Human Activity Recognition in Car Workshop

Omar Magdy

Faculty of Computer Science,
October University for Modern
Sciences and Arts (MSA),
Giza, Egypt

Ayman Atia

HCI-LAB, Faculty of Computers and Artificial Intelligence,
Helwan University.
Faculty of Computer Science,
October University for Modern
Sciences and Arts (MSA), Giza, Egypt

Abstract—Human activity recognition has become so widespread in recent times. Due to the modern advancements of technology, it has become an important solution to many problems in various fields such as medicine, industry, and sports. And this subject got the attention of a lot of researchers. Along with problems like wasted time in maintenance centers, we proposed a system that extracts worker poses from videos by using pose classification. In this paper, we have tested two algorithms to detect worker activity. This system aims to detect and classify positive and negative worker's activities in car maintenance centers such as (changing the tire, changing oil, using the phone, standing without work). We have conducted two experiments, the first experiment was for comparison between algorithms to determine the most accurate algorithm in recognizing the activities performed. The experiment was done using two different algorithms (1 dollar recognizer and Fast Dynamic time warping) on 3 participants in a controlled area. The one-dollar recognizer has achieved a 97% accuracy compared to the fastDTW with 86%. The second experiment was conducted to measure the performance of a one-dollar algorithm with different participants. The results show that a 1 dollar recognizer achieved an accuracy of 94.2% when tested on 420 different videos.

Keywords—*Machine learning; human activity recognition; pose identification; industry analysis*

I. INTRODUCTION

Human activity recognition has gained importance in recent years because of its applications in various fields like health, security and surveillance, entertainment, and intelligent environments. Human activity recognition accounted a lot of researches in different approaches, like wearable devices [1][2][3][4][5],object-tagged[6][7][8], and device-free [9][10][11], to acknowledge human activities. Wasted time has become a prominent problem in various fields of work, and this problem affects the percentage of products and services required in our daily lives. The survey provided by Salary.com [12] found that 89% of workers admitted to wasting time at work every day. The survey showed that 61% claim to waste between 30 minutes to an hour a day. While this may not seem like much, it can add up to 5 hours a week or 260 hours a year/per employee. Maintenance places and factories are the most vulnerable to this issue. The survey depend heavily on human activity, which harms production rates and disrupts many required services. The field of car maintenance has become an important field in our daily life, and car workshops have spread greatly in the recent period, but difficulties have begun to appear inside the maintenance centers. The problem is that workers waste a lot of time while working by doing some negative activities. Fig. 1 shows a worker who uses the

phone at work. Moreover, there are a lot of negative activities such as (eating, drinking, talking to others, standing without work...). Besides, it is difficult to have an employee responsible for monitoring the worker for no less than 12 hours.

Fig. 1. A Negative Wasting Time Activity (using Mobile while being in Work).

The maintenance workshop could have dozens of activities some of them are consider positive and other consider negative. However,some activities look very similar such as shown in Fig. 2(a) and Fig. 2(b). The posture of changing oil activity looks very similar in motion trajectory to engine rebuild activity.
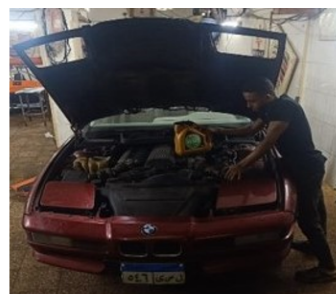
(a) Changing oil   (b) Engine rebuild

Fig. 2. A Sample for Positive Activity in Workshop.

The proposed system works by taking an input of recording videos of the worker's activities inside the maintenance center. The video is classified to determine the number of activities the worker performed such as (Changing the oil, changing the tire, or using the phone... Etc.). The classification was done by taking the path (x, y) of the important points in the skeleton

as shown in Fig.3, Fig. 4 and comparing them with the points that were taken from the collector of the data set. The system can determine the type of activities that the worker performed in the video.
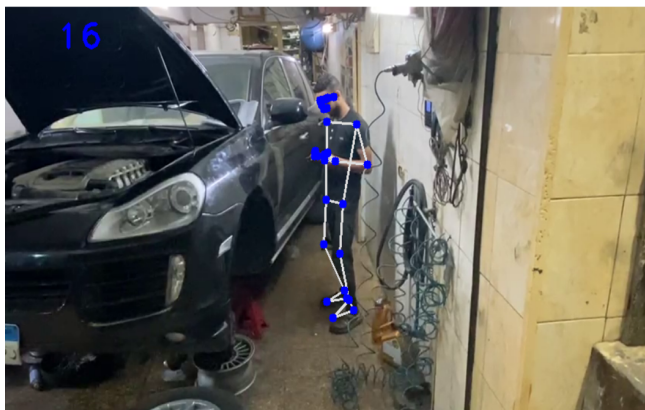

Fig. 3. Worker Changing Oil.

