IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 1, November 2022

Early Detection of Driver Drowsiness using Deep Learning Algorithms

M. Poojitha[1], M. Madhan Sai[2], N. Chandu[3], N. Lokesh[4], N. S. S. K. Dil[5], N. Rahul Babu[6]
GMR Institute of Technology, Rajam, Andhra Pradesh, India
mumminapoojitha@gmail.com[1], madhanmullapudi2003@gmail.com[2]

Abstract: *All India Institute of Medical Science Neurology Indian research has found that more than 20% of accidents are due to insomnia. Exhausted drivers who doze off the wheel are liable for 40% of road accidents. An accident is an event that occurs instantly, out of the blue, due to unplanned incidents. Every day, thousands of accidents occur due to insomnia. According to studies, almost one-quarter of all two-wheeler accidents are due to drivers' handling in a state of heavy drowsiness. It will be dangerous when the driver falls asleep. A driver drowsiness system is a technology that helps to prevent accidents caused by the driver. However, with the help of deep learning and convolutional neural network algorithms, this analysis can predict driver fatigue states through facial video. The algorithms are used to predict the driver's fatigue state based on eye, head, and mouth behavior. In this analysis, the objective of the algorithm is to alert the driver to prevent accidents.*

Keywords: Driver drowsiness, deep learning, Face video, accidents, Convolutional neural network

I. INTRODUCTION

One of the most common causes of road accidents is fatigued driving. According to estimates provided by the National Highway Traffic Safety Administration, approximately 50,000 injuries and nearly 800 fatalities are caused by drunk driving incidents that are reported to the police every year. About one in 25 adult drivers, according to the Centers for Disease Control and Prevention, have admitted to driving while intoxicated, and many more have admitted to doing so. Due to the difficulty of recognizing fatigue-related symptoms, a driver may not even be aware that they are exhausted. Accidents may be reduced if the driver's drowsiness is detected. Various methods for detecting driver drowsiness have been developed as technology has advanced. The driver's physical symptoms are used to determine whether they are tired. Yawning, eye blinking, nodding, and some facial features are all signs of drowsiness. Dwivedi and team proposed computer vision-based drowsiness detection, but it has poor performance and low accuracy. Zhang and co. proposed a deep learning-based system for detecting yawns. This method has a high detection accuracy, but it relies heavily on the equipment's performance, and it is unable to detect all vehicles in real-time. Various methods were used by many authors to identify the drowsiness. However, the majority of methods for detecting eye blinking ignore the other facial regions and concentrate only on the eyes. The results achieved by these methods are still far from ideal. To take care of this issue, a cross-breed profound learning model will be utilizedto recognize sluggishness. To detect drowsiness, a variety of models, including CNN, Inception V3, LSTM, ResNet50, and Faster R-CNN utilized.

II. LITERATURE SURVEY

1. Z. Cui, H. M. Sun, R. N. Yin, L. Gao, H. B. Sun, and R. S. JiaDeep's learning of face video is used to detect driver fatigue in real-time.80(17):25495-25515, Multimedia Tools and Applications.

According to the author of their publication, face video-based drive fatigue detection has received a lot of attention due to its low cost and non-invasive nature.To solve the problem of insufficient memory and limited computing power, they use a light neural network, and the multi-feature fusion judgment algorithm is used to determine the driver's fatigue level. Our method's accuracy for predicting drowsiness and yawning in a real vehicle environment is 98.30 percent, according to the results of our experiments.

2. Muhammad, K., Ullah, A., Lloret, J., Del Ser, J., and de Albuquerque, V. H. C. Safe autonomous driving based on deep learning: Current difficulties and plans for the future22(7): 4316-4336 in IEEE Transactions on Intelligent Transportation Systems.

According to the author of this publication, deep learning approaches to the control procedure for autonomous driving have not been thoroughly studied. As a result, they emphasize the strength of DL architecture for safe autonomous driving in terms of reliability and effective real-time performance. This study identifies pedestrian detection by investigating body part semantics and contextual information with complex handling of occlusions, drowsiness detection by CNN and LSTM, collision detection by TCNN, R-TCNN, and traffic sign detection by R-CNN, and YOLO. This study also identifies vehicle detection by convolutional regression neural network, Faster R-CNN, multi-task deep CNN, and CNN's. The main advantages of DL methods were discussed in this article, and the average human accuracy in the daytime and night time scenarios is only 80%

3. Kumar, V., and Sharma, S.Utilizing a modified deep learning architecture, driver drowsiness are detected.1-10 in Evolutionary Intelligence.

