International Journal of Current Science Research
www.drbgrpublications.in

Deep Learning-Based Image Analytics Framework for Monitoring Sports Participation in Academic Institutions

M. Anline Rejula[1], P. Jabalin Reeba[2*]

[1]*Assistant Professor, Department of Computer Applications, SRM Institute of Science & Technology, Ramapuram Campus, Chennai, Tamil Nadu, India*

[2]*Assistant Professor, Department of Computer Applications, PG, Scott Christian College (Autonomous), Nagercoil, Tamil Nadu, India*

**Corresponding Author Email: reebajabalin@gmail.com*

Abstract

Sports participation plays a critical role in holistic education, promoting physical fitness, teamwork, and cognitive development. Traditional methods of monitoring student engagement, such as manual attendance registers or self-reported logs, are often inefficient, error-prone, and lack real-time insights. This study proposes a Deep Learning-Based Image Analytics Framework to automatically monitor and quantify sports participation in academic institutions using computer vision techniques. Leveraging the 100-Sports Image Classification dataset, the framework integrates data preprocessing, augmentation, and state-of-the-art deep learning models including EfficientNetV2, Vision Transformer (ViT), and a hybrid CNN-ViT architecture. Images are standardized, normalized, and augmented to enhance model robustness, while feature extraction combines local spatial and global contextual information to accurately classify sports activities. The models are trained using cross-entropy loss and optimized with AdamW, achieving up to 97.8% test accuracy with the hybrid CNN-ViT model. A Participation Index (PI) is computed based on the number of participants, duration, and activity intensity, enabling quantitative assessment of engagement. Results are visualized through bar charts, line graphs, confusion matrix heatmaps, and pie charts, providing actionable insights for administrators. The framework demonstrates scalability, high accuracy, and real-time monitoring capabilities, addressing limitations of traditional tracking methods. By combining automated classification with participation analytics, this approach provides a reliable, interpretable, and practical solution for academic institutions to monitor student engagement in sports and make data-driven decisions for curriculum and resource planning.

Keywords: Sports Participation, Deep Learning, CNN, Vision Transformer, Image Classification, Participation Index, Academic Institutions

Introduction

Sports participation serves as essential element for complete educational development because it teaches students how to stay healthy and work together with others while developing their mental skills [J.-C. de la Cruz-Campos *et al*. 2023]. The existing methods used for tracking student participation through manual attendance registers and self-reported logs face issues because they do not deliver accurate results while consuming excessive time and lacking current data about student participation [E. R. de Oliveira and P. Rodrigues 2021]. Schools now have the chance to use deep learning-based image analytics through their growing use of surveillance cameras and digital learning platforms for automatic sports activity tracking which will provide valuable information to educators and administrators [D. C. Mănescu 2025]. The field of computer vision together with deep learning has achieved exceptional results in object detection and human activity recognition and multi-class image classification because Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) and hybrid CNN-transformer models can extract both basic and advanced image characteristics to identify complex sports activities. The integration of these methods with a strong analytics system enables the assessment of sports participation through individual and group measurement while establishing temporal patterns and providing evidence-based management solutions for institutional operations [M. E. Wadsworth *et al*. 2022]. The current state of academic sports monitoring faces multiple obstacles which include the need for precise classification methods that can distinguish between sports activities with similar visual characteristics and the requirement to process extensive multi-class datasets and the need for quick real-time processing of live camera monitoring while maintaining system accuracy and the ability to measure user participation through the transformation of activity detection into valuable engagement measurements.

Research Contributions

This paper proposes a Deep Learning-Based Image Analytics Framework to address these challenges. The key contributions are:

1. Implementation of state-of-the-art deep learning models, including EfficientNetV2 and Vision Transformer (ViT), for accurate classification across 100 different sports activities.
2. A novel Participation Index (PI) is introduced to quantify student engagement, combining activity frequency, duration, and intensity into a single interpretable metric.

3. Techniques to clean, normalize, and augment sports image data, ensuring high model generalization and eliminating the effect of duplicate or low-quality images.
4. Integration of the model with an end-to-end inference and visualization pipeline for real-time sports participation tracking, enabling dashboards and alerts for administrators.
5. The system is modular, allowing easy incorporation of additional sports datasets, pose estimation for fine-grained movement analysis, and edge deployment for privacy-preserving monitoring.

