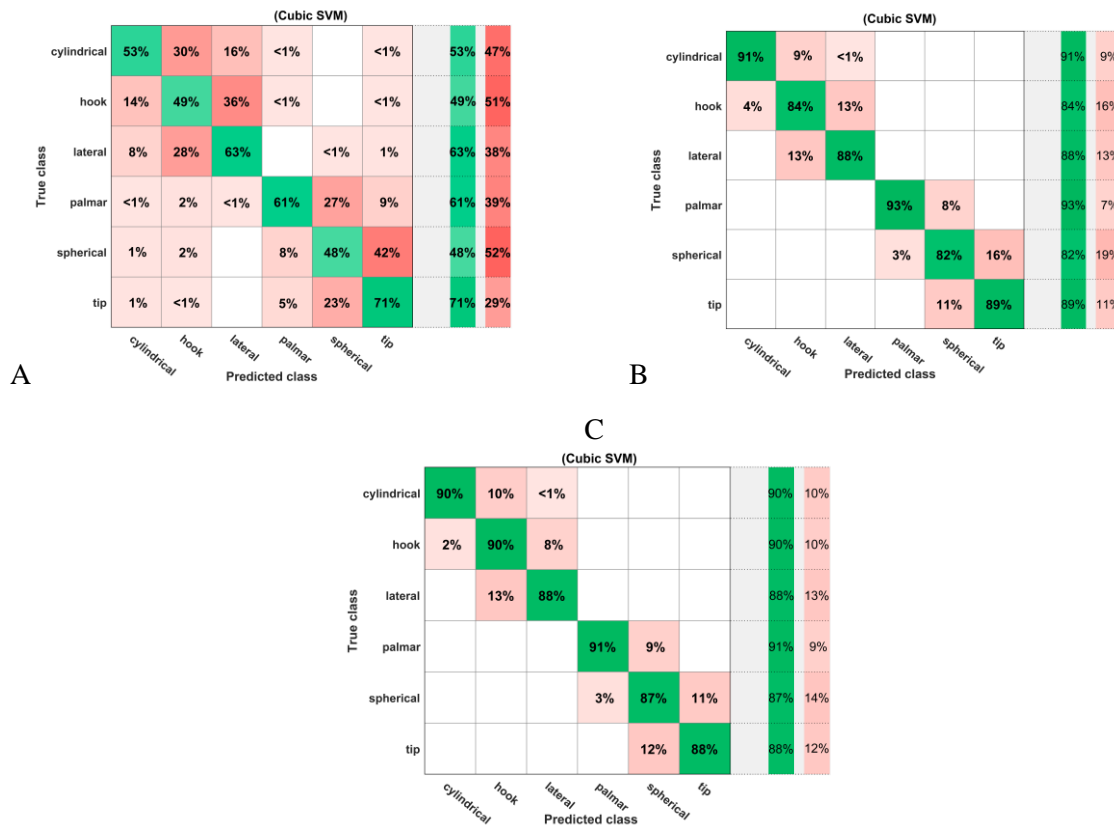


Figure 3. The confusion matrix of SVM (A) One Feature, (B) Two Features, and (B) Three Features.

### 3.2.2. RGB image data

The second phase contains the training session of 723 images for 13 items; these items are categorized according to how the hand will hold the item. That means the shape of the figure movements will differ from one item to another. Therefore, all the items are classified into seven classes, i.e., cylindrical, spherical, palmar, tip, lateral, hook, and full hand.

