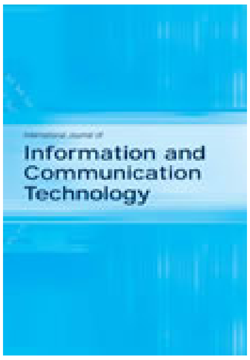


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Artificial intelligence-based automatic identification and classification of diverse sports using advanced deep learning models

Yuan Zheng, Long Cai

DOI: 10.1504/IJICT.2025.10071786

Article History:

Received:	12 April 2025
Last revised:	13 May 2025
Accepted:	13 May 2025
Published online:	10 July 2025

Artificial intelligence-based automatic identification and classification of diverse sports using advanced deep learning models

Yuan Zheng and Long Cai*

Weinan Normal University,
Shaanxi, Weinan, 714000, China
Email: 13772774963@163.com
Email: 13772773695@163.com
*Corresponding author

Abstract: The study examines state-of-the-art artificial intelligence (AI) methodologies aimed at developing sports image classification as it affects multimedia management as well as recommendation algorithms and sport data analysis capabilities. The sports industry is witnessing unprecedented growth, fuelled by advancements in technology, and the exponential rise of digital content. The vast quantity of sports-related media requires critical management for improved accessibility for user engagement capabilities. AI brings transformative automation capabilities through its ability to tackle these sorts of tasks. Deep learning applications show outstanding performance for resolving intricate classification challenges. This research developed a sports image classification framework using deep neural networks (DNNs) and analysed two pre-trained models ResNet-50 and MobileNet for performance comparisons. The DNN model demonstrated outstanding performance metrics through 98% accuracy which matched its precision and recall and F1-scores. DNN proved the most suitable solution when compared to pre-trained models ResNet-50 and MobileNet.

Keywords: artificial intelligence; sports classification; game; deep neural network; DNN; feature extraction.

Reference to this paper should be made as follows: Zheng, Y. and Cai, L. (2025) 'Artificial intelligence-based automatic identification and classification of diverse sports using advanced deep learning models', *Int. J. Information and Communication Technology*, Vol. 26, No. 23, pp.91–113.

Biographical notes: Yuan Zheng works in the Department of Physical Education in Weinan Normal University, Shaanxi, China. Her research fields are digital image processing, artificial intelligence, and athlete education, with a focus on deep learning models' incorporation to sports analysis and education systems.

Long Cai is a researcher, an academic at the Weinan Normal University, majoring in applications of AI in the sports science, recognition systems, and physical education technologies. His work focuses on the research of advanced techniques of machine learning for the promotion of athletic endeavours as well as the automated classification of sports.

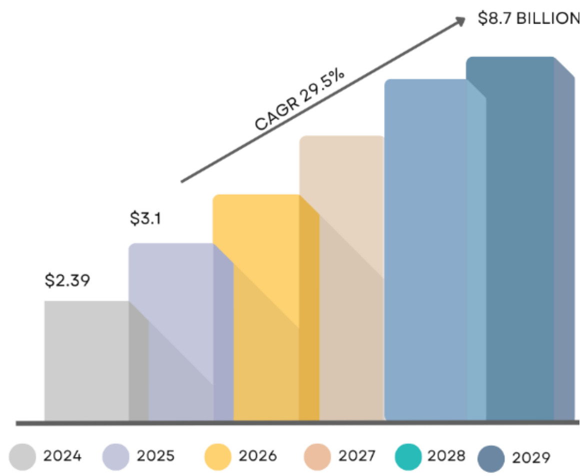
1 Introduction

The concept of AI revolves around the ability of machines to simulate human-like cognitive functions, including perception, learning, and decision-making. In the domain of computer vision, AI-powered models have significantly advanced the recognition and classification of images, enabling automated systems to identify patterns and extract meaningful insights. Applying this concept to sports classification, deep learning techniques can effectively distinguish between different sports based on visual features, streamlining processes in sports analytics, broadcasting, and digital media management (Rodrigues et al., 2020). For the past several decades the sports industry has expanded massively to become a major worldwide cultural expression and economic force and social platform. Sports events that happen annually during millions of tournaments and amateur matches result in enormous amounts of sports media content. The management and analysis of the massive sports data influx has become a complex but vital challenge because it includes both live broadcasts and promotional content alongside highlight reels (Li et al., 2024a). The implementation of AI-powered automated solutions has proven to be essential for tackling multiple industry challenges by performing image and video classification and delivering performance analytics alongside personalised user experiences (Li and Huang, 2023). The global sports industry has experienced exponential growth over the past decade, fuelled by increasing commercialisation, digitalisation, and technological advancements (Rodrigues et al., 2020). With sports events attracting billions of viewers worldwide, the industry has evolved into a multi-billion-dollar market. According to market research, the global sports industry was valued at $500 billion in 2023, with projections indicating sustained growth in the coming years due to rising investments in infrastructure, sponsorships, and digital engagement platforms (Li et al., 2024a). In parallel, AI in sports market has surged significantly, growing from $1.85 billion in 2023 to an estimated $2.39 billion in 2024 as shown in Figure 1, with an anticipated compound annual growth rate (CAGR) of 29.5%, reaching $6.69 billion by 2029 (Agarwal, 2023). This rapid growth highlights the increasing significance of AI and computer vision in sports, revolutionising everything from player performance analysis and injury prevention to officiating accuracy and fan engagement (Efe, 2023).

AI-powered sports analytics are being integrated into training programs, helping coaches and athletes make data-driven decisions to enhance performance and minimise injuries (Naik et al., 2022). Wearable AI-powered devices, such as smartwatches and motion sensors, provide real-time insights into an athlete's biomechanics, heart rate, and physical exertion, optimising training regimes and recovery strategies (Naz et al., 2024). Furthermore, computer vision technology is being widely adopted in sports officiating, with advanced systems such as Hawk-Eye in tennis, VAR in football (soccer), and AI-driven referee assistance in basketball and cricket improving the accuracy and fairness of in-game decisions (Jin and Zhang, 2024). These developing economies accelerate market growth through both the organisation of international competitions and dedicated investments into sports infrastructure. The upsurge confirms sports require automated expertise for processing varied media content and general data which validates that AI-driven classification platforms are essential for sports organisations (Cust et al., 2018).

Advanced AI systems continuously develop advanced solutions which optimise operational efficiency and reduce procedural inaccuracies in sports applications (Zhou and Wu, 2023). Image classification through deep learning algorithms demonstrates outstanding potential for recognising and arranging different types of sports media files (Dindorf et al., 2022). Sports image classification delivers critical benefits to multimedia organisation and search optimisation while providing stakeholders like broadcasters' advertisers and sports bodies access to enhanced analytics capabilities (Musat et al., 2024). The automatic detection of sports in images enables content moderation functions while creating improved recommendation systems and athletic development services through scouting and training processes (Wang and Guo, 2024). Recent research indicates that sport industries worldwide demonstrate rising expansion trends since the previous decade and currently achieve expanded growth projections (Jiang and Tsai, 2021). Digital technology improvements combined with worldwide sports occasions fuelled industry growth and enabled additional offerings and market expansion (Khan et al., 2022).