A non-invasive method for detecting driver drowsiness is proposed in this paper. The majority of the time, facial features are used to detect the driver's drowsiness. The video frame is used to remove the parts of the mouth and eye. In this case, they detected driver drowsiness by employing an LSTM network and a modified InceptionV3 algorithm. The NTHU-DDD dataset is used to evaluate the effectiveness of the proposed method. Over the NTHUDD dataset, the proposed model achieved an accuracy of 91.36 percent.

4. A. A. Minhas, S. Jabbar, M. Farhan, and M. NajamulIslama convolutional neural network-based smart analysis of driver fatigue and drowsiness detection.1-18 of Multimedia Tools and Applications

According to this author, driver drowsiness, such as fatigue and insomnia, is to blame for the rising number of road accidents. The primary goal of this study is to use deep learning models like convolution neural networks to identify fatigue and drowsiness. Modern Convolutional Neural Network models like InceptionV3, VGG16, and ResNet50 are used for detection. To replace the "sleepiness-impact" scenario with more variables and parameters, this research may be an excellent way to open up new horizons in smart textiles and digital textiles. With a 0.6931 loss, the InceptionV3 has an accuracy of 90.70%, the VGG16 has an accuracy of 39.87%, and the ResNet50 has an accuracy of 93.69% with a 0.6931 loss. As a result, the ResNet50 model outperforms all other deep learning models in terms of accuracy.

5. Huang, R., Y. Wang, Z. Li, Z. Lei, and Y. XuRF-DCM: for the purpose of detecting driver fatigue, a multi-granularity deep convolutional model based on feature fusion and recalibration.23(1), 630-640, IEEE Transactions on Intelligent Transportation Systems.

According to this author, previous fatigue detection methods have not produced the desired results in distinguishing actions with similar appearances, such as yawning and speaking, because of the large pose deformations exhibited by the captured global face. A highly optimized multi-granularity deep model based on feature recalibration and fusion is designed, and it contains three sub networks: FFN , MEN, FRN.A multi-granularity Deep Convolutional Model for driver fatigue detection (RF-DCM) is used to detect drowsiness in this instance. The proposed method achieves an accuracy of 89.42 percent, which is significantly higher than that of the works that were compared to it.

III. METHODOLOGIES

3.1 Methodology 1

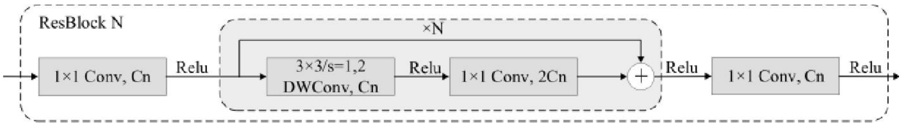
The driver fatigue detection method proposed in this model consists of two parts, namely the object detection model for detecting the driver's eyes and mouth regions and then the EM driver fatigue detection model.

Step 1: Extracting Feature Information

The Backbone network first performs standard convolution on the input image through a 3 × 3 convolution kernel. Secondly, in order to extract eye and mouth region information more quickly, the Lean ResBlock is designed. Then the three sub-networks merge with each other in a fully connected manner according to the CNN results of different models to obtain Concat-A, Concat-B, and Concat-C, which are used to extract feature information of different depths. The backbone network's design principles are:1) Make the input image smaller.2) Don't use the batch normalization (BN) layer anymore.3) Create a Lean ResBlock design.4) Create the module for High-Low Fusion.