The GoogLeNet is used to extract and classify the 724 images into seven classes. In this work, the GoogLeNet is constructed with 144 layers, starting in the input layer representing the image that enters the net with size 224*224*3. Since the GoogLeNet is considered a type of CNN, the net is contained 57 layers as the convolutional layers. Then, the output of each convolutional layer is fed to an activation function layer to generate an activation map. In this work, 56 layers used Rectified Linear Unit (ReLU) as activation function due to the fast rectification to output zero if the input less than zero. After that, 13 max-pooling layers are used to reduce the number of parameters and the spatial size of input by selecting the maximum value from a group of numbers. For more optimization in feature selection, two layers with cross-channel normalization are added. They are also adding nine layers for depth concatenation to extract the features from the third dimension of the image. The last four layers are started with the average pooling to select the final features, then the dropout layer to drop some neurons randomly during the training to overcome the problem of overfitting. After that, the fully connected layer connects each neuron from the previous layer to all the neurons in the next layer. The last layer is the Softmax classifier that satisfied the probability of 1 for each class. Fig. 4 shows the training curve for the GoogLeNet. As noticed from the curve, the training accuracy reaches 99% using six epochs and 390 iterations; that is the mean value of 65 iterations per epochs with a learning rate of 0.0003.
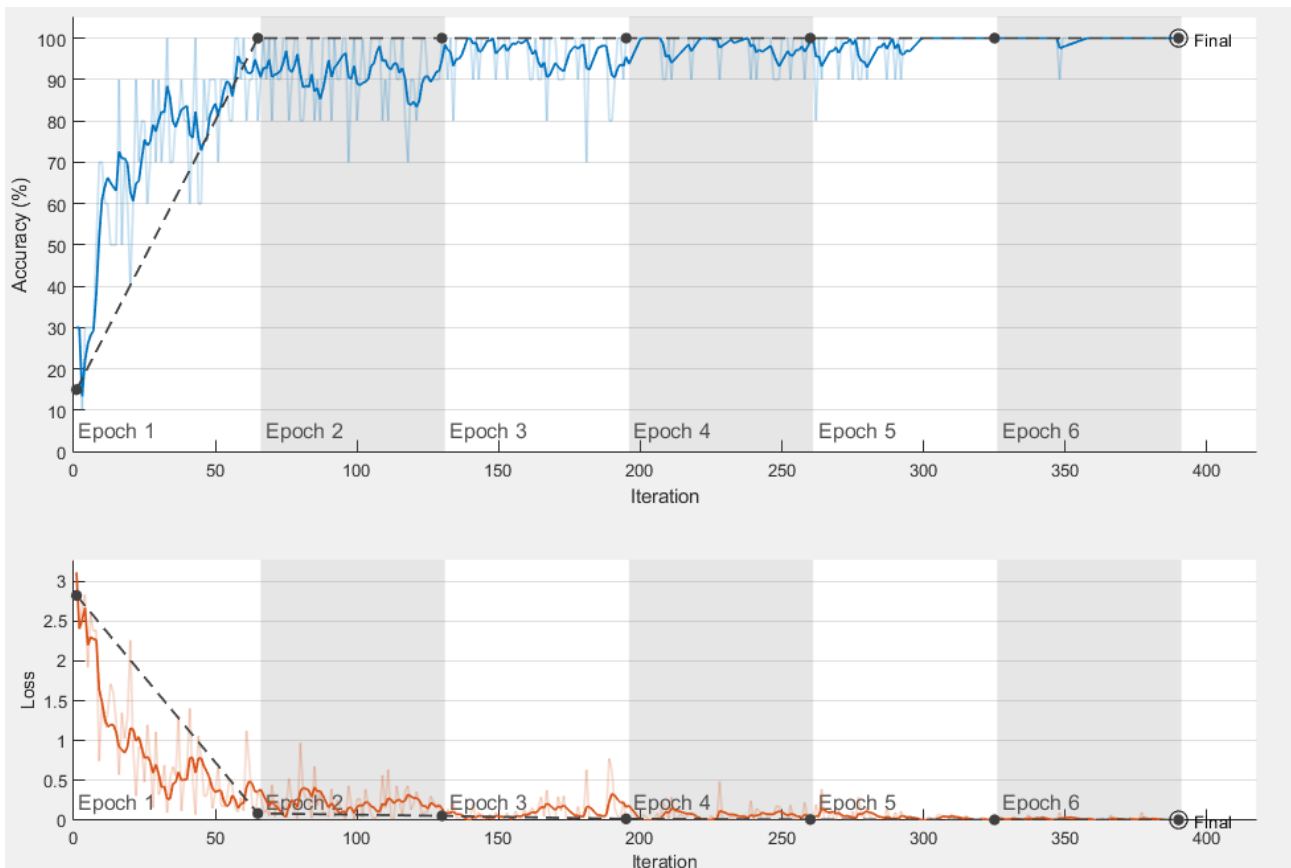
Figure 4. The training curve for the GoogLeNet

### 3.3. Testing dataset

Testing for the proposed system is needed after finishing the training part. Therefore, the test is applied to 13 items with different backgrounds, brightness, contrast, and direction if horizontal or vertical. The testing part aims to clarify the classification for each item and the ability of the system to achieve high recognition accuracy. Table. 4 and fig. 5 show the testing for 13 items and the accuracy of PR above each item.