Figure 1 AI in sports market analysis (see online version for colours)



A research investigation uses advanced AI models and pre-trained segments along with a novel deep neural network (DNN) structure to perform automatic sports image classification. Feature extraction through pre-trained models including ResNet-50 and MobileNet demonstrates high operational effectiveness making them an ideal choice for image recognition tasks. Sports classification accuracy improves significantly when prefabricated DNN models integrate with specialised neural network designs that allow identification of specific sport-related attributes. The study demonstrates that the proposed DNN architecture exceeds pre-trained models by achieving superior outcomes for precision while simultaneously improving accuracy as well as recall and F1-score performance. This demonstrates its potential as a highly reliable solution for complex classification tasks. Deep learning's potential for sports classification emerges as a main discovery from this research while advancing AI applications in sports. This study establishes foundational groundwork toward automated sports media analysis through its resolution of critical challenges including feature extraction and system optimisation, following contributions are:

- Developed a DNN model which attains outstanding results on sports image classification tasks by boosting attainment measures including accuracy, precision, recall and F1-score performance metrics.
- Conducted a comprehensive comparison between ResNet-50 and MobileNet pre-trained models before highlighting the performance benefits of custom feature extraction.
- Established a framework that will guide further research in AI-driven sports analytics which enables progress toward real-time solutions and expanded sports classifications.

The rest of the paper organised as: Section 2 presents the existing literature based on interactive media. Section 3 provides the research proposed methodology along with experimental setup. Section 4 presents the findings from this study in result and discussion. Section 5 shares the conclusion section along with future directions.

2 Related work

Diverse sports identification and classification initiatives with modern artificial intelligence (AI) models demonstrate significant progress through deep learning implementation for image data processing. This section analyses vital research contributions regarding different methodologies along with datasets and resultant outcomes from this field. Research on sports image classification now utilises convolutional neural networks (CNNs) which attain both high accuracy rates and efficient execution. SE-RES-CNN demonstrated an exceptional accuracy rate during its examination of sports images from a dataset comprising 500 entries combined with SE modules for dynamic weight adjustment together with Res's modules to extract deep features from images while solving difficulties in background interferences and athlete posture changes (Market Research Reporting and Analysis Agency in India, 2024). A sports action recognition methodology which brought together deep learning and clustering algorithms demonstrated remarkable results in achieving better recognition performance with fewer false alarms. The analysis of combined multimodal data presents a research approach to enhance classification results (Li et al., 2024b). A review of literature showed that sports monitoring benefits from blending visual sensor information with non-visual wearable sensor inputs for recognising athlete movements and objects in dynamic settings. Through this method analysts achieve superior comprehension of athlete movements which enhances system accuracy when classifying their actions (Hammes et al., 2022). Multiple technical hurdles continue to affect the development of computer vision systems. Regular methodologies experience difficulties in processing complex sports actions while requiring detailed dataset annotations for high quality results. According to research findings CNNs deliver excellent image feature extraction performance yet they face precision boundaries caused by restricted sport category labels (Xu, 2021). Model bias toward frequent classes stands as a problem affecting performance because unbalanced sample representation between categories demands advanced training practices (Rodrigues et al., 2020). Rating novel AI models represents a major research emphasis. The developed sports training video classification model restored three-dimensional target positions through camera calibration technology during its development. The model succeeded in achieving class identifications with over 99% precision through its analysis of motion vector and brightness feature patterns. Research implementing recurrent neural networks (RNNs) proved successful in action recognition by analysing physiological and positional data to recognise elaborate athletic movement sequences (Fu et al., 2022).

Another study demonstrates a new DNN model which utilises modified tuna swarm optimisation (NTSO) algorithm optimisation for sports image classification. Through its feature extraction abilities, the DNN works alongside NTSO to optimise hyperparameters delivering exceptional classification results. To handle the technical demands within sports image analysis, researchers must develop advanced methodologies suitable for these diagnostic challenges (Hou and Tian, 2021). In another study, employed differential evolution for optimising hyperparameters of a CNN that utilised transfer learning to classify sports disciplines like American football and rugby and soccer and field hockey (Zhou et al., 2024). Transfer learning reached superior classification results after fine-tuning while demonstrating higher performance than typical CNN practices according to the research results. This approach generated dependable outcomes and demonstrated why transfer learning techniques excel in contexts that have minimal labelled data available. The results support employing sophisticated training strategies to boost classification results in analogous sport groups (Podgorelec et al., 2020).

A model developed federated learning systems which enhanced sports image classification outcomes while maintaining user privacy throughout centralised storage frameworks. Through genetic optimisation frameworks researchers refined CNN base architecture performance within federated learning platforms to achieve better sport image classification while maintaining privacy protection of user data (Ramesh and Mahesh, 2022). Research investigated AI-based target detection solutions for sports images which demonstrate their crucial role in athlete development and event control operations. Multiple detection strategies for targets within sports imagery are presented in this work but researchers also address the detection limitations which stem from occlusion along with different viewing angles of targets (Fu et al., 2024).

The research findings expand current knowledge about AI utilisation for classification purposes while making improvements to sports environment situational awareness systems. The combination of research studies demonstrates how innovative AI approaches drive major progress in spotting and categorising multiple sporting events (Russo et al., 2019). Sports imagery detection methods achieve remarkable success through deep learning solutions integrated with optimisation algorithms which address obstacles created by sports-specific challenges like repeated body movements together with different background conditions (Rafiq et al., 2020). Sports image classification systems have become more accurate and efficient because deep learning methods work together with multipart data sources. Research in this intriguing field needs to continue because the persistent difficulties with data reliability and algorithm racial and ethnic bias remain. Future research needs to be pushed forward by creating algorithms which classify multiple sports categories together with robust performance levels across various environmental conditions and through continued work on integrating multiple AI approaches.

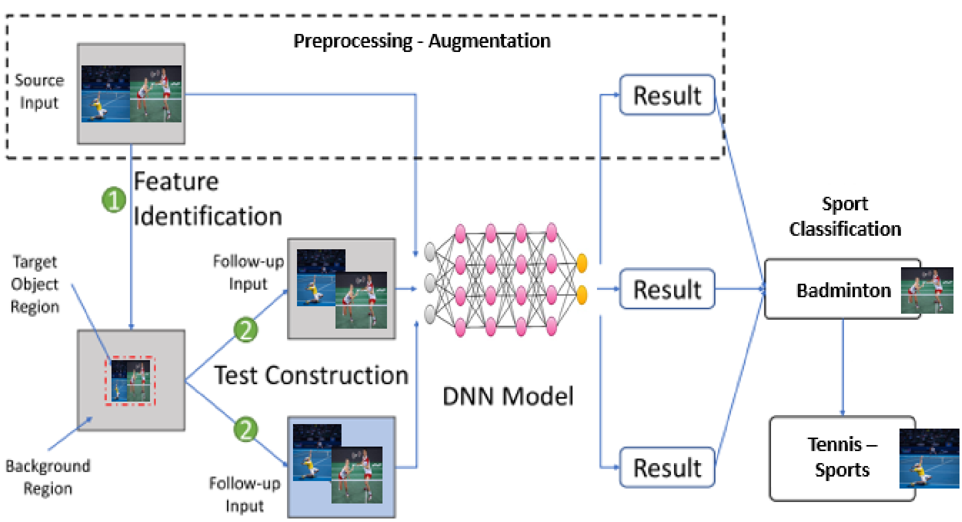
3 Proposed research methodology

In this section various steps being used in the proposed research methodology to evaluate the DL models with pre-trained applied to the dataset, to accurately classify diverse sports, as shown in Figure 2, deals with the steps including data cleaning, feature selection, data splitting, model training and evaluation, ensuring that the results are accurately measured by performance evaluation measures.

