Recognition of Nurse Activities in Endotracheal Suctioning Procedures: A Comparative Analysis Using LightGBM and Other Algorithms

Penglin Jiang , Boyang Dai [2], Bochen Lyu [3], Zeng Fan [4], Gulustan Dogan [5]
Sichuan University, [24]Beijing Institute of Technology, [3]Nanjing University,
[5]University of North Carolina Wilmington

Abstract

This research is based on the 6th ABC Challenge which focuses on leveraging Human Activity Recognition (HAR) systems to enhance Endotracheal Suctioning (ES) procedures. The challenge's objective is to accurately identify the activities performed by nurses based on the dataset. The dataset comprising skeleton data and video recordings of healthcare professionals performing ES procedures is collected and preprocessed. Informative features capturing joint angles, velocities, and spatial relationships are extracted. These features are then used as inputs to three di erent prediction models GBDT, XGBoost, and LightGBM. Our experimental results demonstrate that LightGBM outperforms the other models with the highest accuracy of 0.819, followed by XGBoost (0.807) and GBDT (0.763) on the Nurse Care Activity Recognition Challenge benchmark dataset. These findings contribute to advancing nurse activity recognition and have implications for improving healthcare monitoring and workflow management. Given the outstanding performance of LightGBM, we chose to submit our results using this algorithm for the challenge. The code is available at *https://github.com/mobaaa12/Endotracheal Suctioning Procedure Recognition*.

1 Introduction

The precise execution of Endotracheal Suctioning (ES) is a cornerstone in managing patients requiring mechanical ventilation. As a critical

2021141520184@stu.scu.edu.cn
[2]1120211588@bit.edu.cn
[3]221098027@smail.nju.edu.cn
[4]1120212696@bit.edu.cn
[5]dogang@uncw.edu

procedure, ES demands not only a high degree of skill and knowledge from healthcare professionals but also necessitates innovative approaches to training and monitoring to minimize risks. Integrating technology into healthcare practices offers a pathway to achieving these objectives, mainly through applying Human Activity Recognition (HAR) systems.

This work is part of the 6th ABC Challenge titled Activity Recognition of Nurse Training Activity using Skeleton and Video Dataset with Generative AI on Activity and Behavior Computing[1]. The challenge's focus on ES arises from understanding its critical role in maintaining patient airway patency and preventing ventilator-associated pneumonia[2]. The need for precise movements and adherence to sterile procedures characterizes the complexity of ES. It calls for advanced solutions supporting healthcare professionals in training and practice.

In recent times, there have been significant improvements in sensor technologies, optimization of traditional models, and deep learning algorithms, which have enhanced the capacity of HAR systems to capture and interpret complex human movements with greater accuracy. The advancements in generative AI have further augmented these capabilities by enabling the synthesis of realistic and diverse human activity data, which helps in training HAR systems more effectively, particularly in scenarios where real training data is scarce or difficult to obtain. This progress has opened up new possibilities for applying HAR in the medical field beyond general activity recognition to support specific, high-stakes procedures. By utilizing skeletal data and employing sophisticated computational models, we can develop systems that offer detailed insights into the performance of ES. This innovative approach boosts the model's learning abilities and paves the way for creating more adaptable AI systems for various medical procedures with minimal human intervention.

The main contributions of our work are summarized as follows:

- We have proposed a comprehensive methodology for recognizing activities related to tracheal intubation suctioning. Our approach utilizes advanced machine learning algorithms, including GBDT, XGBoost, and LightGBM, to accurately recognize different activities. The best accuracy degree is in LightGBM(0.819).
- We have used generative AI to enhance our feature extraction processes by leveraging its capability to generate and evaluate diverse feature extraction scenarios. We have employed a comprehensive set of features derived from joint motion data to capture the subtle differences between various activities. These features include 2D coordinates, immediate displacement, immediate motion orientation, and joint angles.
- We have evaluated the performance of multiple machine learning algorithms, including GBDT, XGBoost, LightGBM, Random Forest, Gaussian naive Bayes, decision tree, logistic regression, support vector machine, and k-nearest neighbor. This comparative analysis provides insights into the strengths and weaknesses of different algorithms for activity recognition in the context of tracheal intubation suctioning.