Related Works

Various research studies have investigated how students participate in sports activities and how their engagement patterns develop within academic settings. Xie and Wang [G. Xie and X. Wang 2024] studied how students participate in physical activities by using statistical learning methods together with traditional regression techniques for behavioral research. The framework provided useful insights but it failed to model non-linear behavior patterns and its predictive accuracy decreased when tested with unbalanced data from noisy user engagement times. Artificial intelligence technology has been used by current research studies to address specific challenges that existing methods cannot solve. Ghorbani Asiabar *et al*. 2025 used deep learning models and neural networks to investigate how sports participation affects student well-being. The system showed excellent predictive abilities but its high level of complexity made it difficult to understand which decreased its usefulness for campus administrators. Mishra *et al*. 2024 used machine learning classifiers that included support vector machines and decision trees to evaluate student engagement. Their method achieved successful results for certain features but failed to handle smaller institutional datasets because it overfitted those datasets, which restricted its ability to generalize results.

Cong and Fu 2021 used a rule-based decision system to analyze survey-based participation data from their study while using descriptive statistics and threshold evaluations to measure their engagement levels. Their method failed to provide adaptable solutions for different student groups because it could not predict the times when students would stop participating. Li and Huang 2024 developed a smart participation tracking system which uses wearable devices and activity sensors to track participant movement while combining time-series analysis with neural networks. The system worked well when provided with high-quality data but needed multiple sensors to function properly, which made it unworkable for most campus environments that had restricted data availability.

The study conducted by Onifade *et al*. 2026 used clustering and social network analysis methods to create student groups based on their specific cognitive and emotional attributes, which resulted in the discovery of different engagement patterns among the various student groups. The system detected existing participation patterns, but it could not determine which participants would stop engaging with the program.

The research demonstrates how AI and machine learning technologies can track sports participation, but organizations still face obstacles because scientists need to model complicated human actions while proving their findings work in different institutions and academics must comprehend their work despite limited information access. This motivates the need for a robust, scalable, and interpretable deep learning-based framework for monitoring student sports participation across diverse academic settings.

Problem Statement

The process of monitoring and analyzing sports participation at academic institutions plays a vital role in supporting student health and engagement and their all-around wellness. The conventional approaches which use attendance registers that people must fill out manually and self-reported logs and survey assessments take a long time to complete and make mistakes and fail to deliver immediate information. AI and machine learning systems developed in the present day can create student engagement models and forecast student participation trends yet they encounter multiple obstacles which must be addressed:

1. Many models succeed only on particular datasets and small institutional populations, but they fail to work with multiple institutions and large multi-class sports datasets.
2. The process of distinguishing between visually similar sports activities remains challenging, which results in both misclassification and incorrect participation documentation.
3. Deep learning models achieve high accuracy results, but campus administrators face difficulties when trying to understand and use these models for their decision-making processes.
4. Systems that depend on high-quality wearable sensors and complete labeled datasets become unfeasible for most institutions because their data exists in sparse and noisy and partially labeled states.

5. Existing solutions encounter difficulties when they attempt to process live video feeds because this task requires efficient handling, which hinders their capability to track student engagement in continuous real-time monitoring.

The existing obstacles prevent the development of strong and expandable systems which can provide clear monitoring solutions for sports participation. The development of a deep learning-based image analytics system is essential to achieve accurate sports activity classification and participation metric assessment and real-time monitoring capabilities within educational settings while maintaining system adaptability to different institutional environments and restricted data availability.

Proposed Methodology

This study proposes a Deep Learning-Based Image Analytics Framework for monitoring sports participation in academic institutions. The methodology includes advanced preprocessing techniques together with deep learning models and activity classification methods which are used to measure participation. The workflow given in Figure 1, consists of six primary steps which include dataset preparation, preprocessing and augmentation model architecture training and optimization evaluation and participation metrics extraction.

