Enhancing Nursing Activity Recognition During Endotracheal Suctioning Through Video-based Pose Estimation and Machine Learning

Samiul Islam [1*], S. M. Hozaifa Hossain [2*], Md. Zasim Uddin [3], Shahera Hossain [4], Md Atiqur Rahman Ahad [5]
[12]University of Dhaka, [3]Begum Rokeya University, Rangpur, [45]University of East London

Abstract

Endotracheal suctioning is a crucial medical process where skilled professional nurses are needed. However, there is a lack of research on automatically recognizing nurse activities during this procedure. This paper presents an innovative method to identify nurses' actions during endotracheal suctioning procedures by analyzing skeleton data from image sequences using different traditional machine learning-based (ML) methods. Firstly, we preprocess the skeleton data and extract the feature, then employ the ML method for classification. Moreover, we explored the Generative AI with LLM for feature generation and selection to improve the accuracy. We evaluated our proposed framework using metrics such as accuracy and F1-Score. We demonstrate our proposed framework on the Activity Recognition of Nurse Training Activity of the 6th ABC Challenge dataset and find out that XGBoost obtained the best accuracy. The accuracy and F1-score both are 97.0%. We hope this research contributes to automated nursing activity recognition, potentially benefiting patient care and safety during endotracheal suctioning procedures.

1 Introduction

Endotracheal suctioning (ETS) is a crucial medical procedure, especially for patients with compromised respiratory function and who cannot clear

[1]samiulislam29073@gmail.com
[2]20hozaifa02@gmail.com
[3]zasim@brur.ac.bd
[4]s.hossain2@uel.ac.uk, shaherahossain@gmail.com
[5]mahad@uel.ac.uk, atiqahad@gmail.com
[*]Both authors contributed equally to this work

their airways efficiently [1]. It involves the insertion of a sterile suction catheter through an endotracheal tube (ETT) to remove pulmonary secretions from a patient with an artificial airway in place. The ETS procedure is most commonly performed in the intensive care unit (ICU). When performing this intervention, the ICU nurses should be aware of the potential hazards a patient is exposed to and strive to prevent or minimize these [2].

Nurse activity recognition plays a crucial role in the ETS system as it can help in monitoring the compliance of healthcare plans for individual patients. This, in turn, frees up nurses from the tedious task of manual reporting and documentation. Nurses carry out complex and time-consuming activities that require top-notch nursing care for the safety and improvement of patient's quality of life in the ICU system. During their work time, nurses usually follow a worksheet that lists all the activities they need to perform [2, 3]. To reduce their working hours, nurses need to devise a plan to perform multiple activities collectively, instead of doing them one-by-one. Additionally, they need to reflect on their skills to improve while conducting this complicated procedure. Moreover, activity recognition can assist nurses in better managing and improving the quality of their work, and also evaluate their performance while conducting ETS [4].

Human activity recognition is a heavily studied domain in computer vision, given its practical applications in areas like human-computer interaction and video comprehension. Recently, some studies have been conducted for nursing activity recognition in the literature using different sensors. For example, Ijaz et al. [5] studied a transformer-based network to extract multi-modal features using skeletal and acceleration data. Finally, they performed fusion for nursing activity recognition. In contrast, Takebe et al. [6] explore the uses of accelerometers and radio frequency identification (RFID) for 13 nursing activity recognition. In addition, Momen et al. [7] used an accelerometer to recognize six nursing activities around a patient, in [8] prepared a benchmark dataset including 23 nursing activities from 14 caregivers using five sensors.

In this paper, we study to recognize nursing activities occurring during endotracheal suctioning using skeleton-based data and develop a procedure with the help of generative AI. The core contributions are summarized below:

- We explored using generative AI to select data processing techniques and methodologies for nursing activity recognition using skeleton data. Moreover, we proposed an approach using generative AI for nursing activity recognition, combining all skeleton data for all users to train a robust model.
- Furthermore, we extensively studied nursing activity recognition by considering different input data pipelines with traditional machine learning methods, including single-user skeleton data iterated over each sample and a combination of all skeleton data for all users by incorporating Random Convolutional Kernel Transform (ROCKET).
- We demonstrated our proposed pipeline on the Activity Recognition of Nurse Training Activity of the 6th ABC Challenge dataset, and we

found that multi-user data with the feature suggested by Generative AI obtained the best accuracy.

