2022 19th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 9-11, 2022

Human Upper Limb Motion Recognition Using IMU sensors and Artificial Neural Networks

Lauro Armando Contreras Rodríguez
Electrical Engineering Department, Bioelectronics section
CINVESTAV-IPN
Mexico City, Mexico
lauro.contreras@cinvestav.mx

Eladio Cardiel
Electrical Engineering Department, Bioelectronics section
CINVESTAV-IPN
Mexico City, Mexico
ecardiel@cinvestav.mx

Ángel Llanas Soto
Electrical Engineering Department, Bioelectronics section
CINVESTAV-IPN
Mexico City, Mexico
angel@cinvestav.mx

José Antonio Barraza Madrigal
Departamento de Formación Básica, Academia de Física
ESIQIE-Instituto Politécnico Nacional
Ciudad de México, México
jabarraza@ipn.mx

Pablo Rogelio Hernández Rodríguez
Electrical Engineering Department, Bioelectronics section
CINVESTAV-IPN
Mexico City, Mexico
pablo.rogeli@cinvestav.mx

***Abstract*— Computer vision researchers have been tasked with developing applications to recognize the human body's structure, movements, and tracking. Recognizing human upper limb movements has brought benefits in physical therapy, virtual reality, human-robot systems, sports. This paper presents a multi-layer perceptron (MLP) artificial neural network (ANN) designed to recognize the upper limb movements performance in daily life tasks. For this, inertial measurement units (IMU) were employed to acquire information regarding of the upper limb movement. The performance of the model was assessed through a confusion matrix and the receiver operator characteristic (ROC) curves. Recognition accuracy obtained for the ANN model was 97.39%, the mean area under the curve (AUC) of the ROC curves was 0.973. According to the results, the proposed ANN could recognize upper limb movements in daily life tasks. Further research is suggested to test the model with fine movements.**

Keywords— Upper Limb, Human Motion Recognition, MLP, ANN.

I. Introduction

Research on human motion try to explain how the musculoskeletal system interacts while doing activities of daily living. Since the end of the last century, computer vision researchers have increasingly investigated human motion analysis. The focus of these investigations has been carried out in three areas: (1) Motion analysis of the human body structure, (2) Tracking of the human motion, and (3) Recognition of human movements [1, 2].

For motion performance recognition, artificial neural networks (ANN) have been of great help in detecting patterns that identify activities. EMG signals, vision systems, and wearable sensors are the common data source for this application. However, the average efficiency of EMG signal is around 70% and requires advanced detection, decomposition, processing, and classification methods before using ANN [3]. Vision systems require specific conditions for data acquisition, such as light and several cameras to avoid occlusion and markers. The need for outdoor analysis remains an open challenge for these systems [4]. Finally, wearable sensors have become the most popular option in recent years. These sensors provide information related to human movements in three-dimensional space and they are non-invasive. However, because the devices include a magnetometer, they are susceptible to electromagnetic signals that could affect the sensors' measurements [5-7].

Recent studies have reported the use of wearable sensors and ANN to classify activities such as walking, running, standing, walking upstairs, or walking downstairs [8, 9]. However, recognizing upper limb movements remains challenging since the number of identified joints is limited [10, 11]. Upper limb activities are usually limited to hand gesture recognition or trajectory prediction to compensate for limb tremors for orthoses [12, 13]. Also, some works aim to locomotor intention prediction of the upper limb action [14, 15].

This paper presents the design of a multi-layer perceptron (MLP) ANN model for the classification of seven daily life tasks of the upper limb based on IMU sensors. The MLP architecture (16-7-7) and a backpropagation training algorithm allowed the model to classify the seven upper limb movements in daily life tasks.

II. Methodology

This work consisted in training a multi-layer perceptron (MLP) artificial neural network with data acquired from human upper limb movements by motion capture devices. The data corresponded to the three joints with the greatest range of motion (shoulder, elbow, and wrist, with reference to the chest) when performing seven daily life tasks. These movements included activities performed on the anatomical planes and combined movements for the joints.