2 Related Work

2.1 Previous ABC Challenges

In 2023, ABC Challenges aimed to develop e ective methods utilizing wearable devices to monitor and predict PD patients' wearing-o periods and symptom fluctuations. It will help the doctors to create specific treatment strategies to manage Parkinson's disease and its associated symptoms properly, enabling personalized treatment strategies and improved quality of life[3]. In 2022, the challenge was Nurse Care Activity Recognition aiming to develop a model that can accurately recognize and classify daily nurse care activities at a nursing care facility[4]. According to the result provided, the combination of proper data pre-processing and the two-stage training process resulted in improved performance for user-specific future activity prediction in the healthcare domain. In 2021, the goal was to develop methods for accurately identifying di erent activities performed during Bento-box packaging using motion capture to advance human-robot collaboration in industrial settings[5]. In 2020, the same conference organized a challenge to recognize the macro and micro activities taking place during cooking sessions[6]. The result showed that by employing a clustering approach to identify individuals and then conducting cluster/person-specific feature selection with semi-supervised learning techniques to classify the test data, an internal test accuracy was achieved at 81.01%[7].

Unlike the other years, this year's challenge was to recognize and assess the activities performed during Nurse Training for ES to improve patient safety and enhance nurses' skills.

2.2 Generative AI Models Using in Activity Recog nition

Additionally, with the emergence of generative AI models such as large language models (LLMs) and text-driven motion synthesis models, generative AI such as Inertial Measurement Unit Generative Pre-trained Transformer (IMUGPT) has been used in HAR. Leng et al. (2024) made enhancements to the IMUGPT model, including a motion filter and diversity metrics, to improve the generation and efficiency of virtual IMU data[8]. Another study conducted by Xia et al. (2023) introduces a novel two-stage prompt engineering approach. It uses ChatGPT to generate activity descriptions and classify activities based on object sequences in an unsupervised manner, demonstrating improved performance over existing methods[9]. Their works show an enhancement in HAR by using generative AI models.

2.3 Human Pose Estimation in Activity Recogni tion

Human pose estimation technologies play a key role in Human Activity Recognition(HAR), especially for medical and nursing activities that require analysis and understanding of complex human movements. With technological advancements, especially in artificial intelligence and machine learning, recognizing and monitoring human activities have become possible[10, 11, 12, 13]. Song et al. (2021) demonstrated how human pose estimation technology could improve the performance of activity recognition systems[14]. Their research confirmed that analyzing the movement patterns of human joints improves the capability to recognize specific activities.

2.4 YOLOv7 Using in The Pose Skeleton

The Pose Skeleton (Keypoints) in this challenge were extracted from videos by using YOLOv7. YOLOv7 is an advanced real-time object detection system in the field of computer vision. YOLOv7 adopts a trainable bag-of-freebies approach, which combines efficient training tools, a specific architecture, and a compound scaling method. This combination of techniques enhances the performance and accuracy of the object detection system. Moreover, with a GPU V100, YOLOv7 achieves the highest accuracy among all known real-time object detectors, with an average precision (AP) of 56.8%[15]. Du et al. (2024) implement an AI-generated resource allocation scheme that utilizes an improved YOLOv7-X object detector to extract crucial semantic information from images. It improves data management and transmission efficiency[16].

2.5 Limitations in Prior Research and Our Inno vations

However, traditional activity recognition models often rely on predefined features and manually crafted algorithms, which may not be flexible or accurate enough when dealing with dynamic and complex nursing activities. Additionally, the performance of traditional HAR systems can vary greatly across di erent environments, as they are susceptible to background noise, changes in lighting, and other visual disturbances[17]. This is particularly problematic in various environments such as hospitals and ICUs, where visual conditions can change rapidly.

Despite the deficiency, traditional machine learning models o er several advantages in terms of stability and interpretability. The interpretability of these models allows healthcare professionals to understand and trust the results, enabling better decision-making in patient care. Furthermore, traditional machine learning methods can be optimized for real-time processing, essential for timely medical analysis and diagnosis.

In conclusion, in our research, we have combined the strengths of these traditional approaches with the innovative capabilities of modern techniques, including generative AI, to develop robust and reliable solutions. Specifically, we have introduced a comprehensive methodology for recognizing activities related to tracheal intubation suctioning, utilizing advanced machine learning algorithms. Furthermore, we have utilized generative AI to refine our feature extraction methods, allowing us to identify and utilize a wide range of joint motion features e ectively for activity recognition.

3 Dataset

The 6th ABC Challenge is centered around the recognition of activities related to tracheal intubation suctioning, a critical procedure in patient care within ICU settings. The challenge provides a comprehensive dataset comprised of both skeletal data and video recordings, specifically designed to capture the nuances of healthcare-related activities.

3.1 Data Collection

The dataset was meticulously curated from a controlled environment where professional healthcare workers performed a series of predefined tracheal intubation suctioning tasks. Each activity was captured using high-definition video cameras and depth-sensing equipment to generate both RGB video data and 3D skeletal data. The collection process was conducted with strict adherence to privacy and ethical standards, ensuring the anonymization of all participants involved.