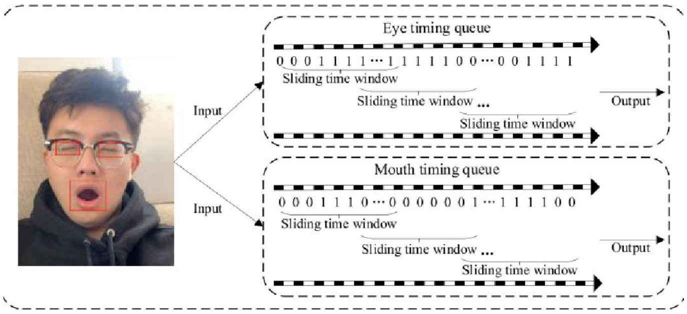
Step 2: Object Detection Network

The detection network receives the Concat-A, Concat-B, and Concat-C results. The output of the three scales is represented by the three branches of the detection network. The Concat-C module is used to combine the feature maps of the three subnets, and then a standard convolution of one-to-one results in a feature map with a size of 7 × 7 × 27. This is the first branch. The Concat-B module combines the feature maps of the three subnets in the second branch, followed by a standard one-to-one convolution to produce a 14 x 14 x 27 feature map. The Concat-A module is used to combine the feature maps of the three subnets, and then a standard convolution of one-to-one results in a feature map with a size of 28 × 28 × 27. This is the third branch.TIn order to shorten the training time of the network, anchor boxes are designed to assist in predicting bounding boxes. A single scale's output corresponds to three anchors'-means clustering is used to calculate the nine anchors in the current dataset because the designed network has a three-scale output: (10 × 16), (18 × 30), (33 × 48), (50 × 62), (78 × 93), (91 × 127), (115 × 142), (137 × 172), (161 × 198).The detection results are output following non-maximum suppression (NMS) processing.



Step 3: EM Driver Fatigue Detection Model

There are two types of fatigue detection feature parameters are extracted. To a certain extent, fatigue can be detected by each feature parameter. The number 0 indicates that the driver's eyes are open, while the number 1 indicates that they are closed. The number 0 indicates that the driver's mouth is open and the number 1 indicates that it is closed in the mouth timing queue.



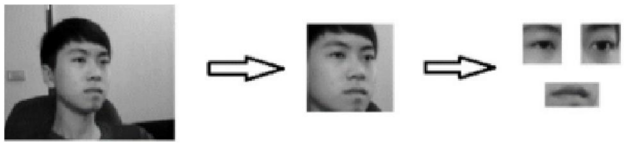
The fatigue determination module receives the driver's PERCLOS and FOM values from the sliding window.FOM=m/M, where M is the total number of frames in the sliding time window and m is the number of frames for opening the mouth, and PERCLOS=n/N, where N is the total number of frames in the sliding window and n is the number of frames with closed eyes. It is determined that the system is fatigued when either the FOM value or the PERCLOS value is greater than or equal to 0.5.

3.2 Methodology 2

The most widely used deep learning architecture is Convolutional Neural Networks (CNN). A convolutional neural network designed specifically for image classification is called InceptionV3. Recurrent neural network architecture is called long short-term memory (LSTM).

Step 1: Data Preprocessing

The videos' facial frames are taken out. The Violajones algorithm is used to determine the face. The Haar cascade method distinguishes the face from the frame. Local binary features are used to identify facial shape landmarks. The figure below depicts the extraction of the mouth and eye regions from the given frame.



Step 2: Modified Deep Learning Model

In order to extract spatial features from the mouth and eye regions, the pre-processed data are first applied to a modified InceptionV3. The pre-prepared InceptionV3 is changed by integrating the worldwide normal pooling layer. Convolutional layers, global average-pooling layers, and dropout layers are present in the modified InceptionV3. The completely associated layer is taken out and the result of altered InceptionV3 is applied on LSTM. The features obtained from the modified InceptionV3 have been used in the LSTM model. The eye region is used to examine the duration of close vision. The duration of the yawn is measured from the mouth area. The driver's state is measured by how long they yawn and close their eyes. From the given frames, the SoftMax layer of LSTM is used to determine the driver's drowsiness.

