Recognition of walking activity and prediction of gait periods with a CNN and first-order MC strategy

Uriel Martinez-Hernandez, Adrian Rubio-Solis and Abbas A. Dehghani-Sanij

Abstract—In this paper, a strategy composed of a Convolutional Neural Network (CNN) and a first-order Markov Chain (MC) is presented for recognition of walking activities and gait periods using wearable sensors. First, a Convolutional Neural Network (CNN) is developed for the recognition of three walking activities (level-ground walking, ramp ascent and descent) and gait periods. Second, a first-order Markov Chain (MC) is employed for the prediction of gait periods, based on the observation of decisions made by the CNN for each walking activity. The validation of the proposed methods is performed using data from three inertial measurement units (IMU) attached to the lower limbs of participants. The results show that the CNN, together with the first-order MC, achieves mean accuracies of 100% and 98.32% for recognition of walking activities and gait periods, respectively. Prediction of gait periods are achieved with mean accuracies of 99.78%, 97.56% and 97.35% during level-ground walking, ramp ascent and descent, respectively. Overall, the benefits of our work for accurate recognition and prediction of walking activity and gait periods, make it a suitable high-level method for the development of intelligent assistive robots.

I. INTRODUCTION

Recognition of activities of daily living (ADLs) is an important capability required in autonomous systems to deliver safe and accurate assistance to humans [1], [2]. Activities such as walking, ramp ascent/descent and sit-to-stand provide independence of living and transportation across different terrains, which make them particularly important for research of computational recognition methods [3], [4], [5].

In recent years, a rapid progress has been observed in sensor technology for collection of multimodal data measurements from human motion. Sensors have become wearable and lightweight, with modules integrated to provide inertial measurements and soft kinematic data [6], [7], [8]. Even though the rapid advancement on sensors, robust and accurate computational methods for the analysis and recognition of human motion are still under development [9], [10].

In this work, a strategy composed of a Convolutional Neural Network (CNN) and a first-order Markov Chain (MC), is presented for both, recognition of walking activity and prediction of gait periods. The recognition of level-ground walking, ramp ascent and descent activities is im-

This work was supported by the EPSRC for the project ‘Wearable soft robotics for independent living’ (EP/M026388/1).

Uriel Martinez-Hernandez is with the Department of Electronic and Electrical Engineering, the University of Bath, Bath, BA2 7AY, U.K. (email: u.martinez@bath.ac.uk)

Adrian Rubio-Solis is with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, S1 3JD, U.K. (email: a.rubiosolis@sheffield.ac.uk)

Abbas A. Dehghani-Sanij is with the School of Mechanical Engineering, the University of Leeds, Leeds, LS2 9JT, U.K. (email: a.a.dehghani-sanij@leeds.ac.uk)

Fig. 1. High-level method, composed of a CNN and first-order MC, for recognition and prediction of walking activities and gait periods. Real data collected from wearable sensors attached to the lower limbs of participants.

The experiments show that the CNN is capable to achieve mean accuracies of 100% for recognition of walking activities, 98.32% for gait periods, 97.42% for stance phase and 99.83% for swing phase. The first-order MC is able to predict gait periods with mean accuracies of 99.78%, 97.56% and 97.35% for level-ground walking, ramp ascent and descent. This information is important to know the probability of the next gait period and phase during the gait cycle. The recognition and prediction functionalities, provided by the proposed combination of high-level methods, are crucial for multi-layer architectures required for learning, interaction and control of autonomous assistive robots [15], [16], [17].

Overall, the results from the experiments show that the strategy composed of a CNN and first-order MC is highly accurate, which makes it suitable for the development of intelligent wearable robots capable to assist humans in ADLs.
Sensor signals

recognition of walking activity and gait periods, using data

II. METHODS

A. Experimental protocol and data collection

Angular velocity, accelerometer and magnetometer signals were employed from three inertial measurement units (IMU), worn by 12 healthy human participants (Figure 1). Anthropometric data from participants are as follows: ages between 24 and 34 years old, heights between 1.70 m and 1.82 m, and weights between 75.5 kg and 88 kg.