A. Materials

Xsens dot (Xsens, Netherland), a wearable platform developed to provide an accurate 3D orientation in any environment, was used to measure the movements [16, 17]. The platform consisted of four wireless measurement units, four straps to attach the units, and the Xsens dot app. The anatomical attachment of the sensors was defined to guarantee to map full range of the movements. The first sensor was located in the thorax, at the sternum level, and serves as a reference to the other sensors. The second sensor was located at the external lateral portion of the arm, between the acromion and the lateral epicondyle of the humerus, referenced at the midpoint between the brachii biceps and brachii triceps muscles. The third sensors was placed on the *extensor digitorum communis* muscle, 5cm from the wrist. While the fourth sensor was located on the back of the hand.

B. *Artificial Neural Network*

a) Database

A database of the sensors' recordings corresponding to the seven activities of the right upper limb was generated to train, validate, and test the artificial neural network (ANN). Ten healthy, right-handed volunteers, six males and four females, age 39±15.9 years, were recruited to conduct the activities. The sensor's sampling rate was 60Hz with a 16-bit resolution using the Xsens' latest Kalman filter core algorithm (XKFCore) for sensor fusion, optimized for human motions [18]. The first activity, called the static position, recorded the n-pose of the volunteer. The following five activities were performed with one degree of motion for a single joint. These movements were recorded by the sensors, while repeating a shoulder flexion-extension, shoulder abduction-adduction, elbow flexion-extension, wrist flexion-extension, or wrist abduction-adduction starting at the n-pose. The seventh activity was a grasp-to-reach test. This test required the volunteer to take a glass from a table, bring it to his mouth and leave it on the table. The activities were performed only once per volunteer and in only one action without constraints in the time required to execute. Each sensor provides a quaternion related to the activity of the measured segments; therefore, four data samples are provided in each sample per sensor. To avoid confusion in the training stage, the movement below 10° from the n-pose was considered static. A total of 42,356 data samples were taken from the seven activities: 12,805 samples for the static position; 5,614 samples for the shoulder flexion-extension activity; 6,110 samples for the shoulder abduction-adduction; 5,057 samples for the elbow flexion-extension; 4,410 samples for the wrist flexion-extension; 4,686 samples for the wrist abduction-adduction; and 4,321 samples for the reach-to-grasp test. The raw data provided by the sensors, quaternions, were used to feed the ANN without preprocessing.

b) MLP Architecture

The ANN architecture was designed and programmed to classify the seven movements as seven different classes. The common approach to fit the model uses one neuron per class at the output layer; therefore, the proposed architecture of the ANN has seven neurons in the output layer. The input layer corresponded to the sensors' information since the output of each sensor gives the orientation measured in quaternions; the input vector for each sampled position has 16 data. The number of neurons in the hidden layer was allowed to encompass the targets for each class in three-dimensional space. Seven neurons fulfilled the previous condition, so the final architecture for the ANN was 16-7-7. Both the hidden layer and the output layer use a sigmoid function due to its derivative properties [19].

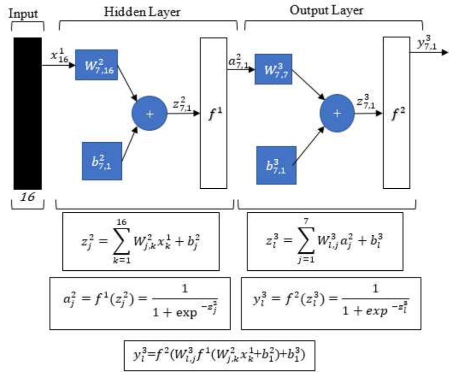

Figure 1. Summarized representation for the architecture of the ANN

c) Cross-Validation

The database was randomized and segmented into two sets before the training and tests processes to avoid a biased dataset. The first set, 75% of the entire database, was used as the training set, while the second one, the remaining 25%, was used for the test. The training set was validated using a cross-validation process which consisted of segmenting the training set in k-segments with equal samples to create k-models. Each model used k-1 segments for training and the kth segment to validate the model. The validation set was different for each model. The performance of the artificial neural network during the training process was calculated with the average of the accuracy of the k-models. The value of K used for this study was 3 since it was sought to have 66.6% data for training and 33.3% data for validation in each model.

d) Learning

The learning process used for this architecture was the backpropagation (BP) algorithm. This algorithm calculates the output error of the architecture on each iteration and propagates the error backwards to each neuron on the network determining the paths with the most influence on the output. This process is carried out on each layer to tune the weight and bias along the network. The BP algorithm evaluates the derivative of the output layer activation function as a product of derivatives between each layer. The gradient of the weight and bias between each layer is a modification of the partial products [20].