Step 3: NTHU-DDD Dataset

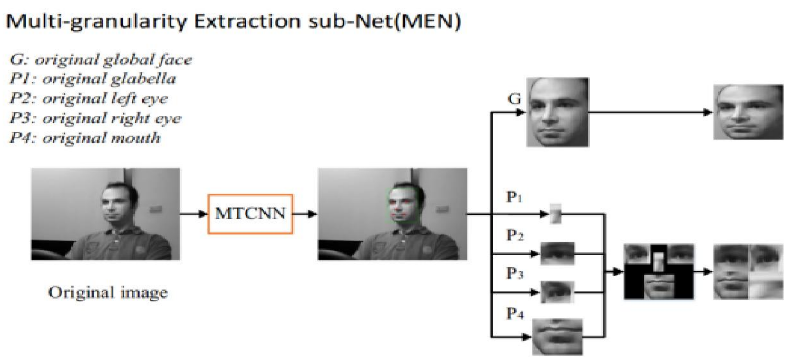
The NTHU-DDD Dataset consists of eighteen sets of distinct individuals, each with five distinct scenarios. The dataset is arranged into three groupings to be specifically 70% for preparation, 20% for approval, and 10% for testing. Understanding the features that the proposed model detects can be improved with the help of feature maps. As a result, the NTHU-DDD dataset can be used to evaluate the performance.

3.3 Methodology 3

The Multi-granularity Extraction sub-Net (MEN), the Feature Recalibration sub-Net (FRN), and the Feature Fusion sub-Net (FFN) are the primary components of the overall framework of the proposed fatigue detection model.

Step 1: Multi-Granularity Extraction Sub-Net

First, in MEN, we use MTCNN to get well-aligned facial boundaries and five landmarks, such as the positions of the left eye pupil, right eye pupil, nose tip, and left and right mouth corners with cascaded structure by choosing a slice with a width of 9/25 face and a height of 1/5 face to serve as the pupil-centered eye area. The left and right pupils, as well as the tip of the nose, determine the subsequent cutting of the mouth patch. Finally, we choose the glabella regions based on the extracted eye patches that can accurately depict states of facial fatigue. We splice multiple local regions, including P1-P4, to save memory and reduce calculation time.



Step 2: Feature Recalibration Sub-Net

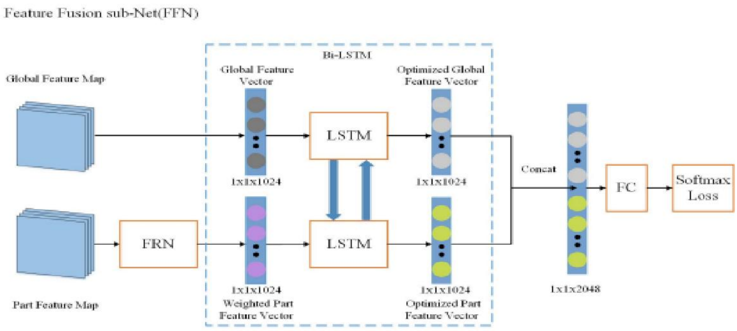
Due to the possibility of noises being introduced by the direct application of convolutional features from local streams, we have implemented FRN to recalibrate local parts features so that these part characteristics can participate in decision-making in a more proportionate manner. To rebalance the weights of fatigue significant region features, the FRN employs a combination of convolution and full connection operations that are weighted directly from the high-

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

level granularity. The network can learn to selectively emphasize informative features and suppress less useful ones through feature recalibration made possible by FRN

Step 3: Feature Fusion Sub Net

To get the spatial characteristics of fatigue detection, we must combine the recalibrated local features with global features after extracting multi-granularity in MEN and recalibrating the weights of local features using FRN. Fusing global and local features may result in noises being introduced. As a result, we suggest using the FFN rather than fully connected layers for feature fusion.



Using two LSTM models, an information interaction module based on Bi-LSTM (Bi-Long Short-Term Memory) updates the output features of each stream in the multi-stream network. A one-dimensional vector is used to represent each stream feature learned by multiple sets of convolution layers. Through the information interaction module, we combine the previously merged features before feeding them to the fully connected layers for final feature fusion

Step 4: Exploring Dynamical Characteristics

After extracting multi-granularity in MEN and recalibrating the weights of local features using FRN, we must combine the recalibrated local features with global features to obtain fatigue detection's spatial characteristics. Combining global and local features may introduce noises. Consequently, we recommend that feature fusion be performed with FFNs rather than fully connected layers.