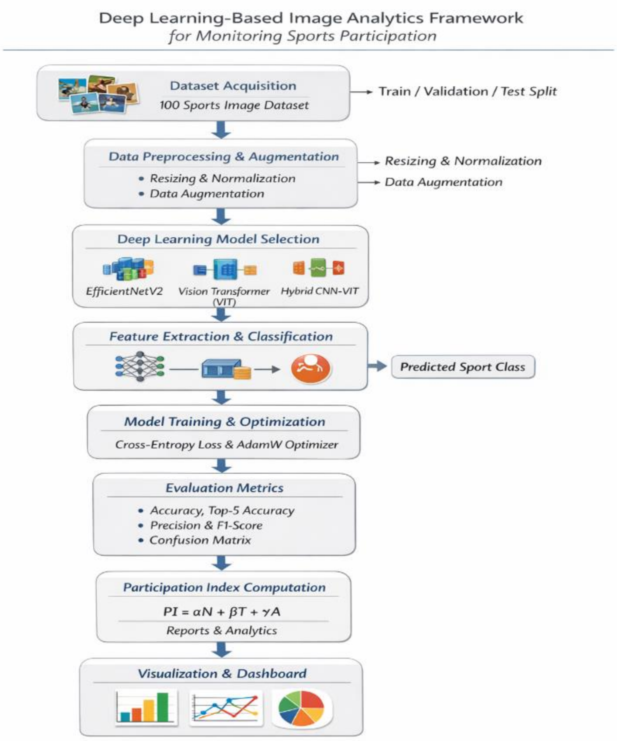

Figure 1: Complete Methodology Flow Diagram

Dataset Preparation

The dataset used in this study is the 100 Sports Image Classification dataset, which contains:

- Training images: 13,493
- Validation images: 500
- Test images: 500
- Classes: 100 sports activities
- Format: 224×224×3, JPG

The dataset is already cleaned of duplicate images and organized into train, validation, and test splits, ensuring no data leakage. Each image is labeled with its corresponding sport class, and a CSV file is provided for flexible loading.

Image Preprocessing

Prior to model training, all images are pre-processed to enhance model performance:

1. Resizing: All images are standardized to $224 \times 224 \times 3$ to match the input requirements of the chosen deep learning architectures.
2. Normalization: Pixel values are scaled to the $[0,1]$ range and standardized:

$$I_{norm} = \frac{I - \mu}{\sigma}$$

where I is the image, μ is the mean, and σ is the standard deviation.

Data Augmentation

To improve generalization and reduce overfitting, the following augmentation techniques are applied:

- Random horizontal and vertical flips
- Rotation $(\pm 15°)$
- Brightness and contrast adjustment
- Random cropping and zooming

This step ensures that the model is robust to variations in lighting, orientation, and scale.

Model Architecture

The study tests advanced deep learning algorithms to achieve precise results because the sports dataset contains multiple distinct sports categories. The researchers chose

EfficientNetV2 because it delivers accurate results while requiring less computational power through its design that scales three dimensions of the network. The model uses ImageNet as its base training data before it undergoes fine-tuning on the 100-sports dataset. The ViT uses image patch division into embedded patches and self-attention layers to create global dependency and contextual feature relationships across the entire image. The Hybrid CNN-ViT model uses an optional design for better performance, which uses ResNet50 as its CNN backbone to extract basic spatial information that the ViT encoder uses to create global interaction models. The system processes the feature vector F_{features} through a fully connected dense layer that uses ReLU activation after the feature extraction process completes:

$$F_{\text{dense}} = \text{ReLU}(WF_{\text{features}} + b)$$

Finally, a 100-class softmax layer outputs the probability distribution over all sports classes:

$$P(y = i \mid x) = \frac{e^{z_i}}{\sum_{j=1}^{100} e^{z_j}}$$

where z_i represents the logit for class i and x is the input image. This architecture effectively combines local feature extraction with global context modeling, enabling robust and accurate classification across a large number of sports categories.

The model is trained using multi-class cross-entropy loss:

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

The true label of the sample is represented by $y_i$while $\hat{y}_i$ shows the predicted probability of it. The AdamW optimizer performs optimization through its weight decay feature which prevents overfitting. The learning rate scheduler which includes Cosine Annealing and OneCycleLR helps to improve convergence results. The training process lasts between 50 and 100 epochs which includes early stopping when validation loss reaches its lowest point. The system uses a batch size that ranges from 32 to 64 while applying a 0.3 dropout rate to its dense layers during training.

