Date of publication JUN-30, 2023, date of current version OCT-14, 2022.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.22.2.3085

Classification of Sprain and Non-sprain Motion using Deep Learning Neural Networks for Ankle Sprain Prevention

Natrisha Francis, Hazwani Suhaimi and Pg Emeroylariffion Abas
Faculty of Integrated Technologies, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei

Corresponding author: Natrisha Francis (e-mail: natrisha.francis@gmail.com).

This work was supported by UBD Research Grant No: UBD/RSCH/1.3/FICBF(b)/2020/012.

ABSTRACT A smart wearable ankle sprain prevention device would require an intelligent monitoring system that can classify data from the sensors as sprain or non-sprain motion. This paper aims to explore Deep Neural Network method, specifically the Long Short Term Memory Fully Convolutional Network (LSTM-FCN) for classifying sprain and non-sprain motion. A study is conducted on 11 participants to record sprain and non-sprain motions, which are used to train and test the LSTM-FCN model and previously used Support Vector Machine (SVM) model. It has been demonstrated that the LSTM-FCN model is more accurate at classifying sprain and non-sprain motion. The LSTM-FCN also proved to be more useful as its architecture allows for the Class Activation Mapping (CAM) method to be employed. The CAM method allows for the identification of temporal regions of the time series that contribute most or least to the classification decision of the LSTMFCN. Visualizing the regions of high or low contribution makes it easy to see patterns in the data correlation with sprain motion and better understand why certain non-sprain data can be misclassified as sprain motion. Overall, LSTM-FCN is found to be a viable method for the classification of sprain and non-sprain motion.

KEYWORDS Ankle Sprain Prevention; Time Series Classification; Long Short Term Memory Fully Convolutional Network; Class Activation Mapping.

I. INTRODUCTION

ANKLE sprain continues to be an area of interest in sports medicine [1]. It is the most common musculoskeletal injury in sports [2], with ankle sprain prevalent in 24 out of 70 sports [2]. Out of the three types of ankle sprains, lateral or inversion ankle sprain is the most common type of ankle sprain, involving about 25% of all injuries to the musculoskeletal system [3]. Lateral ankle sprains result from excessive inversion of the rear foot or a combined plantar flexion and adduction of the foot, also known as supination [4]. Damage to the ligaments of the foot during ankle sprains initiates changes to the joint's biomechanics and modifies neural control of the foot [5]. This creates a negative feedback loop, thus increasing the probability of reoccurrence of ankle sprains and, subsequently, increasing the risk of developing chronic ankle instability. Prevention of ankle sprain is therefore important given the frequency of occurrence of ankle sprain, the associated risk of reinjury, and its long-term effects.

In recent years, there has been an interest in developing smart wearable devices that can identify and correct ankle sprain motion before serious ligament injury occurs. The smart wearable device would require an intelligent ankle sprain monitoring system that can detect different phases of an ankle sprain motion, and more importantly, an identification system that can classify sprain motion or non-sprain motion from its kinematic data.

With the increasing availability, affordability and reduction in size of sensors, it is now possible to attach different sensors onto an ankle to detect its time series motion. A univariate time series is a sequence of real values of a single sensor ordered according to time. When a set of time series data is recorded simultaneously by a set of sensors, it is referred to as a multivariate time series (MTS). Distance-based methods have been used to classify time series data, with Support Vector Machine (SVM) being one such distance-based method. For binary class data classification, i.e., sprain and non-sprain motions, SVM finds the hyper-plane that separates not only the two classes but also maximizes the distance between the separating hyper-plane and the closest data point of each class. Chan et al. [6] suggested using an SVM model to classify three-dimensional linear acceleration and three-dimensional angular velocity data as sprain and non-sprain motion data and obtained an accuracy of 93%. Classification using SVM was also implemented for similar problems, such as fall detection [7],

action classification [8], gait classification [9] and several other biomedical applications [10].

An increasingly popular alternative to a supervised machine learning method is the Deep Neural Network (DNN), and it was successfully implemented in various applications, including in the field of image recognition [11], text classification [12], audio classification [13], health application [14], and speech recognition [15]. These networks were designed to learn hierarchical representations of the data, and they have also been suggested for time series classification [16].

One such deep learning architecture for classifying time series is the Long Short Term Memory-Fully Convolutional Network (LSTM-FCN) proposed by Karim et al. [17]. The LSTM-FCN architecture combines a Fully Convolutional Network (FCN), known for its high performance in time-series classification, with the Long Short Term Memory (LSTM) to learn temporal dependencies in the sequences, making it suitable for classifying multivariate time series data. LSTM-FCNs have been used for other time series classification problems, including driver identification for in-vehicle software systems [18], sound classification [19], hardware monitoring [20], and human motion recognition [21]. However, it has not been used to classify sprain and non-sprain motion.

The LSTM-FCN architecture can be trained to learn hidden discriminative features from raw time series without any feature engineering of the input data. However, the black-box effect of the LSTM-FCN model makes it difficult to understand how the LSTM-FCN decides to classify the data. One of the ways of reducing the black-box effect is by employing a Class Activation Maps (CAM), with the convolutional layers of the CAM used to detect regions of the input data that contribute to the class label.

In this paper, data collection is conducted on 11 participants to record sprain and non-sprain motions, which are used to train and test the LSTM-FCN for the classification of sprain and non-sprain motions. An investigation is also performed to find the optimum features that give the highest classification accuracy. Lastly, the use of CAM to identify temporal regions of interest is also explored.

II. METHODOLOGY

Ankle kinematics data that can be utilized for the classification of sprain and non-sprain motions are not widely available, and hence, data collection is performed to be used for the classification task. These data are then processed and combined into three datasets, to compare the efficacy of the different sensor data for the classification task. These datasets are then used to test and train the classification models: the LSTM-FCN and the popular SVM models. Lastly, the CAM method is implemented on the LSTM-FCN model, to give insights on the different portion of the time-series data that are crucial for the success of the task.

A. DATA COLLECTION AND PROCEDURE

EXPERIMENTAL PROCEDURE

A study is conducted to obtain a suitable dataset for classifying sprain and non-sprain motions, which has been approved by the university ethics committee. Eleven volunteers (five males and six females) were recruited as participants, with the participants required to fill in a questionnaire, including height, weight and history of ankle sprain information, prior to the study. Participants who had previously suffered a sprain on their right ankle were requested to rate the severity of their ankle sprains on a one to five where 1 being not severe and 5 being very severe, with additional questions posed to rate the severity of their ankle sprain. These questions include the following one: if they had experienced any pain or instability and its duration.

Participants were also asked if they were still active in any sports. Selected details of the participants are given in Table 1, with the eleven participants having a mean height of 161 cm (standard deviation of 11cm) and a mean weight of 67kg (standard deviation of 24kg).