Participants were asked to walk at their self-selected speed and perform ten repetitions of three walking activities: level-ground walking, ramp ascent and ramp descent (Figure 2A). Level-ground walking was performed on a flat cement surface. A metallic ramp, with a slope of 8.5 deg, was used for ramp ascent and descent. Sensor signals were systematically collected and filtered with a cut-off frequency of 10 Hz. For this process employed three IMUs (Shimmer Inc.) attached to the thigh, shank and foot of participants. For each IMU, angular velocity, accelerometer and magnetometer signals, in x-y and z axes, were sampled at 100 Hz. These signals were concatenated to form datasets, composed of 27 signals (3 signals × 3 axes × 3 sensors) and 200 sensor samples, from each activity performed by participants. Datasets from 8 and 4 participants were used to train and test the proposed strategy, respectively. Figure 2B shows an example of the signals collected from the wearable sensors during a walking activity. In addition, two foot pressure-insole sensors were used to detect the beginning and end of each gait cycle.

Figure 2C presents the segmentation of the gait cycle into stance phase, swing phase and eight periods (initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing, terminal swing). This segmentation allows the proposed strategy to recognise and predict the state of the human body during a walking activity.

B. CNN for recognition of walking activity and gait period

Convolutional Neural Networks (CNN) have shown their potential for speech recognition and image classification [18], [19], [20]. Here, a CNN is developed for recognition of walking activity and gait periods, using data from wearable sensors. The proposed CNN model is presented in Figure 3A. The first layer uses 32 kernels of sizes $5 \times 5$ and $2 \times 2$ for convolution and max-pooling. The second layer uses 16 kernels of sizes $3 \times 3$ and $2 \times 2$ for convolution and max-pooling. Features from the second layer, which are flattened and fully connected, are used by the softmax layer to estimate the probability of the current walking activity and gait period. The CNN model receives input data from all walking activities, arranged in matrices of 27 signals × 25 samples based on the segmentation into 8 periods of the complete activity matrix ($27 \times 200$, see Figure 2). This approach allows to recognise the walking activity and gait period performed by participants, e.g., level-ground walking and period 5 (pre-swing). The output map from each convolutional layer is obtained as follows:

$$x_{ij}^l = b_j + \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} k_{ab} \ast y_{(i+a)(j+b)}^{l-1}$$  \hspace{1cm} (1)

where $x_{ij}^l$ is the output of the $l$-th layer of the $j$-th feature map on the $i$-th unit, and $b_j$ is the bias. The operator $\ast$ denotes the convolution between the $m \times m$ kernel $k_{ab}$ and the nonlinear output $y_{(i+a)(j+b)}^{l-1}$ from layer $l-1$. The nonlinear function $\sigma$ is applied to the output from Equation (1) as follows:

$$y_{ij}^l = \sigma(x_{ij}^l)$$  \hspace{1cm} (2)

where $y_{ij}^l$ is the nonlinear output from the $l$ convolutional layer and $\sigma$ is the hyperbolic tangent function tanh. A downsampling process is performed with a max-pooling layer after each convolutional layer. This process takes a $u \times u$ region ($2 \times 2$ size in our CNN model) and provides the maximum value from that region as follows:

$$y_{ij}^l = \max_{u \times u}(y_{ij}^{l-1})$$  \hspace{1cm} (3)

where $y_{ij}^l$ contains the maximum values from the nonlinear output $y_{ij}^{l-1}$. The process performed by convolutional and max-pooling layers is known as feature learning. The learned features, connected to a 1-dimensional feature vector $y_c$, are used by a softmax layer for classification, as follows:
Fig. 3. CNN and first-order MC strategy for recognition and prediction of walking activities and gait periods using wearable sensors. (A) The CNN model is composed of two convolutional and max-pooling layers, followed by flatten, fully connected and softmax layers. Input sensor data from all walking activities (level-ground walking (LGW), ramp ascent (RA) and ramp descent (RD)) are segmented into 8 periods (initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing and terminal swing). The CNN estimates the current walking activity and gait period performed by a participant. (B) The first-order MC predicts the next gait period based on the observation of current recognition decisions from the CNN.

The prediction of gait periods is based on the sequential analysis of decisions made by the CNN. Markov models are useful to model sequential data and assume that future predictions are independent of all but the most recent observations. Specifically, here, a first-order Markov Chain (MC) of observations is employed, where the probability distribution \( P(z_t|z_{t-1}) \) of a particular observation \( z_t \) at time \( t \) is conditioned on the observation \( z_{t-1} \) at time \( t-1 \). The directed graph in Figure 3B shows the first-order MC, which receives the gait period, \( z_t \), recognised by the CNN. The joint distribution for a sequence of \( T \) observations is given by:

\[
P(z_1, \ldots, z_T) = P(z_1) \prod_{t=2}^{T} P(z_t|z_{t-1})
\]