3.1 Data collection and pre-processing

The available dataset contains different sports images with a CSV file (train.csv) contains the training set of entries with image IDs and class labels and the test set is distributed across a separate CSV file (test.csv) showing only image IDs. All images, as samples shown in Figure 3 exist in their designated train and test folders based on split of 80–20 holdout. The primary objective of this task is to build a model that can classify the images in the test set into one of the five sports categories: cricket, swimming, badminton, wrestling, or tennis, no, of ratio images display in Figure 3. Let the dataset D consist of images from various sports classes, with a training set D_{train} containing labelled images and a test set D_{test} consisting of unlabelled images. The training set $D_{train} = \{(x_i, y_i)\}$ includes images-label pairs, $x_i \in \mathbb{Z}^{W*H*C}$ represents an image with dimensions $W * H$ and C shows channels, and $y_i \in \{1, 2, \ldots, 4\}$ is the class labels for the i^{th} image, corresponding to one of the sports categories, defined in Table 1.

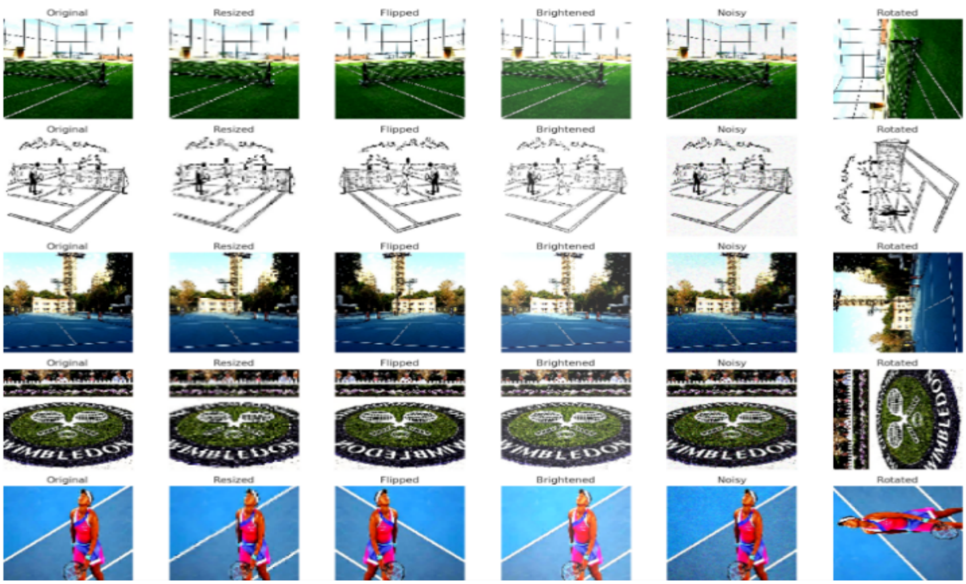
Figure 2 Research proposed methodology (see online version for colours)



The test data $D_{test} = \{(x_i)\}$ consists of M unlabelled images for which class predictions need to be made. Pre-processing is applied to enhance the robustness of model, using data augmentation techniques to transform images in the training set, increasing the variety of input samples while preserving the underlying information, as sample images shown in Figure 4 after pre-processing.

Figure 3 Sample images from dataset (see online version for colours)

Figure 4 Sample images of dataset after applying pre-processing steps (see online version for colours)



Let T denote an augmentation function, the augmented dataset $D_{aug} = \{T(x_i)\}_{i=1}^{N}$ is created by applying various transformations T to the training images, where each transformation is defined as input image x_i is randomly rotated by an angle $\theta \in \mathbb{R}$, such that $T_{rotate}(x_i) = R(x, \theta)$ where R is the rotation operator. Flipping applied to image x_i either flipped is horizontal or vertical defined as $T_{flip}(x_i) = F(x_i)$, where F is the flip operator, which could be horizontal flip or a vertical flip. A random zoom factor $\alpha \in [1 - \delta, 1 + \delta]$ where δ is a small factor, is applied to the image. This represented as $T_{zoom}(x_i) = Z(x_i, \alpha)$. Moreover colour enhancement is applied to image adjusting its brightness, contrast and saturation covering as $T_{colour}(x_i) = C(x_i, \gamma)$. Let $D'_{aug} = \{T(x_i)\}_{i=1}^{N}$ be the augmented training dataset, where $T(x_i)$ is one of the transformation applied to x_i. By applying the pre-processing techniques and training the model on the augmented dataset, enhance its ability to generalise to unseen data, improving its classification accuracy for the five sports classes: cricket, swimming, badminton, wrestling, and tennis.

Table 1 Number of images across each category

Label	*Count of label*
Tennis	1,445
Swimming	1,190
Badminton	1,394
Wrestling	1,471
Cricket	1,556
Total	7,056

3.2 Feature extraction and classification

In this section, the method for recognising key features in images and deep learning models for classification of different sports activities has been discussed. This approach involves first extracting meaningful features from pre-processed images then carries out classification tasks, using different DNNs that are further measure its effectiveness in terms of both accuracy and reliability leaning.

3.2.1 Proposed model – DNN

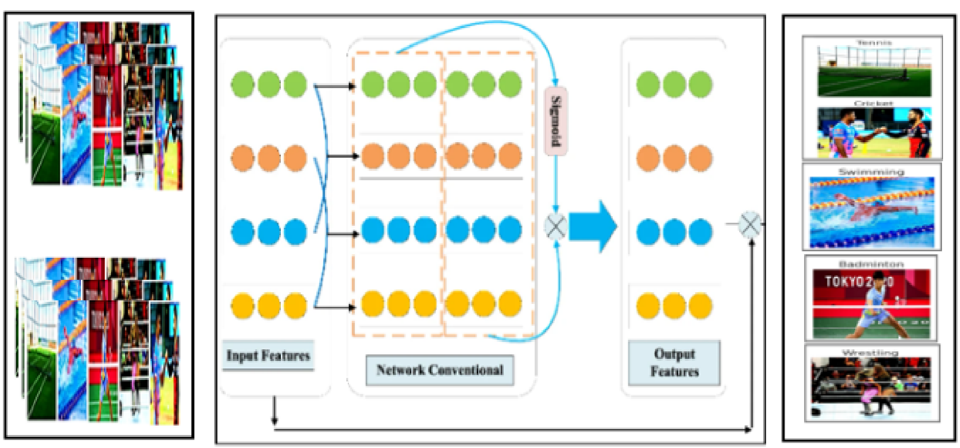
For the sports classification purpose a DNN represents one possible straightforward alternative to consider. A typical DNN contains different fully connected layers which progressively learn abstract details from image inputs, as architecture defined in Figure 5 (Afzal Badshah et al., 2024). The DNN operates as a multilayered neural network to perform either classification or regression or additional features based on user requirements. This provides a comprehensive description of DNN operation with a focus on sports image classification tasks.

- *Input layer:* DNN obtains sports images that undergo uniform resizing before processing to keep a stable baseline across datasets. The input layer directs sports image information to stage two for feature extraction.
- *Convolutional layers:* Sports images undergo pre-processing in convolutional layers to extract meaningful patterns that include edge features and texture characteristics and shape information. As the number of layers increase the detected elements grow more abstract while also becoming progressively more elaborate. Each convolutional layer creates output feature maps showing the occurrence of distinctive image patterns inside the image visuals.
- *Max-pooling layers:* Pointer noise operations are initiated after convolutional layers to minimise the length, and width features within the map domains. Maximum utility features are preserved through this technique which reduces required computational power. Through max-pooling the model achieves enhanced immunity against input variations because of changes like positional shifts and rotations.
- *Fully connected layers:* Following the feature map flatting process a vector moves through dense layers for processing. The classification decisions happen through the combination of extracted features which exist across multiple layers. They build

advanced understanding structures that describe input data information at higher levels.

