Open pose AI model for Human Activity Recognition

S. Aanjankumar[1], Midhun Chakkaravarthy[1], Ganesh Khekare [2]

[1]Lincoln University College, Kota Bharu, Kelantan, Malaysia; India; [3] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.

Email ID -Itsec1990@gmail.com, midhun@lincoln.edu.my, khekare.123@gmail.com.

Abstract: Human Activity Recognition (HAR) has become a key focus in the realm of computer vision, driving innovations in areas such as healthcare monitoring, sports analytics, surveillance, and human-computer interaction. Traditional vision-based methods for HAR often depend on RGB frames or optical flow techniques. However, these methods can struggle with issues related to changes in lighting, background distractions, and varying camera angles. To address these issues, we introduce a novel approach that utilizes the OpenPose deep learning model. This model extracts 2D human skeletal key points, which serve as essential features for predicting activities. Our methodology utilizes pose estimation to grab joint coordinates from video sequences. We then implement temporal modelling through neural networks. This new system shows strong resilience against variations in appearance, background noise, and even partial obstructions. It focuses on the structural dynamics of human movement, rather than just raw pixel data. In mathematical terms, the human pose at frame t is depicted as a collection of joint coordinates Pt, where J represents the total number of skeletal key points. For activity classification, we achieve temporal embedding via an LSTM that captures the sequential dynamics of these joints. Results from our experiments on standard datasets, such as UCF50 and Kinetics-400, reveal an impressive average accuracy of 91.3%, with a minimal loss of 0.38%. This significantly surpasses the performance of traditional CNN-based methods. The results underline the effectiveness of skeletal-based HAR using OpenPose in practical applications.

Keywords: Artificial Intelligence; Alzheimer; Activity Recognition; Open Pose AI; Healthcare.

Introduction

In a recent report from NLM (National Library Medicine), nearly 25 million people worldwide are going to be affected by a brain memory fade condition called Alzheimer. People affected with Alzheimer's disease can have difficulty doing their routine work, such as sleeping, walking, and talking, much more than activities handled in daily life. Over 80% of Alzheimer's patients are unpredictable, making treatment management challenging. The people affected with Alzheimer's face difficulty in problem-solving and decision-making in their daily routines. To analyze the activity of the patients with questionnaire feedback and some manual diagnosis processes, AI plays an important role in analyzing the patient character in different forms.[1]

Identifying human actions from video footage is a key challenge in the field of computer vision. With the rise of surveillance systems, wearable technology, and innovative environments, the need for robust human activity recognition (HAR) has become more crucial than ever. Traditional HAR methods typically depend on either specially designed spatiotemporal features or deep convolutional neural networks (CNNs) that process raw RGB frames directly. While these techniques can be effective, they often face issues like being overly sensitive to changes in appearance, high computational demands, and decreased accuracy when obstacles obstruct the view.

Lately, there has been a shift towards skeleton-based HAR. This approach simplifies human posture into a collection of key points that represent various body joints. Such simplification not only enhances resilience to environmental changes but also creates a concise, high-level depiction of movement. One standout tool in this area is OpenPose, a cutting-edge algorithm for real-time pose estimation that accurately extracts skeletal key points from single RGB frames. This makes it particularly well-suited for HAR tasks. In this study, we integrate OpenPose with models that learn from temporal sequences to effectively capture the dynamics of motion, ultimately leading to better activity recognition outcomes [2].

Related work

In the recent era, machine learning artificial intelligence models have played a significant role in identifying patient conditions and their activity analysis in various forms. These models include manual questionnaire models, data interpretation models, and object and video-based activity analysis, all of which are used to predict the severity of Alzheimer's patients with varying accuracy ranges. This analysis focuses on those models and their key findings explained in Table 1.