Table 1. Selected Details of the Participants

Participant Number	Height (cm)	Weight (kg)	Active in sports ?	History of an ankle sprain?	Severity of Sprain
1	183	78	No	Yes	2
2	154	50	Yes	Yes	2
3	153	59	Yes	Yes	1
4	165	64	No	Yes	3
5	152	35	No	Yes	2
6	149	52	No	Yes	2
7	161	95	No	Yes	2
8	155	96	No	Yes	2
9	163	70	No	Yes	1
10	179	89	Yes	Yes	3
11	159	65	Yes	Yes	4

Each participant performed a total of 100 trials: 50 trials of simulated supination sprain motion and 50 trials of non-sprain motion. Simulated supination sprain motions were performed on the tilting platform which had been previously designed and fabricated [22]. Each participant was required to place their feet on the rotating disc of the supination sprain simulator shoulder-width apart and weight evenly distributed on both feet, whereupon their right foot was then suddenly inverted by 20 degrees. The tilting angle was reduced if the participant showed signs of discomfort. Between each trial, subjects were allowed sufficient rest to prevent fatigue.

Non-sprain motion data collected include walking, jogging, stepping down, jump landing and cutting. Each non-sprain motion contributed to 10 trials, totaling 50 non-sprain trials. These motions were chosen because they are common in sports and daily activities. For walking, participants were requested to walk for approximately four meters at their preferred speed, whilst for jogging, participants were requested to jog on the spot for about 10 seconds.

Fig. 1 shows the participants performing selected non-sprain motions. For stepping down non-sprain motion (Fig. 1(a)), participants were requested to stand on a 30cm platform before stepping down from the platform with their right foot first. For jump landing, participants were requested to jump as high as possible and land on both feet, as depicted in Fig. 1(b). For cutting, participants were asked to run for approximately seven meters and then cut 90 degrees to their right, as shown in Fig. 1(c). For participants who were not active in sports and did not know how to perform the cutting motion correctly, the researcher would stand in front of the participants to 'block' them as they ran forward to force them to cut to the right. Participants were given time to rest in between trials.

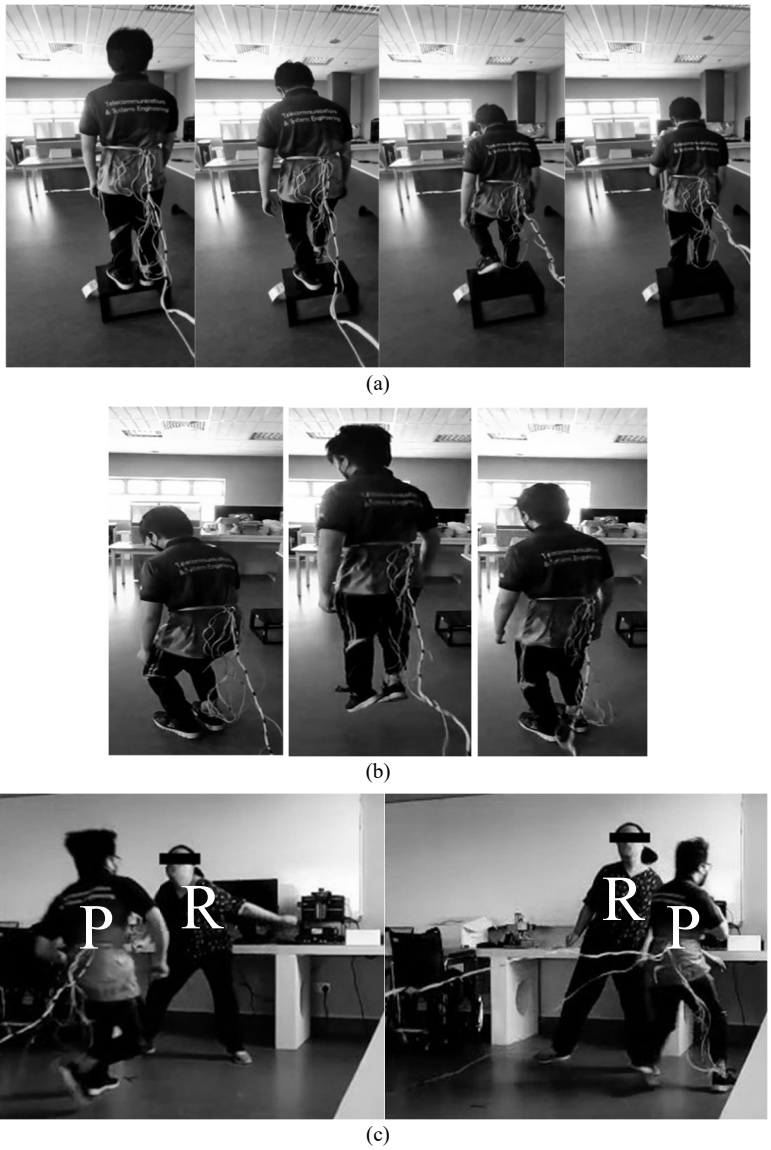


Figure 1. Non-Sprain Motions with (a) Participant stepping down from a 30cm platform, (b) Participant performing jump landing motion, and (c) Participant (P) performing cutting motion with the researcher (R) blocking the participant.

A total of 1,155 trial data were collected, as seen in Fig. 2. Data cleansing was performed to remove data discrepancies, including recording errors and time-syncing errors. 1,122 trials were included in the final dataset (550 non-sprain motion trials and 572 sprain motion trials).

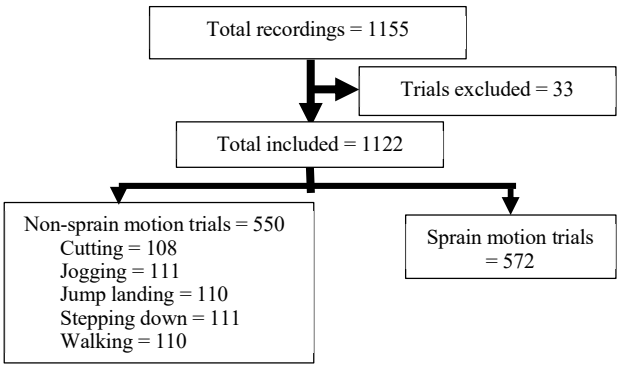


Figure 2. Flowchart of trials recorded and included in the experiment

EXPERIMENTAL SETUP

Three types of sensors: Inertial Measurement Units (IMU), flex sensors and Force Sensitive Resistors (FSR), were used to collect the time-series data. Gyroscopes and accelerometers are the most used sensors for monitoring foot kinematics during supination; whereby Preatoni et al. [23] did a systematic review on wearable sensors in injury studies, and all six studies on ankle sprain monitoring had used accelerometers, gyroscopes or both. Pressure sensors were also used to monitor ankle sprain and other lower limb injuries in some studies, with Fong et al. [24] suggesting the implantation of three pressure sensors in the soles of sports shoes for estimating and monitoring ankle supination. Flex sensors were not used in ankle sprain prevention research; however, they were used in wrist injury research to measure flexion/extension of the wrist [25].

