

Bidirectional Residual LSTM-based Human Activity Recognition

Zohair Malki², Elsayed Atlam^{1,2}, Guesh Dagne³, Ahmad Reda Alzighaibi², Elmarhomy Ghada², Ibrahim Gad¹

¹ Faculty of Science, Tanta University, Egypt

² College of Computer Science and Engineering, Yanbu, Taibah University, Saudi Arabia

³ Department of postgraduates Studies and Research in Computer Science, Mangalore University, India

Correspondence: Elsayed Atlam, Faculty of Science, Tanta University, Egypt.

Received: April 15, 2020 Accepted: May 27, 2020 Online Published: June 8, 2020

doi:10.5539/cis.v13n3p40 URL: <https://doi.org/10.5539/cis.v13n3p40>

Abstract

The Residual Long Short Term Memory (LSTM) deep learning approach is attracting attention of many researchers due to its efficiency when trained on high dimensional datasets. Nowadays, Human Activity Recognition (HAR) has come with enormous challenges that have to be addressed. In addressing such a problem, one can think of developing an application that can help the elderly people as an assistant when it works in collaboration with other timely technologies such as wearable devices with the help of IoT. Many research works are using a standard dataset in evaluating their proposed method in this regard. The dataset comes with its own challenge such as imbalanced classes. In this work, we propose to apply different machine learning techniques to address the specified problems and the method is validated on a standard dataset. To validate the proposed method, we evaluated using different standard metrics such as classification accuracy, precision, recall, f1-score, and Receiver Operating Characteristic (ROC) curve. The proposed method achieves an Area Under Curve (AUC) of 100%, 97.66% of accuracy, 91.59% precision, 93.75% of recall and 92.66% of F1-score respectively.

Keywords: human activity recognition, IoT, deep neural networks, LSTM, imbalanced data

1. Introduction

Fall detection comes in the domain of human activity recognition in simple or complex environments and the source of data is mostly from sensors. The literature so till dates are dominated with human crafted features that enable to handle complex actions based on high dimensional sensor datasets. Nowadays, the direction of such research is shifted towards deep learning executing on high computational power machines with potential applications in health care systems such as assisting elderly people. As the neural networks become deeper and deeper, it becomes challenging to train them. To address this challenge, the residual neural networks come in place as they train by jumping certain layers during training instead of sequentially (He et al., 2016).

One's health condition highly affects the action to be performed. Studying and evaluating human activities in a systematic way enables to study the behaviour of individuals and such study can have several applications such as in health care systems, in surveillance and security (Elbayoudi et al., 2019). Human activities are hierarchical in nature and are transitionally formed with temporal sequences of sub-activities to formulate the main (Kautz et al., 2017).

Human activity recognition has many practical applications such as automatically categorizing human actions task, to train and monitor an employee whether the required task is correctly performed, checking for proper steps and procedures when performing a task, to verifying if the employee follows the organizational cultures (Sun et al., 2018). Information obtained from human activity via sensors provides a vital role in studying the behaviour of human beings from their activities and is served as an input for taking measures in the case of hazardous activities and accidents such as fall of elderly people (Subasi et al., 2018).

Human activity recognition is a vital and open research area in the domains of behavioural analysis of humans and human-computer interaction. Due to the emerging of new machine learning algorithms, the research outcome in human activity recognition enables in recognizing various activities such as falling, jumping, jogging, etc. Recognizing such activities is vital in many ways such as maintaining a healthy lifestyle, patient rehabilitation and provide support for the elderly citizens. The Long short-term memory (LSTM) which is a type of recurrent neural network is used for modelling of sequential time series data (Goel et al., 2019). It has several layers that learn the dynamics in the sequential data. Widely applicable in modelling temporal patterns in activity recognition such as progressive detection of activity levels, fall detection and heart attack in elderly people. In deep learning based human action recognition, the most commonly used commonly used function is softmax, hence we also use the same during training as it handles the probabilities of the

prediction of actions in multi-class classification (Nweke et al., 2018; Agarwal et al., 2019).

In this study, we have proposed to use the residual-based LSTM deep learning to recognize the activity of humans from imbalanced class standard dataset namely mHealth which is freely accessible from UCI repository. The dataset contains 12 classes and 23 features. To measure the performance of our work, we have used standard performance metrics namely classification accuracy, precision, recall, and f1-score.