In Equation (6), the conditional probability distribution for observation \( z_t \), given the d-separation property and all observations up to time \( t \), is obtained as follows:

\[
P(z_t|z_1, \ldots, z_{t-1}) = P(z_t|z_{t-1})
\]

The temporal models in Equation (7) predicts the next observation in a sequence using only the immediately preceding observation [21]. However, there is no standard procedure for an efficient application [22], [23]. Here, online implementation of MC for prediction of the next observation \( z_{t+1} \), is based on the algorithm presented in [24], as follows:

\[
P(z_{t+1}) = A_t P(z_t)
\]

where \( z_t \) is a random variable with \( N \) states, \( P(z_{t+1}) \) is the predicted state, \( P(z_t) = [P(z_t) = 1, \ldots, P(z_t) = N] \) is a stochastic state vector and \( A_t = [a_{ij}] \), with \( i, j = 1 \ldots N \), is a time-dependent transition matrix. The state vector is recursively updated by estimates of \( A_t \), as follows:

\[
a_{ij} = \rho \Gamma_{ij}, \quad i, j = 1, \ldots, N
\]

The output from the CNN allows to know the state of the human body while performing walking activities. This information is needed for control of assistive and rehabilitation robots, but also is important for prediction of gait periods during the current walking cycle. An approach for prediction of gait periods is presented in the next Section II-C.

C. First-order MC for prediction of gait periods

The prediction of gait periods is based on the sequential analysis of decisions made by the CNN. Markov models are useful to model sequential data and assume that future predictions are independent of all but the most recent observations. Specifically, here, a first-order Markov Chain (MC) of observations is employed, where the probability distribution \( P(z_t|z_{t-1}) \) of a particular observation \( z_t \) at time \( t \) is conditioned on the observation \( z_{t-1} \) at time \( t-1 \). The

\[
P(c|y) = \frac{e^{y^T w_c}}{\sum_{n=1}^{N} e^{y^T w_n}}
\]

\[
\hat{c} = \arg \max_c P(c|y)
\]

where \( P(c|y) \) contains the probabilities for all classes (walking activities and gait periods), given the sample vector \( y \). The parameters \( w \) and \( N \) represent the weight vector and total number of classes, respectively. In Equation (5), the recognition of the current walking activity and gait period, \( \hat{c} \), is obtained with the maximum a posterior (MAP) estimate.

![Diagram of CNN and first-order MC strategy](image)

The probability distribution for a sequence of \( T \) observations is given by:

\[
P(z_1, \ldots, z_T) = P(z_1) \prod_{t=2}^{T} P(z_t|z_{t-1})
\]

In Equation (6), the conditional probability distribution for observation \( z_t \), given the d-separation property and all observations up to time \( t \), is obtained as follows:

\[
P(z_t|z_1, \ldots, z_{t-1}) = P(z_t|z_{t-1})
\]

The temporal models in Equation (7) predicts the next observation in a sequence using only the immediately preceding observation [21]. However, there is no standard procedure for an efficient application [22], [23]. Here, online implementation of MC for prediction of the next observation \( z_{t+1} \), is based on the algorithm presented in [24], as follows:

\[
P(z_{t+1}) = A_t P(z_t)
\]

where \( z_t \) is a random variable with \( N \) states, \( P(z_{t+1}) \) is the predicted state, \( P(z_t) = [P(z_t) = 1, \ldots, P(z_t) = N] \) is a stochastic state vector and \( A_t = [a_{ij}] \), with \( i, j = 1 \ldots N \), is a time-dependent transition matrix. The state vector is recursively updated by estimates of \( A_t \), as follows:

\[
a_{ij} = \rho \Gamma_{ij}, \quad i, j = 1, \ldots, N
\]
The model accuracy and error, randomly drawn from the testing datasets, were used to evaluate the CNN with new data. This process achieved an accuracy of 100% for recognition of individual walking activities (Figure 5A). The mean accuracy of 98.32% for recognition of gait periods for all walking activities is shown in Figure 5B. From these results, it is observed that recognition of stance (periods 1 to 5) and swing (periods 6 to 8) phases are 97.42% and 97.83%, respectively (Figure 5C). The recognition of gait periods and phases is important to know the state of the human body during the gait cycle, e.g., heel contact and toe-off. The accuracy recognition of gait periods for individual walking activities are shown in Figures 5D, 5E, and 5F. These results show that the CNN was able to recognise gait periods for level-ground walking, ramp ascent and descent with accuracies of 99.92%, 97.62% and 97.43%, respectively.

