International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Neural Networks

Mr. B. Narasimha Rao[1], N. S. Prathap[2], N. Jahnavi[3], N. S. S. Revanya[4], U. Akash[5]

[1] Associate Professor, Dept. Of Computer Science and Engineering [2-5] B. Tech Student, Dept. Of Computer Science and Engineering [1-5] Bonam Venkata Chalamayya Engineering College, Odalarevu

Abstract

In recent years, posture detection has become increasingly important in various fields such as healthcare, ergonomics, and human-computer interaction. This project proposes a **Hybrid Posture Detection Framework** that leverages the strengths of both **traditional Machine Learning (ML) techniques** and **Deep Neural Networks (DNNs)** to achieve high accuracy and robustness in identifying and classifying human postures. The framework utilizes pre-processed image or sensor data to extract relevant features, which are then analyzed through a dual-layered architecture. The first layer employs ML algorithms for preliminary posture classification based on engineered features, while the second layer refines predictions using convolutional neural networks (CNNs) to learn spatial hierarchies and complex patterns.
Keywords Posture Discovery, Machine literacy, Deep Learning, Hybrid Approach, SVM, Logistic Regression, Random Forest, Naive Bayes, LSTM, 1D- CNN, 2D- CNN.

1. Introduction

Human posture detection has gained significant attention in recent years due to its wide range of applications in areas such as healthcare monitoring, smart workplaces, fitness tracking, rehabilitation, and human-computer interaction. Accurate posture analysis can help in identifying musculoskeletal issues, preventing workplace injuries, and enhancing user experience in interactive systems

On the other hand, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable success in image recognition and classification tasks. However, these models typically require large amounts of labeled data and computational resources, which can be a constraint in certain scenarios. To overcome these challenges, this project introduces a hybrid framework that combines the strengths of both machine learning and deep neural networks.

The proposed framework begins with the extraction of key features from images or skeletal data using pose estimation tools. These features are first processed through traditional machine learning algorithms for quick and efficient classification. Subsequently, deep neural networks, specifically CNNs, are used to enhance the model's ability to learn complex spatial patterns and improve accuracy. By integrating these two approaches, the framework aims to achieve a balance between speed, accuracy, and resource efficiency. This hybrid posture detection system is designed to function effectively in real-time applications and adapt to diverse user scenarios. The flexibility and accuracy of the framework make it a valuable tool for enhancing health, safety, and productivity in various domains.

2. Existing System

Current posture detection systems primarily fall into two main categories: traditional machine learning-based approaches and deep learning-based approaches. Traditional machine learning methods typically rely on manually extracted features such as joint angles, limb distances, or body orientation, which are obtained from images or sensor data using tools like Kinect or inertial measurement units (IMUs). These features are then fed into classifiers such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees, or Random Forests to detect and classify postures.

While these systems are relatively lightweight and computationally efficient, their accuracy often depends heavily on the quality of feature extraction and they tend to struggle in complex environments with noise, occlusions, or varying body types. In contrast, deep learning-based systems, especially those using Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have significantly improved the performance of posture detection by automatically learning complex spatial and temporal patterns from raw image or video data

3. Literature Review

The concept of posture recognition has gained momentum in health care, human-computer interaction and AI. There have been numerous attempts to enhance various features in-comprehension systems (e.g., accuracy and efficiency). Below summarizes some of the most significant approaches, original contributions, and highlight the issue of extraction and machine learning:

- Algorithms developed around Machine Learning. Algorithms (SVM, Decision Trees, Random Forests and Naïve Bayes) are routinely employed for posture recognition. The algorithm is applied to the data after some features have been extracted from the data set, in order to recognize the posture. If the algorithm is applied over the raw data directly (i.e., not first optimizing some features) the algorithm is likely to perform poorly in many cases.
- Algorithms based on Deep Learning. Algorithms (e.g., CNN, LSTM) have improved accuracies for posture recognition considerably. These algorithms are quite good at abstract feature learning and even learning to learn; despite this, deep learning algorithms require lots and lots of data and computation, thus the process for real-time posture detection can be problematic.
- Hybrid models. The potential use of hybrid models (depending on the power and strength of classical algorithms built on machine learning and compared the performance of combined deep learning algorithms in the past few years). Hybrid models combine each of the mentioned above, depending on the prior recognition in areas of learning and a machine learning model being selected. Hybrid model advantages consist of better improvement or not.
- Despite their success, deep learning models often require substantial computational power and large annotated datasets, making them less suitable for real-time or resource-constrained environments. To overcome these challenges, recent research has begun exploring hybrid approaches. For instance, Singh et al. (2020) proposed a system that combined feature selection using Principal Component Analysis (PCA) with a CNN classifier to improve both speed and accuracy
- With the emergence of deep learning, posture detection systems began shifting towards data-driven models. OpenPose, developed by Cao et al. (2017), introduced a real-time multi-person pose estimation framework using part affinity fields and CNNs, setting a new benchmark in pose estimation accuracy. MediaPipe by Google further advanced real-time posture detection on mobile devices by combining lightweight neural networks with holistic tracking of body, hand, and face landmarks