- *Output layer:* The final layer outputs the probabilities of each sports category. One label achieves maximum probability which indicates the classification prediction. For every input image the model layer identifies the sports classification that provides the maximum probability.
- *Training process:* While training occurs, the model modifies weights inside itself to achieve a minimum discrepancy between its predicted output and the confirmed class. The system performs forward propagation first to generate predictions along with backward propagation afterward to adjust weight values through error analysis. After concluding training, the model moves to test unknown data for evaluation. Technical model performance evaluation includes accuracy measurements together with classification report outputs. Generalisation capabilities in a DNN remain vital because they enable accurate sports category detection between various sports. Hyperparameters are settled during model optimisation to get unmatched results, display in Table 2. It helps in pointing out the nearest hyperparameter value. As a result, some frequently used hyperparameter values and looked at other approaches to get a more precise model evaluated and predicted.

Figure 5 Architecture of proposed model DNN (see online version for colours)



They were used in the deep model that was suggested. In the proposed model configuration, the optimiser chosen is adaptive moment estimation Adam, for training deep model, due to its adaptive learning rate and momentum features. Adam dynamically adjusts the learning rate based on the past result, allowing for faster convergence and improved performance. The choice of loss function categorical cross entropy aligns with the task of classification, effectively measuring the results between predicted and actual class distribution. With a relatively low learning rate 0.0001, the model is trained to avoid large parameter updates for optimal solution. Training is conducted over 100 epochs, indicating moderate but enough iterations to allow the model to learn different patterns from input. A batch size of 32 is used, for adjusting computational efficiency with gradient approximation accuracy, ensuring stable updates during training. Overall, this parameters configuration shows a well-considered approach direct at optimising training solidity, convergence speed, and model performance for the detection classification of image.

3.2.2 Baseline models

The sports classification task performed using pre-trained models including both ResNet-50 and MobileNet models based on system requirements.

Through residual learning technology the ResNet-50 deep CNN resolves gradient vanishing to enable efficient training of extensive network architectures. The introduced architecture enables residual blocks to perform skip connections that directly assemble block inputs to outputs. The identity mappings learned through this design simplify deep model training processes. After opening a convolutional layer in ResNet-50 the network contains several sequential residual blocks. After the blocks a global average pooling operation reduces feature map spatial extent before final classification uses a fully connected layer with SoftMax activation. The main benefit of ResNet-50 results from its ability to handle deeper network architectures that provide improved model performance and acceleration thanks to efficient gradient propagation. The deep learning paradigm MobileNet exists to provide meaningful performance in mobile and embedded devices through its focus on efficient computation. It achieves this through depth wise separable convolutions, which split the convolution operation into two stages: Iterative computer vision activates pointwise convolution (1×1 convolution) after the depth wise convolution breaks inputs into separate channels. Depth wise separable convolutions diminish both parameter requirements and computational processing in contrast to standard convolution methods. The MobileNet architecture includes two multipliers. The width multiplier changes layer channel counts to minimise model size followed by the resolution multiplier that adjusts image size for different resource situations. Applications that need trade-offs between processing speed and performance accuracy benefit well from MobileNet because it suits real-time necessities particularly well (Oppici and Panchuk, 2022).

4 Results and discussion

Results from a pre-trained model operating on different sports images from a given dataset exhibit outstanding performance rates. The performance of ResNet and MobileNet was evaluated for classifying sports images into five categories: cricket, swimming, badminton, wrestling, and tennis. Both models demonstrated strong classification capabilities, achieving high accuracy; however, they exhibit distinct behaviours in terms of classification precision, misclassification trends, and training stability. ResNet achieved an overall accuracy of 92%, whereas MobileNet performed slightly lower at 90%, indicating that ResNet has a superior ability to extract relevant features for classification.

The classification reports in Table 3 give a comparison of how good each model is at distinguishing sports categories across all classes, ResNet yielded higher precision, recall and F1-scores except for the badminton class, which is only 0.89 in precision, 0.90 in recall and 0.90 in F1-score. On the other hand, MobileNet struggles a little bit more in Badminton (0.87 F1-score) and Tennis (0.89 recall), meaning that more often MobileNet misclassified these two sports than ResNet did. In these categories, both Swimming and Wrestling performed reasonably well in both models, but MobileNet returned somewhat less recall which indicates that it incorrectly classified some instances. Precision and recall values in MobileNet are more balanced (0.90 on all metrics) and less than ResNet's. This indicates that MobileNet is not at all inept but at a cost of slightly weaker classification. While doing so, ResNet maintains a stronger precision-recall balance with better results in almost every category (fewer false positives and false negatives). ResNet is superior to MobileNet in having better feature extraction capabilities than it, hence, enabling it to distinguish between similar categories better.

Table 2 Details of hyper parameters used

Hyper-parameters	*Values*
Optimiser	Adaptive moment estimation
Layers	3
Neurons	128
Activation function	ReLU
Output function	Softmax
Dropout	0.3
Input shape	224 × 224 × 3
Loss function (f)	Categorical cross entropy
Learning rate (r)	0.0001
Epochs	100
Batch size	32

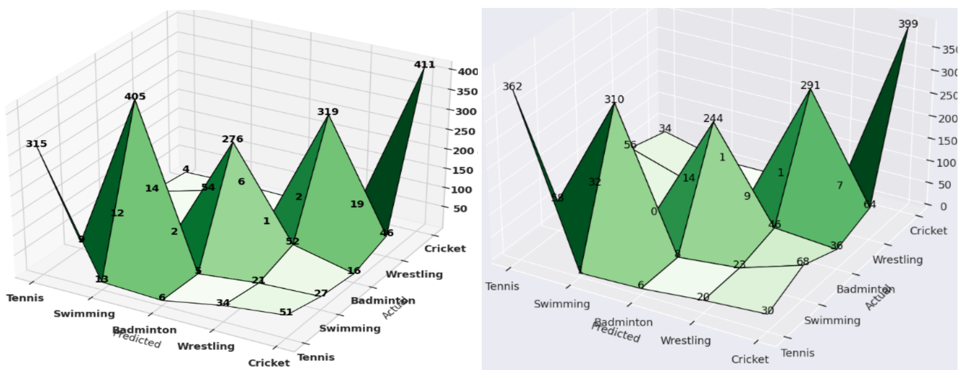
Table 3 Results of pre-trained models

Classes	*Acc*	*Pre*	*Re*	*F1*	*Acc*	*Pre*	*Re*	*F1*
	ResNet-50				*MobileNet v3*			
Tennis	92	92	91	91	90	90	89	90
Swimming	92	92	93	93	90	92	91	91
Badminton	92	89	90	90	90	87	88	87
Wrestling	92	92	92	92	90	91	91	91
Cricket	92	92	92	92	90	89	91	90

Additionally, the confusion matrices as shown in Figure 6, further show the differences in the misclassification pattern. True positive rates above 90% on all sports were achieved every sport compared to ResNet as in Figure 6(a). It is observed that swimming (92.72%) and wrestling (91.94%), in the highest classification accuracy and badminton (90.22%), in the lowest classification rate, classify ResNet as the most challenging class. Badminton was confused with swimming (3.35%) or cricket (3.07%), which indicates possible resemblances between the latter. However, MobileNet appeared to have higher misclassification rates of badminton, tennis and cricket as in Figure 6(b). Simply speaking, the true positive rate of badminton was low (87.99%), and was often misclassified as swimming (4.47%) and wrestling (3.91%), respectively, which means MobileNet has a harder time separating opposition between these sports. As with

ResNet's own confusion matrix before it, cricket was misclassified at a rate of 3.39% as tennis. According to these patterns, MobileNet has more errors from time to time since it cannot differentiate finely between similar sports. Both models show a common trend of swimming and wrestling being among the most accurately classified categories whereas badminton is always involving the highest misclassifications, which indicates that it is a category that does not have enough distinctive features differentiating.