Table 1: Description of some columns in CSV les

Column name	Description of column
nose_x	X coordinate value of nose
nose_conf	Con dence value of nose
left_eye_y	Y coordinate value of left eye
left_eye_conf	Con dence value of left eye
right_shoulder_conf	Con dence value of right shoulder
left_elbow_x	X coordinate value of left elbow

3.2 Dataset Structure

The dataset is divided into two primary components: Video Data and Skeletal Data.

Video Data There are 32 videos in MTS format of subjects in the Training set. High-definition videos showcasing the execution of suctioning tasks from multiple angles. Each video is annotated with timestamps indicating the start and end of specific activities, providing a visual reference for activity recognition. The frame per second of videos is 30 and the image size is 1920×1080.

Skeletal Data 3D skeletal data extracted from the depth-sensing equipment, representing the movements of healthcare workers in a structured format. Each video frame contains 17 positions on the subject's body with x and y coordinates and confidence scores as seen in Table1. These files enable detailed analysis of motion patterns associated with suctioning activity. Post-processing steps were applied to solve the problem that other people were included in the frame while passing by in the background, therefore only the skeleton of the main nurse was kept. The sampling rate of the key points is 30.

3.3 Annotations

The dataset includes detailed annotations for each activity. Annotations were performed by medical experts and included labels for each specific suctioning task, as well as sub-tasks that are critical to the procedure (Shown in Table2). The annotations provide a ground truth for developing and benchmarking activity recognition models.

Table 2: Activities in endotracheal suctioning and their id

Activity class id	Activity name
0	Catheter preparation
1	Temporal movement of an arti cial airway
2	Suctioning phlegm
3	Re tting the arti cial airway
4	Catheter disinfection
5	Discarding gloves
6	Positioning
7	Auscultation
8	Others

4 Methodology

Firstly, we eliminated nonsensical We first processed the joint coordinate data obtained from YOLOv7 by applying smoothing techniques. Following that, generative AI was utilized to determine the appropriate features to extract from the smoothed data. These extracted features were then fed into various prediction models for activity recognition. In our research, we specifically employed the Light Gradient Boosting Machine (LightGBM) algorithm and compared its performance with two other algorithms: Gradient Boosting Decision Tree (GBDT) and Extreme Gradient Boosting (XGBoost). The modeling and prediction process is implemented in Python 3.7, utilizing the respective libraries: LightGBM, scikit-learn[18], and XGBoost. Figure 1 illustrates the experimental procedure we followed.

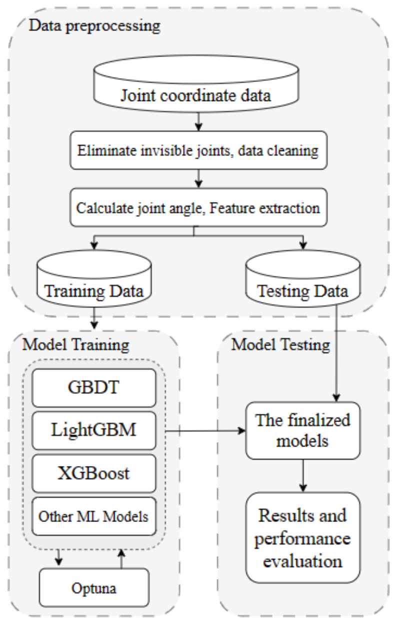


Figure 1: The architecture of Recognizing endotracheal suctioning procedures with ML models

4.1 Joint Selection and Data Preprocessing

For our analysis, we focused on 13 joints above the hip. The knee and ankle data were excluded due to the limited visibility of the lower part of the nurse's body in the frame. To eliminate noises and facilitate feature extraction, we applied interpolation to missing values (represented as zeros) occurring within a time frame of 3 seconds.

4.2 Generative AI Utilization

Although there have been numerous research studies on human activity recognition, there is a scarcity of prior work focusing specifically on identifying Endotracheal Suctioning procedures. In this situation, we utilize Generative AI to identify the essential features of the Endotracheal Suctioning procedure. Figure 2 shows our prompt and the features suggested by the GPT-3.5-Turbo model[19]. GPT-3.5-Turbo is an enhanced version of the natural language processing model developed by OpenAI, belonging to the GPT (Generative Pre-trained Transformer) series. GPT-3.5-Turbo utilizes deep learning and extensive data training, enhancing the model's ability to comprehend complex texts, generate coherent content, and handle subtle language nuances[20]. Notably, this version is particularly optimized for applications requiring faster response times and reduced latency, making it highly suitable for real-time applications such as conversational AI, interactive tools, and other scenarios where quick processing is crucial. Out of the 8 responses generated by the AI, 6 pertained to feature extraction, with 5 of these methods being implemented in our study. As depicted in Figure 2, the highlighted section signifies our utilization or adaptation of the approach suggested by generative AI, which will be discussed in the next section.

