10-25-2019

A deep learning approach for final grasping state determination from motion trajectory of a prosthetic hand

Cihan Uyanik
Tennessee State University

Syed F. Hussaini
Tennessee State University

Erdem Erdemir
Tennessee State University

Erkan Kaplanoglu
Marmara University

Ali Sekmen
Tennessee State University

Recommended Citation

Cihan Uyanik, Syed F. Hussaini, Erdem Erdemir, Erkan Kaplanoglu, Ali Sekmen, "A deep learning approach for final grasping state determination from motion trajectory of a prosthetic hand", Procedia Computer Science, Volume 158, 2019, Pages 19-26, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.09.023.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 158 (2019) 19–26

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd World Conference on Technology, Innovation and Entrepreneurship (WOCTINE)

A deep learning approach for final grasping state determination from motion trajectory of a prosthetic hand

Cihan Uyanik[α], Syed F. Hussaini[b], Erdem Erdemir[b], Erkan Kaplanoglu[c], Ali Sekmen[b]

[a]*College of Engineering, Tennessee State University, Nashville 37011, USA*
[b]*Department of Computer Science, Tennessee State University, Nashville 37011, USA*
[c]*Department of Mechatronics Engineering, Marmara University, Istanbul 34000, Turkey*

Abstract

Deep Learning has been gaining popularity due to its numerous implementations and continuous growing capabilities, including the prosthetics industry which has trend of evaluation towards the smart operational decision. The aim of this study is to develop a reliable decision-making system for prosthetic hands which is responsible to grasp or point an object located in the interaction area. In order to achieve this goal, we have exploited the measurements taken from a low-cost inertial measurement unit (IMU) and proposed a convolutional neural network-based decision-making system, which utilizes 9 distinct measurement variables as input, 3 axis accelerometer, 3 axis gyroscope and 3 axis magnetometer. The given experiments on this paper successfully identify that the deep learning approach produces reliably 99.2% accuracy on deciding the final prosthetic hand action by analyzing IMU sensor readings that represents motion trajectory of the hand.

Keywords: Inertial Measurement Unit, Deep Learning, Convolutional Neural Network, Smart Prosthetic Hand

1. Introduction

This paper expands on the potential of a deep learning-based prediction system on motion path determination by using multiple sensors from one low cost Inertial Measurement Unit (IMU). The IMU device motion trajectory is captured by Cartesian space motion tracing application which allows distinguish among trajectories with the help of acquired accelerometer, gyroscope, and magnetometer data. The main focus of this paper is to construct a stable prediction methodology which utilizes Convolutional Neural Network (CNN) as the classifier for the observed 9

* Corresponding author. *E-mail address:* cuyanik@my.tnstate.edu

1877-0509 © 2019 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 3rd World Conference on Technology, Innovation and Entrepreneurship
10.1016/j.procs.2019.09.023

channel measurement (accelerometer, gyroscope, and magnetometer) data of the performed motion for final hand action.

1.1. Deep Learning in Machine Learning

In contemporary, deep learning abstractly can be defined a type of machine learning that takes advantage of networks of multiple layers of nonlinear processing units for feature extraction where each successive layer input the output from the previous layers [1]. In a network architecture where convolutional networks are responsible to process of feature extraction while fully connected network are tasked with classification are known as deep neural network. Data fed into a deep neural network is converted into abstract meaningful representation [2]. The idea of multi-tier presentation and transformation of data has propelled researchers and the industries to design more capable neural networks models [3],[4],[5],[6], for various fields such as computer vision, speech recognition, medical investigation [7],[8],[9],[10] robotics etc.

1.2. Medical Investigations with Deep Learning

In [11], University of Michigan in 2017 for gait identification utilized deep learning with data collected from IMUs. Gait refers to manner of steeping or walking of an individual, gait can be used to diagnosis various disorders in early stages [10]. The subject had five IMU placed on them; chest, lower-back, right hand wrist, right knee, and right ankle. The approach for human gait identification uses data captures by the IMU which feeds into the deep convolution neural network which synchronously produces results which can be used to investigate the impact of the sensor locations [10]. They achieved 91% subject identification accuracy.