### 4. Discussion

The intelligent prosthetic hand design requires adding multifunctionality characteristics. Hence, this work is divided into two parts; the EMG signal part to support the muscles' action and the RGB image part to support the visual section. The reason for choosing the basic hand movements because it considers the essential step toward implementing the prosthetic hand by covering the following motions (cylindrical, spherical, hook, palmar, tip, and lateral).
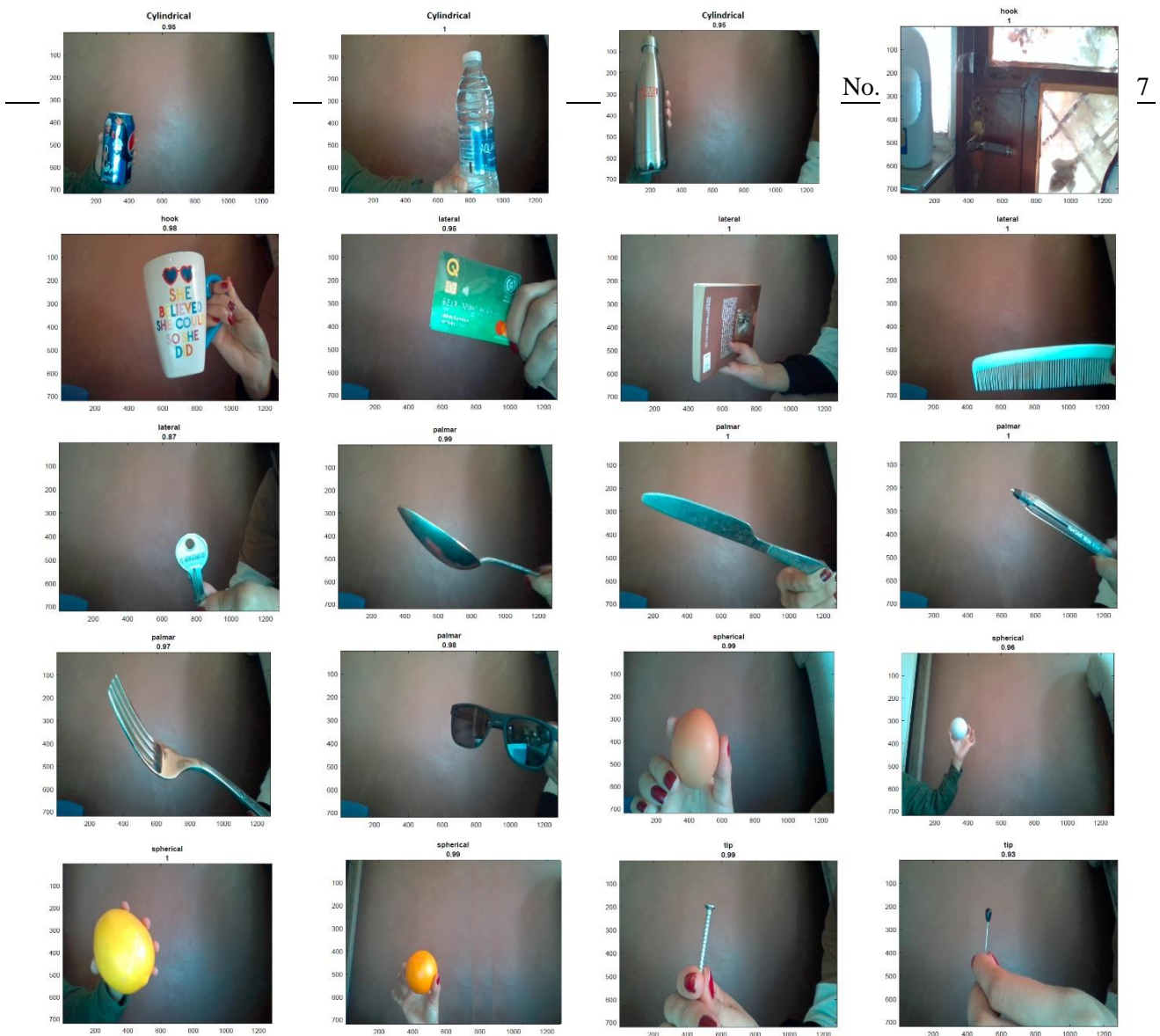
Figure 5. The testing for random items.