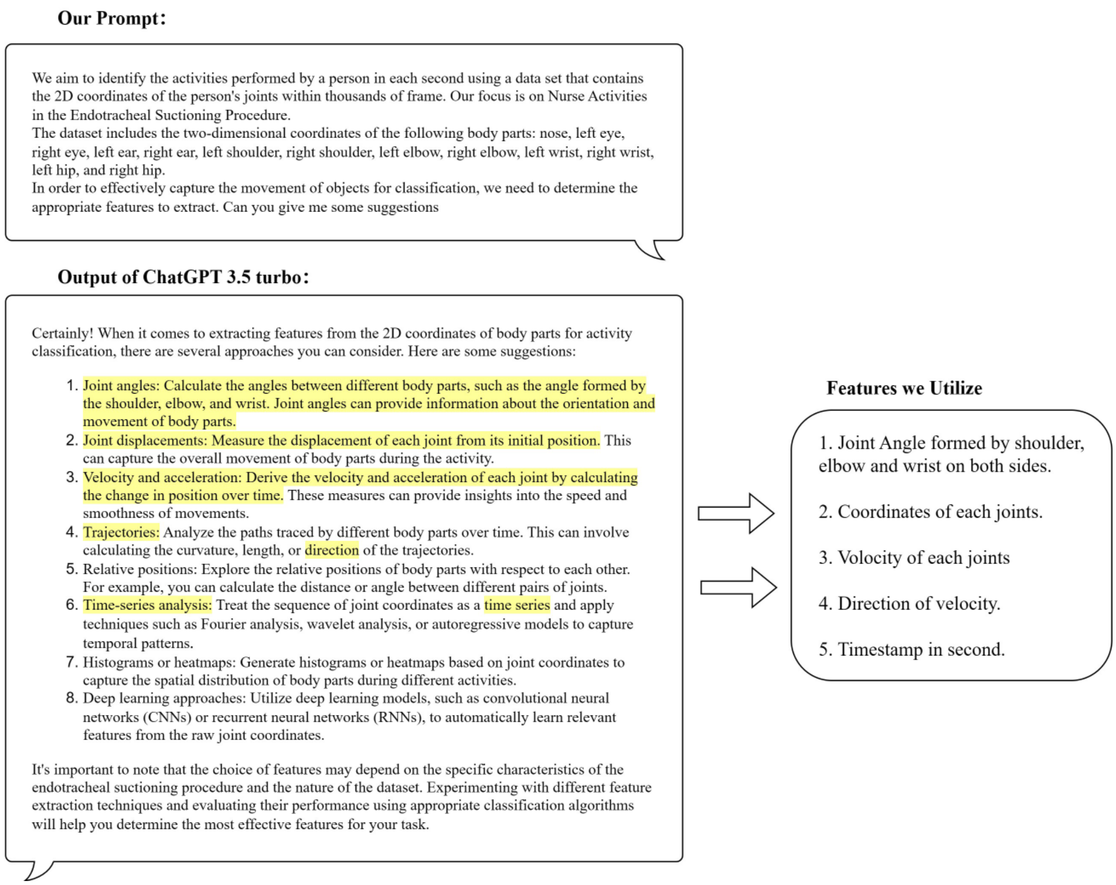


Figure 2: Output from GPT-3.5 Turbo [19]

4.3 Feature Extraction

We have fully adopted the 1st response generated by Chat-GPT 3.5 turbo, which suggests using the angles formed by the shoulders, elbows, and wrists on both sides as key input features. We believe that in the context of endotracheal suctioning activities, these joints are engaged more actively, and the angles between them provide a more holistic representation of the interplay among these joints. The joint angle we select is illustrated in Figure 3.

In addition, we have partially integrated the suggestions from the 2nd, 3rd, and 4th generated responses, incorporating the 2-D coordinates, immediate displacement, and immediate motion orientation of each joint as features. This aligns with the classical physics theory's description of object motion, which involves distance, velocity, and direction.

Furthermore, we have referred to the 6th response, treating the joint coordinate data as a time series. From the 2D coordinates, immediate displacement, immediate motion orientation, and joint angles, we extracted seven time-domain features per second, encompassing mean, standard deviation, minimum, maximum, variance, median, and sum. To improve our model's ability to recognize a sequence of continuous activities, we added the timestamp as an extra feature. Lastly, we performed data normalization, scaling all 498 features to a range between 0 and 1.