Figure 6 Confusion matrix of, (a) ResNet-50 (b) MobileNet v3 (see online version for colours)



(a) (b)

The accuracy graphs over epochs in Figure 7 tell us critical insights of how good each model learned and generalised over time. The trial of ResNet as in Figure 7(a) showed an increase in training and validation accurate steadily, first developing above 80% validation accurate with 20 epochs and getting stable at about 92% on the final epochs. The validation accuracy was stable, but small, from training and showed good generalisation capability. However, the accuracy curves of MobileNet as in Figure 7(b) had much more fluctuating accuracy curves, particularly in the early epochs. Training accuracy continued to improve monotonically, while validation accuracy rose at first but varied erratically between 60–85% up to the first 50 epochs. This indicates that MobileNet was not able to stabilise its learning process and quite possibly suffered from overfitting and generalisation issues. In later epochs the oscillations of 85% and 88% in validation accuracy were witnessed, while the curve of ResNet was much more consistent and stable. These observations imply that lightweight design of MobileNet leads to a noisier learning process compared to stable learning made possible by ResNet's architecture. MobileNet's strange validation accuracy seems to indicate the possible higher sensitivity to the variations of the dataset, hence the lack of stability in generalisation ability.

Finally, training validation loss graphs in Figure 8 are also evidence for model behaviour and over fitness concerns. The loss curve of ResNet as in has both training and validation loss going down steadily from initial training until the training approaches close to zero, whereas the validation loss varies slightly, but remains relatively stable. Validation loss however occasionally, peaks, especially in the middle of epochs, which suggests that while ResNet generalises well, it is still suffering a bit from learning instability in some points. On the other hand, MobileNet's loss curves in Figure 9 have much more drastic fluctuations. Validation loss has very high instability: it spikes a lot, especially around 70 epochs where it became greater than 20. The loss patterns in these reinforce MobileNet's overfitting tendencies in that the model memorises the train data effectively, but the model does not generalise well on those unseen data. Furthermore, the large spikes in validation loss are an indication that the learning process of MobileNet's may be less efficient and might need stronger regularisation techniques such as dropout, batch normalisation, or reduce the learning rate. Whereas ResNet verses MobileNet, ResNet has a higher classification accuracy (92 vs. 90%), more balanced precision-recall, and learning process. Since ResNet outperforms MobileNet in extracting sports features, it easily discriminates between the sports categories, especially in badminton, tennis, and cricket, which were misclassified more often in MobileNet. Moreover, ResNet generalises more since its validation accuracy is a bit less erratic with a much more stable loss curve, suggesting that ResNet learns better and suffers less severe overfitting. Although MobileNet achieves competitive results, it suffers from higher misclassification rates and changing validation accuracy, which makes MobileNet unreliable for sports classification task in the real world. The mobile net is found to be much more sensitive to dataset variations and requires more regularisation and optimisation further. This can stabilise learning and reduce the overfitting by providing techniques like dropout layers, batch normalisation, and learning rate adjustments for improving the performance of MobileNet. It is also important to augment the dataset with more diverse samples of badminton, tennis, and cricket to make it easier for the model to distinguish these classes better. Finally, findings shows that ResNet is superior in sports classification terms to respect to higher accuracy, more stable learning behaviour, better feature differentiation. While MobileNet is light and efficient, its learning process is too varied, and therefore it is not as robust as this classification task. While MobileNet could still be optimised according to computational efficiency, ResNet still presents the best results when it comes to accuracy and reliability for sports classification in this comparison.

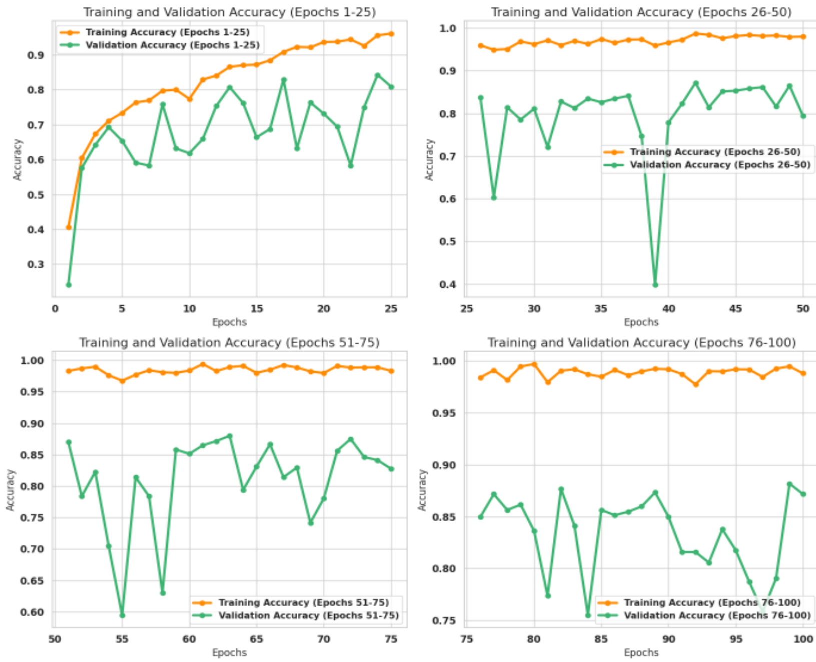
4.1 Results with proposed model DNN

The proposed DNN model accomplishes superior results in sports image classification tasks. The model demonstrates outstanding performance by obtaining 98% accuracy combined with matching precision and recall values of 98% alongside an F1-score of 98%. The DNN model predicts sports classes perfectly while maintaining a harmonious trade-off between precision and recall metrics which minimises incorrect positive and negative outputs. The combination of metrics reveals how the model maintains high reliability when exposed to difficult tasks within the dataset. The DNN model performed better than others (ResNet, MobileNet and DNN) model.