2 Related work

2.1 Nursing activity recognition using deep learning

Computer vision and deep learning were widely used for human activity recognition [9]. Recently, researchers explored the computer vision and deep learning technique for nursing activity recognition [10, 11]. For example, Rasul et al. [10] employed Convolutional Neural Networks (CNNs) to recognize various nursing activities in intensive care units (ICUs) using wearable sensor data. It highlights the potential of CNN for analyzing nurse movements and improving workflow efficiency. Kadir et al. [12] delve into using Gated Recurrent Units (GRUs) for recognizing nurse care activities. It demonstrates the effectiveness of GRUs in handling sequential data, like nurse movements during care procedures.

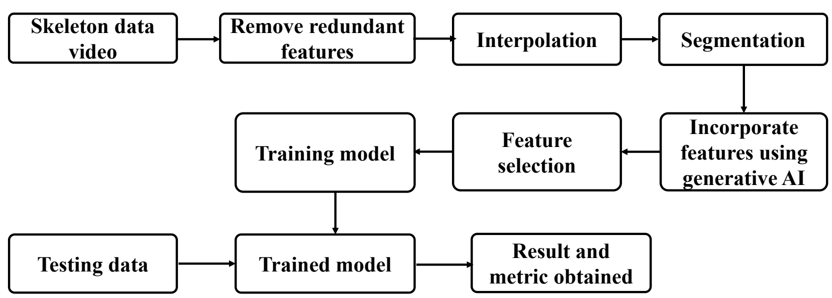

Figure 1: Block diagram of our proposed work.

In addition, Ascioglu et al. [11] explores utilizing deep learning with various sensor modalities, like accelerometers and gyroscopes to recognize human activities. While not specific to nursing, it showcases the versatility of deep learning for activity recognition with different data sources. Gu et al. [13] offers a comprehensive overview of deep learning techniques used for Human Activity Recognition (HAR). It provides valuable insights into various deep-learning architectures that could be adapted for pose estimation and activity recognition during nursing tasks. Werner et al. [14] explore machine learning for airway management. It utilizes machine learning algorithms to analyze audio data from the ventilator and identify potential endotracheal tube dislodgement. This concept demonstrates the feasibility of machine learning for tasks surrounding nurse interventions. Kennedy-Metz et al. [15] explore using computer vision for activity recognition in operating rooms. It demonstrates the feasibility of video analysis for identifying specific surgical procedures. This concept can be adapted to analyze nurse movements during various tasks, differentiating between different care activities. Yoon et.al. [16] discusses the potential of machine learning and artificial intelligence (AI) in critical care medicine. It highlights areas like patient monitoring and risk prediction, which could be indirectly relevant.

2.2 Nursing activity recognition using generative AI

Recently, researchers have explored the use of generative AI for activity recognition, demonstrating the potential for future applications in nursing. Radford et al. [17] delves into using GANs for unsupervised activity discovery. It demonstrates how GANs can identify patterns in unlabeled data, which could be valuable for discovering new or unseen nursing activities from real-world video recordings. Ghosheh et al. [18] provides a broader overview of various applications of generative AI in healthcare. While not explicitly mentioning nursing activity recognition, it highlights the potential of GANs for tasks like data augmentation and anomaly detection, which could be adapted for nursing activity recognition in the future.

3 Methodology

Our proposed working pipeline overview is shown in Fig. 1. The skeleton data points were initially extracted using YOLOv7 from the video dataset. During preprocessing, we analyzed nursing activity key points and identified knee and ankle key points irrelevant to our activity recognition. This finding was based on two factors: firstly, our study focuses on upper body and arm movements for nursing activity recognition, with lower body key points offering limited distinction. Removing irrelevant key points simplifies the feature space, enhancing model training efficiency. Afterwards, we performed interpolation and segmentation. Next, we scaled the data and performed feature extraction with the assistance of Generative AI, e.g., GPT-4, Gemini, and Llama 7B, contributing to improved results. Additionally, Generative AI also suggested different machine learning (ML) methods: Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gaussian Naive Bayes, Decision Tree (DT), Adaboost, and XGBoost to train our pipeline to nursing activity recognition.