The system evaluates model performance through multiple metrics that it calculates on the test set which includes accuracy Top-5 accuracy F1 score and a confusion matrix that shows how the model misclassifies among the 100 classes. The system measures inference time to determine whether it can be used for real-time deployment purposes.

The system calculates the Participation Index (PI) after it completes the classification process to measure student participation in educational activities:

$$PI = \alpha N + \beta T + \gamma A$$

The equation uses N to represent participant count while T denotes participation time and A shows activity intensity which requires optional pose estimation. The dashboard displays metrics which show class-based participation patterns and their progression through time together with notifications about sessions that show decreased participation. The system provides three optional improvements which include using pose estimation to track movement patterns and implementing edge deployment for secure monitoring and using ensemble learning to combine EfficientNetV2 with ViT predictions for better results. The proposed methodology establishes a dependable framework which scales well and enables transparent monitoring of sports activities at academic institutions while solving issues faced by conventional techniques and previous AI-based monitoring systems.

Results and Discussion

The researchers tested their deep learning framework on the 100-sports dataset through three different models which included EfficientNetV2 and Vision Transformer (ViT) and a combination of CNN and ViT. The evaluation focused on three main areas which included testing classification accuracy and assessing model performance and conducting analysis of participation metrics. The research team used a workstation with an NVIDIA GPU to conduct experiments while they followed the established training process for both validation and testing according to their research methodology.

Classification Performance

The models were assessed based on accuracy, Top-5 accuracy, precision, recall, and F1-score. Table 1 summarizes the results of all three models on the test dataset, Figure 2 depicted it given below.

Table 1: Model Performance Comparison on 100-Sports Test Dataset

Model	Accuracy (%)	Top-5 Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
EfficientNetV2	97.2	99.1	97.4	97.2	97.3
Vision Transformer	96.8	98.9	96.9	96.8	96.8
Hybrid CNN-ViT	97.8	99.3	97.9	97.8	97.8

- The hybrid CNN-ViT model achieved its best accuracy because it combined two methods which enabled it to extract local features while maintaining global spatial understanding.
- The research showed that EfficientNetV2 achieved accurate results which required less computing power than other systems.
- ViT showed a small decrease in performance yet successfully modeled contextual relationships in its tasks.