The rest of the paper is organized as follows. Section 2 covers state-of-the-art literature in the domain of human activity recognition specifically in fall detection. Section 3 specifies the proposed residual Long-Short Term Memory (LSTM) deep learning. Sections 4 experimental results of the proposed method and the concluding remarks are presented in Section 5.

2. Related Works

In this section, we presented a survey of recent developments in human activity recognition. Nweke et al. (Nweke et al., 2018) conducted a detailed survey of recent works and highlighted the state of the art and research challenges in the domain of mobile and wearable sensor-based human activity recognition pipeline. Their main focus is to give an idea on deep learning algorithms for mobile and wearable sensor-based human activity recognition. In human activity recognition, deep learning is used in various tasks such as labelling of human activity sequence, estimating pattern of movement of humans, recognizing the feelings with human beings, and in health care systems such as diagnosis of patients using physiological signals. Jobanputra et al. (Jobanputra et al., 2019) have conducted an intensive review on recent advancement in the domain of human activity recognition. Ji et al (Ji et al., 2012) has developed an end-to-end automated human activity recognition in an uncontrolled environment with the help of Convolutional neural networks (CNNs) deep learning model that enables to perform automatic features generation from a 3D datasets captured from motion. Zhang et al. (Zhang et al., 2020) have attempted to motion information in HAR systems to address the influence of bad samples using Motion-patch-based Siamese Convolutional Neural Network (MSCNN). To validate the method, they carried out several experiments on UCF-101 and HMDB-51 datasets. Moreover, Xie et al. (Xie et al., 2019) design an end-to-end residual stochastic model to describe spatio-temporal disparities. The experimentation is carried out on the datasets NTU RGB+D, SYSU-3D, and UT-Kinect. Subasi et al (Subasi et al., 2018) proposed an ensemble learning based human activity recognition using Adaboost classifier on data taken from body parts via sensors. The experimentation carried out in this study depicts that the Adaboost ensemble classifier enables in recognizing human activities in a higher and acceptable rate.

Wen and Wang (Wen and Wang, 2016) proposed real-time activity recognition system from dynamically available data sources. They have used ensemble classifiers that enable to automatically pick discriminatory features to enhance the recognition rate. Chen et al. (Chen et al., 2016) introduced the LSTM-based feature extraction method to recognize human actions by employing the tri-axial accelerometers datasets. Their experimentation exhibited that the LSTM-based method is applicable to the task of activity recognition.

3. Methodology

In our work, we proposed residual bidirectional based LSTM deep learning approach to recognize the action of humans based on sensor datasets namely mHealth that is freely accessible from the famous UCI dataset repository.

The residual LSTM learns the residual nodes with reference to the hidden state. The residual LSTM delivers efficient training and validation when compared with an ordinary LSTM model by providing a shortcut path during training (Yue et al., 2018; Kim et al., 2017). Table 1 presents the description of the layers of the proposed model. Figure 1 shows the architecture of the proposed method. It indicates the residual bidirectional LSTM deep learning with fully connected layers.

3.1 Performance Measures

To assess the classification performance the proposed method, we have applied several standard performance measures such as classification accuracy, precision, recall, f1-score, ROC curve and confusion matrix.

The ROC curve is used to test the classification performance of classification models. It is particularly useful in assessing predictive models since it records the trade-off between specificity and sensitivity. According to Hajian et al. (Hajian-Tilaki, 2013) the closer the ROC curve to the upper left corner, the better overall performance of the model. The AUC is used to visualize the performance of a classifier in terms of numerical value. The ROC curve plots provides two parameters namely True positive rate (TPR) and False positive rate (FPR). True positive rate (TPR): measures of the whole positive part of a dataset whereas the False positive rate (FPR) is used to measure of the whole negative part of a dataset. The classification accuracy measures the effectiveness of a predictive model by considering the correctly classified samples to the ratio of the total number of test samples. The classification accuracy is not the best measure in the case of class imbalanced datasets, hence we are using additional performance metrics to alleviate the limitation.