B. Prediction of gait periods

Prediction of gait periods allows to know the probability for the next gait period during the walking cycle, which is important for a better control of assistive robots. The results from prediction of gait periods, using the first-order MC temporal model, are shown in the confusion matrices of Figure 6. In this process, the MC model used the recognition output from the CNN during the walking activity. Then, the transition over time of gait periods was observed and learned by the MC model to estimate the next probable gait period. In Figure 6, the gait periods recognised by the CNN are shown in orange colour, while the estimated predictions from the MC are shown in grey scale colours. The confusion matrix in Figure 6A shows the prediction of gait periods for level-ground walking activity, where a mean accuracy of 99.78% was achieved. Similarly, Figures 6B and 6C show the mean gait period prediction accuracies of 97.56% and 97.35% for ramp ascent and ramp descent activities, respectively. Then, the first-order MC model was able to predict gait periods for all walking activities with a mean accuracy of 98.23%.

The results show that the performance of the prediction process, using the first-order MC, depends on the recognition accuracy of the CNN model. For that reason, prediction of gait periods for level-ground walking is more accurate than the results achieved for ramp ascent and descent. This prediction process is important to allow assistive robots to understand not only the current state of the human body, but also to perceive what is the next expected event during the walking activity. For instance, prediction of the swing phase is possible based on the recognition of gait periods during the stance phase. This predictive capability allows robots to be prepared for an expected event, and thus, delivering the needed assistance at the appropriate time in ADLs.
A comparison of the performance between the proposed strategy and state-of-the-art methods is presented in Table I. All methods are able to recognise walking activities with high accuracies ranging from 98% (DBN [30]) to 100% (GMM [28] and our CNN+first-order MC). Only a few methods are able to recognise gait periods and events with accuracies of 95.25% (DBN [30]), 97% (SVM [29]) and 98.32% (CNN + first-order MC), where the highest accuracy is achieved by our work. Our proposed strategy allows the prediction of gait periods (98.23%), which contrasts with the capabilities offered by all methods in Table I. This comparative analysis shows the benefits of combining convolutional neural networks and Markov Chain models, together with wearable sensors, for robust recognition and prediction of walking activities and gait periods.

Overall, the results from all experiments showed the capability of the CNN and first-order MC strategy to perform high-level recognition and prediction processes. Thus, this work offers a suitable computational approach that, connected to mid- and low-level processes, has the potential to allow wearable robots to not only identify human motion, but also to provide reliable assistance in ADLs.
TABLE I
COMPARISON OF STATE-OF-THE-ART METHODS FOR RECOGNITION OF WALKING ACTIVITIES AND PREDICTION OF GAIT PERIODS

<table>
<thead>
<tr>
<th>Method</th>
<th>Activity</th>
<th># Sensors</th>
<th>Recognition accuracy (%)</th>
<th>Recognition gait period accuracy (%)</th>
<th>Prediction gait period accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-sim distance [25]</td>
<td>Level walking, ramps, sitting</td>
<td>9</td>
<td>99.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN [26]</td>
<td>Level walking</td>
<td>32</td>
<td>98.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I4AN T [27]</td>
<td>Level walking, ramps, stair</td>
<td>13</td>
<td>99.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM [28]</td>
<td>Level walking, standing, sitting</td>
<td>4</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM [29]</td>
<td>Level walking, ramps</td>
<td>9</td>
<td>99</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>DBN [30]</td>
<td>Level walking, ramps, stair</td>
<td>13</td>
<td>98</td>
<td>97.27</td>
<td></td>
</tr>
<tr>
<td>CNN + first-order MC</td>
<td>Level walking, ramps</td>
<td>3</td>
<td>100</td>
<td>98.32</td>
<td>98.23</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

In this work we presented a strategy, composed of a CNN and first-order MC, for recognition of walking activities and prediction of gait periods using data from wearable sensors. The CNN was implemented for recognition of walking activities (level-ground walking, ramp ascent and descent) and gait periods (initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing and terminal swing), achieving mean accuracies of 100%, and 98.32%, respectively. The first-order MC was developed to predict the next probable or expected gait period during the gait cycle. This temporal model achieved mean accuracies of 99.78%, 97.56% and 97.35% for prediction during level-ground walking, ramp ascent and descent, respectively. All the experiments, for validation of the proposed method, employed real data collected from three wearable sensors attached to the thigh, shank and foot of participants while performing walking activities. The results showed that our method is able to identify the state of the human body and estimate the next expected event during the walking activity. Overall, the recognition and predictive functionalities, offered by the proposed strategy, are essential for control of intelligent wearable devices capable to provide safe and reliable assistance to humans in ADLs.

REFERENCES