Three FSRs were attached approximately at the fourth or fifth metatarsal-phalangeal joint (FSR1), the third metatarsal-phalangeal joint (FSR2), and the second or third distal phalange (Position 98) on the right sole of the participants' shoes (FSR3),

to give PFSR1, PFSR2, and PFSR3 as shown in Fig. 3. The three force-sensitive resistors were attached to the right shoe sole of the participants, with the FSRs calibrated to measure pressure (N/cm2).

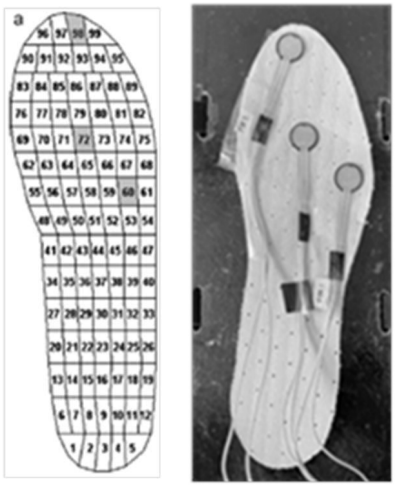


Figure 3. Location of the three FSRs on the right sole of the shoe

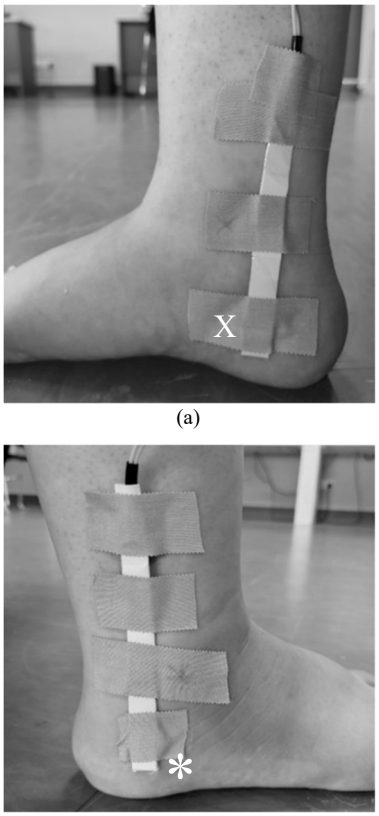
In addition to recording pressure, the pressure values at these three points were also used to calculate supination torque $SupT$ (Nm) [24]:

$$SupT\ (Nm) = -2.068 + 0.910\ (P_{FSR1}) + 1.318\ (P_{FSR2}) + 1.549\ (P_{FSR3}). \quad (1)$$

Additionally, two flex sensors, which detect the contraction of the leg muscles, were used to measure the inversion angle of the foot, with the flex sensors calibrated to measure the inversion angle in degrees. The flex sensors were placed in a paper sleeve, with the sleeve attached behind the medial malleus and the lateral malleus of the participants' foot, as seen in Fig. 4. Positions of the medial and lateral malleus were determined by palpitating the foot of the participants. The flex sensors were not directly attached to the participant's foot to allow free movement of the foot during inversion.

An IMU (x-io technologies) was used to obtain angular velocity and linear acceleration during the trials. Chan et al. [6] suggested that positioning the motion sensors on the medial malleus would give the highest signal strength. However, attaching the housed IMU directly to the medial malleolus would have caused discomfort and restricted movement of the foot, and as such, the IMU was slipped into the side of the participant's shoe below the medial malleolus, as seen in Fig. 5, with tape used to secure and prevent the IMU from slipping out of the shoe during movement.

A total of 11 features were recorded: angular velocity in the X-axis (GyroX), Y-axis (GyroY) and Z-axis (GyroZ), linear acceleration in the X-axis (AccX), Y-axis (AccY), and Z-axis(AccZ), inversion angle from the flex sensors on the medial side (Flex1) and lateral side (Flex2), pressures at FSR1, FSR2 and FSR3. Supination Torque (SupT) was calculated from FSR1, FSR2 and FSR3.



(a)

(b)

Figure 4. Flex sensor attached to participant's foot (a) medial view of the foot. X is the position of the medial malleolus. (b) lateral view. The asterisk (*) is the position of the lateral malleolus.

Figure 5. Position of the housed IMU below the medial malleolus of the foot

Data from the FSRs and flex sensors were collected using an Arduino microcontroller, whilst data from the IMU were collected using the provided IMU Graphical User Interface (GUI). Recorded data were manually synchronized according to the Local Coordinated Time. All data from a particular trial were combined into one comma-separated value file (.csv), with the time-series data segmented to 1 second.

Three datasets were constructed, which are shown in Table 2. Dataset A has $F_A = 4$ features consisting of kinematic data: angular velocity in the y-axis, inversion angle and supination torque. These data were used in our previous work for the demonstration of the tilting platform [22]. Dataset B has $F_B = 4$ features corresponding to data from the tri-axial gyroscope and accelerometer of the IMU. Finally, dataset C consists of all $F_C = 11$ features recorded, with the exception of Supination Torque, as it is a function of FSR1, FSR2 and FSR3, which have been included in the 11 features of dataset C.

Table 2. The Three Datasets and Their Features

Dataset	Number of Features ($F_i, i \in \{A, B, C\}$)	Features included
A	4	GyroY, Flex1, Flex2, SupT
B	6	GyroX, GyroY, GyroZ, AccX, AccY, AccZ
C	11	GyroX, GyroY, GyroZ, AccX, AccY, AccZ, Flex1, Flex2, FSR1, FSR2, FSR3

B. CLASSIFICATION MODELS

Two classification models are considered: SVM and LSTM-FCN models. SVM was used in previous studies on the classification of sprain and non-sprain motions. LSTM-FCN was shown to give high accuracy in classifying time series, albeit not for ankle sprain classification, as well as its structure which allows for the implementation of the CAM method.

SUPPORT VECTOR MACHINE (SVM)

A binary SVM model is trained to classify either sprain or non-sprain motions from sensor data. SVM maps data points onto a high-dimensional space and finds the optimal hyperplane that divides the data into two classes. A radial basis function kernel was used, defined by,

$$RBF(x_i, x_i') = e^{\left(-\frac{\|x_i - x_i'\|^2}{2\sigma^2}\right)}, \quad (2)$$

where $\|x_i - x_i'\|^2$ is the squared Euclidean distance between feature i of different classes, and σ2 is the variance of the data array. For the SVM model, time-series training and testing data $x_i(t), \forall i, t$, were first converted into the frequency domain representation, $X_i(f)$ using Discrete Fourier Transform (DFT). The frequency domain representations display important spectral information related to the motion. A similar pre-processing procedure was adopted in previous fall detection studies [7] and ankle sprain classification [6].

LONG SHORT TERM MEMORY-FULLY CONVOLUTIONAL NETWORK

Karim et al. [17] proposed the LSTM-FCN, with general architecture as shown in Fig. 6, with the FCN block augmented by an LSTM block [17]. The datasets were reshaped in a 3D array (s, T_{window}, F_i), where s represents the sample, T_{window} is the lookback period and F_i is the number of features in dataset $i \in \{A, B, C\}$. A lookback period of $T_{window} = 400$ is taken. The input sizes onto the LSTM-FCN model for datasets A, B and C are (400, 4), (400, 6) and (400, 11), respectively.