Fig. 4. Worker use Mobile.

The main contribution of this paper is to create a human activity recognition system that extract worker poses from the videos by using pose classification. The system used a dollar recognizer (1$) to detect worker positive and negative activities in maintenance workshop. As a result of the presence of many maintenance workshops, a problem occurred which is many workers neglect their work resulting in the occurrence of negative activities. Therefore, this paper is needed to detect the positive and negative activities and differentiate between them without the need for a person to monitor these activities and depend on the computer as an alternative. The purpose of this paper as well as to facilitate the reduction of the waste time done by each worker so they can be more productive. Moreover, positive activity is detected to reward the committed workers and to evaluate the performance of the workers in general.

II. RELATED WORK

Due to the popularity of human activity recognition systems and with the rapid advancement of computer technology, lots of research efforts were dedicated to this subject. Congcong Liu et al. [6], has proposed a activity recognition method that can identify abnormal human activity recognition in surveillance video using a combination of Bayes Classifier and CNN (Convolution Neural Network) to detect the activities and also use KTH dataset as the input of Bayes Classifier and CNN. Another system that is able to recognize human activity is proposed by Bagate et al. [1], which identifies the activity using RGB-D sensors that is developed with deep learning model CNN (Convolution Neural Network) and using knight depth camera for capturing 3-D skeleton data. For Human detection and Motion tracking, Sandar et al. [8], used frame wise displacement and recognition is based on the skeletal model with the deep learning framework to understand human behavior in the indoor and outdoor environment. And in the field of Human activity recognition by sensors, Murat et al. [13], automatically identifies human activity using joint coordinates skeletons and uses two types of deep learning to make classification and use data set of multiple people in the images. Song et al. [5], propose (1D) Convolution Neural Network (CNN) -based method for recognizing the activities using collected accelerometer data from smartphones and this method gave high accuracy of 92.71. And using wearable devices, Tahera et al. [2], used the eSense accelerometer sensor to detect the matching of activity between the head and the mouth, from this collected data some activities of the head and mouth were identified and using the machine learning and deep learning for data classification. Nitin et al. [14], propose to use the TCN (temporal Convolutional Network) to recognize the activities because it better than other deep learning methods , it has strong ability to capture long-term dependencies. Godwin et al. [9], combined gyroscope sensors with accelerometers to detect human activity and perform analysis and recognition using ANN (artificial neural networks). Tsokov et al. [15], use of the 1D synaptic neural network (CNN) with accelerometer data to make recognition of human activity more accurate.

Isah et al. [3] collected hip motion from the different waist mounted sensors, and convert each signal into spectrum image and use them as input to the CNN (Convolution Neural Network). Nacer et al. [16], use entropy point estimate for 1D heat map to separate between human maps and animal maps to give high accuracy in human activity recognition. Selçuk et al. [10] used a novel design to reduce the number of sensors used for human activity recognition and detection by using (EMD) empirical mode decomposition.And Jiewen et al. [4], identified and interacted by focusing on two wearable cameras and the interactive activities that involve only two people. Peter Washington et al. [11], addressed the topic of identifying human activity in the treatment of autism and grouped movements with a handheld camera and used the classifier CNN (Convolution Neural Network) for detecting headbanging in home videos. Hristov et al. [17]proposed a method that classifies human activity by using 3D skeleton data and normalizing it beforehand and it was represented in 2 forms. They applied this method to the UTDMHAD dataset, the system has achieved a 92.4% accuracy rate. Heilym et al. [18] was opposed to the idea of wearing devices and sensors to determine human activities and pointed out that these devices could cause inconvenience to the bearer and could give false results if used in crowded places, so he relied on the camera and determining the activity through the human skeleton features. Salahuddin Saddar et al. [19]have an objective which is to compare some machine learning algorithms