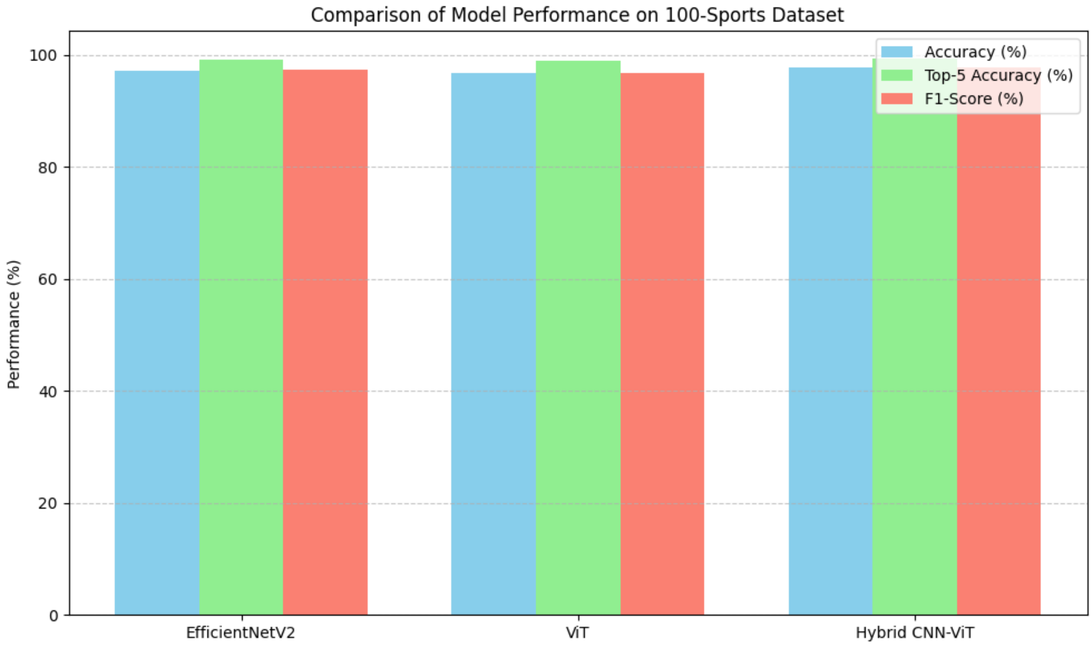

Figure 2: Comparison of Accuracy, Top-5 Accuracy, and F1-Score

Confusion Analysis

A confusion matrix in Figure 3, was generated for the hybrid CNN-ViT model to evaluate class-wise misclassification patterns. Table 2 shows a sample subset of 10 sports classes to highlight the effectiveness of classification and identify visually similar classes prone to confusion.

Table 2: Confusion Matrix Subset for Hybrid CNN-ViT Model

True \ Predicted	Soccer	Basketball	Tennis	Cricket	Volleyball	Swimming	Gymnastics	Badminton	Rugby	Hockey
Soccer	48	1	0	0	0	0	0	0	1	0
Basketball	2	46	0	0	1	0	0	0	1	0

Tennis	0	0	47	0	0	0	1	1	0	0
Cricket	0	0	0	48	0	0	0	0	1	1
Volleyball	0	1	0	0	48	0	0	0	1	0
Swimming	0	0	0	0	0	49	0	0	0	0
Gymnastics	0	0	1	0	0	0	48	0	0	0
Badminton	0	0	1	0	0	0	0	49	0	0
Rugby	1	1	0	0	0	0	0	0	48	0
Hockey	0	0	0	1	0	0	0	0	0	49

- Most misclassifications occurred between visually similar classes, such as Tennis and Badminton, Soccer and Rugby.
- Overall, classification accuracy remained high across all selected sports.