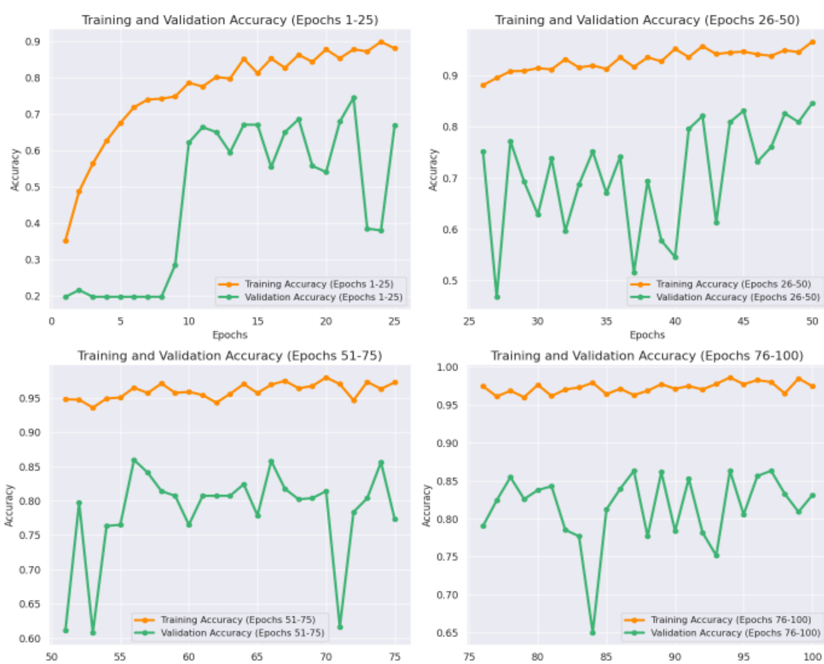
Overall, the DNN achieved an accuracy of 98% (above 92% for ResNet and 90% for MobileNet), and had remarkably good precision, recall, and F1 among all five categories: cricket, swimming, badminton, wrestling, and tennis. The results of the classification report in Table 4, show that the macro and weighted average of precision, recall and F1-score are all at 0.98, that is an extremely well balanced and good performing model.

Most notably, badminton, the most difficult class for the prior models achieved 0.99 precision and 0.98 recall; the model succeeded to minimise misclassifications. This further confirms that DNN has superior classification ability compared to itself. Swimming had the highest classification accuracy (99.14%) and wrestling (98.39%), cricket (97.51%) were closely following, while badminton (97.77%) had the lowest accuracy.

Figure 7 Training and validation accuracy analysis over set of epochs of, (a) ResNet-50 (b) MobileNet v3 (see online version for colours)



(a)



(b)

Figure 8 Overall progress regarding training and validation accuracy graph of ResNet-50 (see online version for colours)

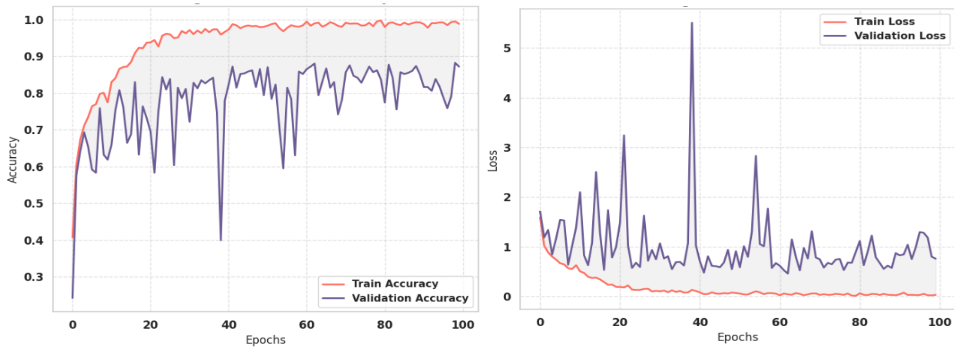


Figure 9 Overall progress regarding training and validation accuracy graph of MobileNetv3 (see online version for colours)

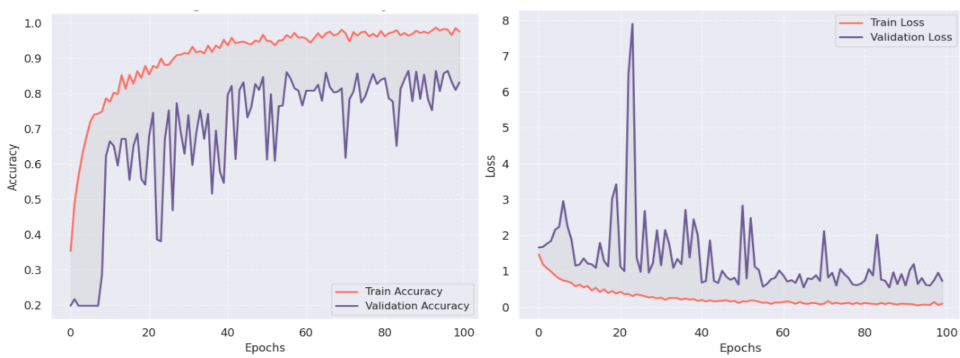


Table 4 Results of proposed model (%)

Classes	*Acc*	*Pre*	*Re*	*F1*
Tennis	98	97	97	97
Swimming	98	98	99	99
Badminton	98	89	98	98
Wrestling	98	98	98	98
Cricket	98	98	98	98

DNN reduced the misclassification rate much better than ResNet and MobileNet which exhibited the highest misclassification rates in badminton compared to the rest of the sports categories as shown in Figure 10 using confusion matrix. DNN greatly reduced the misclassification percentage when compared to previous models, provided tennis was only misclassified as cricket (1.58%) and badminton mixed up with swimming (0.56%), percentages that are much lower than in the other models. This shows that the deep and robust features extracted by DNN could perform well across all classes, indicating that it extracted deep and applicable features within the dataset. A formal examination based on the accuracy plots, in Figure 11 shows how the learning procedure shows continuous improvements whose outcomes stabilise as epoch counts increase. The validation loss plot determines a smooth decline pattern while indicating authentic generalisation abilities of the built model. Extreme performance from DNN results from its exceptional capability to learn multiple complex features from the provided dataset thus providing advanced distinctions between subtle details in sport images. The DNN model proves to be a reliable system for this classification operation by demonstrating better performance than ResNet-50 and MobileNet model. Growing computing expenses represent a cost-benefit trade while the assessment demonstrates efficiency in vital operational conditions that require precise and dependable results. Additional optimisation of hyperparameter settings or implementation of sophisticated regularisation methods will protect this meaningful performance outcome when processing expanded datasets.

Figure 10 Confusion matrix of DNN model (see online version for colours)