A study done at Samsung in 2017, used data collected from IMUs wrist-worn wearables to provide insights on players performance during the game for swing-based sports like Tennis, Badminton, and Golf [12]. The research developed a generalized framework for shot detection in the swing-based sports with novel concept they termed as Jerk-effective. The concept revolves around attempting to accurately calculate point of impact from a captured shot sensor signal. Followed by creating effective shot classification via two proposed techniques: correlation-based feature selection and neural network based feature classification. The IMU consisted of gyroscope and accelerator produced six axes of data: x, y, z axes for acceleration and orientation. 500 samples of the six axes of over 400 units of time were inserted into the CNN, 400 by 6 feature map. Which resulted in classification accuracy of 92% for squash and 93.8% for tennis [12].

In addition, a 2017 study in United Kingdom used a webcam as data collector for their i-limb prosthetic arm. Main objective was to classify objects with regards to grasp patterns without explicitly identifying or measuring their dimensions [11]. The deep learning architecture was trained with images of over 500 objects. The objects where divided into four categories: pinch, tripod, palmar wrist pronated, and palmar wrist neutral. After training, the volunteer subjects who are amputee controlled the i-limb ultra and prosthetic wrist with the augmented of webcam produced success results up to 88%.

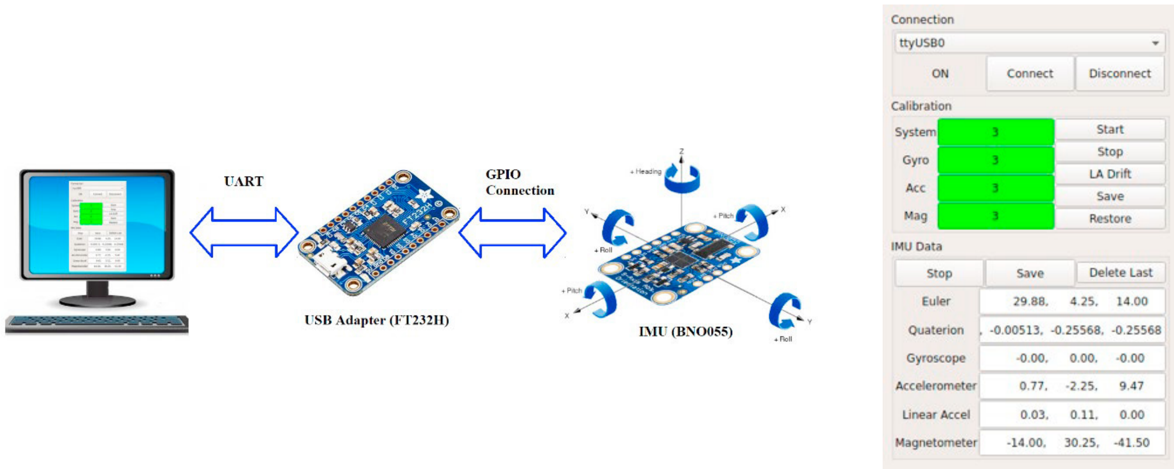
2. Experimental Study

2.1. Hardware Structure

To develop a motion classification system based on deep learning, the initial steps taken in process was the hardware setup. For this research, one example of cost effective IMU was utilized. Adafruit BNO055 Absolute Orientation Sensor connected to an Adafruit FT232H [13] Breakout which converts the General-Purpose Input Output (GPIO) signals to be interpreted through a USB connection (See Fig. 1a). According the device documentations, there is not a defined perfect calibration technique [14]. After the IMU has been calibrated, it is capable of outputting absolute orientation (euler and quaternion vector), three axis orientation data based on 360 degree sphere, angular velocity vector, three axis rotation speed, acceleration vector, three axis of acceleration, magnetic field strength vector, three axis of magnetic field sensing and ambient temperature in Celsius.