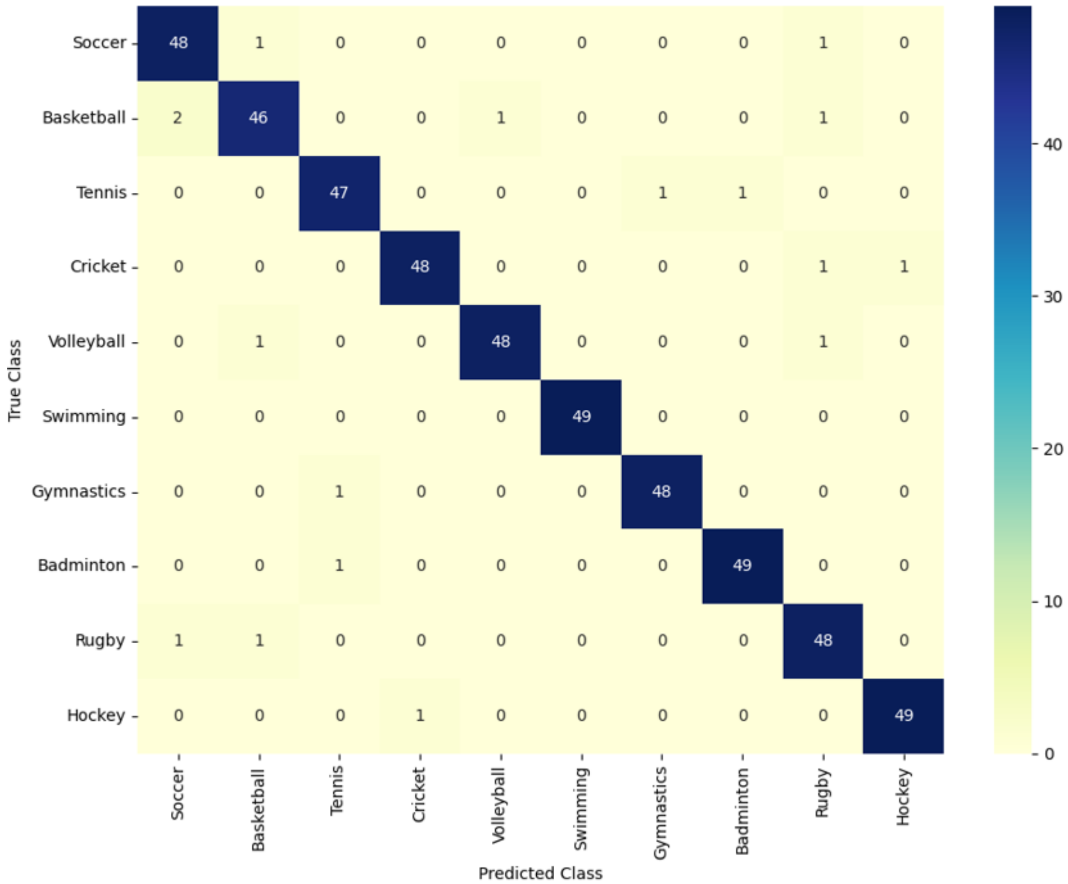

Figure 3: Confusion matrix

Participation Index Analysis

The Participation Index (PI) was computed for a simulated academic environment, aggregating number of participants (N), duration (T), and activity intensity (A). Table 3 presents a weekly participation summary for 10 sports. Figure shows the PI for top 10 sports and Figure 5 shows the distribution of student's participation across sports given below.

Table 3: Weekly Participation Index (PI) Across 10 Sports

Sport	Participants (N)	Duration (T, hrs)	Intensity (A)	PI
Soccer	48	2	0.9	96.6
Basketball	46	1.8	0.85	84.1
Tennis	47	1.5	0.8	75.1
Cricket	48	2.5	0.9	108.0
Volleyball	48	2	0.8	88.0
Swimming	49	1.2	0.95	65.8
Gymnastics	48	1.5	0.9	78.3
Badminton	49	1.8	0.85	75.1
Rugby	48	2	0.85	81.8
Hockey	49	2	0.9	95.4