TABLE I. COMPARISON WITH SIMILAR SYSTEMS

Paper	Activities	Algorithms	Methods	Area	Accuracy
[7]	stopping the furnace operating checking the solid fuel tank checking the gear motor and auger tightening the mounting screws of the gear motor	CNN CNN+SVM Yolov3	Image classification	controlled area	95.7%
[22]	grab tools hammers nail wrench use screwdriver	CNN	Image classification	controlled area	87%
This paper	changing tire changing oil use mobile stand without work	one-dollar Fastdtw	Pose classification	uncontrolled area	95%

that were used in human activity recognition such as (SVM, Decision Trees, Random Forests, XGBoost). They tested these algorithms with measurement sensor data that was recently released from the LARA dataset. The XGBoost has achieved the best accuracy with a rate of 78.6%. Ismael et al. [20] took the topic of identifying human activity in terms of reducing aggressive actions inside prisons and on the streets to reduce aggression and used "handcrafted/learned" as a hybrid feature framework that gave it very high accuracy rates. Yusuf Erkan et al. [21]used depth sensor to classify 27 different activities and by using long-short term memory, they analyzed skeleton data. It has achieved an accuracy rate 93%. Halikowski et al. [7], presented a system for monitoring activities inside the factory using (CNN, CNN+SVM, Yolov3) algorithms. They used some activities such as (stopping the furnace operating, checking the solid fuel tank, checking the gear motor and auger, tightening the mounting screws of the gear motor) and achieved an accuracy 94%. The work presented utilize deep learning for extracting features without considering human post estimation. Zhaozheng et al. [22] presented a system for detecting activities in smart manufacturing. They used some activities (grab tools, hammer nail, wrench use, rest lever, screwdriver). They captured these activities using IMU and sEMG signals obtained from a MYO armband. They extract feature using a convolutional neural network (CNN) model. The CNN model is evaluated on this data set and achieves 98% and 87% recognition accuracy in the half-half and leave-one-out experiments. All of the previous explained the importance of identifying human activity in solving some problems in various fields. Alghyaline et al. [23] has proposed system that detects different actions in the street such as (walking, running, stopping). They measure the movement type by using three different techniques which are (Yolo, Kalman filter, Homography). The method was tested by CCTV camera and BEHAVE dataset, it has achieved an accuracy of 96.9% for the Behave dataset and achieved 88.4% for the dataset that was collected by CCTV camera.Arzani et al. [24]proposed a structural prediction strategy proposed by this system to recognize the simple and complex actions by using probabilistic graphical models (PGMs). These activities require various model parametrization to be spanned, category-switching scheme is used to deal with this parametrization. Three datasets were used to cover the two action types which are (CAD-60, UT-Kinect, and Florence 3-D). This system could recognize simple and complex activities while the previous systems focused on only one type of these two. The system proposed by Archana et al.[25] recognized human activity with Resnet and 3D CNN without using the LSTM- attention model as the 3D CNN

is achieved by modifying the 2D Resnet in order to achieve better accuracy, so that the development of detecting, and recognizing real-time human motion has been achieved.The system proposed by Zheng Dong et al. [26]resolves the issue of incomplete feature extraction by a new framework called CapsGaNet which proposed multi-feature extraction, and gated recurrent units (GRU) with attention mechanisms. The constructed dataset was a daily and aggressive activity dataset (DAAD). Moreover, the paper approved that CapsGaNet has efficiently improved the accuracy of recognition. Radhika V. et al. [27]proposed a system that used Random Forest Algorithm(RFA) to recognize human activity using Smartphones. RFA algorithm has different decision trees that is used in classification of the dataset. There were four various evaluation parameters used to measure the performance of the system such as F1 score, accuracy, precision, and sensitivity. The accuracy of the system achieved 98.34%. The system proposed by Navita et al. [28] detects the activity of aged people using the Internet of Things (IoT) monitoring model to monitor the activity of their health state. The SVM has attained 98.03%. The proposed system by Yin Tang et al. [29], a new CNN model that used hierarchical-split (HS) for a huge number of varieties in human activity recognition. Each one feature layer uses multi-scale feature representation by capturing a wide range of receptive fields of human activities. The proposed HS model can achieve high recognition performance compared to similar models complexities. The system achieved 94.10% SOTA accuracy on human activity recognition dataset.The proposed system by Maciej A. Noras et al. [30] discussed the topic of far-field electric field sensors, which accompany different physical events. The determination of activities in the proximity of the sensor is done by field signature signals. Moreover, the paper provided enhancements for electric field sensor usage and signal processing in human and animal motion recognition, perimeter monitoring, moving objects recognition, and electric power faults detection.