In our proposed pipeline, we combined all samples, resulting in a very large dataset, X. We downsampled it using stratified and reservoir sampling procedures and later trained our model with the downsampling and stratified dataset. Finally, the best model was selected based on the validation set and average F1-score. The overview of the training pipeline for our proposed method is shown in Fig. 2.

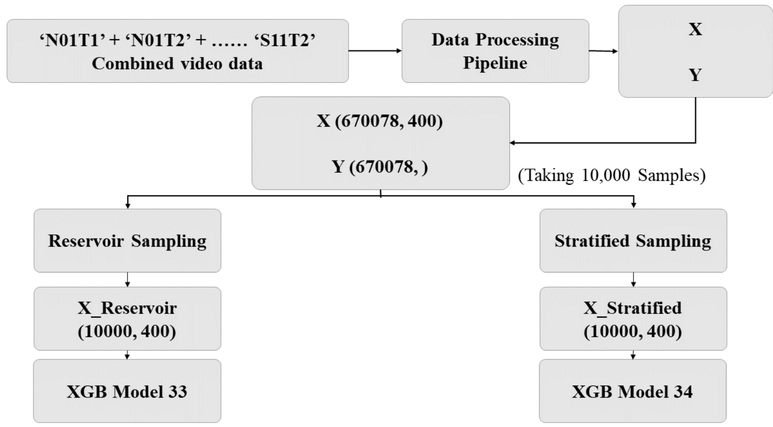

Figure 2: Training pipeline of our proposed method.

4 Experiment

4.1 Dataset description

We employed the Activity Recognition of Nurse Training Activity of 6th ABC Challenge dataset [19, 20]. The dataset consists of activities of nurses and nursing students performing Endotracheal Suctioning (ES) procedures. A total of ten experienced nurses, each with over three years of clinical suctioning experience, and twelve nursing students participated, performing the ES procedure twice, resulting in a dataset divided into a training set of 32 videos and a submission set of 12 videos. Each of the keypoint and annotation data has been labeled as N01T1.csv, N01T2.csv, S01T1.csv and so on where 'N' and 'S' indicate Nurse and Student respectively with N01 to N12 referring to 12 nurses and T1 and T2 indicating the two trials of each user. It encompasses two main data types: videos and skeleton key points. The videos, intended solely for training purposes, were recorded from the front side of the nurse, beyond the patient mannequin, while the pose skeletons, extracted using YOLOv7, were provided for both training and testing. The dataset includes annotations specifying the start and stop times of activities along with their corresponding IDs, and in total, nine actions were performed in the dataset as shown in Table 1. We employed the skeleton data point for our study.

4.1.1 Preprocessing

Preprocessing begins with feature removal, targeting redundant key points such as confidence scores and knee and ankle positions, which are considered irrelevant for the analysis since these parts don't appear frequently in the video. Subsequently, interpolation addresses missing data points in the key points by employing linear interpolation, thereby ensuring the dataset's completeness. The overview of the data preprocessing is shown

Table 1: Activities in Endotracheal Suctioning and their IDs

Activity Name	Activity ID
Catheter Preparation	0
Temporal Removal of an Artificial Airway	1
Suctioning Phlegm	2
Refitting the Artificial Airway	3
Catheter Disinfection	4
Discarding Gloves	5
Positioning	6
Auscultation	7
Others	8

in Fig. 3. These preprocessing steps play a crucial role in refining the dataset and optimizing it for subsequent feature engineering and modeling tasks. By eliminating irrelevant features and filling in missing values, the dataset is primed for more accurate and meaningful analysis, laying the foundation for robust model development and evaluation. Afterwards, we segmented our data with a window size of 3 seconds and an overlap rate of 0.6 (these values were experimentally obtained for best results).