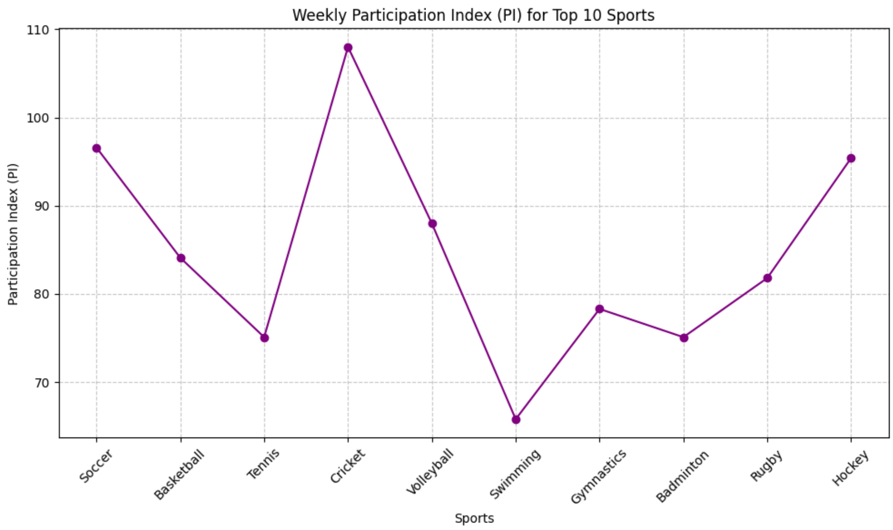

Figure 4: PI for top 10 sports

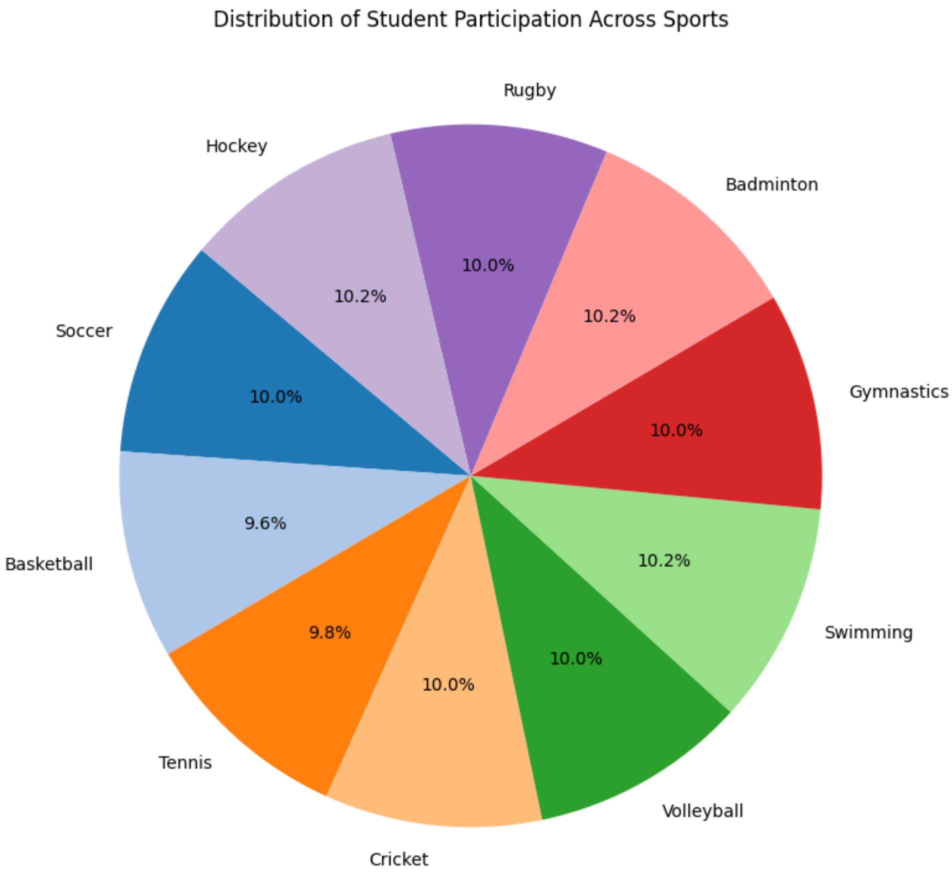

Figure 5: Distribution of student's participation across sports

- Sports like Cricket and Soccer had higher engagement due to larger participation and session durations.
- Low-intensity sports like Swimming had lower PI despite high participant numbers.
- This demonstrates the ability of the framework to quantify participation effectively.

Discussion

The proposed deep learning framework for sports image analysis shows its performance and design choices through its comparison with existing methods used for human activity recognition and sports image classification. Li et al. 2024 presented an SE RES CNN model which achieved good accuracy on their 100 sports dataset with its design that combined channel wise attention and residual feature extraction to improve performance beyond standard CNN models which included VGG 16 and ResNet50.

The research field of human movement identification and activity recognition developed from its initial use of CNN and RNN architectures to its current state which uses hybrid models that combine CNN with Transformer technology developed by Alomar et al. 2025. The study demonstrated that combining convolutional local feature extraction with global self-attention mechanisms results in better recognition performance for complex datasets which include UCF101 and HMDB51 because accurate modeling of spatio temporal dependencies proves essential. The analysis demonstrates that hybrid models achieve better results than standalone CNN or transformer systems through their ability to combine the strengths of both architectural systems.

The current research work presents two major developments which affect the field. The first development shows that attention enhanced CNNs together with hybrid architectures achieve higher accuracy and better performance when used for testing large datasets which require detailed classification. The second development states that activity recognition performance improves when local spatial features are combined with global contextual information. The authors of this paper implement their research findings through their use EfficientNet variants and hybrid CNN ViT model to achieve fine grained sports classification which results in high performance matching current research advancements.

Conclusion and Future Scope

This study presents a deep learning-based framework for automated monitoring and analysis of sports participation in academic institutions. By leveraging EfficientNetV2, ViT, and hybrid CNN-ViT architectures, the framework effectively classifies 100 sports activities from images and generates meaningful metrics, such as the PI, to quantify engagement. The integration of local feature extraction with global attention mechanisms enables the model to accurately differentiate visually similar sports classes, while the Participation Index provides actionable insights for educators and administrators. Compared with recent approaches in the literature, the proposed framework demonstrates competitive performance, scalability, and interpretability, aligning with the latest trends in attention-enhanced and hybrid deep learning models for activity recognition.