Table I shows a comparison between our system and different systems. The first system, Halikowski et al. [7] proposed this system to measure the performance of the worker in the factory, they recognized four different activities which are (stopping the furnace operating, checking the solid fuel tank, checking the gear motor and auger, tightening the mounting screws of the gear motor) in a controlled area using image classification method. They used more than one algorithm such as (CNN, CNN+SVM, Yolov3), their system has achieved a 95.7% accuracy rate. The second system which is proposed by Zhaozheng et al. [22] was used for qualification and evaluation of the workers. They also used

image classification to detect four activities which are (grab tools, hammer nail, wrench use, rest lever, screwdriver) in a controlled area, and they achieved an accuracy rate of 87%. Our system was proposed to detect the activities of the worker inside the car maintenance center and differentiate between the negative and the positive activities. To help the workshop owner in measuring the worker performance, the system has used the pose classification method to detect four different activities which are (changing oil, changing tire, use mobile, stand without work). The system was used in an uncontrolled area and on different body characteristics such as(height and weight).

III. PROPOSED SYSTEM

We presented a method that recognized and classified human activity in car workshops performed and captured from videos. It differentiated between the positive activities and the negative activities based on a comparison between input video and dataset stored in the templates. The proposed system used mediapipe for collecting key points of the skeletal joints and used the one-dollar and Fastdtw algorithms to classify the poses.

Fig. 5 shows the system overview, the system has two different way to input which is videos or live camera. The processing part starts with face recognition to differentiate between the workers because there is a large number of workers in the maintenance center. Then the mediapipe starts to extract the poses by calculating the path of each point in the skeleton. The mediapipe can extract 32 points, but this system focuses on extracting five important points which are (shoulder, elbow, wrist, hip, knee). The path of points was saved in a file to be ready for classification by the algorithms. The algorithms start to match these points with the points stored in the data set and send the results to the database to create a report that the managers and workshop owners can see.
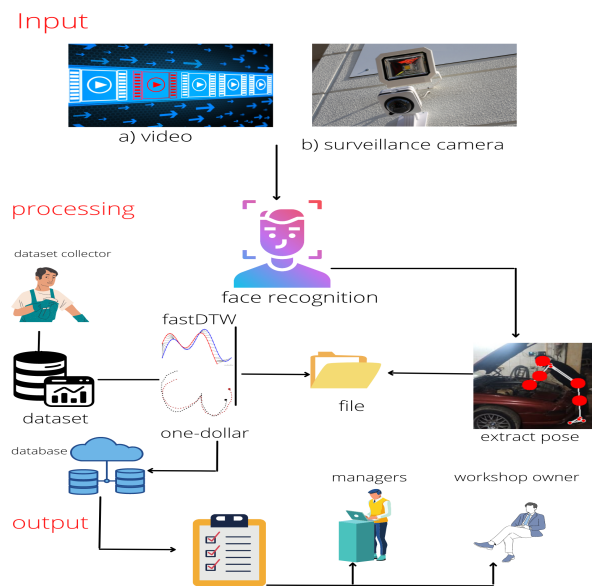

Fig. 5. System Overview.

The dataset was collected by recording videos of a specialist in the activities of the worker inside the maintenance

workshop. These videos were entered into the mediapipe and OpenCV to extract the human pose from the videos showing the path of each point during the video. Each list of points has been saved in a file contains the points and the name of the activity. With the aid of a partner in the industry, we managed to collect 560 videos including the four activities which are (changing tires, changing oil, using a mobile, and standing without work), to start testing them and in order to have a huge number of data to help in testing.