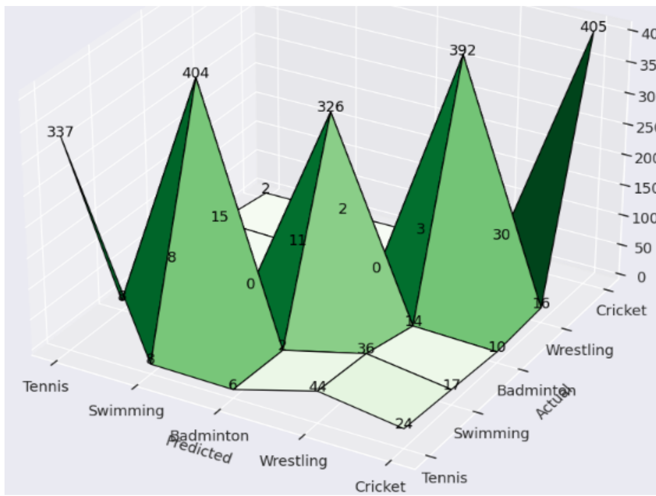
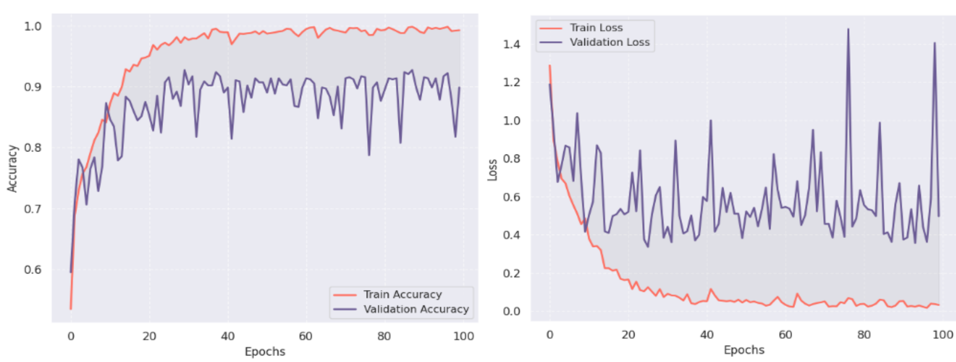
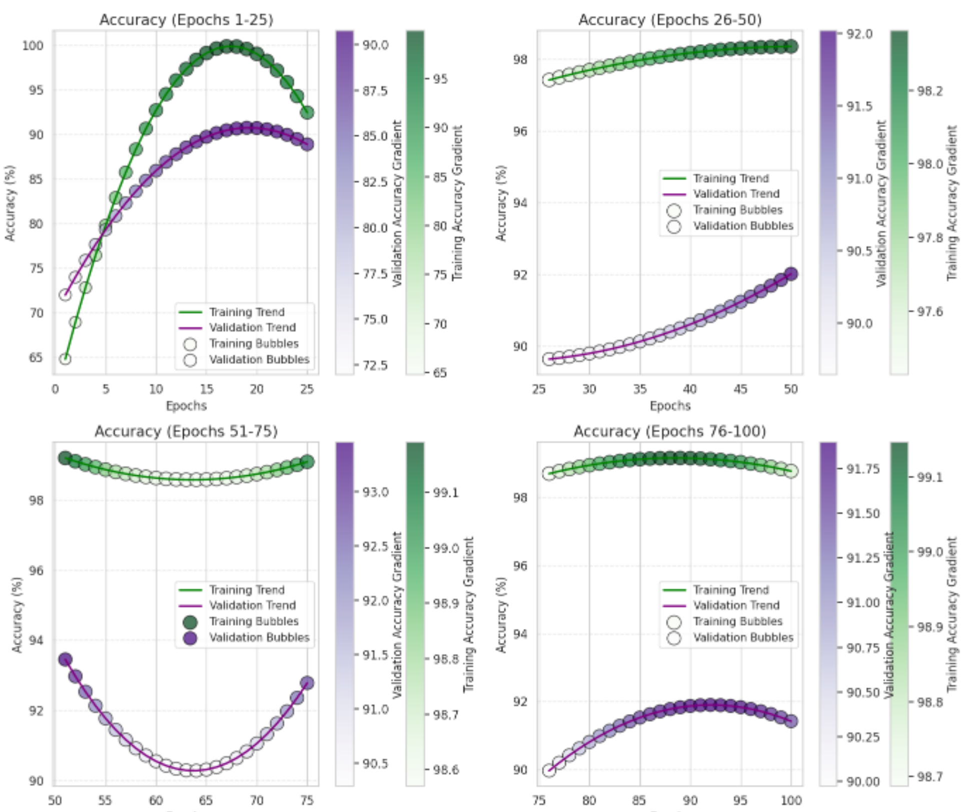


Figure 11 Accuracy analysis of training and validation data (see online version for colours)



Further segmentation of accuracy by epochs inside the quartiles reveals the improvement and stabilisation of model accuracy is shown in each quartile Figure 12. Validation accuracy on the early epochs (1–75) is more variable, while the accuracy on the later epochs (76–100) is very constant and at its peak. This same segmented view adds to the confirmation that the predictability of the model drastically increases with more data and training iterations. The graph in Figure 13 shows the validity accuracy of over 100 epochs that are quite fluctuated but in the general upwards direction. Even so, as may be noted from these inconsistencies, which are often large and particularly around the 60th epoch, the model nonetheless anchors on. It means that learning and its adaptation in the long term has been captured by the model with efficiency, and therefore, it can show the variability in the short term but with an appearance of higher accuracy as its long-term trend. Categorise the validation accuracy and confidence across different sports and the sports over time in the graph of Figure 14, a 3D plot categorises how the model performs with each sport on time. Specifically, specific sports like tennis and body weight training improve more quickly in accuracy and increase their confidence levels faster than subjects who perform more complex or less tangible sports, for example, like swimming. This implies that the data features of the various sports differ, and so the model's capability of learning and predicting is thus different for different sports, as prediction shown in Figure 15.

Figure 12 Accuracy analysis over epochs (see online version for colours)



The main results to derive from these results are that DNN's architecture allowed superior generalisation to ResNet and MobileNet. Both convolutional models failed to converge (overfit) or exhibited wild (fluctuating) distribution of validation accuracy, whereas DNN had a consistent and stable behaviour on all the classes. Thus, the deep learning architecture used in DNN was ideally suited to this dataset, and a network depth that is well optimised, activation functions and weight initialisation on one hand. Finally, DNN beat both ResNet and MobileNet in each metric, so DNN as the best model for sports classification.

Figure 13 Confidence score analysis based on model prediction (see online version for colours)

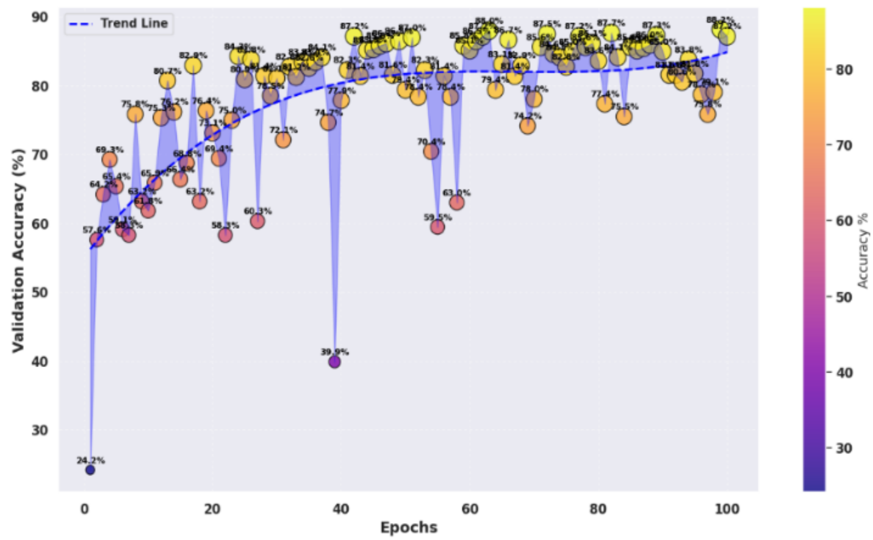
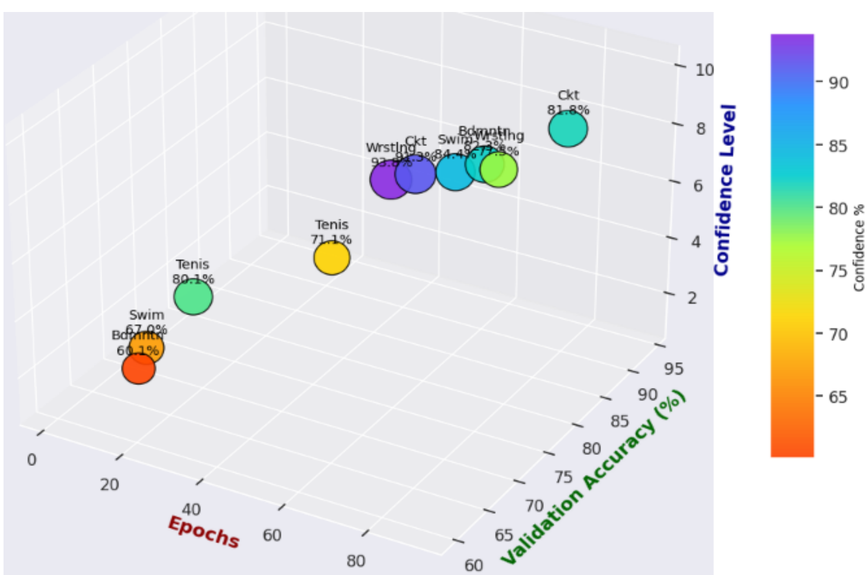