2.2. Data Collection Application

After the stage of hardware setup is executed. Data collection application is programmed to capture data efficiently from measurement module and correctly store the data. The data collection application is developed on python with QT user interface development package as displayed in Fig. 1b. The user interface executes connection procedure in between computer and the device, also displays the connection state. The GUI presents opportunity to initiate a simple calibration process for the IMU device for each individual sensor: accelerometer, gyroscope, and magnetometer. It supports to reload previously stored calibration data, too. Finally, the IMU Data section contains start and stop buttons for the data collection process activation and termination. Following sections are real time measurements received from communication channel. When the recording is terminated, the captured data is stored into the related storage system.



(a) (b)

Fig. 1. (a): Connection schema for hardware setup, (b): Data collection helper user interface

2.3. Convolutional Deep Neural Network Structure

Completing the stage of data collection, the data acquired by utilizing the data collection application could be utilized as training set of the proposed model which is responsible to classify the motion of prosthetic hand. The optimized CNN model is given in Table 1. The model contains series of convolution layers following by fully connected layers. The convolutional layers are responsible to identify the features on the data and the fully connected layers are utilized as classification layer. For the constructed model, 9-channel $48x48$ image stack data is utilized as input. This data is processed through the network. In the first layer, a convolutional layer analyzes the image and generates 32 filtered data by utilizing $5x5$ kernel and *Same* padding option. The obtained data non-linearized by following *ReLU* activation. The represent the processing data smaller in size and to keep on the important sections on the processing pipe, maximum pooling layer is utilized at the end of the convolution activation output. Additionally, to decrease the overfit possibility and allow the network to find more optimal solutions in the training phase, dropout is applied with 0.25 drop-out rate. Similar procedure is repeated for following two convolution group, details are given in Table 1. When feature detection part generates the filtered output from the given input data, series of fully connected layers executes classification phase of the pipeline. These parts contain 3 fully connected layers with 1024, 128, 7 units respectively. *ReLU* activation and dropout layer are included for the same purposes explained above. As a final layer, a *Softmax* operation is carried to produce confidence score for each 7 output.

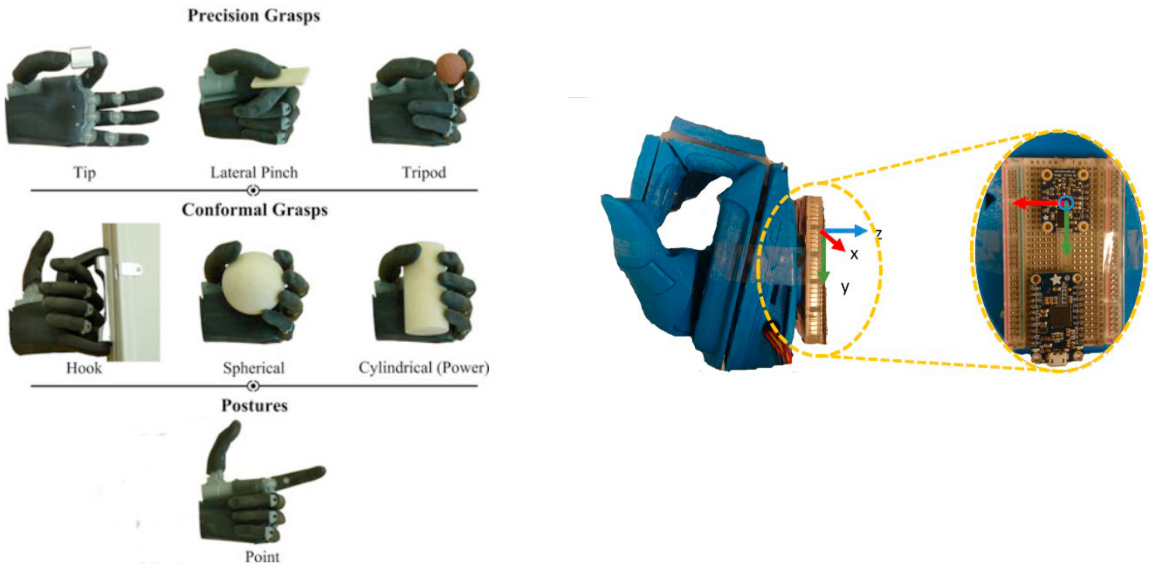
Table 1. Convolutional deep neural network structure