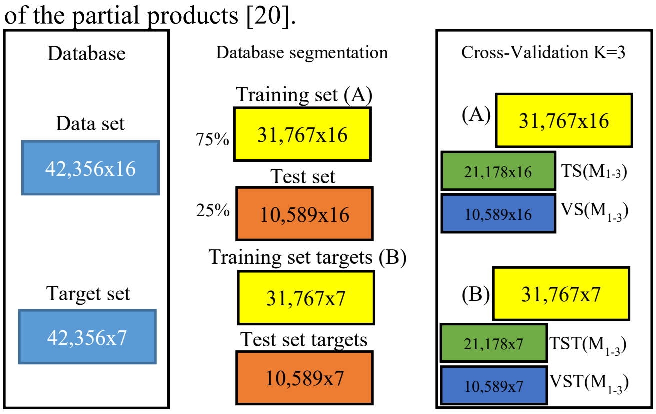

Figure 2. Segmentation of the database for the training and test sets and segmentation of the training set using cross-validation with k=3. TS(M_{1-3}): Training sets for model 1, 2 and 3, VS(M_{1-3}): Validation sets for model 1, 2 and 3, TST(M_{1-3}): Training sets targets for model 1, 2 and 3, VST(M_{1-3}): Validation sets targets for model 1, 2 and 3

e) Statistical methods

The results presented in this work, learning curve, confusion matrix, and receiver operating characteristic (ROC) curves, are the tools used for classifiers. The learning curve allowed observing the model learning process to reduce the output error for all classes in each iteration. The confusion matrix allowed visualizing the performance of the ANN by comparing the predicted values respect to the output values after training the ANN. The ROC curve illustrated the discrimination capacity of a classifier according to the threshold; these curves will be plotted in a one-vs-all graph for each class. Another ROC curve parameter is the area under the curve (AUC) which determines the model's capacity to distinguish between classes.

Besides, the confusion matrix identifies the model's accuracy, sensitivity, and specificity. The model accuracy was calculated with Eq. (1).

$$Accuracy = TP + TN / TP + TN + FP + FN \quad (1)$$

Where, TP stands for true positive, a test result that correctly indicates the presence of a condition or characteristic; TN means true negative, a test result that correctly indicates the absence of a condition or characteristic; FP, false positive, a test result which wrongly indicates that a particular condition or characteristic is present; FN, false negative, a test result which wrongly indicates that a specific condition or characteristic is absent.

Sensibility, the probability of a positive test, conditioned on truly being positive, can be calculated with Eq. (2)

$$Sensibility = TP / TP + FN \quad (2)$$

Specificity, the probability of a negative test, conditioned on truly being negative, was calculated with Eq. (3)

$$Specificity = TN / TN + FP \quad (3)$$

III. Results

A. Learning Curve

Six iterations were required to achieve the proposed error, 0.01, to identify the seven classes, figure 3.

B. Confusion Matrix

The accuracy, calculated through the main diagonal of the confusion matrix, Table 1, indicates the correctly classified classes giving a 97.39% with a mean performance of 97.84%. The last row of the confusion matrix indicated a sample mistake.

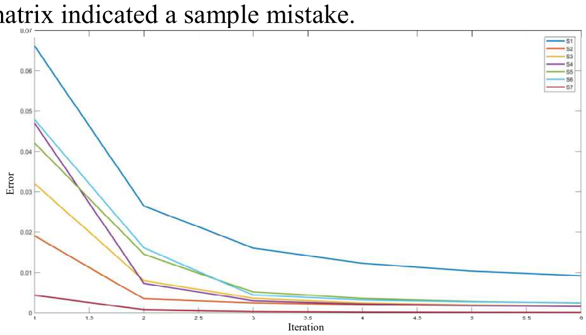

Figure 3. The learning curve for the ANN architecture.