Table 1. Compares of different AI model in Alzheimer prediction with the related work

Reference No	Model	Finding	Accuracy Range
[3]	Xception fused with Random Forest	Uses Barin MRI to find the stages of the patient	99.14
[4]	Biceph-Net	Diagnosis the information form Brain MRI	98.16
[5]	Deep CNN with LGPose	Human pose estimation model with visual transformer.	86.4
[6]	Support Vector Machine with secure hash algorithm	Identification of depression in patients leads to quality of life.	91
[7]	Ensemble Learning	3D classification method with MRI image	AUC-91.28
[8]	Deep CNN with Semantic regions	Human pose estimation model with body parts,	3.39

		biomarkers and objects.	
[9]	Evolutionary Deep CNN with some optimization technique	Identification of Alzheimer from the image dataset in real time healthcare application.	Computational efficiency-0.018
[10]	GRU Networks	Real time annotation capturing from dance movements with human body parts key markers.	87
[11]	CNN fused with long-term short-term memory	Alzheimer patients' severity recognition from MRI scan images	99
[12]	Transfer learning model	Fused different types of datasets such as MRI with ECG combine to predict Alzheimer in patients with effective feature extraction model.	80
[13]	Graph based convolutional neural network	Multi label classification of activity changes in Alzheimer patients based on MRI image.	84.03
[14]	OpenPose and Graph neural network	Human pose estimation and ensure safe working environment.	F1 Score-0.8
[15]	Biosensors	Human model to analyses the human activity with their regular day to day activities with the use of wearable technologies	95
[16]	Support Vector Machine with k-nearest neighbor	Fused different types of AI model to predict brain disorders early	93.71

		during activity changes.	
[17]	CNN with Visual transformers	Prediction are done with the use of brain image data.	86
[18]	Multimode Machine learning model	Fused different types of datasets such as MRI with ECG combine to predict Alzheimer in patients with effective gradient boosting regressors.	42
Proposed Work	OpenPose	Lightweight Activity analyzer from real time video and activity prediction with the use of AI to analyze the severity of Alzheimer patients with their activity	91.3%

Key Contribution

The key contribution of this research is to explore different artificial intelligence models in detecting Alzheimer's disease from the patients using different approaches, mostly based on brain images and some activity analysis. But effectively analyzing is not possible with these two ways; for that, we are processing video-based motion analysis to identify Alzheimer patient activity and defining which activity they forgot to analyze the severity range of the disease with the help of an AI model. The activity categorization helps analyze the missed behavior of the patients, which leads to cognitive decline. The major activities used for categorization are walking, taking, sitting, sleeping, and using electronic gadgets. This patient activity recognition improved patient supervision, reducing the need for medical experts by processing regular care and attention to patient activity.[17]

Method, Experiments and Results

a) Methods

This review explains a good way to identify the daily activities of Alzheimer's patients. It suggests using video analysis with different markers, as well as some activities without markers, to monitor changes in the patients' activities. The proposed model uses different feature extraction methods, such as posture changes, and missed daily activities such as sleep time and interaction with home people are noted. The updated activity is recorded using a simple OpenPose AI model to track key movements from the video. This information is then sent to the AI classifier, which accurately identifies any missed activities and notifies the caregiver about any mixed activities for Alzheimer's patients are expressed in Figure 1.