An information interaction module based on Bi-LSTM (Bi-Long Short-Term Memory) brings the output features of each stream in the multi-stream network up to date by utilizing two LSTM models. Each stream feature that has been learned by multiple sets of convolution layers is represented by a one-dimensional vector. Before feeding them to the fully connected layers for final feature fusion, we combine the previously merged features using the information interaction module.

IV. RESULTS & DISCUSSION

4.1 Performance Metrics and Evaluation

The main performance metrics are Accuracy, Recall, F1 Score. The formulas for respective are:

Accuracy: The proportion of true positives and true negatives to all positive and negative observations is the definition of accuracy, which is a performance metric for machine learning classification models. In other words, accuracy tells us how often, out of the total number of predictions it made, our machine learning model will correctly predict an outcome.

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + FalsePositive + TrueNegative + FalseNegative}$$

F1 Score: An alternative to accuracy metrics, the F1-score is a performance metric for machine learning models that gives equal weight to Precision and Recall when evaluating their accuracy performance.

$$F1Score = 2 * \frac{Recall * Precision}{Recall + Precision}$$

IJARSCT

ISSN (Online) 2581-9429

IJARSCT
Impact Factor: 6.252

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 1, November 2022

Recall: It can be expressed mathematically as a ratio of true positives to the total of true positives and false negatives.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

Precision: It can be expressed mathematically as a ratio of True positive to False Positive and a ratio of True positive to False Positive.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

In Methodology 1, the results shows that the precision rate is 98.2% and the recall rate is 89.1%. The results are obtained based on two values known as PERCLOS and FOM values. When the PERCLOS value is greater than or equal to 0.4 and the FOM value is greater than or equal to 0.6, the system is determined to be fatigued. In Methodology 2, the precision and recall obtained from the proposed model are 74% and 92%, respectively. The value of F1-measure obtained from the proposed model is 82%, which is very promising result. LSTM attained the F1-score of 79%. The values of precision and recall obtained from InceptionV3 are 64% and 85%, respectively. CNN attained 62% precision and 83% recall.

Table 1 Performance evaluation of different deep learning models on NTHUDD dataset

Model	Accuracy (in %)	Validation (in %)	Testing (in %)
Proposed Model	93.69	91.36	81.04
LSTM	89.85	87.40	79.83
InceptionV3	82.34	79.18	72.91
CNN	80.12	77.44	71.37
MLP	70.71	72.42	61.89

Table 2 Precision, recall, and F1-measure obtained from different deep learning models

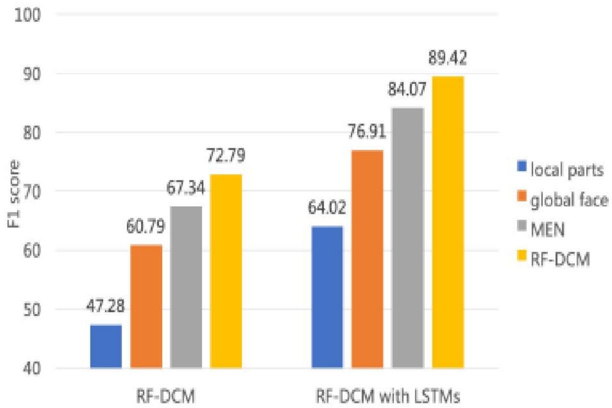
Model	Precision (in %)	Recall (in %)	F1-measure (in %)
Proposed Model	74	92	82
LSTM	71	89	79
InceptionV3	64	85	73
CNN	62	83	70
MLP	59	100	74

Table 3 Performance comparison between the proposed model and other driver drowsiness models

Model	Accuracy (in %)	F1-measure (in %)
Proposed Model	91.36	82.0
DAM [41]	78.79	44.6
ESC [42]	82.51	44.9
HTDBN [43]	84.95	79.5

The proposed model attained the accuracy of 91.36% over NTHUDD dataset. F1-measure from the proposed model is 82%, which is better than the other existing models. In Methodology 3, the proposed method achieves 89.42%

accuracy, RF-DCM achieves 72.79% accuracy, while RFDCM with LSTMs obtains 16.63% more than that by only RF-DCM.