4.1.2 Feature extraction

Firstly, we computed mean, standard deviation, min, max, variance, median, and the sum of pose skeleton data as features. Then, we explore the generative AI to get suggestions for spatial and temporal features as follows to improve the accuracy.

- Temporal features: Velocity and Acceleration are used as a temporal feature. We calculated the first and second derivatives of the skeleton data to obtain velocity and acceleration. This can provide insights into the speed and smoothness of movements.
- Spatial Features: Distances and Angles for skeleton data points are used as spatial features. We computed distances between specific key points or angles formed by keypoint triplets. This can capture spatial relationships and geometric properties of the pose. The procedure for computing the angle is below:
  1. Left Elbow-Shoulder-Hip Angle:
     Represents the angle formed between the left elbow, shoulder, and hip.
  2. Right Elbow-Shoulder-Hip Angle:
     Represents the angle formed between the right elbow, shoulder, and hip.
  3. Left Wrist-Elbow-Shoulder Angle:
     Represents the angle formed between the left wrist, elbow, and shoulder.
  4. Right Wrist-Elbow-Shoulder Angle:

Represents the angle formed between the right wrist, elbow, and shoulder.

5. Right Elbow-Shoulder Angle:
   Represents the angle formed between the right elbow and shoulder.
6. Left Elbow-Shoulder Angle:
   Represents the angle formed between the left elbow and shoulder.
7. Left Elbow-Wrist-Shoulder Angle:
   Represents the angle formed between the left elbow, wrist, and shoulder.
8. Right Elbow-Wrist-Shoulder Angle:
   Represents the angle formed between the right elbow, wrist, and shoulder.
9. Left Shoulder-Wrist-Hip Angle:
   Represents the angle formed between the left shoulder, wrist, and hip.
10. Right Shoulder-Wrist-Hip Angle:
    Represents the angle formed between the right shoulder, wrist, and hip.

The features mentioned above represent the final selection after incorporating suggestions from various Large Language Models (LLMs). Notably, the utilization of different LLMs yielded distinct sets of features, highlighting the variability in feature extraction methodologies for this nursing activity recognition.

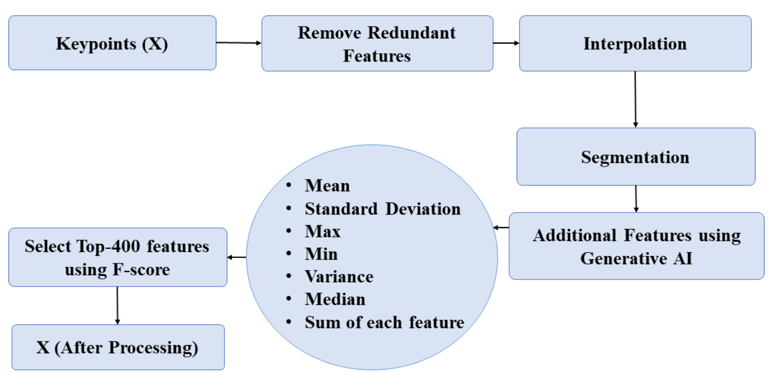

Figure 3: Data preprocessing pipeline.

4.1.3 Scaling and segmentation

Scaling was implemented to standardize the data, ensuring that all features were on a comparable scale for modeling. This step is essential for many machine learning algorithms, as it helps to improve performance accuracy [21]. Additionally, segmentation was applied using a 3-second window size and an overlapping rate of 0.6, allowing the data to be divided into smaller segments for analysis. Segmentation aids in capturing temporal patterns and relationships within the data, which is crucial for accurately identifying and classifying activities.

4.1.4 Sampling and features selection

For our proposed method, we combined all the skeleton data from all 32 videos to enhance the range of activity patterns fed into our model. To address the large size of the combined skeleton data, a representative subset of 10,000 samples was created using stratified and reservoir sampling.

4.1.5 Feature vector generation

We combined keypoint and annotation datasets from multiple videos, resulting in a comprehensive dataset comprising 923,474 samples. From this, we utilized reservoir and stratified sampling techniques to extract a subset of 10,000 samples, with reservoir sampling yielding superior results [22, 23].