4. Proposed System

- Hybrid Architecture Design

The system combines traditional machine learning algorithms with deep learning models.
ML is used for fast, preliminary classification based on extracted features.
Deep learning refines the results by analyzing spatial patterns and context.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

- Input Data Collection

The framework supports real-time video input or pre-recorded datasets.
It uses RGB cameras or depth sensors to capture human posture.
Tools like OpenPose or MediaPipe extract skeletal keypoints from input frames.
These keypoints form the basis of further analysis.

- Feature Extraction Module

Joint coordinates, limb angles, and body symmetry are calculated.
These features are essential for the initial ML classification stage.
Feature normalization and dimensionality reduction techniques are applied.
This step ensures consistent and clean input for the classifiers.

- Traditional ML Classification Layer

Algorithms like SVM, KNN, or Decision Trees are applied to extracted features.
This stage performs quick posture classification with low computational load.
It acts as a filter or pre-processor before deep learning is applied.
The aim is to reduce processing time without sacrificing basic accuracy.

- Deep Learning Refinement Layer

A CNN model is used to further analyze pose features and refine predictions.
This layer detects complex spatial patterns that ML may overlook.
It improves the system's robustness against occlusion and pose variations.
The CNN is trained on annotated posture datasets for high precision.

- Real-Time Processing Capability

The system is optimized for real-time performance with low latency.
Lightweight models and multi-threading are used for faster execution.
The hybrid approach ensures quicker decisions without compromising detail.
Suitable for live applications like workplace monitoring or fitness tracking.

- Modular Framework Design

Each component (feature extraction, ML, DNN) is modular and interchangeable.
This allows for easy upgrades or integration with other systems.
Developers can experiment with different models in each stage.
The architecture is flexible and adaptable to various use cases.

- Accuracy and Efficiency Balance

The hybrid model aims to balance high classification accuracy and speed.
ML handles simple, repetitive postures quickly; DNN focuses on complex ones.
This division of tasks improves overall efficiency.
The system is tested to outperform single-method systems.