V. CONCLUSION

The main theme is to decrease the accidents that are caused by fatigue detection. So, we used different algorithms to detect fatigue detection like CNN, modified Inception V3, RF-DCM, etc. The first methodology solves the problem of vehicle-mounted embedded devices with limited computing power and insufficient memory. Although it has obvious advantages in speed and accuracy, it cannot predict fatigue detection at the night. The Second Methodology's accuracy is quite high and it can also detect night scenes. Although the LSTMs require more time to train, it is a good approach to detect drive fatigue other than two. The third methodology's accuracy is low compared to the other two methods. So, Modified Inception V3 and LSTM is the best approach to detect the fatigue detection model. The experimental results reveal that the merging of the feature maps enables us to extract efficient features.

REFERENCES

[1]. Cui, Z., Sun, H. M., Yin, R. N., Gao, L., Sun, H. B., &Jia, R. S. (2021). Real-time detection method of driver fatigue state based on deep learning of face video. Multimedia Tools and Applications, 80(17), 25495-25515.

[2]. Muhammad, K., Ullah, A., Lloret, J., Del Ser, J., & de Albuquerque, V. H. C. (2020). Deep learning for safe autonomous driving: Current challenges and future directions. IEEE Transactions on Intelligent Transportation Systems, 22(7), 4316-4336.

[3]. Kumar, V., & Sharma, S. (2022). Driver drowsiness detection using modified deep learning architecture. Evolutionary Intelligence, 1-10.

[4]. Minhas, A. A., Jabbar, S., Farhan, M., & Najamul Islam, M. (2022). A smart analysis of driver fatigue and drowsiness detection using convolutional neural networks. Multimedia Tools and Applications, 1-18.

[5]. Huang, R., Wang, Y., Li, Z., Lei, Z., &Xu, Y. (2020). RF-DCM: multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection. IEEE Transactions on Intelligent Transportation Systems, 23(1), 630-640.

[6]. Alkinani, M. H., Khan, W. Z., & Arshad, Q. (2020). Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges. Ieee Access, 8, 105008-105030.

[7]. Magan, E., Sesmero, M. P., Alonso-Weber, J. M., & Sanchis, A. (2022). Driver Drowsiness Detection by Applying Deep Learning Techniques to Sequences of Images. Applied Sciences, 12(3), 1145.

[8]. Nguyen, D. L., Putro, M. D., & Jo, K. H. (2022). Driver Behaviors Recognizer Based on Light-Weight Convolutional Neural Network Architecture and Attention Mechanism. IEEE Access, 10, 71019-71029.

[9]. Kassem, H. A., Chowdhury, M., &Abawajy, J. H. (2021). Drivers Fatigue Level Prediction Using Facial, and Head Behavior Information. IEEE Access, 9, 121686-121697.

[10]. Yang, C., Wang, X., & Mao, S. (2020). Unsupervised drowsy driving detection with RFID. IEEE transactions

on vehicular technology, 69(8), 8151-8163.

[11]. Pandey, N. N., & Muppalaneni, N. B. (2021). Temporal and spatial feature based approaches in drowsiness detection using deep learning technique. Journal of Real-Time Image Processing, 18(6), 2287-2299.

[12]. Zhao, L., Wang, Z., Zhang, G., & Gao, H. (2020). Driver drowsiness recognition via transferred deep 3D convolutional network and state probability vector. Multimedia Tools and Applications, 79(35), 26683-26701.

[13]. Alotaibi, M., & Alotaibi, B. (2020). Distracted driver classification using deep learning. Signal, Image and Video Processing, 14(3), 617-624.

[14]. deNaurois, C. J., Bourdin, C., Stratulat, A., Diaz, E., & Vercher, J. L. (2019). Detection and prediction of driver drowsiness using artificial neural network models. Accident Analysis & Prevention, 126, 95-104.

[15]. Paulo, J. R., Pires, G., &Nunes, U. J. (2021). Cross-subject zero calibration driver's drowsiness detection: Exploring spatiotemporal image encoding of EEG signals for convolutional neural network classification. IEEE transactions on neural systems and rehabilitation engineering, 29, 905-915.