C. ROC curve

The ROC curve analysis results for seven different combinations are shown in figure 4. Each combination contrasts the results obtained for a particular class against the rest in a one-vs-all classification.

Note that the 97.39% accuracy obtained for all comparisons is due to the fact that the sum of the TP and TN are the same for all the confusion matrices in each comparison.

IV. Discussion

The present work explores the possibility to recognize movements in the human upper limb using IMU sensors combined with the designed artificial neural network. The low number of iterations required to achieve the proposed error during training indicated that the architecture adequately trained the model. The recognition accuracy of 97.39% showed that the model could classify according to the patterns of the seven daily life tasks. The 2.61% model error may be due to the threshold defined for identifying movements as static positions. This could be improved by extending the boundary used to distinguish between movements, however, the purpose of the threshold allowed the decision line to be placed as close as possible between two or more classes.

The average value of the area under the curve (AUC) from the seven comparisons was 0.973; the result suggest a good ANN capacity to distinguish between classes, since a value close to 1 indicates that the model can predict between classes with a high-reliability [21].

The accuracy value obtained for this work of 97.39% using four IMU sensors and executing five degrees of freedom movements for the upper limb in daily task activities. This accuracy is in the range of the performance reported for intention prediction of the upper limb movement [14, 15].

The limitation of this work focus on the lack of activities with similar movement patterns that could compromise the performance of the ANN, such as shaking hands, opening a doorknob, and receiving an object. Future studies are pending to test the model's capacity for distinguishing these activities. Furthermore, deep learning neural networks have shown better generalization for these applications, so their use in future works may improve the performance.

Table 1. Confusion matrix for the test set. The overall performance is determined by the average of the correctly classified samples of each class divided by the total of samples. C1-7: Classes 1-7.

Confusion Matrix	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**
C1	3146	29	29	28	1	14	0
C2	0	1286	0	0	0	0	0
C3	0	0	1434	0	0	0	0
C4	0	0	0	1183	0	0	0
C5	21	0	0	0	1127	0	0
C6	32	0	0	0	0	1204	0
C7	0	0	0	0	0	0	897
MC	50	21	11	20	2	2	0
Performance (%)	96.83	96.26	97.29	96.10	99.73	98.67	100

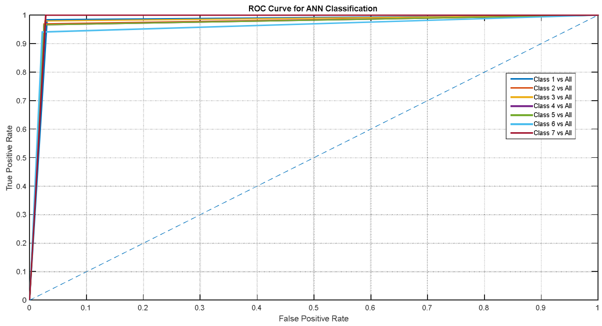

Figure 4. ROC curves for the one-vs-all comparison for each class.

V. CONCLUSION

The results of this work indicated that it is possible to recognize specific movements of the human upper limb in daily life tasks with IMU sensors and ANN. The recognition of these movements can provide information about a person's condition that could be useful for medical diagnosis and rehabilitation or the prevention of injuries and performance improvements in sports applications by predicting the outcome of the move according to the patterns in the initial stages of the activity. Additionally, this information can be used in industry and virtual environments to determine optimal trajectories required by robots for a specific application.

ACKNOWLEDGMENT

Authors express thanks to the Consejo Nacional de Ciencia y Tecnología (CONACyT, México), for the scholarship granted to Lauro Armando Contreras-Rodríguez for the PD studies. The authors would like to acknowledge the valuable contributions of Dr. Roberto Muñoz Guerrero (1960-2019) to this work.