Table 4. The testing accuracy

| Items | No. of Images | Items | No. of Images | Items | No. of Images | Items | No. of Images |
|---|---|---|---|---|---|---|---|
| Cylindrical | | Tip | | Hook | | Full Hand | |
| Cans | 95% | Pins | 93% | Mug | 98% | | |
| Water bottles | 100% | Coins | 89% | Door's hand | 100% | Soft things | 91% |
| Liquid bottles | 95% | Nails | 99% | | | | |
| Lateral | | Palmar | | Spherical | | | |
| Book | 100% | Spon | 99% | Balls | 96% | | |
| Key | 87% | Knife | 100% | Egg | 99% | | |
| Flash memory | 98% | Fork | 97% | Orang | 99% | | |
| Card | 95% | Pen | 100% | Apple | 98% | | |
| Comb | 100% | Glasses | 98% | Balls | 100% | | |

## 4.1. The prosthetic based on EMG signals

**I**n the first part of the work, the Myo Armband sensor is used instead of the traditional sensors to solve noise issues and avoid the DC component in the raw signals; there is no need to use any filter. Also, the Myo armband covers the full forearm muscles. The data in each session were collected for six EMG movements from five subjects. The time-domain analysis is used instead of frequency domain analysis to extract the raw EMG signals' features because it shows better performance in classifying the EMG signals.

In the basic hand movements, the SVM was used and achieved acceptable training accuracy. Also, show a fast average training time reaching 7.55 sec. The obtained classification results show that the cubic support vector machine achieves the highest training accuracy among other SVM types of ML due to the multi- hyper lines in this algorithm.

Three experiments are done, as noticed from the results; the first experiment has the lowest accuracy of 57.4% among all the experiments due to using only one feature (Mean) for signals analysis. The second experiment used two features (Mean and standard deviation) to reach a training accuracy of 87%. The results achieve the highest accuracy reaching 89%, when using three features (Mean, standard deviation, and the principal component analysis). At the same time, the PCA is working on signal dimensions and increasing the accuracy. Therefore, the increasing number of features means increasing training accuracy because feature extraction helps to analyze signals. Fig. 6 shows the classification performance using different features.
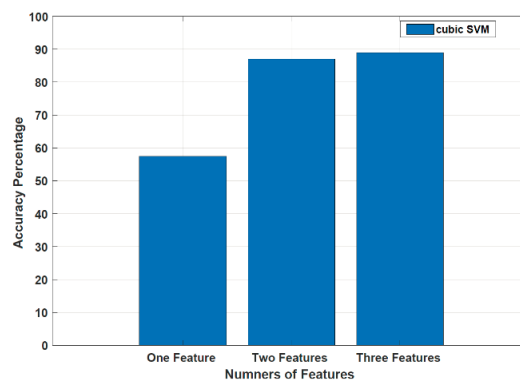


Figure 6. The classification performance using different features

## 4.2. The prosthetic based on RGB images

In the second part, the camera is planted into a prosthetic hand; this configuration enriches the system transmission information. Besides, combining the EMG signals and the visual part using artificial intelligence to implement the prosthetic hand can significantly expand the bandwidth of the feedback information to enhance the interactivity control. If there are multi objects, the direction of the prosthetic hand will be toward the closest one.