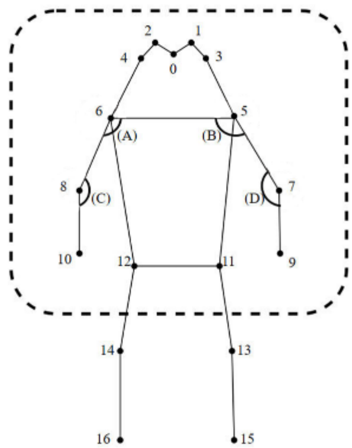


Figure 3: Joints and angles selected (Inside the enclosed dashed line)

4.4 Algorithm Overview

The Gradient Boosting Decision Trees (GBDT) is an ensemble learning algorithm that combines multiple decision trees (DTs)[21]. DTs are machine learning models with tree-like structures that can handle di erent types of data and explore all possible paths to prediction results[22]. However, DTs are prone to overfitting and sensitive to noise in the dataset. GBDT tackles these issues by integrating multiple DTs, enhancing prediction performance by compensating for individual tree errors. Based on the framework of GBDT, XGBoost[23] and LightGBM[24] received wide attention for their excellent performance as extensions. These algorithms e ectively address overfitting, ensuring reliable predictions for future data[22].

Table 3: Hyperparameters provided by Optuna.

lgorithms	**Hyperparameters**	**Values**
GBDT	learning_rate	0.07620383307456176
	n_estimators	384
	max_depth	12
	min_samples_leaf	28
	min_samples_split	206
	subsample	0.9287588738032913
	max_features	14
XGBoost	alpha	5.902013633952215
	lambda	0.3928176818610557
	learning_rate	0.04214635827710827
	n_estimators	1159
	colsample_bytree	0.6527299591714091
	subsample	0.665359655991066
	max_depth	8
LightGBM	learning_rate	0.0952265688751135
	n_estimators	301
	lambda_l1	0.014924041633401866
	lambda_l2	0.02068744601075835
	max_depth	14
	colsample_bytree	0.8950907663814264
	subsample	0.9972956871793692
	min_child_samples	49

4.5 Hyperparameter Optimization

The selection of hyperparameters for tree-based machine learning models has a significant impact on their forecasting performance[25, 26]. It is crucial to adjust these hyperparameters based on the dataset rather than manually specifying them. In our study, we employ Optuna[27] to identify the most suitable hyperparameters. Optuna is a software framework that automates the tuning of hyperparameters in models. It o ers flexibility by allowing users to dynamically construct the search space during runtime using methods and trial objects[28]. By formalizing the process of minimizing or maximizing an objective function that takes hyperparameters as inputs, Optuna optimizes hyperparameter values and returns a validation score that represents the model's performance. The efficiency of Optuna's dynamic searching strategy is crucial in determining the set of hyperparameters, while the performance estimation strategy estimates the value of the hyperparameter set based on discarded learning curves and hyperparameters. These dynamic strategies contribute to a cost-e ective solution for model optimization[27]. Table 3 displays the hyperparameters obtained through Optuna for GBDT, XGBoost, and LightGBM.

4.6 Model Evaluation Indexes

The accuracy, precision, recall, and F1 score are widely used metrics for evaluating the performance of machine learning (ML) algorithms[29]. Accuracy measures the proportion of correctly predicted samples, while precision focuses on accurately predicting positive samples. Recall emphasizes the ability to correctly predict as many positive samples as possible. The F1 score provides a comprehensive metric that balances both precision and recall. In this study, these metrics are adopted to evaluate the performance of the models, providing a comprehensive understanding of their predictive capabilities and the trade-o between precision and recall.

5 Results

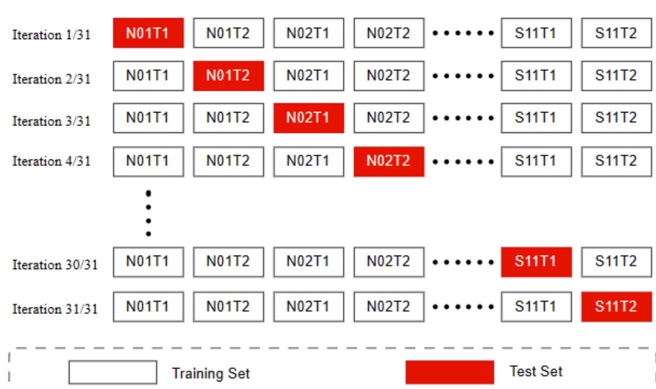
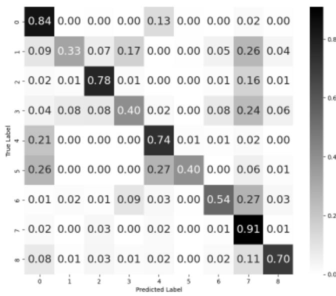
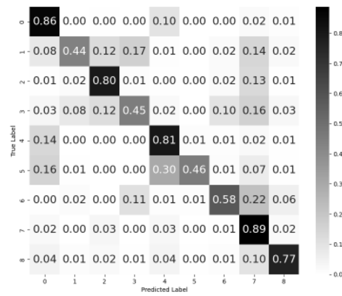