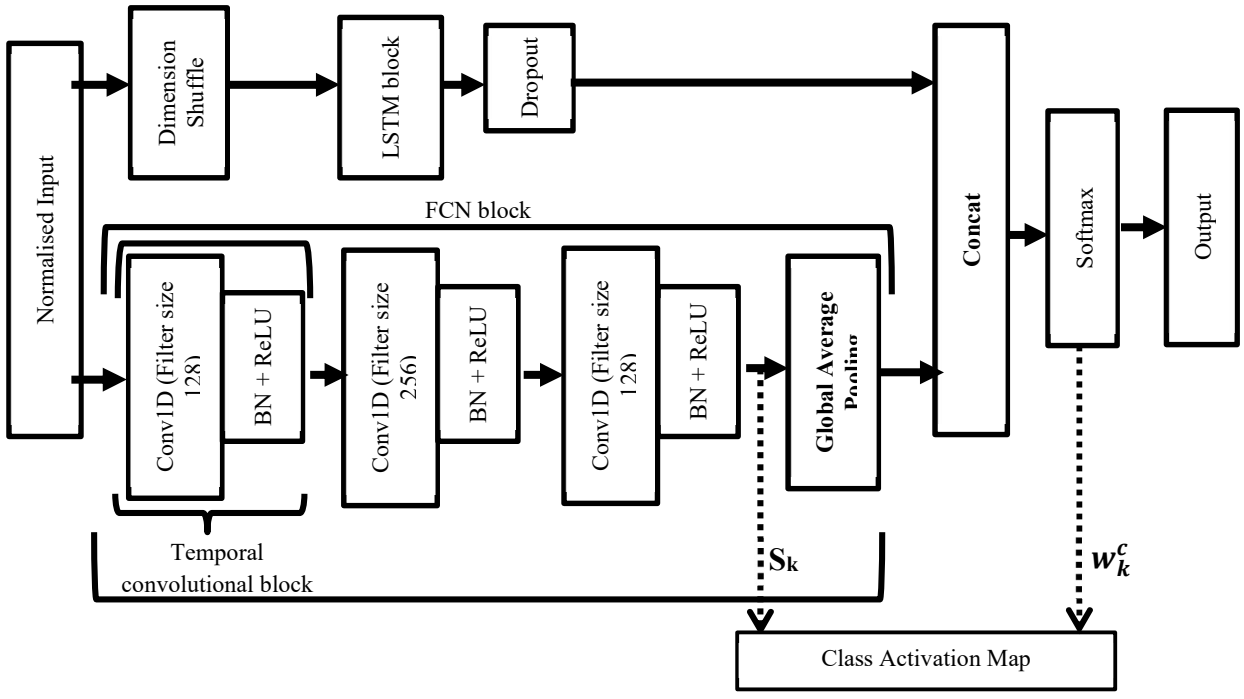


Figure 6. Architecture of LSTM-FCN model (adapted from [17]) and the CAM method, which uses the activation of the filters of the last convolutional layer, Sk, and the weights of the output, w_k^c

Similar to other deep learning models, the LSTM-FCN model does not require any feature engineering. However, input data is commonly normalized, by transforming its value to be between 0 and 1, using the Min-Max scaler in equation (3), before feeding into the LSTM-FCN model [26]. The normalization equation is given as

$$y_j(t) = \frac{x_j(t) - x_{i,min}}{x_{i,max} - x_{i,min}}, \quad (3)$$

where $x_j(t)$ and $y_j(t)$ are the original and normalised values of feature j at time t, respectively, whilst $x_{i,min}$ and $x_{i,max}$ are the minimum and maximum values of feature i for all t, respectively. After pre-processing, the normalized time series input is fed onto the LSTM-FCN model as multivariate time series data. A dropout rate of the LSTM block is set to 0.8, as suggested by Karim et al. [17].

The FCN block views the time series as a univariate time series with multiple time steps. In the LSTM-FCN architecture, the FCN block consists of three stacked temporal convolutional

blocks with filter sizes of 128, 256, and 128 and kernel sizes of 8, 5 and 3, respectively. Each temporal convolutional block consists of a temporal convolutional layer, which is accompanied by batch normalization (BN) followed by a rectified linear (ReLU) activation function. Batch normalization is a method of normalizing the layers' inputs by re-centering and rescaling [27]. Finally, global average pooling (GAP) is applied after the last temporal convolutional block, with the GAP layer outputting the spatial average of the feature map of each unit at the last convolutional layer.

The output of the global pooling layer and the LSTM block are then concatenated and passed onto a softmax layer, with the softmax function normalizing the network's output to a probability distribution over the predicted output class. The training epoch of the LSTM-FCN is set to 20, with batch size set to 128.

CLASS ACTIVATION MAPPING (CAM) METHOD

The use of CAM was first introduced by Zhou et al. [28] for image classification. CAM highlights parts of an image that contribute the most to a given class identification. Wang et al. [29] introduced a one-dimensional CAM for time series classification by using it to find the contributing temporal regions in the raw data for the specific label. CAM is implemented after the training and testing of the LSTM-FCN model have been completed. The CAM method uses the averaged feature map of each unit after global average pooling (GAP) and the weights of the final softmax function, as shown in Fig. 6. In the LSTM-FCN, GAP outputs the spatial average of the feature map of each unit at the last convolutional layer. For a time series, let $S_k(t)$ represents the activation of filter k in the last convolutional layer at temporal location t. The output of the following GAP (fk) was given by [29].

$$f_k = \sum_t S_k(t). \quad (4)$$

A weighted sum of these values is inputted into a softmax layer to generate the final output, with the input to the softmax function given by [29]:

$$g_c = \sum_k w_k^c \sum_t S_k(t), \quad (5)$$

where w_k^c indicates the weight of the final softmax function for the output from filter k and class c. The CAM (CAMc) for class c was given by [29]:

$$CAM_c = \sum_t w_k^c S_k(t). \quad (6)$$

Hence, CAMc indicates the importance of the activation at temporal location ti leading to the classification of the time series to a certain class (c), in this case, either sprain or non-sprain. The CAM outputs, CAMc = {z1,z2,…,zn}, were rescaled so that the CAM values would fall between -1 and 1 using (7), where min(CAMc) and max(CAMc) are the minimum, and maximum values of CAM outputs, respectively, whilst zi and Zi are the original and normalized CAM, respectively,

$$Z_i = 2\left(\frac{z_i - \min(CAM_c)}{\max(CAM_c) - \min(CAM_c)}\right) - 1 \quad (7)$$

III. RESULTS AND DISCUSSIONS

A. CLASSIFICATION OF SPRAIN AND NON-SPRAIN MOTIONS

The data were manually labelled to correspond to either non-sprain or sprain motions. After pre-processing, 70% (785 trials) of the dataset was randomly set aside for training the classification models to identify either sprain or non-sprain motions, whilst 30% (337 trials) were used for testing. The testing dataset was assigned an index number to ease the identification of trial data (whether it was sprain or non-sprain, type of non-sprain motion) for analysis.