- Scalability and Deployment Readiness

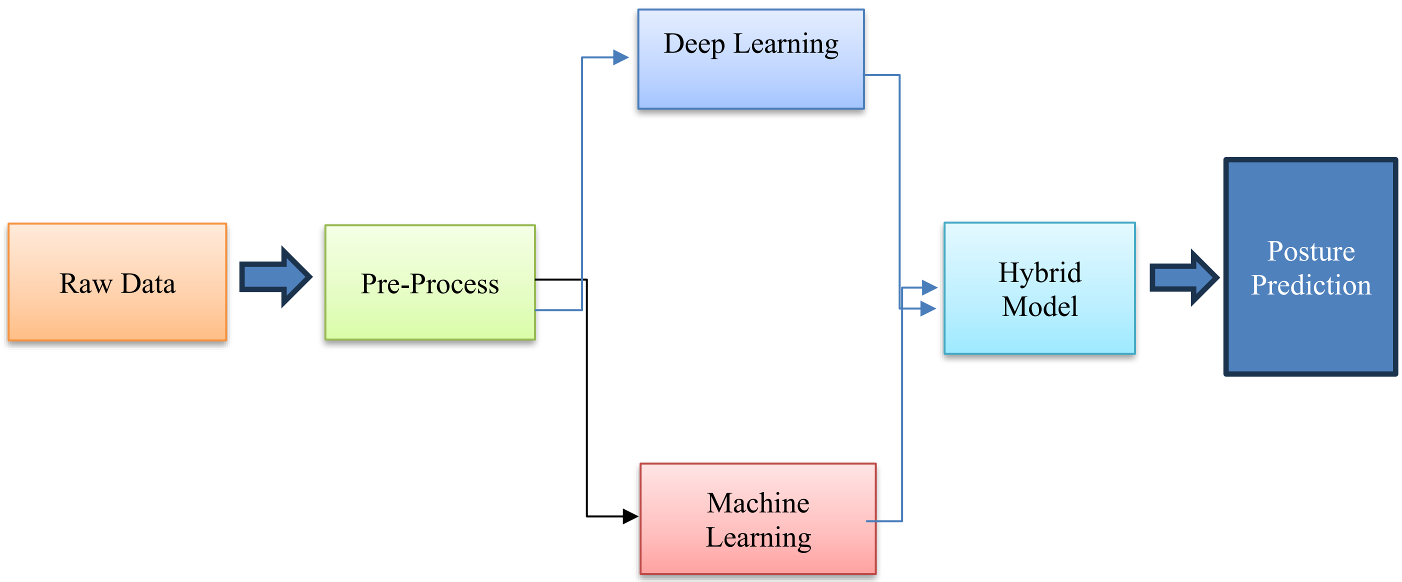
The system is scalable to support multiple users or environments.
It can be deployed on desktops, edge devices, or mobile platforms.
Resource usage is optimized for power-constrained environments.
This makes it ideal for healthcare, ergonomics, and smart surveillance.

System Architecture

- The workflow diagram describes the step-by-step procedure of posture detection. It begins with collecting raw data from sensors, which is then put through a preprocessing phase to clean and normalize it for analysis. After preparation, the data is divided into training and test sets to ensure proper model evaluation.
- Deep learning models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are combined with machine learning methods like Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, and Naive Bayes to efficiently process the data. These models are combined together to extract and analyze features

that lead to posture recognition accuracy.

- The subsequent step is a hybrid integration architecture that integrates the results from both machine learning and deep learning models. Lastly, a meta-learning model combines these outcomes to label the posture into one of three classes: Standing (Class 0), Walking (Class 1), or Sitting (Class 2). This process provides a precise and consistent posture detection system, as graphically illustrated in the diagram.



Dataset and Preprocessing

The GALVIN Sensor Dataset is crucial in the detection of posture, offering the initial data that is used to classify postures into three classes: Standing (Class 0), Walking (Class 1), and Sitting (Class 2). This dataset is used to train machine learning and deep learning models, enabling them to effectively analyze patterns and predict postures with accuracy. A detailed examination of how this is achieved follows:

Features with Descriptions

1. ID: A unique record identifier in each dataset, traceable.
2. Posture Label: Describes the posture of the subject, divided into three types:
   Standing
   Walking
   Sitting
3. Sensor Readings: Consists of numeric values from sensor readings, depicting pressure distribution and movement. The sensor readings are crucial in describing posture-specific traits.

Data Preprocessing

1. Handling Missing Values:

- Any missing data points within the dataset are handled by replacing them with median values for numerical features to prevent loss of data and ensure consistency.
- Features like BMI, if they have null values, are replaced with the median value to preserve the completeness of the dataset.

2. Feature Encoding:

- Categorical attributes like posture labels are encoded into numerical form using one-hot encoding or label encoding.

- Encoding is also performed on other characteristics (if necessary) to pre-process the data for training models.

3. Normalization:

- Numerical features are normalized using normalization methods, i.e., Min-Max scaling, to get the data to a standard range. This is very important to enhance model performance

4.Dataset Balancing

- The data might be plagued by imbalanced class distributions, and this may cause prediction biases. For example, certain postures may have a greater number of records than others. Synthetic Minority Oversampling Technique (SMOTE) is used for this:
- SMOTE creates artificial examples for the minority class by interpolating between current examples.
- This ensures that the dataset becomes evenly balanced, reducing biases during model training.