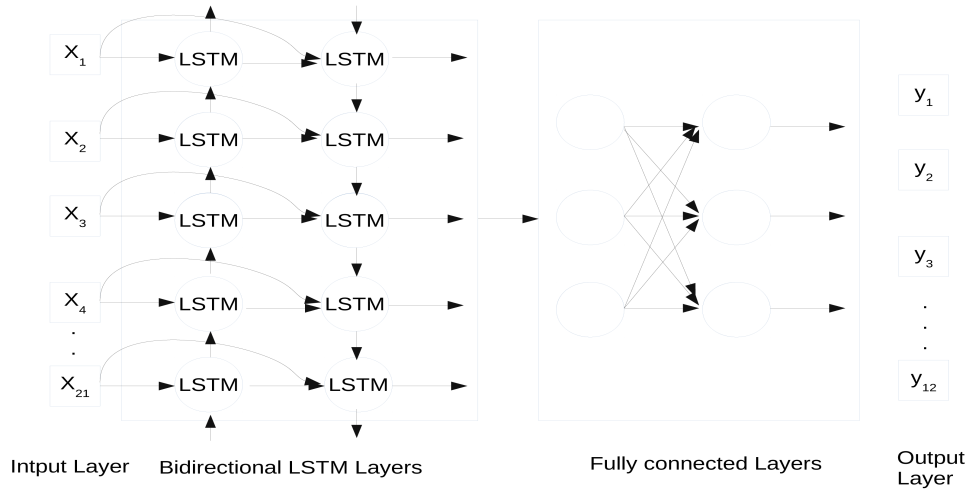


Figure 1. The architecture of residual bidirectional LSTM

Therefore, four more performance metrics namely specificity, sensitivity (recall), precision, F1-score and ROC curve are used as additional performance measures. Equations 1, 2, 3, and 4 presents the mathematical expressions of classification accuracy (CA), f1-score, recall and precision respectively which are calculated based on True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) samples on the test data. (Gad et al., 2018).

$$CA = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$F1 - score = 2 \frac{precision \star recall}{precision + recall} \tag{2}$$

$$Recall = Specificity = \frac{TP}{TP + FP} \tag{3}$$

$$Precision = \frac{TP}{TP + FN} \tag{4}$$

3.2 Dataset

The mHealth dataset consists of two main components namely Human activities and Vital signs and the total number of participants are 10 persons and they did 12 distinct phases of physical activities. We have used 21 from 24 features of the datasets avoiding the person ID, weight, height and age as they have less impact on the recognition of the action. The dataset is downloaded from the UCI repository which is free access in this repository. The sensors are used to record the activity of each candidate such that the positions of these sensors are Chest, Left Ankle, and Right Wrist. Table 2 shows the different human physical activities and the size for the training set and testing set for each class label. The class distribution for each class label is presented in Figure 2

4. Experimental Results

In our work, we have used Google Colab to execute the proposed model using python 3.6 and Keras for deep learning and Scikit-learn library. We have used the mHealth dataset which has 12 classes and 24 features. The dataset is normalized so as to make easily fit to the model during training using the min-max approach $Z = \frac{X-\mu}{\sigma}$. We have validated the proposed method using different performance metrics.

In general, to achieve unbiased prediction of performance, the collected data of participants are randomly divided into two sets namely training and testing sets in 70:30 ratios respectively. In the training stage, enhancing the performance of the model is challenging due to the imbalance data. To address the problem of imbalance of class labels, we use a

Table 1. Description of the layers of the proposed model

Layer (type)	Output Shape	#Param	Connected to
input_1 (InputLayer)	(None, 21, 1)	0	
lstm_1 (LSTM)	(None, 21, 10)	480	input_1[0][0]
lstm_2 (LSTM)	(None, 21, 10)	840	lstm_1[0][0]
add_1 (Add)	(None, 21, 10)	0	lstm_1[0][0] lstm_2[0][0]
lstm_3 (LSTM)	(None, 21, 10)	840	add_1[0][0]
add_2 (Add)	(None, 21, 10)	0	add_1[0][0] lstm_3[0][0]
lstm_4 (LSTM)	(None, 21, 10)	840	add_2[0][0]
add_3 (Add)	(None, 21, 10)	0	add_2[0][0] lstm_4[0][0]
lstm_5 (LSTM)	(None, 21, 10)	840	add_3[0][0]
add_4 (Add)	(None, 21, 10)	0	add_3[0][0] lstm_5[0][0]
lstm_6 (LSTM)	(None, 21, 10)	840	add_4[0][0]
add_5 (Add)	(None, 21, 10)	0	add_4[0][0] lstm_6[0][0]
lstm_7 (LSTM)	(None, 21, 10)	840	add_5[0][0]
add_6 (Add)	(None, 21, 10)	0	add_5[0][0] lstm_7[0][0]
lambda_4 (Lambda)	(None, 10)	0	add_6[0][0]
lstm_8 (LSTM)	(None, 10)	840	add_6[0][0]
add_7 (Add)	(None, 10)	0	lambda_4[0][0] lstm_8[0][0]
dense_1 (Dense)	(None, 1024)	11264	add_7[0][0]
dense_2 (Dense)	(None, 256)	262400	dense_1[0][0]
dense_3 (Dense)	(None, 13)	3341	dense_2[0][0]
Total params		283,365	
Trainable params		283,365	
Non-trainable params		0	