A. FastDTW Algorithm

Fast Dynamic Temporal Warping (FastDTW) is a time series alignment algorithm that was initially designed for speech recognition. Its goal is to align two sequences of feature vectors by warping the time axis iteratively until an optimal match is discovered. The two sequences can be placed on the sides of a grid, one on top of the other, with the Y time-series axis on the top and the X time-series axis on the left. In each cell, a distance metric can be set, comparing the relevant elements of the two sequences. Path is detected through the grid by determining the best match alignment between the two sequences which are "I", and "j", and results in the minimum overall distance between each cell. Finding all different routes possibilities through the grid and computing the total distance regarding each one is the process for computing this total distance (D) as mentioned in Equation 1. The total distance is calculated by dividing the sum of the distances between individual items on the path by the weighting function's sum. This is achieved by making an equal number of points in each of the two series, then calculating the Euclidean distance between the first and each subsequent point in the first series.

$$D(i,j) = |t(i) - r(j)| + min \begin{cases} D(i-1,j) \\ D(i-1,j-1) \\ D(i,j-1) \end{cases} \quad (1)$$

B. 1 Dollar Algorithm

The one-dollar is a geometric template matcher, the previously stored templates (T) are compared to the candidate strokes (C) resulting in the match that is the closest in 2-D Euclidean space as mentioned in Equation 2. Thus, we have exactly N points that will allow us to calculate the distance between C[k] to T[k] in which k= 1 to N. The most used pairwise point comparisons in the one-dollar algorithm are scale, rotation, and position invariant. One-dollar cannot differentiate between gestures whose identities rely on special orientations, ratios, and locations. One-dollar algorithm does not contain usage of time, as a result, gestures cannot be separated with respect to the speed.

$$d_i = \frac{\sum_{K=1}^{N} \sqrt{(C[K_x] - T_i[K_x])^2 + (C[K_y] - T_i[K_y])^2}}{N}$$

$$(2)$$

IV. Experiment

A. Experiment 1

The objective from conducted this experiment was to find out the most accurate algorithm in determining the activity performed by the worker inside the maintenance center. Besides, how it is estimated to differentiate between positive and negative activities.

This experiment was conducted on 3 participants by recording 3 video streams with a duration of 7 minutes for each video. The system has 84 videos of four different activities, including the positive and the negative activities in the different sequences. Each activity was repeated 7 times in each video 4 from the right side and 3 from the left side. The input to our system is split videos for each sequence of activities, the duration is 25 seconds for each video. consequently, After that, the videos were inserted into the system and extracted the important points in the skeleton to determine the path of each point (the X-coordinate and the Y-coordinate) using mediapipe, and save them in a file to be ready to be tested by the algorithms. The system calculated the accuracy of each algorithm in identifying and distinguishing between activities.

In this experiment, we were able to choose the best algorithm by testing each one of them separately, the one-dollar recognizer was able to give the best accuracy rate of 97% regarding the following activities: (changing tires, changing oil, using mobile, and standing without work). The rate of the fastDtw algorithm was 86% in the same previously mentioned activities as shown in Table II.

TABLE II. The Results of Experiment 1

Algorithms	Activites	Accuracy
One-dollar	changing tire changing oil use mobile stand without work	97%
Fastdtw	changing tire changing oil use mobile stand without work	86%

B. Experiment 2

This experiment was conducted to measure the performance of one-dollar algorithm with an increase in the number of participants.

We asked 15 participants to do a sequence of activities in different stream videos. The scenario of the activities was changing tire, use mobile, changing oil, stand without work and it was changeable from one to another. The average duration of the activity in the video was 30 seconds, Each activity was repeated 7 times 4 from the right side and 3 from the left side. The average age of the participants ranged from 19 to 23 years, and the characteristics of the body were also different. Afterwards, the videos were entered into the system, and the system was able to extract 420 videos for a range of different activities in uncontrolled environments. Through this experiment, we were able to measure the accuracy of the system in identifying activities in different conditions.

In this experiment, with large number of participants, experts in the field of mechanics, as well as workers from maintenance centers, the system was able to extract 420 videos, and the one-dollar algorithm started the stage of identifying activities. It achieved an accuracy rate of 94.2%.