Table 2: Performance of the proposed model with different machine learning methods before incorporating the features suggested by Generative AI.

Model Name	**Accuracy** [%]	**F1-Score** [%]
Random Forest	62.3	47.0
SVM	60.0	45.4
K-Nearest Neighbors	58.8	42.6
Gaussian Naive Bayes	55.2	40.0
Decision Trees	63.1	48.7
Extra Trees	61.3	46.0
Adaboost	59.0	44.5
XGBoost	70.0	52.9

4.2 Result and discussion

The performance for our proposed method is shown in Table 2 using different ML methods without incorporating the features suggested by Generative AI and Table 3, after adding the features suggested by Generative AI. We can observe that, after incorporating features suggested by AI, the accuracy and F1-scores of the models significantly improved. This demonstrates the effectiveness of incorporating generative AI techniques to enhance model performance. Furthermore, XGboost obtained the best method. XGBoost is known for its robustness and efficiency in handling

Table 3: Performance of our proposed framework with different machine learning methods after incorporating the features suggested by Generative AI.

Model Name	**Accuracy** [%]	**F1-Score** [%]
Random Forest	80.5	70.7
SVM	72.1	63.4
K-Nearest Neighbors	75.7	68.8
Gaussian Naive Bayes	70.1	60.0
Decision Trees	78.6	68.3
Extra Trees	65.0	55.2
Adaboost	65.9	55.3
XGBoost	97.0	97.0

complex datasets with high dimensionality and varied feature distributions [24]. It was further optimized using grid search to find the optimal hyperparameter configuration. In addition, the result for the model (i.e., XGboost) using different features suggested by Generative AI is shown in Table 4. The Confusion Matrix of our proposed method after using the feature of Generative AI with XGBoost. is shown Fig. 4.

The integration of multiple user data, the rigorous sampling procedures and the utilization of features suggested by generative AI collectively contributed to notable enhancements in the model's performance for nursing activity recognition. By incorporating data from multiple users, our model gained access to a wider range of activity patterns, improving its ability to generalize and classify nursing activities accurately. Moreover, using features extracted through generative AI methods provided valuable insights into the underlying patterns and nuances of the pose skeleton data, further enhancing the model's discriminative power.

4.3 Ablation studies

We conducted thorough analyses of our proposed framework using two other input data pipelines. Firstly, we considered using single-user skeleton data and iterated over each user sample. The evaluation result for this method (i.e. single-user data) using different ML classification methods is shown in Table 5 and Table 6 respectively with and without added feature with the help of Generative AI. Here we also observe a similar tendency that, the accuracy is improved when we added the feature with the help of Generative AI, and XGBoost obtained the best-performing benchmark. However, the accuracy was degraded when compared with our proposed method, where we employed multiple user data. This observation underscores the significance of incorporating diverse user data for model training. Additionally, the inclusion of more user data not only enhances model performance but also provides a deeper understanding of the dataset, thereby contributing to improved overall performance.

Furthermore, we incorporate the Random Convolutional Kernel Transform (ROCKET) to capture complex temporal patterns and efficiently encode information for downstream analysis tasks like activity recognition

Table 4: Performance of different feature sets suggested by different Generative AI using XGboost for our proposed method.