weighted-class approach that performs balancing for the training set by automatically increasing the number of samples in the minority classes that ultimately enhances the performance of classification models.

Table 3 presents classification report of the proposed method. For the minority class label Jump front & back, the best performance achieved is 89%, 82%, and 85% respectively for precision, recall and f1-score respectively. For the same class label, the accuracy and AUC is 97% and 100% respectively. When we consider one of the majority class label namely jogging, its performance results in terms of precision, recall and f1-score are given as 90%, 0.93%, and 0.92% respectively. In the case of Cycling class label the performance achieved by the proposed method is given as 100%, 100%, and 100% respectively.

As can be seen from Table 4, the proposed method achieves an AUC of 100%, 97.66% of accuracy, 91.59% precision, 93.75% of recall and 92.66% of F1-score respectively. Therefore, it is clearly shown that for all classification metrics, our proposed method has achieved higher performance level.

Figure 3 shows the ROC-AUC curves corresponding with the performance of the proposed method for each label. The ROC for each class label is displayed in dashed lines. From Figure 3, it is clear that the ROC curve is close to the top left corner to show maximum performance by the proposed method. Figure 4 shows the confusion matrix generated due to the the proposed method and the Figure 5 presents the precision-recall curve indicating maximum possible performance due to the proposed method.

Moreover, Table 5 presents a comparison of the proposed model with state-of-the art method and our work outperforms when it is evaluated in terms of the specified performance measures namely accuracy and f1-score.

Table 2. Size of each class label in Train set and Test set

Class Label	All	Train	Test	Weight	New Weight
Climbing stairs	30720	21577	9143	1.0	1
Cycling	30720	21576	9144	1.0	1
Frontal elevation of arms	29799	20873	8926	1.031	1
Jogging	30720	21603	9117	1.0	1
Jump front & back	10444	7430	3014	2.94	30720
Knees bending	29848	19807	9041	1.029	1
Lying down	27646	19202	8444	1.111	1
Running	30720	21550	9170	1.0	1
Sitting and relaxing	27646	19283	8363	1.111	1
Standing still	25256	17664	7592	1.216	1
Waist bends forward	28519	19917	8602	1.077	1
Walking	28850	20139	8711	1.065	1