Layer	Description
Input	48x48x9
Convolution	32 filters of size 5x5
Activation	ReLU
Max Pooling	2x2 Pooling with 2x2 stride
Dropout	0.25 cut ratio to increase generality
Convolution	64 filters of size 5x5
Activation	ReLU
Max Pooling	2x2 Pooling with 2x2 stride
Dropout	0.25 cut ratio to increase generality
Convolution	128 filters of size 3x3
Activation	ReLU
Max Pooling	2x2 Pooling with 2x2 stride
Dropout	0.25 cut ratio to increase generality
Fully Connected	Fully connected dense layer with 1024 units
Activation	ReLU
Dropout	0.5 cut ratio to increase generality
Fully Connected	Fully connected dense layer with 128 units
Activation	ReLU
Dropout	0.5 cut ratio to increase generality
Fully Connected	Fully connected dense layer with 7 units
Activation	Softmax

2.4. Prosthetic Hand Grasping Detection by Motion Trajectory

For human-being six different separable final grasping states and a pointing state exist which construct 80% of daily life [15],[16],[17] (See Fig. 2a). For the point of observation of hand movement trajectory, any action that is performed by the user of the prosthetic has characteristic attributes that allows us to distinguish from each other. This outcome leads to develop a controller system, which is responsible to decide which type of hand action is going to be performed beforehand of the actual operation, by observing the hand motion trajectory.

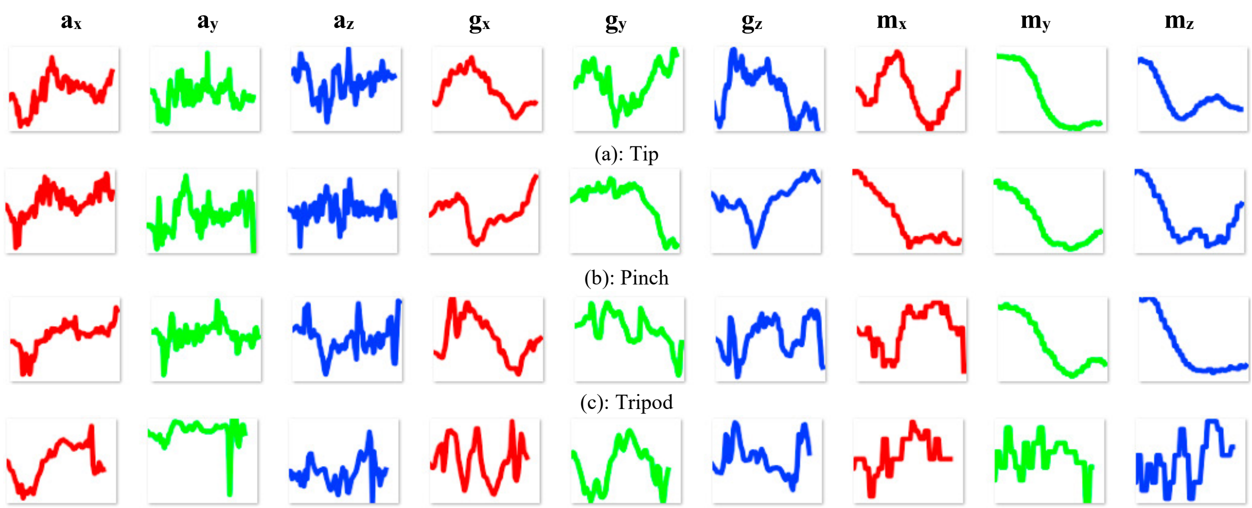
The approach is directly based on the recording the IMU measurements (accelerometer, gyroscope, and magnetometer) at the time interval of motion and generating time graphs of each 3-axis of these sensors which yields 9 time graphs for the prosthetic hand action. When the developed system identifies the final user action (one of the 6 grasping action or pointing) beforehand the action itself, the prosthetic hand could be configured and preadjusted to carry the action (hand orientation, finger states etc.). The enabled pre-adjustment feature, without doubt, not only enrich the user experience of the prosthetic but also, affects the behavior of the device positively in speed, reliability, usability perspective.