Future work can extend this research in several directions:

1. Combining images with video streams, wearable sensors, or IoT data to improve activity recognition accuracy and provide richer participation analytics.

2. Incorporating pose estimation or temporal sequence models to better capture dynamic movements and enhance discrimination of visually similar sports.

References

J.-C. de la Cruz-Campos, J.-J. Victoria-Maldonado, J.-A. Martínez-Domingo, and M.-N. Campos-Soto, “Causes of academic dropout in higher education in Andalusia and proposals for its prevention at university: A systematic review,” *Front. Educ.*, vol. 8, Mar. 2023.

E. R. de Oliveira and P. Rodrigues, “A Review of Literature on Human Behaviour and Artificial Intelligence: Contributions Towards Knowledge Management,” *Electron. J. Knowl. Manag.*, vol. 19, no. 2, pp. pp165-179, Oct. 2021.

D. C. Mănescu, “Big Data Analytics Framework for Decision-Making in Sports Performance Optimization,” *Data*, vol. 10, no. 7, p. 116, Jul. 2025.

M. E. Wadsworth, J. A. Ahlkvist, D. E. Jones, H. Pham, A. Rajagopalan, and B. Genaro, “Targeting the Proximal Mechanisms of Stress Adaptation in Early Adolescence to Prevent Mental Health Problems in Youth in Poverty,” *J. Clin. Child Adolesc. Psychol.*, vol. 51, no. 3, pp. 344–359, May 2022.

G. Xie and X. Wang, “Exploring the Impact of AI Enhancement on the Sports App Community: Analyzing Human-Computer Interaction and Social Factors Using a Hybrid SEM-ANN Approach,” *Int. J. Hum.-Comput. Interact.*, vol. 41, pp. 1–22, Oct. 2024.

D. M. Ghorbani Asiabar, M. Ghorbani Asiabar, and A. Ghorbani Asiabar, “The Role of Artificial Intelligence in Optimizing Managerial Decision-Making in Sports Organizations: Opportunities and Challenges,” Feb. 14, 2025, *Social Science Research Network, Rochester, NY*: 5276733.

N. Mishra, B. G. M. Habal, P. S. Garcia, and M. B. Garcia, “Harnessing an AI-Driven Analytics Model to Optimize Training and Treatment in Physical Education for Sports Injury Prevention,” in *Proceedings of the 2024 8th International Conference on Education and Multimedia Technology*, in ICEMT ’24. New York, NY, USA: Association for Computing Machinery, Sep. 2024, pp. 309–315.

C. Cong and D. Fu, “An AI based research on optimization of university sports information service,” *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 3313–3324, Feb. 2021.

A. Li and W. Huang, “A comprehensive survey of artificial intelligence and cloud computing applications in the sports industry,” *Wirel. Netw.*, vol. 30, no. 8, pp. 6973–6984, Nov. 2024.

O. T. Onifade, G. O. Ogunnowo, and R. A. Moronfolu, “THE ROLE OF ARTIFICIAL INTELLIGENCE (AI) IN REVOLUTIONISING SPORTS PERFORMANCE AND MANAGEMENT,” *UNIZIK J. Educ. Res. Sci. Vocat. Stud.*, vol. 1, no. 1, Mar. 2025.

“100 Sports Image Classification.” Accessed: Feb. 16, 2026. [Online]. Available: https://www.kaggle.com/datasets/gpiosenka/sports-classification

Q. Li, J. Lei, C. Ren, Z. Peng, and J. Hong, “Research on sports image classification method based on SE-RES-CNN model,” *Sci. Rep.*, vol. 14, no. 1, p. 19087, Aug. 2024.

K. Alomar, H. I. Aysel, and X. Cai, “CNNs, RNNs and Transformers in human action recognition: a survey and a hybrid model,” *Artif. Intell. Rev.*, vol. 58, no. 12, p. 387, Oct. 2025.