Of the 337 trials in the testing dataset, 183 and 154 were sprain and non-sprain motions, respectively. The accuracy, precision, recall, and f1-score of the SVM and LSTM-FCN models on the three datasets are given in Table 3, with confusion matrices of the SVM and LSTM-FCN models on dataset A, dataset B, and dataset C, tabulated in Table 4(a), 4(b) and 4(c), respectively.

Table 3. Accuracy, precision, recall, and f1-score of SVM and LSTM-FCN models on the three datasets

	Dataset A		Dataset B		Dataset C	
	SVM	LSTM -FCN	SVM	LSTM -FCN	SVM	LSTM -FCN
Accuracy	0.976	0.976	0.982	0.997	0.988	0.997
Precision	0.989	0.983	0.978	0.995	0.984	0.995
Recall	0.967	0.973	0.989	1.000	0.995	1.000
F1 score	0.978	0.978	0.984	0.997	0.989	0.997

Table 4. Confusion matrices for the SVM and LSTM-FCN models on (a) Dataset A, (b) Dataset B and (c) Dataset C

(a)			Class Predicted by SVM		Class Predicted by LSTM-FCN	
			Non-Sprain	Sprain	Non-Sprain	Sprain
Actua Class	Non-sprain	Stepping down	29	1	30	0
		Cutting	24	1	22	3
		Jogging	33	0	33	0
		Walking	35	0	35	0
		Jump Landing	31	0	31	0
	Sprain		6	177	5	178

(b)			Class Predicted by SVM		Class Predicted by LSTM-FCN	
			Non-Sprain	Sprain	Non-Sprain	Sprain
Actua Class	Non-sprain	Stepping down	27	3	30	0
		Cutting	24	1	24	1
		Jogging	33	0	33	0
		Walking	35	0	35	0
		Jump Landing	31	0	31	0
	Sprain		2	181	0	183

(c)			Class Predicted by SVM		Class Predicted by LSTM-FCN	
			Non-Sprain	Sprain	Non-Sprain	Sprain
Actua Class	Non-sprain	Stepping down	28	2	30	0
		Cutting	24	1	24	1
		Jogging	33	0	33	0
		Walking	35	0	35	0
		Jump Landing	31	0	31	0
	Sprain		1	182	0	183

Accuracy, defined in (8), refers to the ratio of correctly classified trials to the total trials, with TN and TP referring to

the number of non-sprain and sprain motions, respectively, that were correctly classified, and FP and FN referring to the number of non-sprain and sprain motions, respectively, that were incorrectly classified.

$$Accuracy = \frac{TP+}{TP+TN+FP+FN}. \quad (8)$$

Precision is the ratio of the number of trials correctly classified as sprain motion to the total number of trials predicted as sprain motion, whilst recall is the ratio of the number of trials correctly classified as sprain motion to the total number of sprain motion.

$$Precision = \frac{TP}{TP+FP}, \quad (9)$$

$$Recall = \frac{TP}{TP+F}. \quad (10)$$

F1 score gives the harmonic average of precision and recall scores.

$$F1\ score = 2(\frac{Precision*Recall}{Precision+Rec}) \quad (11)$$

It can be seen that the LSTM-FCN model gives the highest accuracy score of 0.997 on both datasets B and C, as compared to that of the SVM model, with accuracy scores of 0.982 and 0.988 on datasets B and C, respectively. However, both LSTM-FCN and SVM models give the same accuracy score of 0.976 on dataset A. Despite the same accuracy score on dataset A, the LSTM-FCN model has a better recall, whilst the SVM model has better precision. This is because whilst the LSTM-FCN model performs better in predicting sprain motions (wrongly predicting only 5 sprain motions as non-sprain motions, as can be seen from Table 4(a)), the model wrongly predicted more non-sprain motions as sprain motions (wrongly predicting 3 non-sprain motions as sprain motions).

On datasets B and C, the LSTM-FCN model performs better in predicting both sprain and non-sprain motions, giving superior accuracy, precision, recall, and f1-score as compared to the SVM model. Accuracy, precision, recall and f1-score of the LSTM-FCN model on both datasets B and C are 0.997, 0.995, 1 and 0.997, respectively. High recall indicates that the model is able to accurately identify the majority of the sprain motions, and hence, within the context of an intelligent ankle sprain prevention device, it is important when the device solely relies on the model to detect ankle sprain motion for the activation of its preventive mechanism. On the other hand, precision is important if the preventive mechanism is overly restrictive to the activity of the wearer or if the preventive mechanism is troublesome to be deactivated, in which case, reducing the misidentification of non-sprain motions as sprain motions is of importance.

Comparing results between datasets A, B and C, it can be seen that the accuracy of the SVM model is highest on dataset C with 0.988 and lowest on dataset A with 0.976. These indicate that 11 features need to be utilized (in the case of dataset C) for the SVM model, which necessitates using the IMU, FSRs and flex sensors, to extract the best performance. On the other hand, it can be observed that using only 6 features (in the case of dataset B) gives the best performance for the LSTM-FCN model. The 6 features can be obtained from the IMU sensor only. This is significant as only a single IMU device is required for the LSTM-FCN model to classify sprain and non-sprain motions with a high 0.997 accuracy.

As can be seen from the confusion matrices in Table 4, non-sprain motion trials that were wrongly classified as sprain motions were either cutting or stepping down motions; with the LSTM-FCN model wrongly classifying cutting motion only, whilst the SVM model wrongly classifying both stepping down and cutting motions. These are similar to Chan et al. [102], with stepping down and cutting contributing to most of the false positives detected by their SVM model. Additionally, Chan et al. [102] also reported false classifications of jump landing motion as sprain motion. These false classifications on the cutting, stepping down and jump landing motions may be due to the sudden changes in ankle motion on these 3 motions, almost similar to ankle sprain motion, albeit with slight differences. As can be seen from the result, the LSTM-FCN model is able to effectively differentiate between sprain motion and the other non-sprain motions, with the exception of cutting motion.

Overall, it is noted that the LSTM-FCN model performs better in identifying both sprain and non-sprain motions on all 3 datasets, with the exception of the identification of non-sprain motions on dataset A. On dataset A, the LSTM-FCN model wrongly predicted 3 non-sprain motions as sprain motions and 5 sprain motions as non-sprain motions. This is in contrast to the SVM model, which wrongly predicted only 2 non-sprain motions as sprain motions, but 6 sprain motions as non-sprain motions. Overall, however, both the LSTM-FCN and SVM models demonstrate comparable performance in terms of accuracy on dataset A.