C. Discussion

After the two experiments have been done, there's a discussion to explain why these results appear in both. The first experiment was to determine the most accurate algorithm, and the result was that one dollar was more accurate than the fastdtw, which is because the One dollar recognizer has been built over fast dynamic time warping (fastDTW), and both can determine path differences between two trajectories. But the difference between them is that one dollar recognizer reduces the noise in orientation by taking the angles into its calculations, which gives preference to one dollar in the accuracy ratios.The second experiment was to test the performance of the one-dollar algorithm in determining the activity of a large number of participants. The results showed that there was a significant difference in accuracy ratios between the activities. The activity of changing the tire was more subtle because this activity differed from the other activities in the motion trajectory, and the activity of using mobile and standing without work had high accuracy, but there were few similarities between them. The oil change activity has a similar motion trajectory with the two activities (use mobile, standing without work), which reduced its accuracy as shown in Fig. 6.
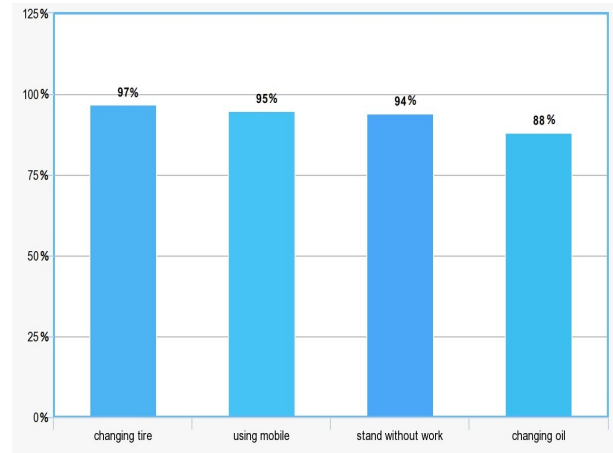

Fig. 6. The Results of the Activities.

The aim of this paper was achieved by using the method of pose classification to extract the pose points from the input video, after extracting the points, a one-dollar algorithm is used to match the extracted points with the collected dataset. As a result, we detect the activities and classify them whether they are positive activities or negative activities. We tested the previous method on 18 participants to prove the accuracy of the mentioned method which reached 94.2%.

V. Conclusion and Future Work

In this paper, we proposed a human activity recognition system in a car workshop that would help maintenance center owners to identify positive or negative worker activities correctly and efficiently. Our system used pose classification to extract workers poses and used two different algorithms to

detect the activities. The system achieved an accuracy rate of 94.2% using this method. However, we are confident that the accuracy of this system will improve with more test videos and that we will increase the activities on which the experiment was conducted in the future. Our future work focus on calculating the wasted time of each worker. Besides, solving problems such as obstacles (engine hood) that appear while identifying activities and affecting the accuracy of the system.

REFERENCES

[1] A. Bagate and M. Shah, "Human activity recognition using rgb-d sensors," in *2019 international conference on intelligent computing and control systems (ICCS)*. IEEE, 2019, pp. 902–905.

[2] T. Hossain, M. S. Islam, M. A. R. Ahad, and S. Inoue, "Human activity recognition using earable device," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 2019, pp. 81–84.

[3] I. A. Lawal and S. Bano, "Deep human activity recognition using wearable sensors," in *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2019, pp. 45–48.

[4] J. Zhao, R. Han, Y. Gan, L. Wan, W. Feng, and S. Wang, "Human identification and interaction detection in cross-view multi-person videos with wearable cameras," in *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, pp. 2608–2616.

[5] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *2017 ieee international conference on big data and smart computing (bigcomp)*. IEEE, 2017, pp. 131–134.

[6] C. Liu, J. Ying, F. Han, and M. Ruan, "Abnormal human activity recognition using bayes classifier and convolutional neural network," in *2018 IEEE 3rd international conference on signal and image processing (ICSIP)*. IEEE, 2018, pp. 33–37.

[7] J. Patalas-Maliszewska, D. Halikowski, and R. Damaševičius, "An automated recognition of work activity in industrial manufacturing using convolutional neural networks," *Electronics*, vol. 10, no. 23, p. 2946, 2021.

[8] S. Win and T. L. L. Thein, "Real-time human motion detection, tracking and activity recognition with skeletal model," in *2020 IEEE Conference on Computer Applications (ICCA)*. IEEE, 2020, pp. 1–5.

[9] G. Ogbuabor and R. La, "Human activity recognition for healthcare using smartphones," in *Proceedings of the 2018 10th international conference on machine learning and computing*, 2018, pp. 41–46.