LLM	**Features suggested**	**Accuracy [%]**	**F1-Score [%]**
Meta-Llama-3-70B-Instruct	Joint angles, Joint distances, Body part orientation, Skeletal pose similarity metrics Velocity and acceleration, Jerk, Movement direction, Frequency analysis Distance (nurse-patient), Distance (body parts) Mean, Median, Standard deviation, Variance and co-variance	95.2	90.0
Claude-2.1	Velocity and acceleration, Relative motion, Periodicity and rhythm of keypoint movements Angles between different body parts, Distances between keypoints, Body part ratios Relative positions of keypoints, Skeleton structure and connectivity between keypoints Trajectory and path of keypoints, Direction of motion	85.5	75.7
Microsoft Copilot (GPT-4)	Joint Angles, Relative Positions, Velocity and Acceleration, Temporal Features, Symmetry, Body Part Interaction, Segment Duration, Segment Velocity and Acceleration, Segmental Symmetry, Distance from Camera, Relative Distance, Hand Movements, Spatial Trajectories, Statistical Features, Frequency Domain Features	74.0	68.3
Mixtral-8x11b-instruct	Keypoint Velocity, Keypoint Acceleration, Keypoint Angles, Keypoint Distances, Frame Difference, Optical Flow, Histogram of Oriented Gradients, Histogram of Optical Flow, Sequence Length	83.2	79.0
Finally Selected Features based on Performance	Velocity and Acceleration - Computed Angles: 1) Left Elbow-Shoulder-Hip Angle 2) Right Elbow-Shoulder-Hip Angle 3) Left Wrist-Elbow-Shoulder Angle 4) Right Wrist-Elbow-Shoulder Angle 5) Right Elbow-Shoulder Angle 6) Left Elbow-Shoulder Angle 7) Left Elbow-Wrist-Shoulder Angle 8) Right Elbow-Wrist-Shoulder Angle 9) Left Shoulder-Wrist-Hip Angle 10) Right Shoulder-Wrist-Hip Angle - Statistical Features: 11) Mean, 12) Variance, 13) Maximum, 14) Minimum, and 15) Standard Deviation	97.0	97.0

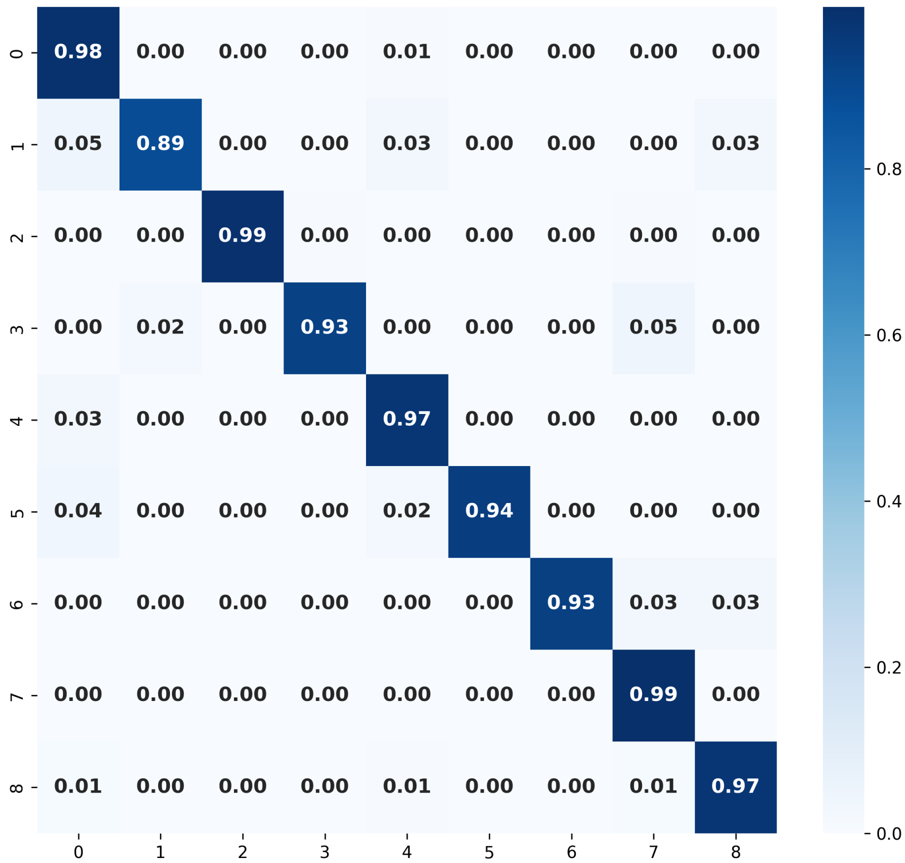

Figure 4: Confusion Matrix of our proposed method after using the feature of Generative AI with XGBoost.

Table 5: Performance of our proposed framework using single-user data with different ML methods before incorporating the features suggested by Generative AI. Train and test split was performed with single-user data.