Figure 4: Schematic diagram for "leave-one-subject-out" cross-validation (LOSO CV)

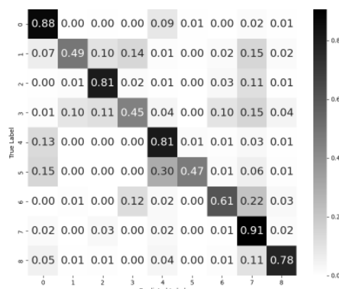
Based on the feature extraction framework o ered by generative AI, the prediction results of GBDT, XGBoost, and LightGBM algorithms were obtained using the "leave-one-subject-out" cross-validation (LOSO



a) Confusion matrix for GBDT



b) Confusion matrix for XGBoost



c) Confusion matrix for LGBM

Figure 5: Confusion matrix for GBDT, XGBoost and LightGBM

CV) method. In this method, as shown in Figure4, we select one set of operation data as the test set and the remaining data as the training set at a time. The process was repeated 31 times. Following that, the confusion matrix for each algorithm was generated, as depicted in Figure5. The values along the main diagonal represent the number of samples that were correctly predicted. It can be seen that actions assigned with labels 1, 3, and 5 exhibit noticeably lower accuracy rates. Conversely, actions labelled 0, 2, 4, 7, and 8 showcase a relatively high level of accuracy across all three models. Based on Figure 5, the accuracy, precision, recall, and F1 score are calculated and shown in Table 4.

According to the results, all three algorithms had good overall performances in recognizing the activities. Furthermore, the accuracy degrees of LightGBM were largest and up to 0.819, followed by XGBoost and GBDT with an accuracy of 0.807 and 0.763. Therefore, based on their overall prediction performances, the rank was LightGBM XGBoost GBDT. Due to the excellent performance of LightGBM, we submitted our results using the LightGBM algorithm.

Table 4: Model Evaluation Results

Model	Accuracy	Precision	Recall	F1 Score
LightGBM	0.82	0.83	0.82	0.82
XGBoost	0.81	0.82	0.81	0.81
GBDT	0.78	0.79	0.78	0.78

6 Discussions

While the overall performance of LightGBM, GBDT, and XGBoost algorithms in recognizing tracheal intubation suctioning activities is satisfactory, the accuracy rates for di erent activities vary. Specifically, activities labeled as 1, 3, and 5 exhibit lower accuracy rates. This can be attributed to two main reasons. Firstly, the number of samples available for activities 1, 3, and 5 is smaller compared to other activities. This is evident from Figure 6, which displays the label distribution in the dataset, showing a significant disparity in the number of occurrences for activities 1, 3, and 5. Secondly, the demarcation boundary on activities labelled 1, 3, and 5 could exhibit a higher degree of uncertainty, thereby exerting a potential influence on the data's overall quality. As data-driven approaches, the prediction performance of GBDT, XGBoost, and LightGBM algorithms is greatly influenced by the quantity and quality of the supporting data. Consequently, the accuracy of activities labelled 1, 3, and 5 is comparatively lower when compared to other activities.

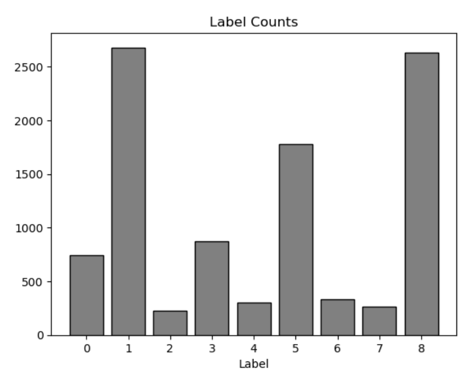


Figure 6: Label counts

It is noteworthy that the features o ered by generative AI significantly enhance the performance of the model. We input the time domain features of the origional data into the model, optimize the hyperparameters using Optuna, and then evaluate the final outcomes by comparing them to the results obtained from feature extraction aided by generative AI. As illustrated in Figure 7, through AI generated feature extraction method, the recognition accuracy is enhanced across all categories. The overall accuracy is augmented by approximately 0.2. This result proves that feature extraction through generative AI is dependable in the realm of human activity recognition.