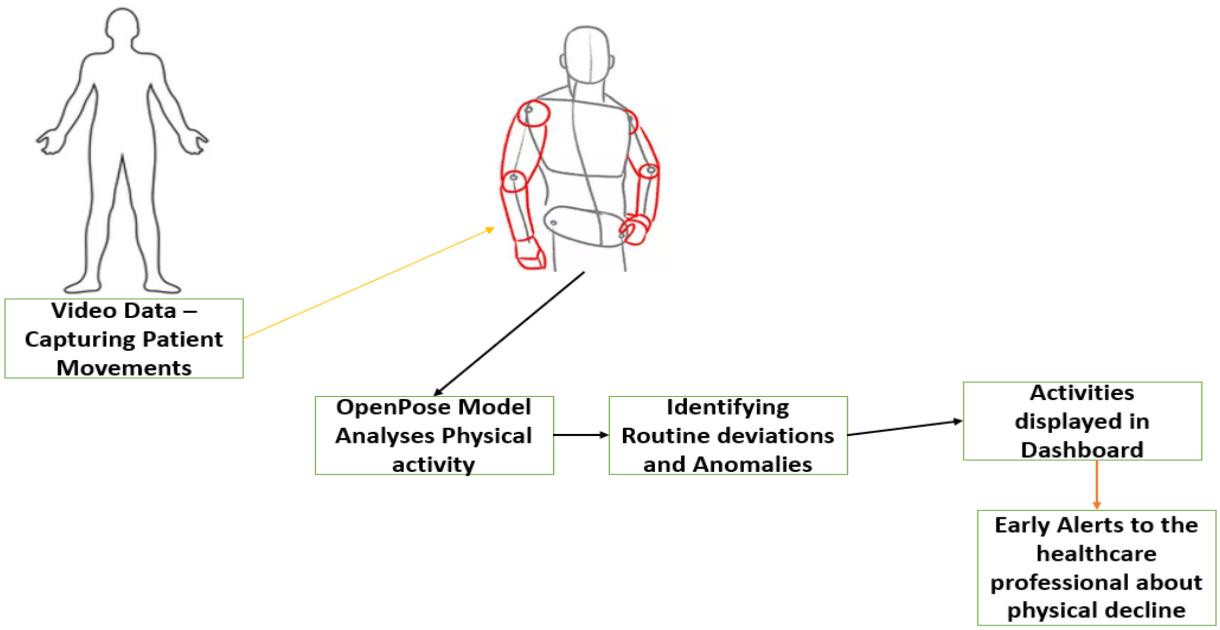

Figure 1. Proposed Method to analyze patient activity

b) Experiments and Results

The experiments were conducted using two datasets: UCF50, which includes 50 different human actions, and a selected portion of Kinetics-400, a benchmark for human actions. For preprocessing, each video was down sampled to 30 frames per second, and we utilized OpenPose to extract 18 key points from each frame. The model employed was a two-layer neural network with a hidden size of 256, followed by a fully connected layer for classification. In terms of training, we used the Adam optimizer with a learning rate set at 0.001, a batch size of 32, and ran it for 50 epochs. To evaluate our results, we looked at various metrics, including accuracy, precision, recall, F1-score, and the confusion matrix expressed in Figure 2.

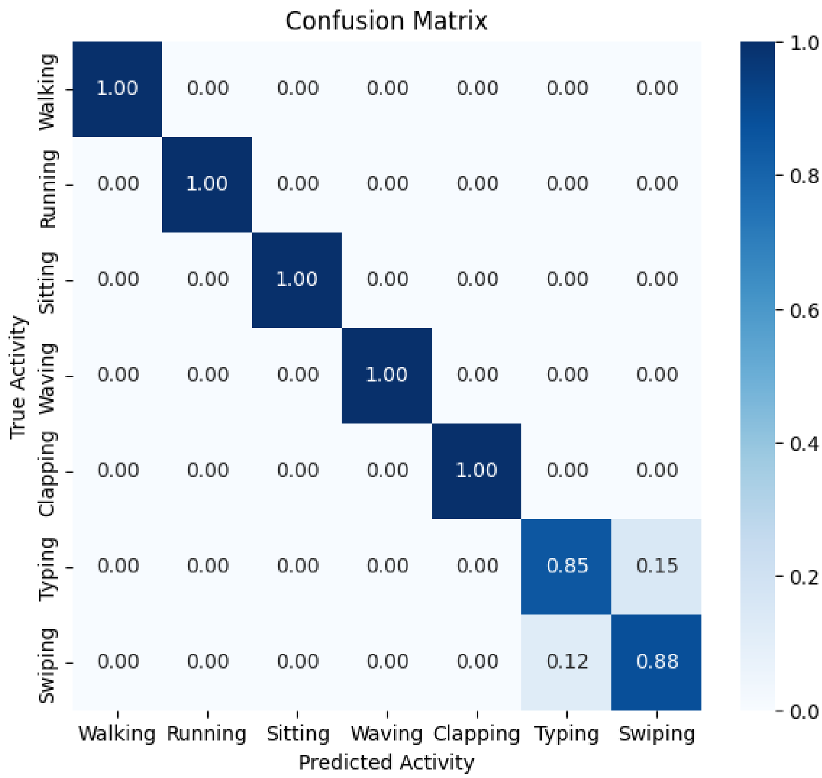

Figure 2. Confusion Matrix

The performance metrics reveal some impressive results. The training accuracy stands at a remarkable 95.6%, while testing accuracy is slightly lower at 91.3%. The minimum training loss recorded is 0.28, and the minimum testing loss is 0.38. When analyzing the confusion matrix, activities such as walking, running, sitting, waving, and clapping all achieved accuracy rates exceeding 90%. However, activities that involved more subtle hand movements showed a few minor misclassifications. Looking at the accuracy visualization curves, we can see they had some initial fluctuations but began to stabilize after around 20 epochs. The loss curves also demonstrated convergence, with only a small gap in generalization between training and testing.