Model Name	**Accuracy** [%]	**F1-Score** [%]
Random Forest	40.5	21.8
SVM	36.7	19.7
K-Nearest Neighbors	42.3	16.6
Gaussian Naive Bayes	41.3	18.9
Decision Trees	40.9	21.8
Extra Trees	38.7	19.1
Adaboost	38.6	16.9
XGBoost	54.0	30.6

Table 6: Performance of our proposed framework using single-user data with different ML methods after incorporating the features suggested by Generative AI. Train and test split was performed with single-user data.

Model Name	**Accuracy** [%]	**F1-Score** [%]
Random Forest	48.8	24.7
SVM	31.0	8.4
K-Nearest Neighbors	36.5	15.4
Gaussian Naive Bayes	22.8	4.9
Decision Trees	46.7	27.6
Extra Trees	12.5	2.4
Adaboost	25.1	13.2
XGBoost	62.2	26.4

[25]. The original feature is the same as our proposed method but processed with ROCKET. The evaluation results are shown in Table 7 and Table 8, without and with added features (respectively) with the help of Generative AI. We can observe that the accuracy is improved from single user data; however, it degrades from our proposed method. We think that the reduction in performance may be attributed to the complexity introduced by the ROCKET features, which might not have been effectively captured by the models.

Table 7: Performance of our proposed framework with different Machine learning methods before incorporating the features suggested by Generative AI. Here, ROCKET is employed to capture temporal patterns.

Model Name	**Accuracy** [%]	**F1-Score** [%]
Random Forest	46.1	38.0
SVM	31.0	25.8
K-Nearest Neighbors	21.0	15.9
Gaussian Naive Bayes	53.8	45.1
Decision Trees	61.7	54.8
Extra Trees	62.0	54.9
Adaboost	45.1	40.5
XGBoost	76.4	68.1

5 Conclusion and future work

This research endeavor has effectively developed a framework to recognize nursing activities utilizing skeleton data. Despite obstacles such as differentiating similar activities, background interference, missing key points, and nurses occasionally being out of frame, we developed a robust model capable of classifying nine suctioning activities from video recordings. Employing pose estimation and feature engineering techniques, includ-

Table 8: Performance of our proposed framework with different Machine learning methods after incorporating the features suggested by Generative AI. Here, ROCKET is employed to capture temporal patterns.

Model Name	**Accuracy [%]**	**F1-Score [%]**
Random Forest	58.0	41.6
SVM	31.0	29.7
K-Nearest Neighbors	45.0	35.8
Gaussian Naive Bayes	68.7	65.0
Decision Trees	81.2	68.7
Extra Trees	73.2	61.9
Adaboost	65.2	61.0
XGBoost	91.0	89.0

ing segmentation, feature extraction, and statistical analysis, proved vital in capturing relevant information from the pose data. Furthermore, by incorporating data from multiple users and integrating features with the help of generative AI, the method significantly enhanced the performance, showcasing the potential of utilizing pose skeleton data for nursing activity recognition. Our findings highlight the promise of machine learning approaches in automating healthcare tasks, with implications for enhancing patient care and safety in clinical settings.

Future research may explore additional data augmentation techniques and refine the model to address diverse clinical scenarios and environments. Moreover, convolutional neural networks (CNNs), such as ResNet and XResNet1D, can be explored for further improvement.

References

[1] Caroline J Wood. Endotracheal suctioning: a literature review. *Intensive and Critical Care Nursing*, 14(3):124–136, 1998.

[2] Sameer A Alkubati, Khaled M Al-Sayaghi, Gamil G Alrubaiee, Mokhtar Abdu Hamid, Khalil A Saleh, Talal Al-Qalah, and Ahmad K Al-Sadi. Adherence of critical care nurses to endotracheal suctioning guidelines: a cross-sectional study. *BMC nursing*, 21(1):312, 2022.

[3] 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, 2024.

[4] Sozo Inoue, Naonori Ueda, Yasunobu Nohara, and Naoki Nakashima. Understanding nursing activities with long-term mobile activity recognition with big dataset. In *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications*, volume 2016, pages 1–11. The ISCIE Symposium on Stochastic Systems Theory and Its Applications, 2016.