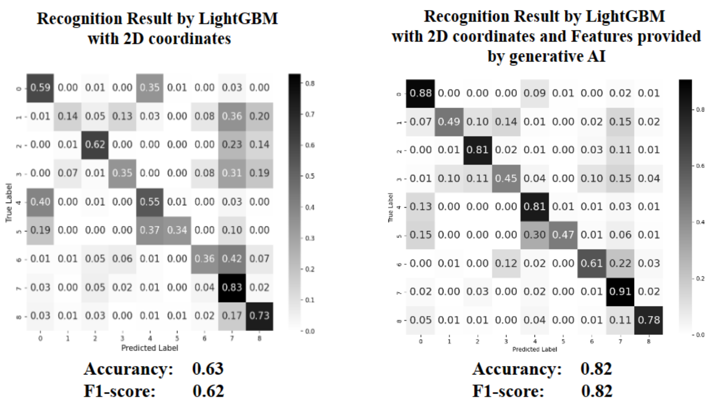


Figure 7: Comparison of the model's performance before and after incorporating features supplied by generative AI

In the study, several machine learning (ML) algorithms were also utilized for activity recognition, including Random Forest (RF), Gaussian naive Bayes (GNB), decision tree (DT), logistic regression (LR), support vector machine (SVM), and k-nearest neighbour (KNN). Similarly, the hyperparameters of these algorithms were optimized using Optuna. The performance of each algorithm was evaluated using Leave-One-Subject-Out Cross-Validation (LOSO-CV) and measured in terms of accuracy. Figure 8 presents the accuracy results for each algorithm. The accuracies of these algorithms were consistently lower compared to the GBDT, XGBoost, and LightGBM algorithms. Most algorithms, except for RF, achieved accuracies below 0.7, while LightGBM algorithms surpassed 0.8 accuracy. It demonstrated that compared with other ML algorithms, LightGBM was better for recognizing tracheal intubation suctioning activities.

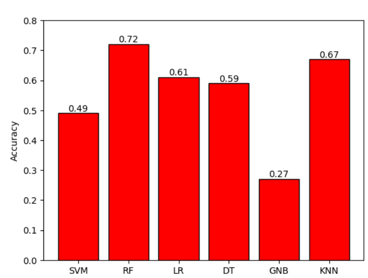


Figure 8: Accuracy of each comparison algorithm

7 Future Work

In future work, several areas can be explored to enhance the activity recognition of ES activities. Due to the limited number of samples available for activities labeled 1, 3, and 5, data augmentation techniques can be employed to artificially increase the dataset size. This can involve generating synthetic samples or applying transformations to existing samples to create variations in the data. Transfer learning techniques can be investigated to leverage pre-trained models on related activity recognition tasks or large-scale datasets. The addition of a regularization term can be explored to limit the complexity of the model and improve its generalization performance.

By incorporating these approaches and techniques, it is possible to e ectively improve the generalization capabilities of the activity recognition models, leading to more accurate and reliable recognition of tracheal intubation suctioning activities in healthcare settings.

8 Acknowledgment

This work has been done as part of the GEARS program by NC State University. We express our sincere gratitude to the organizers of the 6th ABC challenge for providing the research idea and dataset, which were crucial for our study. Additionally, we would like to extend our special thanks to Dr.Gulustan Dogan from the University of North Carolina Wilmington for her invaluable guidance and insights. Her expertise significantly enriched our work and facilitated the smooth progression of our research.

9 Author Contributions Statement

Penglin Jiang, Boyang Dai, and Bochen Lv contributed equally to this work as co-first authors. Penglin Jiang and Boyang Dai were responsible for model construction, conducting experiments, and manuscript writing. Bochen Lv participated in conducting experiments and adjusting hyperparameters. Zeng Fan contributed to the manuscript writing, ensuring the clarity and accuracy of the presentation. All authors have read and agreed to the published version of the manuscript.

References

[1] Hoang Anh Vy Ngo, Haru Kaneko, Iqbal Hassan, Elsen Ronando, Milyun Ni'ma Shoumi, Ryuya Munemoto, Tahera Hossain, and Sozo Inoue. Summary of the nurse care activity recognition challenge using skelton data from video with generative ai. *International Journal of Activity and Behavior Computing*, 2024, 2024.

[2] Hoang Anh Vy Ngo, Quynh N Phuong Vu, Noriyo Colley, Shinji Ninomiya, Satoshi Kanai, Shunsuke Komizunai, Atsushi Konno, Misuzu Nakamura, and Sozo Inoue. Toward recognizing nursing activity

in endotracheal suctioning using video-based pose estimation. *International Journal of Activity and Behavior Computing*, 2024(1):1–20, 2024.

[3] Victorino, John Noel, Haru Kaneko, Christina Garcia, Defry Hamdhana, Muhammad Fikry, Nazmun Nahid, Tahera Hossain, and Tomohiro Shibata; Sozo Inoue. 5th abc challenge: Forecasting parkinson's disease patients' wearing-o phenomenon datasets, 2023.