The proposed vision-based PR method selects the GoogLeNet algorithm to classify RGB images of daily used objects into the six basic hand movement patterns for controlling the prosthetic hand. Some studies focus on image classification using CCN to recognize the objects without categorized them [29]. In comparison, other studies classified the image according to color, texture, Etc. [30]. Differently, this research work is categorized the RGB images into seven categories according to how the prosthetic hand will hold the item—also, considering the object's detail such as color and texture, and increasing the recognition ability by putting the objects at different distances from the camera using different backgrounds.

The GoogLeNet was used for Deep Learning to achieve high training accuracy, reaching 99%, and high testing accuracy with a 95% average. Although some studies on RGB-D image classification are using CNN, most of them focused on object classification [30]. Object classification mainly depends on the detailed characteristics of objects (such as texture, color, etc.).

However, the visual part has many challenges, like the multi-objects, the intricate backgrounds, and the distance between the camera and the objects; all these factors can affect image recognition. Working on these challenges can achieve revolution in the intelligence prostheses.

## 5.  Conclusion

This paper proposes a novel prosthetic hand control system for liver classification using the GoogLeNet and the SVM classifiers. The system also collaborates in building a visual-based object classifier that is conventional with an EMG-based motion classifier; this cooperation controls the prosthetic hand in six hand movements (cylindrical, tip, spherical, lateral, palmar, and hook). The goal of this work has been accomplished by improving the operability of the prosthetic hand. Therefore, the hand can recognize the target object using the camera that builds in the prosthetic hand. After that, the prosthetic hand will change the posture separately according to the shape of the object.

## References

[1] X.,Yang, X. Sun, D. Zhou, Y. Li, and H. Liu, "Towards wearable A-mode ultrasound sensing for real-time finger motion recognition," IEEE Transactions on Neural Systems and Rehabilitation Engineering , vol. 26, pp.1199-1208, April 2018.

[2] M. A. Oskoei, H .Hu, "Myoelectric control systems—A survey," Biomedical signal processing and control, vol. 2, pp.275-294., October 2007.

[3] P. McCool, G. D. Fraser, A. D. Chan, L. Petropoulakis, and J. J. Soraghan, "Identification of contaminant type in surface electromyography (EMG) signals," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 22, pp.774-783, January 2014.

[4] S. M. Rissanen, M. Kankaanpää, Tarvainen, M. P. Nuutinen, J., I. Tarkka, A. Meigal, and P. A. Karjalainen, "Extraction of typical features from surface EMG signals in Parkinson's disease," In 11th International Congress of Parkinson's Disease and Movement Disorders, Istanbul, Turkey, 2007.

[5] L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, and G. Ramkumar, "GoogLeNet-Based Ensemble FCNet Classifier for Focal Liver Lesion Diagnosis," IEEE journal of biomedical and health informatics, vol. 24, pp.1686-1694, September 2019.

[6] E. Chinellato, and A. P. Del Pobil, "The neuroscience of vision-based grasping: a functional review for computational modeling and bio-inspired robotics," Journal of Integrative Neuroscience, vol. 8, pp.223-254, June 2009.

[7] M. Marković, S. Došen, D. Popović, and B. Graimann, and D. Farina " Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis," Journal of Neural Engineering, vol. 12, p.066022., November 2015.

[8] S. Došen, C. Cipriani, M. Kostić, M. Controzzi, M. C. Carrozza, and D. B. Popović, "Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation," Journal of neuroengineering and rehabilitation, vol. 7, pp. 1-14, Decemberr 2010.

[9] Y. Fang, N. Hettiarachchi, D. Zhou, and H. Liu, "Multi-modal sensing techniques for interfacing hand prostheses: A review," IEEE Sensors Journal, vol.15, pp.6065-6076, July 2015.

[10] D. G. K. Madusanka, L. N. S. Wijayasingha, R. A. R. C. Gopura, Y. W. R. Amarasinghe, and G. K. I. Mann, "A review on hybrid myoelectric control systems for upper limb prosthesis," Moratuwa Engineering Research Conference (MERCon), April 2015, pp. 136-141.

[11] C. Shi, D. Yang, J. Zhao, and H. Liu, "Computer vision-based grasp pattern recognition with application to myoelectric control of dexterous hand prosthesis," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 28, pp.2090-2099, July 2020.