[10] S. Sezer and E. Surer, "Information augmentation for human activity recognition and fall detection using empirical mode decomposition on smartphone data," in *Proceedings of the 6th International Conference on Movement and Computing*, 2019, pp. 1–8.

[11] P. Washington, A. Kline, O. C. Mutlu, E. Leblanc, C. Hou, N. Stockham, K. Paskov, B. Chrisman, and D. Wall, "Activity recognition with moving cameras and few training examples: Applications for detection of autism-related headbanging," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–7.

[12] Z. Hussain, Q. Sheng, and W. E. Zhang, "Different approaches for human activity recognition: A survey," 06 2019.

[13] M. Yalçın, N. Tüfek, and H. Yalcin, "Activity recognition of interacting people," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2018, pp. 1–4.

[14] N. Nair, C. Thomas, and D. B. Jayagopi, "Human activity recognition using temporal convolutional network," in *Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction*, 2018, pp. 1–8.

[15] S. Tsokov, M. Lazarova, and A. Aleksieva-Petrova, "Evolving 1d convolutional neural networks for human activity recognition," in *International Conference on Computer Systems and Technologies' 21*, 2021, pp. 49–54.

[16] A. Naser, A. Lotfi, J. Zhong, and J. He, "Human activity of daily living recognition in presence of an animal pet using thermal sensor array," in *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2020, pp. 1–6.

[17] P. Hristov, A. Manolova, and O. Boumbarov, "Deep learning and svm-based method for human activity recognition with skeleton data," in *2020 28th National Conference with International Participation (TELECOM)*. IEEE, 2020, pp. 49–52.

[18] H. Ramirez, S. A. Velastin, I. Meza, E. Fabregas, D. Makris, and G. Farias, "Fall detection and activity recognition using human skeleton features," *IEEE Access*, vol. 9, pp. 33 532–33 542, 2021.

[19] A. S. Syed, Z. S. Syed, M. S. Shah, and S. Saddar, "Using wearable sensors for human activity recognition in logistics: A comparison of different feature sets and machine learning algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2020.0110976

[20] I. Serrano, O. Deniz, J. L. Espinosa-Aranda, and G. Bueno, "Fight recognition in video using hough forests and 2d convolutional neural network," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 4787–4797, 2018.

[21] Y. E. Görgülü and K. Taşdelen, "Human activity recognition and temporal action localization based on depth sensor skeletal data," in *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*. IEEE, 2020, pp. 1–5.

[22] W. Tao, Z.-H. Lai, M. C. Leu, and Z. Yin, "Worker activity recognition in smart manufacturing using imu and semg signals with convolutional neural networks," *Procedia Manufacturing*, vol. 26, pp. 1159–1166, 2018.

[23] S. Alghyaline, "A real-time street actions detection," *International Journal of Advanced Computer Science and Applications*, vol. 10, 01 2019.

[24] M. Arzani, M. Fathy, A. Azirani, and E. Adeli, "Switching structured prediction for simple and complex human activity recognition," *IEEE Transactions on Cybernetics*, vol. PP, pp. 1–12, 01 2020.

[25] N. Archana and K. Hareesh, "Real-time human activity recognition using resnet and 3d convolutional neural networks," in *2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)*. IEEE, 2021, pp. 173–177.

[26] X. Sun, H. Xu, Z. Dong, L. Shi, Q. Liu, J. Li, T. Li, S. Fan, and Y. Wang, "Capsganet: Deep neural network based on capsule and gru for human activity recognition," *IEEE Systems Journal*, 2022.

[27] V. Radhika, C. R. Prasad, and A. Chakradhar, "Smartphone-based human activities recognition system using random forest algorithm," in *2022 International Conference for Advancement in Technology (ICONAT)*. IEEE, 2022, pp. 1–4.

[28] P. Mittal *et al.*, "Machine learning (ml) based human activity recognition model using smart sensors in iot environment," in *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE, 2022, pp. 330–334.

[29] Y. Tang, L. Zhang, F. Min, and J. He, "Multi-scale deep feature learning for human activity recognition using wearable sensors," *IEEE Transactions on Industrial Electronics*, 2022.

[30] M. Noras, "Activity detection and recognition with passive electric field sensors," *IEEE Transactions on Industry Applications*, 2021.