[5] Momal Ijaz, Renato Diaz, and Chen Chen. Multimodal transformer for nursing activity recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2065–2074, 2022.

[6] Yoshihiro Takebe, Masako Kanai-Pak, Noriaki Kuwahara, Jukai Maeda, Miwa Hirata, Yasuko Kitajima, and Jun Ota. Recognition of nursing activity with accelerometers and rfid. *Kybernetes*, 42(7):1059–1071, 2013.

[7] Kaveh Momen and Geoff R Fernie. Nursing activity recognition using an inexpensive game controller: An application to infection control. *Technology and Health Care*, 18(6):393–408, 2010.

[8] Orhan Konak, Valentin Döring, Tobias Fiedler, Lucas Liebe, Leander Masopust, Kirill Postnov, Franz Sauerwald, Felix Treykorn, Alexander Wischmann, Stefan Kalabakov, et al. Sonar, a nursing activity dataset with inertial sensors. *Scientific data*, 10(1):727, 2023.

[9] Charmi Jobanputra, Jatna Bavishi, and Nishant Doshi. Human activity recognition: A survey. *Procedia Computer Science*, 155:698–703, 2019.

[10] Md Golam Rasul, Mashrur Hossain Khan, and Lutfun Nahar Lota. Nurse care activity recognition based on convolution neural network for accelerometer 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*, pages 425–430, 2020.

[11] Gokmen Ascioglu and Yavuz Senol. Activity recognition using different sensor modalities and deep learning. *Applied Sciences*, 13(19):10931, 2023.

[12] Md Eusha Kadir, Pritom Saha Akash, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. Can a simple approach identify complex nurse care activity? 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*, pages 736–740, 2019.

[13] Fuqiang Gu, Mu-Huan Chung, Mark Chignell, Shahrokh Valaee, Baoding Zhou, and Xue Liu. A survey on deep learning for human activity recognition. *ACM Computing Surveys (CSUR)*, 54(8):1–34, 2021.

[14] Sandra L Werner, Charles E Smith, Jessica R Goldstein, Robert A Jones, and Rita K Cydulka. Pilot study to evaluate the accuracy of ultrasonography in confirming endotracheal tube placement. *Annals of emergency medicine*, 49(1):75–80, 2007.

[15] Lauren R Kennedy-Metz, Pietro Mascagni, Antonio Torralba, Roger D Dias, Pietro Perona, Julie A Shah, Nicolas Padoy, and Marco A Zenati. Computer vision in the operating room: Opportunities and caveats. *IEEE transactions on medical robotics and bionics*, 3(1):2–10, 2020.

[16] Joo Heung Yoon, Michael R Pinsky, and Gilles Clermont. Artificial intelligence in critical care medicine. *Annual Update in Intensive Care and Emergency Medicine 2022*, pages 353–367, 2022.

[17] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

[18] Ghadeer Ghosheh, Jin Li, and Tingting Zhu. A review of generative adversarial networks for electronic health records: applications, evaluation measures and data sources. *arXiv preprint arXiv:2203.07018*, 2022.

[19] ABC Research. ABC Research Dataset 2024. https://abc-research.github.io/challenge2024/data/, 2024. [Online; accessed 27-April-2024].

[20] 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.

[21] Andreas Maier, Christopher Syben, Tobias Lasser, and Christian Riess. A gentle introduction to deep learning in medical image processing. *Zeitschrift für Medizinische Physik*, 29(2):86–101, 2019.

[22] Kim-Hung Li. Reservoir-sampling algorithms of time complexity o (n (1+ log (n/n))). *ACM Transactions on Mathematical Software (TOMS)*, 20(4):481–493, 1994.

[23] Jeyabharathy Sadaiyandi, Padmapriya Arumugam, Arun Kumar Sangaiah, and Chao Zhang. Stratified sampling-based deep learning approach to increase prediction accuracy of unbalanced dataset. *Electronics*, 12(21):4423, 2023.

[24] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.

[25] Angus Dempster, François Petitjean, and Geoffrey I Webb. Rocket: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, 34(5):1454–1495, 2020.