[12] G. Alba, "Remote Sensing Classification Alghoritms Analysis Applied to Land Cover Change," Mario Gulich Institute, Argentina, October 2014.

[13] M. E.  Benalcázar, A. G. Jaramillo, A. Zea, A. Páez, and V. H. Andaluz, "Hand gesture recognition using machine learning and the Myo armband," The 25th European Signal Processing Conference (EUSIPCO), August 2017, pp. 1040-1044.

[14] J. Zhang, C. Ling, and S. Li, Human Movements Classification Using Multi-channel Surface EMG Signals and Deep Learning Technique. The International Conference on Cyberworlds (CW), October 2019, pp. 267-273.

[15] M. A. Wani, F. A.,Bhat, S. Afzal, and A. I. Khan. Advances in deep learning. Springer, 2020.

[16] J., Fu, and Y., Rui, "Advances in deep learning approaches for image tagging," APSIPA Transactions on Signal and Information Processing, vol. 6, 2017.

[17] N. Nazmi, M. A. Abdul Rahman, S. I. Yamamoto, S. A. Ahmad, H. Zamzuri, and S. A. Mazlan, "A review of classification techniques of EMG signals during isotonic and isometric contractions," Sensors, vol .16, p.1304, August 2016.

[18] C. Kendell, E. D. Lemaire, Y. Losier, A. Wilson, A. Chan, and B. Hudgins, "A novel approach to surface electromyography: an exploratory study of electrode-pair selection based on signal characteristics," Journal of neuroengineering and rehabilitation, vol. 9, pp.1-8, December 2012.

[19] T. Matsubara, and J. Morimoto, "Bilinear modeling of EMG signals to extract user-independent features for multiuser myoelectric interface," IEEE Transactions on Biomedical Engineering, vol. 60, pp.2205-2213, March 2013.

[20] M. Awad, and R. Khanna, "Support vector machines for classification," In Efficient Learning Machines Apress, Berkeley, CA, pp. 39-66, 2015.

[21] Y. L. Lin, and G. Wei, Speech emotion recognition based on HMM and SVM. In 2005 international conference on machine learning and cybernetics, vol. 8, August 2005, pp. 4898-4901.

[22] Y. Hu, Z. Li, G. Li, P. Yuan, C. Yang, and R. Song, "Development of sensory-motor fusion-based manipulation and grasping control for a robotic hand-eye system," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 47 pp.1169-1180, May 2016.

[23] F. Demir, V. Bajaj, M. C. Ince, S. Taran, and A. Şengür, "Surface EMG signals and deep transfer learning-based physical action classification," Neural Computing and Applications, vol. 31, pp. 8455-8462, December 2019.

[24] R. Roy, A., Kumar, M. Mahadevappa, and C. S. Kumar, "Deep learning based object shape identification from EOG controlled vision system," In 2018 IEEE SENSORS, pp. 1-4, October 2018.

[25] L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, and G. Ramkumar, "GoogLeNet-Based Ensemble FCNet Classifier for Focal Liver Lesion Diagnosis," IEEE journal of biomedical and health informatics, vol. 24, pp.1686-1694, September 2019.

[26] X. Yao, X. Wang, Y. Karaca, J. Xie, and S. Wang, "Glomerulus Classification via an Improved GoogLeNet," IEEE Access, vol. 8, pp. 176916-176923, September 2020.

[27] J. Ma, J. Rao, Y. Qiao, and W. Liu, "Sprouting Potato Recognition Based on Deep Neural Network GoogLeNet," In 2018 IEEE 3rd International Conference on Cloud Computing and Internet of Things (CCIOT), October 2018, pp. 502-505.

[28] P. Panday, and N. Godara, "Decision Support System for Cardiovascular Heart Disease Diagnosis using Improved Multilayer Perceptron," International Journal of Computer Applications, vol. 45, pp. 12–20. 2012.

[29] S. Zia, B. Yuksel, D., Yuret, and Y. Yemez, "RGB-D object recognition using deep convolutional neural networks," In Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017, pp. 896-903.

[30] L. Bo, X. Ren, and D. Fox, "Unsupervised feature learning for RGB-D based object recognition," In Experimental robotics, pp. 387-402, 2013.