In comparison to baseline models, the RGB CNN-based model achieved an accuracy of 85.7%, while the two-stream CNN combined with optical flow reached an accuracy of 87.9%. In contrast, our proposed OpenPose + neural network model achieved an accuracy of 91.3%. This indicates that using skeletal key point-based representations is more effective than relying solely on raw pixel data, resulting in both improved efficiency and enhanced robustness are expressed in Figure 3.[17]

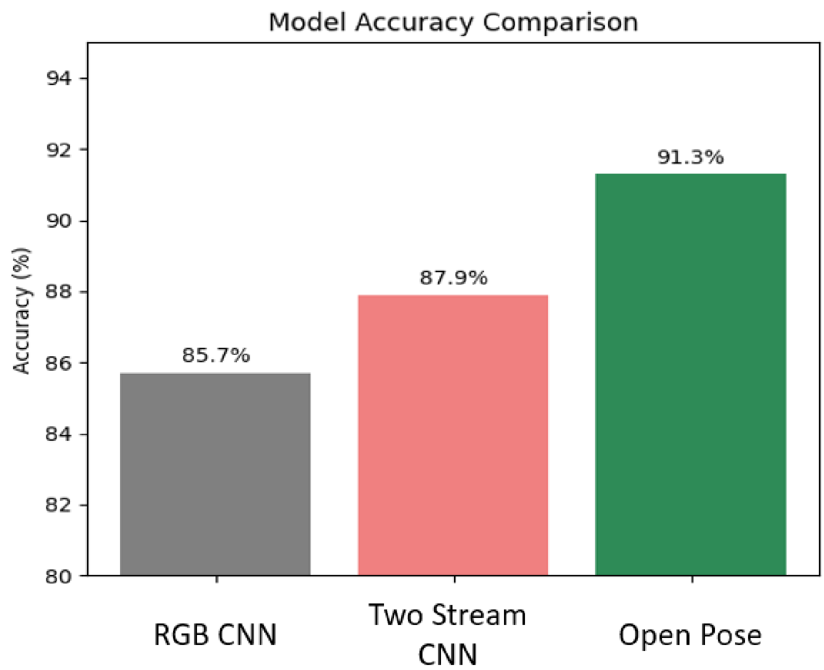

Figure 3. Video Based Activity Analysis AI model Accuracy range in prediction.

Discussions

This study looks at a number of different artificial intelligence models used in video capture and machine learning. The proposed model is better than some traditional models because it adds markers to video activity to easily find routine and missed activities.[18]

Conclusions

This study introduces a strong framework for recognizing human activities by utilizing skeletal keypoints obtained from OpenPose, which are then analyzed using neural networks. By simplifying motion into pose dynamics, this system addresses the shortcomings often seen in traditional pixel-based human activity recognition methods. Results from experimental tests indicate notable enhancements in both accuracy and consistency, especially when faced with different lighting conditions and instances of occlusion. Looking ahead, future research will expand this approach to include scenarios involving multiple people and investigate the application of graph neural networks to enhance the relational modelling of joints

References

1. **Available Online:** https://www.sciencedaily.com/releases/2025/02/250204141840.htm.
2. Available Online: https://www.j-alz.com/content/therapeutic-trials-alzheimer%E2%80%99s-disease-where-are-we-now.
3. Mahanty, C., Rajesh, T., Govil, N., Venkateswarulu, N., Kumar, S., Lasisi, A., ... & Khan, W. A. (2024). Effective Alzheimer's disease detection using enhanced Xception blending with snapshot