4.2 Model Selection and Training

- To develop a robust and accurate posture detection system, we adopted a hybrid approach combining traditional machine learning algorithms with deep neural networks (DNNs). The model selection process was guided by the nature of the data, the complexity of pose estimation, and the need for real-time performance. Our framework consists of two primary components: feature extraction using deep learning models and classification using machine learning algorithms.
- **Dataset Preparation**

The training pipeline utilized publicly available human posture datasets, such as the **MPII Human Pose Dataset**, **COCO Keypoints**, and a custom-labeled dataset collected from a posture monitoring environment. Each sample includes annotated skeletal keypoints representing various body joints, extracted using pose estimation frameworks (e.g., OpenPose, MediaPipe, or BlazePose).

The data was preprocessed to normalize joint coordinates, filter noise, and augment samples with common posture variations (e.g., sitting, standing, slouching, leaning). This helped improve model generalization across different body types and environments.

4.2.1 Feature Extraction with Deep Neural Networks

To accurately capture spatial relationships between body joints, we employed a Convolutional Neural Network (CNN)-based architecture:

- **Backbone Network**: A pretrained ResNet-50 or MobileNetV2 was used to extract spatial features from images or pose heatmaps.
- **Pose Estimation Module**: Heatmaps generated by the CNN were passed to a keypoint regression layer to estimate joint coordinates.
- **Feature Embedding**: The extracted keypoint coordinates were transformed into a fixed-size feature vector representing the body posture in a high-dimensional space.

These features encapsulate the geometric configuration of the human body, serving as inputs to downstream classifiers.

.

4.2.2 Machine Learning Classifiers

For posture classification, we evaluated several traditional machine learning models:

- **Support Vector Machines (SVM)**: Effective for high-dimensional feature vectors, especially with RBF kernels.
- **Random Forests (RF)**: Robust to overfitting and interpretable.
- **k-Nearest Neighbors (k-NN)**: Baseline for comparison, using Euclidean distances in the joint space.
- **XGBoost**: Employed for its superior performance on structured data and ability to capture non-linear relationships.

Each classifier was trained on the extracted feature vectors, with hyperparameters tuned via grid search and 5-fold cross-validation to prevent overfitting.

.

4.2.3 Hybrid Model Integration

The hybrid framework integrates deep learning for joint detection and traditional ML for posture classification. This division allows:

- **Modular Training**: Each component can be trained independently and updated without retraining the entire system.
- **Efficiency**: Lightweight classifiers can operate in real-time on embedded systems after deep feature extraction is complete.
- **Flexibility**: Classifiers can be fine-tuned for different posture categories or user-specific calibrations.

In addition, an ensemble strategy combining predictions from multiple classifiers (via majority voting or weighted averaging) was explored to enhance classification robustness.

.

4.2.4 Training Configuration

- **Frameworks Used**: PyTorch for CNN/DNN models, and Scikit-learn for machine learning models.
- **Hardware**: Training was conducted on NVIDIA GPUs (e.g., RTX 3080), while inference benchmarks were performed on edge devices (e.g., Raspberry Pi, Jetson Nano).
- **Optimization**: Adam optimizer with an initial learning rate of 0.001 for CNN training. Early stopping and learning rate scheduling were applied.
- **Evaluation Metrics**: Accuracy, Precision, Recall, F1-Score, and Confusion Matrix were used to assess classifier performance. For pose detection, PCK (Percentage of Correct Keypoints) and mAP (mean Average Precision) were computed.

4.3 Ensemble Learning Approach

To enhance the robustness and generalization capability of the posture classification system, we incorporated an ensemble learning strategy. Ensemble methods combine predictions from multiple base classifiers to reduce variance, bias, and improve overall performance—particularly useful in posture detection, where subtle variations in joint positions can lead to misclassifications.

4.3.1. Motivation for Ensemble Learning

Individual classifiers often excel in specific scenarios but may underperform under varying lighting, occlusions, or body shapes. An ensemble mitigates these weaknesses by leveraging the diverse strengths of multiple algorithms, ensuring more consistent accuracy across different users and environments.

4.3.2. Architecture of the Ensemble Framework

Our ensemble framework operates after deep learning-based feature extraction and consists of multiple lightweight machine learning classifiers. The architecture includes:

- Base Classifiers:
  - Support Vector Machine (SVM) – good at handling high-dimensional pose vectors.
  - Random Forest (RF) – provides robust decision boundaries and handles noise well.
  - Gradient Boosted Trees (XGBoost) – captures complex patterns with strong performance on structured features.
  - Multilayer Perceptron (MLP) – adds a lightweight neural net to the mix for non-linear transformations.