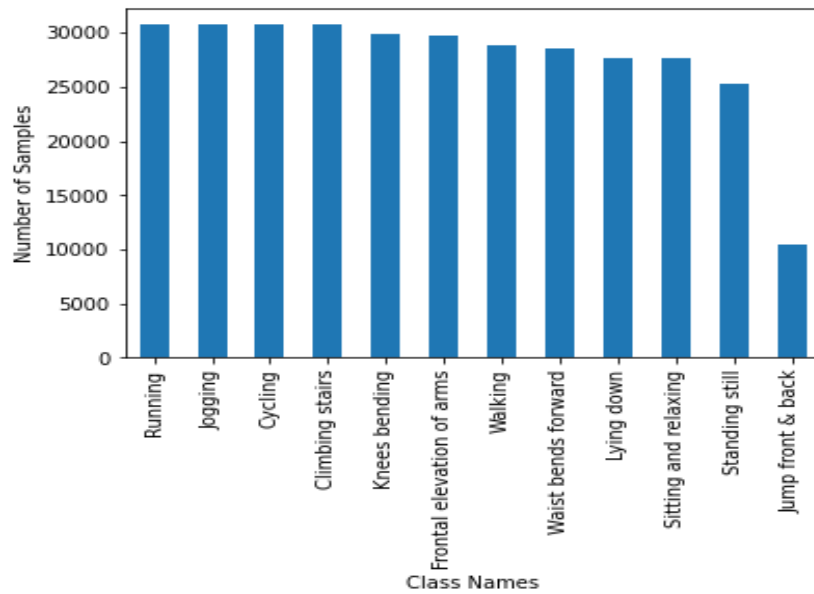


Figure 2. The number of samples for each class label

Table 3. The classification report for each class label

	Precision	Recall	F1-score	Support
Climbing stairs	0.92	0.94	0.93	9143
Cycling	1.00	1.00	1.00	9144
Frontal elevation of arms	0.99	0.99	0.99	8926
Jogging	0.90	0.93	0.92	9117
Jump front & back	0.89	0.82	0.85	3014
Knees bending	0.96	0.96	0.96	9041
Lying down	1.00	1.00	1.00	8444
Running	0.97	0.91	0.94	9170
Sitting and relaxing	1.00	1.00	1.00	8363
Standing still	0.99	1.00	1.00	7592
Waist bends forward	0.96	0.99	0.97	8602
Walking	0.96	0.97	0.97	8711
macro avg	0.96	0.96	0.96	99267
weighted avg	0.97	0.97	0.97	99267

Table 4. The performance of the proposed model

	MSE	Accuracy	Precision	Recall	F1 Score
Residual Bidir-LSTM	0.0088	0.97	0.9159	0.9375	0.9266

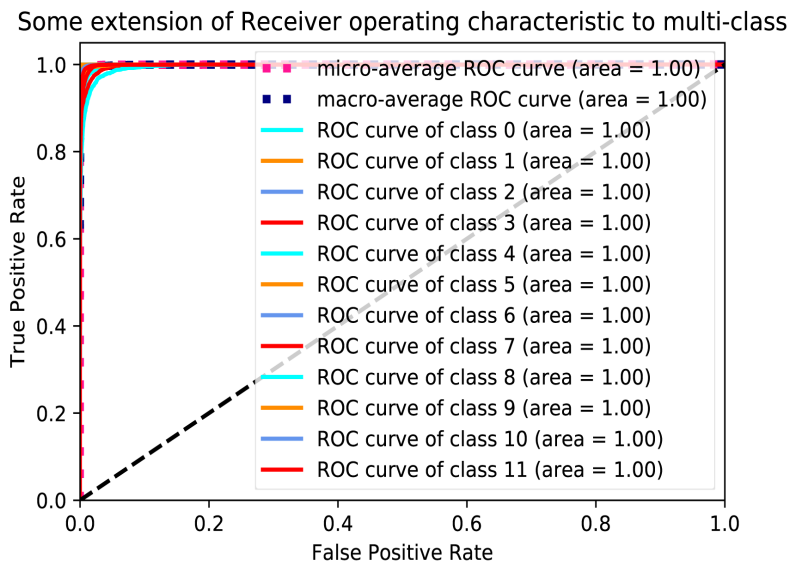


Figure 3. The ROC-AUC curves for each class

Table 5. Comparative study of our proposed method with state-of-the-art models

Model	Accuracy	F1-score
The proposed model	97.02%	92.66%
LSTM (Sansano et al., 2020)	96.41%	96.04%
Residual-LSTM (Zhao et al., 2018)	93.60%	93.50%

5. Conclusion

In this work, we proposed residual bidirectional LSTM deep learning method to address the challenge with imbalanced class problems on mHealth datasets for human activity recognition. The existing studies were focused more on multi-class classification of different human activities. In this paper, we used the residual bi-directional LSTM with weighted-class technique to improve the performance of the minority classes. The experimental results show that each label has better classification accuracy. The experimental results demonstrate that a combination of residual bi-directional LSTM and the weighted-class is an efficient technique for the prediction of human activity based on high-dimensional data. The proposed work is evaluated in terms of standard metrics and experimental results are reported as AUC of 100%, 97.66% of accuracy, 91.59% precision, 93.75% of recall and 92.66% of F1-score respectively. Thus, the performance of the proposed model shows better performance when it is compared with the state of the art work. In our future work, we are planning to apply the proposed model in the area of natural language processing (NLP) such as Arabic language.