(a) (b)

Fig. 2. (a): Grasping and pointing hand states, (b): Prosthetic hand and IMU board attached

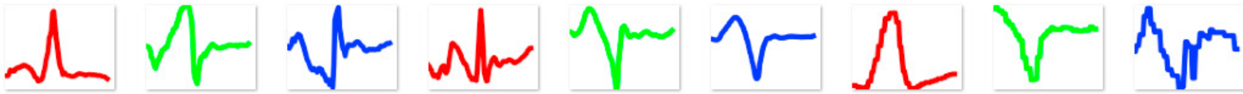
To obtain the training data, same IMU hardware setup described previously is utilized by attaching the device on the dorsal side of the hand (See Fig. 2b). For this application x, y, z axis accelerometer, gyroscope and magnetometer measurements are required to generate a time graph. The measured raw data from each sensor and their 3 axis values are required to be represented in image plane separately. Each image plane representations are used as image channels for final data which is CNN input. This 9-channel data representation contains valuable information which demonstrates the motion specific knowledge and each different type of action has separable characteristic as mentioned previously.



(d): Hook

(e): Spherical

(f): Cylindrical



(g): Point

Fig. 3. Samples of Time Graphs for individual measurements for each class

The pipeline of the explained procedures is summarized in block diagram as shown in Fig. 4. Here the raw measurements in 3 dimensional space taken from accelerometer (a_x , a_y , a_z), gyroscope (g_x , g_y , g_z), magnetometer (m_x , m_y , m_z) are utilized to generate a time graph image for each measurements in the time interval of motion until current time instance. The generated time graphs are utilized to create $9D$ representation which is constructed by the stacking each image data. Fig. 3 demonstrates examples of time graphs acquired for each prostatic action.

For the proposed prosthetic hand grasping state detection system given CNN model (Table 1) should be trained for the explained action space with a training data. The data set is constructed with 120 examples for each motion type which yields 840 samples in total and this dataset randomly divided into training, validation and test sets. Training procedure is executed by feed training set into CNN and validation set is applied as training guider set. At the end of the training procedure, test set is considered as benchmark data. The trained model successfully classifies the actions by analyzing the motion trajectories of actions and results with 99.2% accuracy on randomly selected test data. The resulting performance information is shown in Table 2 and Table 3.

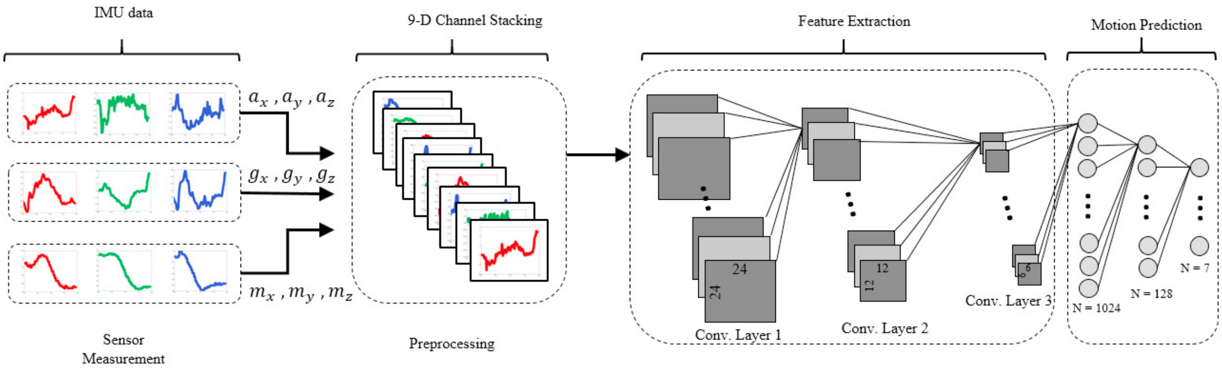


Fig. 4. Prosthetic hand grasping detection system

Table 2. Confusion matrix

	Cylind.	Hook	Pinch	Point	Spher.	Tip	Tripod
Cylind.	**20**	0	0	0	0	0	0
Hook	0	**19**	0	0	0	0	2
Pinch	0	0	**19**	0	0	0	0
Point	0	0	0	**19**	0	0	0
Spher.	0	0	0	0	**20**	0	1
Tip	0	0	0	0	0	**19**	1
Tripod	0	0	0	0	0	1	**18**