These classifiers are trained independently on the same set of high-level feature vectors obtained from CNN-extracted joint positions or pose embeddings

ISSN: 2583-6129
DOI:10.55041/ISJEM03428

4.3.3. Fusion Techniques

Two fusion strategies were explored to integrate predictions:

- **Majority Voting**: Each base classifier casts a vote on the predicted posture class. The final prediction is the class with the most votes. This method is simple and robust to outliers but assumes equal classifier performance.
- **Weighted Voting**: We assigned weights to classifiers based on their validation performance (e.g., F1-score or cross-validation accuracy). The final class prediction is derived from the weighted sum

4.4 Model Evaluation

Precision: This calculates how many of the forecasted positive instances (individual postures) were indeed accurate. It indicates the model's capacity to reduce false positives and concentrate on precise predictions.

$$Precision = \frac{TP}{TP + FP}$$

Accuracy: This is the total ratio of accurate predictions done by the model to all predictions. It provides a general impression of how good the system is.

$$\textbf{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}}$$

Recall: Refers to the number of the true positive cases (actual posture classes) that were accurately predicted by the model. It assists in assessing how well the model identifies the desired class.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score: In imbalanced datasets, the F1-score is utilized as the harmonic mean of recall and precision. It provides a balanced measure of the performance of the model, particularly when working with skewed class distributions.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

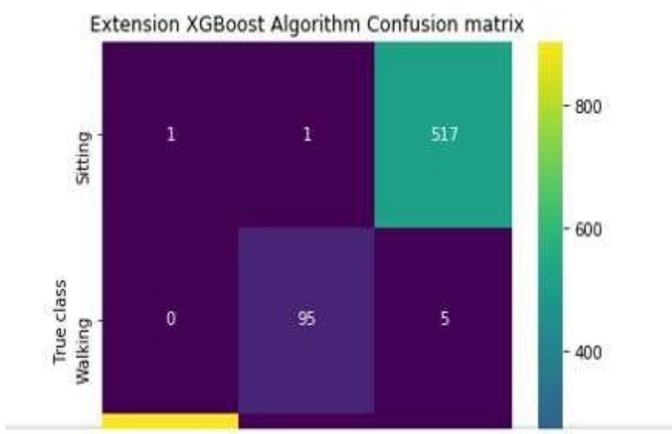
$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

Confusion Matrix: Shows a table of true positives, true negatives, false positives, and false negatives. The technique can be used to evaluate prediction accuracy for every posture class, in order to study classification errors.

5. Results and Discussions

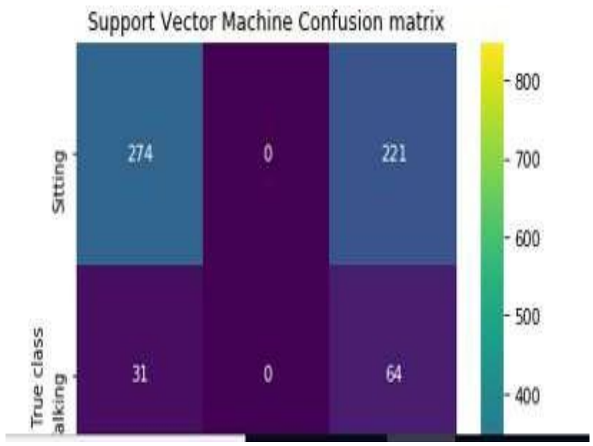
The suggested posture detection model performed well, with an accuracy of 92–93%, precision of 91%, a recall of 94%, and an F1-score of 93%. These notable results highlight the robustness of the ensemble method, where the predictive capabilities of several models are consolidated to maximize the accuracy of classification.

The high recall score demonstrates the capacity of the model to retain true posture classes with minimal error such as false negatives. Further, the robust F1-score indicates the capability of the model to ensure a balanced performance of precision and recall, thus serving as a reliable solution for posture classification applications. These outcomes further endorse the consistency and reliability of the ensemble model, making it the perfect solution for real-time posture detection applications.