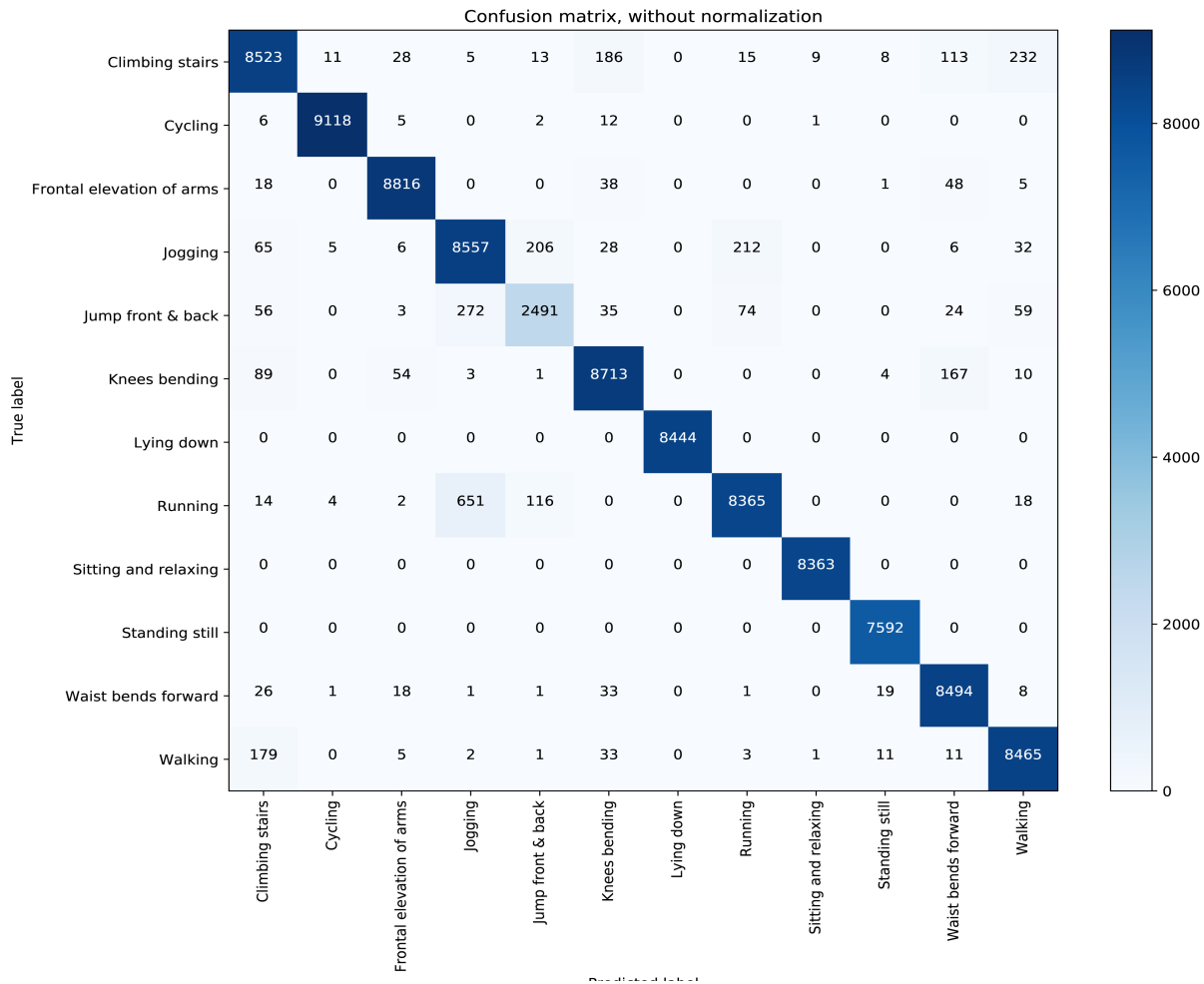


Figure 4. The Confusion matrix of the proposed model

Average precision score, micro-averaged over all classes: AP=1.00

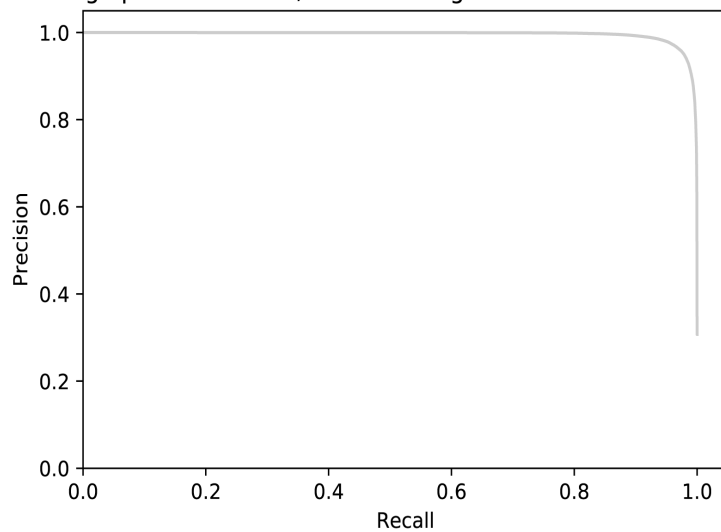


Figure 5. The Precision Recall plot of the proposed model

References

- Agarwal, M., Bohat, V. K., Ansari, M. D., Sinha, A., Gupta, S. K., & Garg, D. (2019). A convolution neural network based approach to detect the disease in corn crop. In *2019 IEEE 9th International Conference on Advanced Computing (IACC)*. IEEE.
- Chen, Y., Zhong, K., Zhang, J., Sun, Q., & Zhao, X. (2016). Lstm networks for mobile human activity recognition. In *2016 International Conference on Artificial Intelligence: Technologies and Applications*. Atlantis Press.
- Elbayoudi, A., Lotfi, A., & Langensiepen, C. (2019). The human behaviour indicator: A measure of behavioural evolution. *Expert Systems with Applications, 118*, 493-505.
- Gad, I., & Manjunatha, B., et al. (2018). Multi-label classification of big ncdc weather data using deep learning model. In *International Conference on Soft Computing Systems*, pp. 232-241. Springer.
- Goel, A. K., Chakraborty, R., Agarwal, M., Ansari, M. D., Gupta, S. K., & Garg, D. (2019). Profit or loss: A long short term memory based model for the prediction of share price of DLF group in india. In *2019 IEEE 9th International Conference on Advanced Computing (IACC)*. IEEE.
- Hajian-Tilaki, K. (2013). Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine, 4*(2), 627.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778.
- Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3d convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence, 35*(1), 221-231.
- Jobanputra, C., Bavishi, J., & Doshi, N. (2019). Human activity recognition: A survey. *Procedia Computer Science, 155*, 698-703.
- Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B. M. (2017). Activity recognition in beach volleyball using a deep convolutional neural network. *Data Mining and Knowledge Discovery, 31*(6), 1678-1705.