B. CLASS ACTIVATION MAPS

Generated CAM graph shows the contribution of the time point of the data to the LSTM-FCN model's decision in classifying the multivariate time series as a sprain or non-sprain motion. Since CAM values are rescaled to range from -1 to 1, for the experiment, values above 0.5 indicate a high contribution of the time point to the classification decision, whilst values below -0.5 indicate a low or no contribution to the classification decision. The CAM graph is then used to highlight temporal regions of the input data that contribute the most or the least to the classification system. Fig. 7(a) demonstrates a CAM graph, with Fig. 7(b) and (c) depicting the corresponding GyroX and Flex1 sensor readings during one of the sprain motion trials from dataset A. CAM values were used to assign a color code to the time points of the data, with red colored plot indicating high contribution (CAM above 0.5) points and dark blue colored plot indicating low contribution (CAM below -0.5) points. The time points with the higher CAM value (Fig. 7(a)) are indicated with dark red color with a pointing black arrow. Coloring the time points allows us to visually see temporal regions of high and low contributions in the classification of the trial as a sprain or non-sprain motion. For instance, the highest contribution in classifying the data as sprain motion is when initial changes in angular velocity in X-axis occur (as seen in Fig. 7(b)) and a few moment before the inversion angle recorded by the medial flex sensor increases (as seen in Fig. 7(c)).

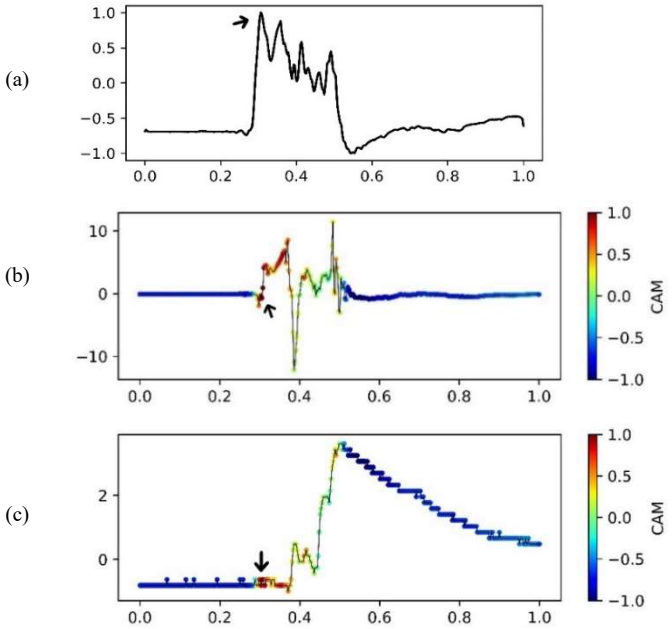


Figure 7. (a) Plot of CAM values (b) Plot of GyroX with colored points corresponding to CAM value (c) Plot of Flex1 with colored points corresponding to CAM value. Black arrows show dark red points of GyroX and Flex1 plot corresponding to the highest CAM value

The number of CAM points above 0.5 were recorded onto a spreadsheet for each sprain motion testing trial that was correctly classified as sprain motion in all three datasets. The average number of CAM points above 0.5 for datasets A, B and C are 19.0, 12.8 and 20.2, respectively (with variances of 8.6, 6.4 and 10.1, respectively), for a total of 100 points for every trial. A smaller number of CAM values above 0.5 indicates fewer time points are needed to be considered in the classification decision, with a possible implication that in future works, a shorter time series can be inputted into the LSTM-FCN model for classification purposes. Comparing datasets B and C, dataset C has a larger average number of CAM points above 0.5 with a larger variance. The addition of data from the flex sensors and FSRs may have forced the LSTM-FCN model to consider more time points during the classification process. However, although dataset B had more features than dataset A, dataset A had more CAM points above 0.5. This suggests that one or more features found in dataset B and not in dataset A could have reduced the number of time points contributing highly to the LSTM-FCN's classification decision.

Fig. 8 shows the contributing temporal regions for the LSTM-FCN model on datasets A, B and C for one of the correctly classified sprain trials (trial index no 217). The trial shown in Fig. 8 is not representative of all sprain trials; however, it was chosen to illustrate how CAM values may differ between the three datasets. Only GyroY is compared as it is the only common feature in all three datasets. Fig. 8 shows that the temporal regions that have positive CAM values (time points colored yellow and red) are similar across the 3 datasets, where there is a large decrease and increase in angular velocity on the Y-axis. This is in line with how angular velocity is expected to change when supination motion occurs. There are visible differences in the amount of interest and disinterest at certain time points. The high contributing time points are at peak angular velocity before the sharp decrease for dataset A (Fig 8(a)) and dataset B (Fig 8(b)), but for dataset C (Fig 8(c)), high contributing time points are when the angular velocity is lowest before the sharp increase.

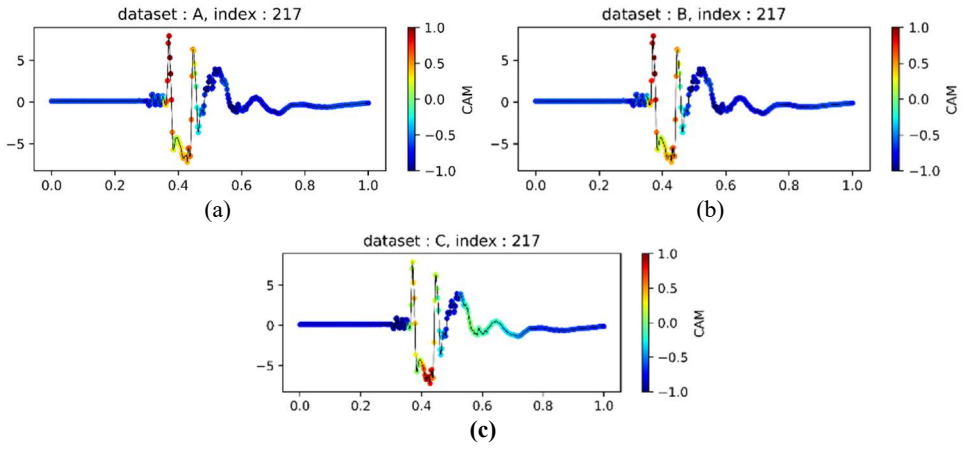


Figure 8. Plots of GyroY showing points of interest to the LSTM-FCN model for (a) Dataset A, (b) Dataset B and (c) Dataset C

The CAM method also allows us to see which region of the data contributed to the incorrect classification of sprain data. Fig. 9 shows the CAM results of cutting motion trials (trials 6 and 274) from dataset A that were incorrectly identified as sprain motion as examples. The time points that highly contribute to the misclassification as sprain motion are colored red. For trial 6 (Fig. 9(a)), it can be seen that a decrease followed by an increase in angular velocity (GyroY), and an increase in inversion angle (Flex1 and Flex2) recorded during the cutting motion, are very similar to what is expected in sprain motions, leading to the misclassification. For trial 274 (Fig. 9(b)), the CAM values are high when there is a small spike in angular velocity (GyroY), inversion angle (Flex 2) and supination torque (SupT), which led to the misclassification of the cutting motion. Further analysis of CAM values of cutting motions can help better understand which temporal region causes the misclassification and determine if there is a pattern.