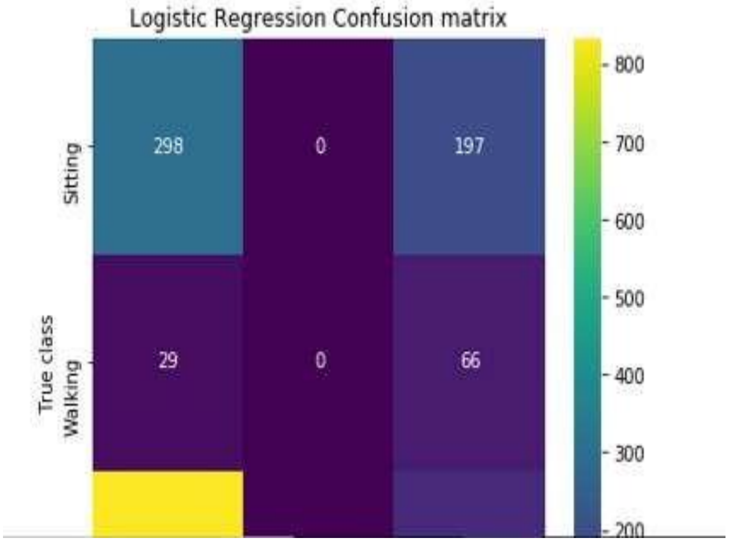


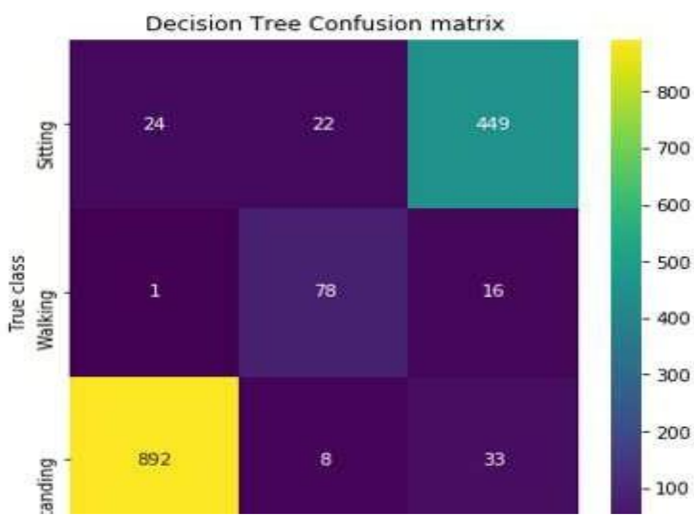
.