Table 2. Summary of the defined statistical parameters for the one-vs-all confusion matrix for each class according to the equations (1-3) and the ROC curves' AUC.

Table 2. One-vs-all statistical comparison for each class.

Class	AUC	Sensitivity	Specificity	Accuracy
1-vs-All	0.977	0.935	0.993	97.39%
2-vs-All	0.972	0.853	0.995	97.39%
3-vs-All	0.976	0.859	0.997	97.39%
4-vs-All	0.971	0.841	0.995	97.39%
5-vs-All	0.971	0.818	0.996	97.39%
6-vs-All	0.96	0.842	0.993	97.39%
7-vs-All	0.986	0.753	1.000	97.39%

REFERENCES

[1] Wang L, Hu W, Tan T. Recent developments in human motion analysis. Pattern recognition. 2003;36:585-601.

[2] Patrona F, Chatzitofis A, Zarpalas D, Daras P. Motion analysis: Action detection, recognition, and evaluation based on motion capture data. Pattern Recognition. 2018;76:612-22.

[3] Gohel V, Mehendale N. Review on electromyography signal acquisition and processing. Biophysical Reviews. 2020;12:1361-7.

[4] Colyer SL, Evans M, Cosker DP, Salo AI. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports medicine-open. 2018;4:1-15.

[5] Iosa M, Picerno P, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis. Expert review of medical devices. 2016;13:641-59.

[6] Picerno P, Iosa M, D'Souza C, Benedetti MG, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis: A five-year update. Expert review of medical devices. 2021;18:79-94.

[7] Filippeschi A, Schmitz N, Miezal M, Bleser G, Ruffaldi E, Stricker D. Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion. Sensors. 2017;17:1257.

[8] Ramasamy Ramamurthy S, Roy N. Recent trends in machine learning for human activity recognition—A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018;8:e1254.

[9] Xu Y, Qiu TT. Human activity recognition and embedded application based on convolutional neural network. Journal of Artificial Intelligence and Technology. 2021;1:51-60.

[10] Elkholy HA, Azar AT, Magd A, Marzouk H, Ammar HH. Classifying upper limb activities using deep neural networks. The International Conference on Artificial Intelligence and Computer Vision: Springer; 2020. p. 268-82.

[11] Wei W, Kurita K, Kuang J, Gao A. Real-Time Limb Motion Tracking with a Single IMU Sensor for Physical Therapy Exercises. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC): IEEE; 2021. p. 7152-7.

[12] Oudah M, Al-Naji A, Chahl J. Hand gesture recognition based on computer vision: a review of techniques. journal of Imaging. 2020;6:73.

[13] Shahtalebi S, Atashzar SF, Samotus O, Patel RV, Jog MS, Mohammadi A. Phtnet: Characterization and deep mining of involuntary pathological hand tremor using recurrent neural network models. Scientific reports. 2020;10:1-19.

[14] Cui J-W, Li Z-G, Du H, Yan B-Y, Lu P-D. Recognition of Upper Limb Action Intention Based on IMU. Sensors. 2022;22:1954.

[15] Lu H, Schomaker LR, Carloni R. IMU-based deep neural networks for locomotor intention prediction. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS): IEEE; 2020. p. 4134-9.

[16] Schepers M, Giuberti M, Bellusci G. Xsens mvn: Consistent tracking of human motion using inertial sensing. Xsens Technol. 2018;1.

[17] Rana M, Mittal V. Wearable sensors for real-time kinematics analysis in sports: a review. IEEE Sensors Journal. 2020;21:1187-207.

[18] Xsense. Xsens DOT User Manual. Netherlands: XUF; 2022. (available on: Xsense.com) view date: august 30, 2022.

[19] Nayak J, Naik B, Behera HS. Solving nonlinear classification problems with black hole optimisation and higher order Jordan Pi-sigma neural network: a novel approach. International Journal of Computational Systems Engineering. 2016;2:236-51.

[20] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. nature. 1986;323:533-6.

[21] Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery. 2016;159:1638-45.