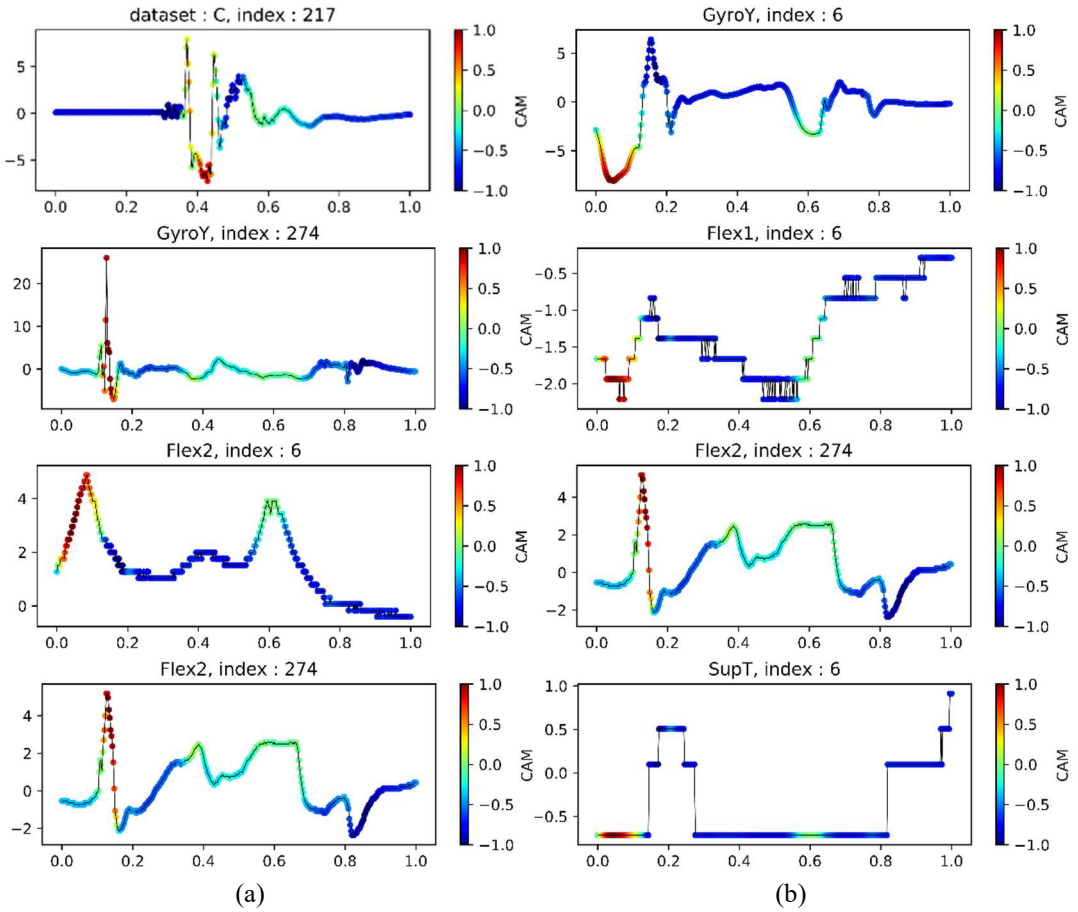


Figure 9. GyroY, Flex1, Flex2 and SupT of cutting motion trials and CAM value for classification as sprain motion for (a) trial 6 and (b) trial 274

As discussed, the CAM method is very helpful in understanding and visualizing the temporal regions of the time series that contribute most or least to the classification by the LSTM-FCN model. However, there are drawbacks to the CAM method. This method only shows the temporal regions of high and low contribution in the data. It does not explain why or how the region contributes to the classification decision. We can only infer why and how it contributes if a visible pattern in the regions of interest is what defines the motion as a sprain motion. It does not show if the LSTM-FCN considers just one feature or if multiple dependent features are considered in the decision-making process. Knowing the feature/features that contribute most to the classification would reduce the size of the input and improve the efficiency of the monitoring system. Further analysis with more data or different DNN methods would be needed to determine which feature/features contribute the most to the classification decision.

IV. CONCLUSION

An SVM model was used in a previous study [6] to classify sprain and non-sprain motion successfully. However, DNNs have recently been considered for classifying time series data rather successfully, albeit not specifically for the classification of an ankle sprain. In this paper, data collection is made on 11 participants performing different trials on sprain and non-sprain motions by collecting different ankle kinematics data. Both SVM and LSTM-FCN models are trained to classify sprain and non-sprain motions on the three different datasets obtained from the experiment. Comparison results have shown that the LSTM-FCN model is better at classifying sprain and non-sprain motions, with maximum accuracy of 0.997. It has also been shown that the use of features from a single IMU

device to track 6 parameters related to ankle kinematics is sufficient to give high accuracy using the LSTM-FCN model. Additionally, the use of the LSTM-FCN model also allows the use of the CAM method to identify temporal regions in the inputted data that contribute to the classification system and patterns associated with sprain motion in the data. The CAM method can also be used to visualize data points that contribute to the misclassification of non-sprain motion as sprain motion, allowing us to better understand what leads to the misclassification for future works.

References

[1] C. Doherty, E. Delahunt, B. Caulfield, J. Hertel, J. Ryan, and C. Bleakley, "The incidence and prevalence of ankle sprain injury: A systematic review and meta-analysis of prospective epidemiological studies," *Sport. Med.*, vol. 44, no. 1, pp. 123–140, 2014, https://doi.org/10.1007/s40279-013-0102-5.

[2] D. T. Fong, Y. Hong, L. Chan, P. S. Yung, and K. Chan, "A systematic review on ankle injury and ankle sprain in sports," *Sport. Med.*, vol. 37, no. 1, pp. 73–94, 2007. https://doi.org/10.2165/00007256-200737010-00006.

[3] M. P. J. van den Bekerom, G. M. M. J. Kerkhoffs, G. A. McCollum, J. D. F. Calder, and C. N. van Dijk, "Management of acute lateral ankle ligament injury in the athlete," *Knee Surgery, Sport. Traumatol. Arthrosc.*, vol. 21, no. 6, pp. 1390–1395, 2013, https://doi.org/10.1007/s00167-012-2252-7.

[4] E. Delahunt, G. F. Coughlan, B. Caulfield, E. J. Nightingale, C. W. C. Lin, and C. E. Hiller, "Inclusion criteria when investigating insufficiencies in chronic ankle instability," *Med. Sci. Sports Exerc.*, vol. 42, no. 11, pp. 2106–2121, 2010, https://doi.org/10.1249/MSS.0b013e3181de7a8a.

[5] J. Hertel, "Sensorimotor deficits with ankle sprains and chronic ankle instability," *Clin. Sports Med.*, vol. 27, no. 3, pp. 353–370, 2008, https://doi.org/10.1016/j.csm.2008.03.006.

[6] Y. Y. Chan et al., "Identification of ankle sprain motion from common sporting activities by dorsal foot kinematics data," *J. Biomech.*, vol. 43, no. 10, pp. 1965–1969, 2010, https://doi.org/10.1016/j.jbiomech.2010.03.014.