Table 3. Precision, Recall & F1-Score

	Precision	Recall	F1-Score	Support
Cylind.	1	1	1	20
Hook	1	1	1	19
Pinch	1	1	1	19
Point	1	1	1	19
Spher.	1	1	1	20
Tip	0.95	1	0.97	19
Tripod	1	0.95	0.97	19

3. Conclusion

Deep learning has gained popularity far and wide as a machine learning technique nowadays and it is applied to many problems in diverse fields. The discussion above has illustrated how a prosthetic hand final user action determination system is developed by applying CNN classifier on the observed motion trajectory of the prosthetic. The motion characteristic is observed by inertial measurement unit (IMU) and acquired raw observation data is converted into time graph representations of each measurement axis of data source sensor's, accelerometer, gyroscope, magnetometer. Although the taken measurements are relatively noisy and the device measurement quality is quite low, the proposed methodology successfully overcome these issues without applying measurement correction or application specialized filters.

Acknowledgements

This work is partially supported by the National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA) Capacity Building Grant under award number CBGP-005696.

References

[1] A. Sekmen, A. B. Koku, M. Parlaktuna, A. Abdul-Malek, and N. Vanamala, "Unsupervised deep learning for subspace clustering," in *Big Data (Big Data), 2017 IEEE International Conference on*, 2017, pp. 2089–2094.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. **521**, no. 7553, p. 436, 2015.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[4] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, 2014, pp. 818–833.

[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Prepr. arXiv1409.1556*, 2014.

[6] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[7] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. **542,** no. 7639, p. 115, 2017.

[8] N. R. O'Shea K., "An Introduction to Convolutional Neural Networks," 2015.

[9] C. M. M. H. [Atzori M., "Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands," *Front Neurorobot*, vol. **10**, 2016.

[10] T. M. C. R. Dehzangi O., "IMU-Based Gait Recognition Using Convolutional Neural Networks and Multi-Sensor Fusion," vol. **17**, 2017.

[11] G. Ghazaei, A. Alameer, P. Degenaar, G. Morgan, and K. Nazarpour, "Deep learning-based artificial vision for grasp classification in myoelectric hands," *J. Neural Eng.*, vol. **14**, pp. 36025–36043, 2017.

[12] A. Anand, M. Sharma, R. Srivastava, L. Kaligounder, and D. Prakash, "Wearable Motion Sensor Based Analysis of Swing Sports," in *Machine Learning and Applications (ICMLA), 2017 16th IEEE International Conference on*, 2017, pp. 261–267.

[13] "FT232 Overview." [Online]. Available: https://learn.adafruit.com/adafruit-ft232h-breakout/overview. [Accessed: 08-Apr-2019].

[14] "BNO055 Overview." [Online]. Available: https://learn.adafruit.com/adafruit-bno055-absolute-orientation-sensor/overview. [Accessed: 08-Apr-2019].

[15] C. Sollerman and A. Ejeskär, "Sollerman hand function test: a standardised method and its use in tetraplegic patients," *Scand. J. Plast. Reconstr. Surg. Hand Surg.*, vol. **29**, no. 2, pp. 167–176, 1995.

[16] R. J. Schwarz and C. L. Taylor, "The anatomy and mechanics of the human hand," *Artif. Limbs*, vol. **2**, no. 2, pp. 22–35, 1955.

[17] D. A. Bennett, S. A. Dalley, D. Truex, and M. Goldfarb, "A multigrasp hand prosthesis for providing precision and conformal grasps," *IEEE/ASME Trans. Mechatronics*, vol. **20**, no. 4, pp. 1697–1704, 2015.