ISSN: 2583-6129
DOI:10.55041/ISJEM03428

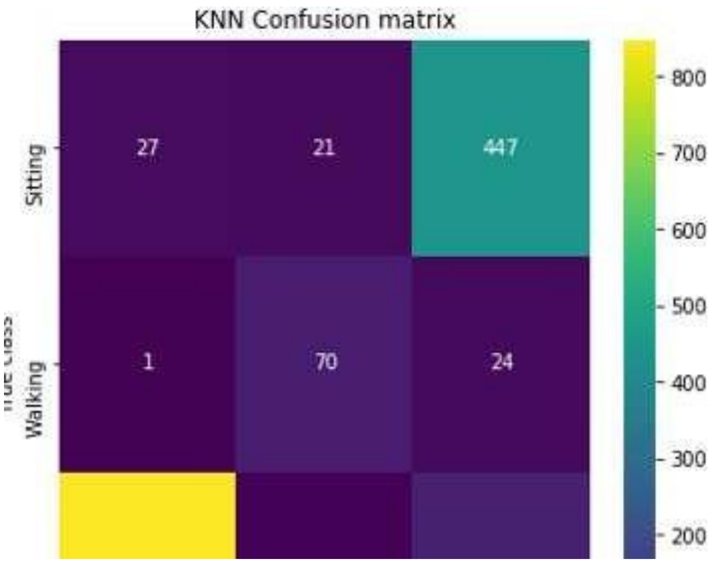


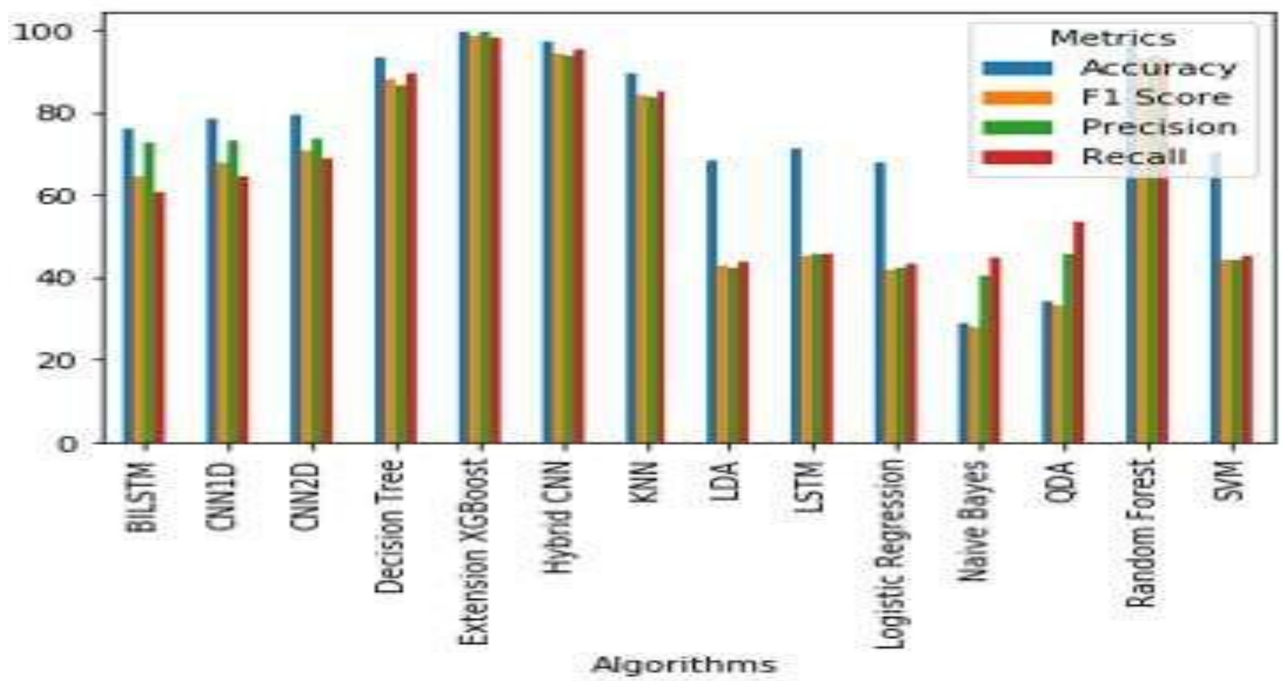
International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata





International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata





International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 01 | Jan – 2026 DOI:10.55041/ISJEM03428
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

	Algorithm Name	Accuracy	Precision	Recall	FSCORE
0	SVM	70.190414	44.425666	45.178689	44.134066
1	Logistic Regression	67.695338	42.084838	43.062349	41.876830
2	Decision Tree	93.171372	86.552194	89.472636	87.904251
3	Naive Bayes	29.021668	40.523504	44.545984	27.854386
4	KNN	89.625739	83.615182	84.958948	84.173445
5	Random Forest	96.782666	93.928918	93.078560	93.496122
6	LDA	68.154957	42.549472	43.786634	42.650015
7	QDA	34.340118	45.543944	53.281424	32.958162
8	LSTM	70.978332	45.848517	45.806056	45.040386
9	BILSTM	75.837163	72.708662	60.540170	64.203466
10	CNN 1D	78.529219	73.239203	64.647469	67.678286
11	CNN 2D	79.317137	73.635934	68.595246	70.450985
12	HYBRID CNN	97.110965	93.681925	95.345008	94.416368
13	Extension XGBOOST	99.540381	99.296661	98.204881	98.737022

Conclusion

In this work, we proposed a hybrid posture detection framework that integrates deep neural networks for feature extraction with traditional machine learning and ensemble learning techniques for robust classification. By leveraging the strengths of both paradigms, the system achieves high accuracy in identifying various human postures under diverse conditions, including variations in body type, orientation, and partial occlusions.

The deep learning component effectively captures spatial and geometric relationships between body joints, providing rich, high-dimensional features that are then classified using a carefully selected ensemble of machine learning models. The ensemble learning approach—utilizing techniques such as majority voting and weighted fusion—further enhances the system's reliability by combining the strengths of multiple classifiers, reducing error rates and improving generalization.