[4] Sozo Inoue, Defry Hamdhana, Christina Garcia, Haru Kaneko, Nazmun Nahid, Tahera Hossain, Sayeda Shamma Alia, and Paula Lago. 4th nurse care activity recognition challenge datasets, 2022.

[5] Sayeda Shamma Alia, Kohei Adachi, Nazmun Nahid, Haru Kaneko, Paula Lago, and Sozo Inoue. Bento packaging activity recognition challenge, 2021.

[6] Paula Lago, Shingo Takeda, Kohei Adachi, Sayeda Shamma Alia, Moe Matsuki, Brahim Benai, Sozo Inoue, and Francois Charpillet. Cooking activity dataset with macro and micro activities, 2020.

[7] Stefan Kalabakov, Simon Stankoski, Nina Reščič, Ivana Kiprijanovska, Andrejaana Andova, Clement Picard, Vito Janko, Martin Gjoreski, and Mitja Luštrek. Tackling the shl challenge 2020 with person-specific classifiers and semi-supervised learning. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '20 Adjunct, page 323–328, New York, NY, USA, 2020. Association for Computing Machinery.

[8] Zikang Leng, Amitrajit Bhattacharjee, Hrudhai Rajasekhar, Lizhe Zhang, Elizabeth Bruda, Hyeokhyen Kwon, and Thomas Plötz. Imugpt 2.0: Language-based cross modality transfer for sensor-based human activity recognition, 2024.

[9] Qingxin Xia, Takuya Maekawa, and Takahiro Hara. Unsupervised human activity recognition through two-stage prompting with chatgpt, 2023.

[10] Gulustan Dogan, Iremnaz Cay, Sinem Sena Ertas, Şeref Recep Keskin, Nouran Alotaibi, and Elif Sahin. Where are you? human activity recognition with smartphone sensor data. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '20 Adjunct, page 301–304, New York, NY, USA, 2020. Association for Computing Machinery.

[11] Şeref Recep Keskin, Ayşenur Gençdoğmuş, Buse Yıldırım, Gülüstan Doğan, and Yusuf Öztürk. Dnn and cnn approach for human activity recognition. In *2020 7th International Conference on Electrical and Electronics Engineering (ICEEE)*, pages 254–258, Antalya, Turkey, 2020.

[12] Gulustan Dogan, Sinem Sena Ertas, and İremnaz Cay. Human activity recognition using convolutional neural networks. In *2021*

IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pages 1–5, 2021.

[13] Massinissa Hamidi, Aomar Osmani, Lukmon Rasaq, Gulustan Dogan, and Nouran Alotaibi. Multiview representation learning for human activity recognition. In *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pages 1–5, Chemnitz, Germany, 2022.

[14] Liangchen Song, Gang Yu, Junsong Yuan, and Zicheng Liu. Human pose estimation and its application to action recognition: A survey. *Journal of Visual Communication and Image Representation*, 76:103055, 2021.

[15] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7464–7475, Vancouver, BC, Canada, June 2023.

[16] Baoxia Du, Hongyang Du, Haifeng Liu, Dusit Niyato, Peng Xin, Jun Yu, Mingyang Qi, and You Tang. Yolo-based semantic communication with generative ai-aided resource allocation for digital twins construction. *IEEE Internet of Things Journal*, 11(5):7664–7678, 2024.

[17] Pranjal Kumar and Siddhartha Chauhan. Human activity recognition with deep learning: Overview, challenges & possibilities. *CCF Transactions on Pervasive Computing and Interaction*, 339(3):1–29, 2021.

[18] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[19] OpenAI. ChatGPT [Large language model]. `https://chat openai com/chat`, 2024.

[20] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[21] Shruti Sachdeva, Tarunpreet Bhatia, and AK Verma. Gis-based evolutionary optimized gradient boosted decision trees for forest fire susceptibility mapping. *Natural Hazards*, 92:1399–1418, 2018.

[22] Sotiris B Kotsiantis. Decision trees: a recent overview. *Arti cial Intelligence Review*, 39:261–283, 2013.

[23] Tianqi Chen and Carlos Guestrin. Xgboost: a scalable tree boosting system proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016: 785-794. *ACM, New York, NY*, 2016.

[24] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

[25] Li Yang and Abdallah Shami. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415:295–316, 2020.

[26] Răzvan Andonie. Hyperparameter optimization in learning systems. *Journal of Membrane Computing*, 1(4):279–291, 2019.

[27] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, Anchorage, AK, USA, 2019.

[28] Satvik Garg and Pradyumn Pundir. Mofit: A framework to reduce obesity using machine learning and iot. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pages 1733–1740, Opatija, Croatia, 2021. IEEE.

[29] Rahul Kumar. *Machine Learning Quick Reference: Quick and essential machine learning hacks for training smart data models.* Packt Publishing Ltd, Birmingham, UK, 2019.