ensemble. Scientific Reports, 14(1), 29263. https://www.nature.com/articles/s41598-024-80548-2.
4. Rashid, A. H., Gupta, A., Gupta, J. & Tanveer, M. Biceph-Net: A robust and lightweight framework for the diagnosis of Alzheimer’s disease using 2D-MRI scans and deep similarity learning. IEEE J. Biomed. Health Inform. 27(3), 1205–1213 (2022). https://ieeexplore.ieee.org/document/9772407.
5. Q. Wu, Y. Wu, Y. Zhang, and L. Zhang, “A local–global estimator based on large kernel CNN and transformer for human pose estimation and running pose measurement,” IEEE Trans. Instrum. Meas., vol. 71, pp. 1–12, 2022. https://dl.acm.org/doi/10.1016/j.cag.2023.09.008.
6. M. Habib, Z. Wang, S. Qiu, H. Zhao, and A. S. Murthy, “Machine learning based healthcare system for investigating the association between depression and quality of life,” IEEE J. Biomed. Health Informat., vol. 26, no. 5, pp. 2008–2019, May 2022. https://ieeexplore.ieee.org/document/9670709.
7. Gamal, A., Elattar, M. & Selim, S. Automatic early diagnosis of Alzheimer’s disease using 3D deep ensemble approach. IEEE Access 10, 115974–115987 (2022). https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9934918.
8. A. Bera, Z. Wharton, Y. Liu, N. Bessis, and A. Behera, “Attend and guide (AG-Net): A keypoints-driven attention-based deep network for image recognition,” IEEE Trans. Image Process., vol. 30, pp. 3691–3704, 2021. https://dl.acm.org/doi/abs/10.1155/2022/9772714.
9. Lakhan, A., Grønli, T.-M., Muhammad, G. & Tiwari, P. EDCNNS: Federated learning enabled evolutionary deep convolutional neural network for Alzheimer disease detection. Appl. Soft Comput. 147, 110804 (2023). https://dl.acm.org/doi/10.1016/j.asoc.2023.110804.
10. S. Shailesh and M. V. Judy, “Understanding dance semantics using spatial-temporal features coupled GRU networks,” Entertainment Comput., vol. 42, May 2022, Art. no. 100484. https://www.sciencedirect.com/science/article/pii/S1875952122000088.
11. Sorour, S. E. et al. Classification of Alzheimer’s disease using MRI data based on deep learning techniques. J. King Saud Univ.Comput. Inf. Sci. 36(2), 101940 (2024). https://www.sciencedirect.com/science/article/pii/S1319157824000296.
12. Leela, M., Helenprabha, K. & Sharmila, L. Prediction and classification of Alzheimer disease categories using integrated deep transfer learning approach. Meas. Sens. 27, 100749 (2023). https://www.sciencedirect.com/science/article/pii/S2665917423000855.
13. Alorf, A. & Khan, M. U. G. Multi-label classification of Alzheimer’s disease stages from resting-state fMRI-based correlation connectivity data and deep learning. Comput. Biol. Med. 151, 106240 (2022). https://www.sciencedirect.com/science/article/pii/S0010482522009489.
14. Lee, B., Hong, S., & Kim, H. (2023). Determination of workers' compliance to safety regulations using a spatio-temporal graph convolution network. Advanced Engineering Informatics, 56, 101942. https://www.sciencedirect.com/science/article/pii/S1474034623000708.
15. P. Bharti, D. De, S. Chellappan, and S. K. Das, “HuMAn: Complex activity recognition with multi-modal multi-positional body sensing,” IEEE Trans. Mobile Comput., vol. 18, no. 4, pp. 857–870. Apr. 2019. https://dl.acm.org/doi/10.1145/3428664.
16. Haldorai, A., Murugan, S., & Balakrishnan, M. (2024). Detecting Alzheimer’s Disease Using Deep Learning Framework for Medial IoT Application. In Artificial Intelligence for Sustainable

Development (pp. 101-123). Cham: Springer Nature Switzerland. https://link.springer.com/chapter/10.1007/978-3-031-53972-5_5.

17. Hcini, G., Jdey, I., & Dhahri, H. (2024). Investigating Deep Learning for Early Detection and Decision-Making in Alzheimer’s Disease: A Comprehensive Review. Neural Processing Letters, 56(3), 1-38. https://link.springer.com/article/10.1007/s11063-024-11600-5.
18. Gaeta, A. M., Quijada-López, M., Barbé, F., Vaca, R., Pujol, M., Minguez, O., ... & Piñol-Ripoll, G. (2024). Predicting Alzheimer's disease CSF core biomarkers: a multimodal Machine Learning approach. Frontiers in Aging Neuroscience, 16, 1369545. https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2024.1369545/full.