- Kim, J., El-Khamy, M., & Lee, J. (2017). Residual lstm: Design of a deep recurrent architecture for distant speech recognition. *arXiv preprint arXiv:1701.03360*.
- Nweke, H. F., Teh, Y. W., Al-Garadi, M. A., & Alo, U. R. (2018). Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications, 105*, 233-261.
- Sansano, E., Montoliu, R., & Fernández, Ó. B. (2020). A study of deep neural networks for human activity recognition. *Computational Intelligence*.
- Subasi, A., Dammas, D. H., Alghamdi, R. D., Makawi, R. A., Albiety, E. A., Brahimi, T., & Sarirete, A. (2018). Sensor based human activity recognition using adaboost ensemble classifier. *Procedia Computer Science, 140*, 104-111.
- Sun, J., Fu, Y., Li, S., He, J., Xu, C., & Tan, L. (2018). Sequential human activity recognition based on deep convolutional network and extreme learning machine using wearable sensors. *Journal of Sensors, 2018*.
- Wen, J., & Wang, Z. (2016). Sensor-based adaptive activity recognition with dynamically available sensors. *Neurocomputing, 218*, 307-317.
- Xie, C., Li, C., Zhang, B., Pan, L., Ye, Q., & Chen, W. (2019). Hierarchical residual stochastic networks for time series recognition. *Information Sciences, 471*, 52-63.
- Yue, B., Fu, J., & Liang, J. (2018). Residual recurrent neural networks for learning sequential representations. *Information, 9*(3), 62756.
- Zhang, Y., Po, L. M., Liu, M., Rehman, Y. A. U., Ou, W., & Zhao, Y. (2020). Data-level information enhancement: Motion-patch-based siamese convolutional neural networks for human activity recognition in videos. *Expert Systems with Applications*, pp. 113-203.

Zhao, Y., Yang, R., Chevalier, G., Xu, X., & Zhang, Z. (2018). Deep residual bidir-lstm for human activity recognition using wearable sensors. *Mathematical Problems in Engineering*, 2018.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).