[7] G. Shi, C. S. Chan, W. J. Li, K. S. Leung, Y. Zou, and Y. Jin, "Mobile human airbag system for fall protection using mems sensors and embedded SVM classifier," *IEEE Sens. J.*, vol. 9, no. 5, pp. 495–503, 2009, https://doi.org/10.1109/JSEN.2008.2012212.

[8] A. Sadiq, S. G. Khawaja, M. U. Akram, N. S. Alghamdi, A. Khan, and A. Shaukat, "Machine learning and signal processing based analysis of sEMG signals for daily action classification," *IEEE Access*, vol. 10, pp. 40506–40516, 2022, https://doi.org/10.1109/ACCESS.2022.3166885.

[9] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 5, pp. 828–838, 2005, https://doi.org/10.1109/TBME.2005.845241.

[10] K. R. Foster, R. Koprowski, and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research – Commentary," *BioMedical Engineering OnLine*, vol. 13, no. 1, p. 94, 2014, https://doi.org/10.1186/1475-925X-13-94.

[11] F. A. Azis, H. Suhaimi, and E. Abas, "Waste classification using convolutional neural network," *Proceedings of the ACM International Conference Proceeding Series*, pp. 9-13, 2020, https://doi.org/10.1145/3417473.3417474.

[12] M. A. Humayun, H. Yassin, and P. E. Abas, "Native language identification for Indian-speakers by an ensemble of phoneme-specific, and text-independent convolutions," *Speech Commun.*, vol. 139, pp. 92–101, 2022, https://doi.org/10.1016/j.specom.2022.03.007.

[13] X. Fan, T. Sun, W. Chen, and Q. Fan, "Deep neural network based environment sound classification and its implementation on hearing aid app," *Measurement*, vol. 159, p. 107790, 2020, https://doi.org/10.1016/j.measurement.2020.107790.

[14] D. Gupta, U. Kose, and O. Castillo, "Special issue on deep neural networks for biomedical data and imaging," *Expert Systems*, vol. 39, no. 3, 2022, https://doi.org/10.1111/exsy.12943.

[15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, https://doi.org/10.1038/nature14539.

[16] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "Deep learning for time series classification: A review," *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019, https://doi.org/10.1007/s10618-019-00619-1.

[17] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2017, https://doi.org/10.1109/ACCESS.2017.2779939.

[18] A. El Mekki, A. Bouhoute, and I. Berrada, "Improving driver identification for the next-generation of in-vehicle software systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7406–7415, 2019, https://doi.org/10.1109/TVT.2019.2924906.

[19] Y. Kim, J. Sa, Y. Chung, D. Park, and S. Lee, "Resource-efficient pet dog sound events classification using LSTM-FCN based on time-series data," *Sensors* (Switzerland), vol. 18, no. 11, pp. 4019, 2018, https://doi.org/10.3390/s18114019.

[20] N. Rücker, L. Pflüger, and A. Maier, "Hardware failure prediction on imbalanced times series data: Generation of artificial data using Gaussian process and applying LSTMFCN to predict broken hardware," *J. Digit. Imaging*, vol. 34, no. 1, pp. 182–189, 2021, https://doi.org/10.1007/s10278-020-00411-4.

[21] Z. Chen, X. Chen, Y. Ma, S. Guo, Y. Qin, and M. Liao, "Human posture tracking with flexible sensors for motion recognition," *Comput. Animat. Virtual Worlds*, vol. 32, no. 5, pp. 1–13, 2021, https://doi.org/10.1002/cav.1993.

[22] N. Francis, A. Ong, H. Suhaimi, and P. E. Abas, "A tilting platform as a sub-injury motion for ankle sprain studies," *Proceedings of the 2021 IEEE National Biomedical Engineering Conference (NBEC)*, 2021, pp. 117–121, https://doi.org/10.1109/NBEC53282.2021.9618764.

[23] E. Preatoni et al., "The use of wearable sensors for preventing, assessing, and informing recovery from sport-related musculoskeletal injuries: A systematic scoping review," *Sensors*, vol. 22, no. 9, p. 3225, 2022, https://doi.org/10.3390/s22093225.

[24] D. T. P. Fong, Y. Y. Chan, Y. Hong, P. S. H. Yung, K. Y. Fung, and K. M. Chan, "A three-pressure-sensor (3PS) system for monitoring ankle supination torque during sport motions," *J. Biomech.*, vol. 41, no. 11, pp. 2562–2566, 2008, https://doi.org/10.1016/j.jbiomech.2008.05.035.

[25] R. M. Greenwald, F. H. Simpson, and F. I. Michel, "Wrist biomechanics during snowboard falls," *Proc. Inst. Mech. Eng. Part P J. Sport. Eng. Technol.*, vol. 227, no. 4, pp. 244–254, 2013, https://doi.org/10.1016/j.jbiomech.2008.05.035.

[26] F. Karim, S. Majumdar, and H. Darabi, "Insights into lstm fully convolutional networks for time series classification," *IEEE Access*, vol. 7, pp. 67718–67725, 2019, https://doi.org/10.1109/ACCESS.2019.2916828.

[27] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *Proceedings of the 32nd International Conference on International Conference on Machine Learning (ICML'15)*, volume 37, 2015, pp. 448–456, https://doi.org/10.48550/arXiv.1502.03167.

[28] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," Proceedings of the IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2016, pp. 2921–2929, https://doi.org/10.1109/CVPR.2016.319.

[29] Z. Wang, W. Yan, and T. Oates, "|Time series classification from scratch with deep neural networks: A strong baseline," *Proceedings of the 2017 IEEE International Joint Conference on Neural Networks (IJCNN)*, 2017, vol. 57, no. 7, pp. 1578–1585, https://doi.org/10.1109/IJCNN.2017.7966039.

***NATRISHA FRANCIS** graduated with a Bachelors in Biomedical Enigneering in 2019 at the University of Glasgow. She is currently pursuing the MEng degree in Chemical and Process Engineering with the University Brunei Darussalam. Her research topic is ankle sprain prevention.*

HAZWANI SUHAIMI *obtained her BEng in Chemical Engineering (2009), MSc in Advanced Chemical Engineering with IT and Management (2010), and PhD in Chemical Engineering (2015) from Loughborough University. In January 2016, Hazwani joined Universiti Brunei Darussalam as a lecturer in Chemical and Process Engineering, and was promoted to Assistant Professor in December 2021. She has been the Programme Leader of the Chemical and Process Engineering since August 2016. Her major research is on the nutrient diffusion in tissue engineering membranes and scaffolds. Over the years, she has expanded her research activities and interests including in the area of ankle injury, rehabilitation, sustainability, computer modeling and simulation, and membrane science.*

PG EMEROYLARIFFION ABAS *received his B.Eng. Information Systems Engineering from Imperial College, London in 2001, before obtaining his PhD Communication Systems in 2005 from the same institution. He is now working as a Senior Assistant Professor in System Engineering, Faculty of Integrated Technologies, Universiti Brunei Darussalam. His present research interest are data analysis, IoT and Sensor Network, photonics, and energy system.*