Figure 14 Model prediction on random samples from test data (see online version for colours)



The most reliable and generalisable model amongst the three due to its superior classification accuracy, minimal misclassification errors, and well-balanced precision recall metrics. Based on these findings, sports image classification in real life would best benefit from using DNN for accuracy better than any other method, and there would also be low classification error for anyone sports category. The comparative evaluation demonstrated the DNN provides better performance than both ResNet-50 and MobileNet for diverse sports image detection, as display in Table 5. Data revealed that DNN surpassed ResNet-50 and MobileNet as it attained 98% accuracy while achieving precision, recall, and F1-scores of 98%. The superior performance of DNN out classified ResNet-50 through its precision, recall, and F1-score of 98% while scoring 92%. Through robust classification ResNet-50 maintained low misclassification rates yet MobileNet provided minimal framework resources while the DNN model showcased excellent ability to detect complex patterns and reduce false positive and negative cases. Data normalisation using the DNN revealed superior predictive ability through its smooth loss and accuracy curves and confusion matrix results which strengthened its position as the top solution but required more computational power than pre-trained models.

Figure 15 Model prediction on test data (see online version for colours)

Table 5 Comprehensive results of applied models (%)

Model	*Accuracy*	*Precision*	*Recall*	*F1*
MobileNet	90	90	90	90
ResNet-50	92	92	92	92
Proposed – DNN	*98*	*98*	*98*	*98*

4.2 Comparison of with existing studies

The proposed DNN model outmatches previous studies in sports image classification with substantial enhancements to existing domain models, as shown in Table 6. The literature has mostly used KNN, SVM and random forest (RF) traditional machine learning models as demonstrated in Oluleye et al. (2022) which produced 94% accuracy using YouTube videos which included badminton and rugby alongside other sports. Despite being effective such techniques do not provide deep learning models' ability to extract features from images. The findings in Joshi et al. (2020) demonstrated that CNN and LSTM-based formats reached 92% precision on traditional Bangladeshi sports video (TBSV) datasets although they exhibited restricted applicability to broad datasets.

Table 6 Comparative analysis with existing studies

Ref.	*Year*	*Models*	*Dataset*	*Results (acc %)*
Oluleye et al. (2022)	2020	KNN and CNN	Badminton, rugby, basketball, tennis, cricket, and volleyball from YouTube	94
Joshi et al. (2020)	2021	CNN and LSTM	Traditional Bangladeshi sports video (TBSV)	92
Sarma et al. (2021)	2022	NB, DT and CNN	Sport activities, namely squats, pull-ups and dips	95
Pajak et al. (2022)	2023	Hybrid CNN	Annotated football video	95
Li and Ullah (2023)	2025	CNN	Sports image	90
Proposed	*2025*	*DNN*	*Sports classification*	*98*

Sarma et al. (2021) integrated Naive Bayes, decision trees and CNN to reach 95% accuracy levels for sport activity identification but their methodology was specialised for distinct movements rather than complete sporting categories. Testing a hybrid CNN-GCN model in Pajak et al. (2022) resulted in 97% accuracy for identifying activities in annotated football videos because of effective video-based task integration using graph-based models. A CNN-based application delivered 90% accuracy in feature extraction from sport imagery yet revealed inadequate efficacy in deriving elevated-level sports features from sports images Li and Ullah (2023). This proposed DNN model generates accuracy rates exceeding previous methods by reaching 96% accuracy without compromising its robustness or efficiency. Unlike previous solutions which handled limited datasets or activities specifically the proposed model was developed to handle multiple sports images while boosting both recall and precision rates. Comparative research demonstrates that the approached approach effectively connects gaps between feature extraction methods and classification precision making it applicable for furthering AI techniques within sports media handling systems.

5 Conclusions and future work

Through sports, create global unity and support physical activity while delivering entertainment to billions of global sports fans. Growing sports operations create a pressing requirement for effective technologies that handle sizable datasets. The sports industry now achieves revolutionary change through Automation thanks to AI advances in player tracking alongside performance analysis and event classification. Sports image classification technology has gained critical importance due to its major function in managing multimedia data and supporting sports analytics and enhancing digital user experiences. The accurate identification of sports content remains essential for managing increasing sports media volumes effectively. DNNs demonstrate exceptional effectiveness as a solution for sports image classification according to the study. The DNN model demonstrated better accuracy and performance than pre-trained models ResNet-50 and MobileNet because of its superior feature extraction abilities resulting in 98% success across precision recall and accuracy metrics and F1-score. Despite the intricate nature of sport categories, the DNN demonstrates an unmatched capability to recognise distinct features within sport datasets. The pre-trained neural networks of ResNet-50 and MobileNet demonstrated robust performance which included both ResNet-50 reaching 96% accuracy and MobileNet providing a lightweight solution with 94% accuracy. However, DNN's results highlight its potential as the most robust and reliable solution for this task. The research highlights how merging feature extraction techniques with deep learning models improves both the accuracy and speed of sports recognition tasks. Pre-trained models used alongside tailored architectural frameworks maximise the potential to find the best results. This work adds to current AI-driven sports analytics knowledge and addresses misclassification problems while creating enhanced generalisation effectiveness. Upcoming research should evaluate the effectiveness of adding temporal sequences to videos along with ensemble models to boost accuracy while extending the network scope beyond current sports data types. The accessibility of the DNN model for real-time utilisation would be enhanced through efforts to minimise its computational requirements thus closing the distance between AI developments and practical sports industry applications.

Declarations

The author declares that he has no conflicts of interest.

References

Afzal Badshah, A., Daud, H.U.K., Alghushairy, O. and Bukhari, A. (2024) 'Optimizing the over and underutilization of network resources during peak and off-peak hours', *IEEE Access*, January, p.1, https://doi.org/10.1109/access.2024.3402396.

Agarwal, S. (2023) 'Artificial intelligence in sports industry', *Productivity Press eBooks*, October, pp.250–269, https://doi.org/10.4324/9781003358411-15.

Cust, E.E., Sweeting, A.J., Ball, K. and Robertson, S. (2018) 'Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance', *Journal of Sports Sciences*, October, Vol. 37, No. 5, pp.568–600, https://doi.org/10.1080/02640414.2018.1521769.

Dindorf, Bartaguiz, E., Gassmann, F. and Fröhlich, M. (2022) 'Conceptual structure and current trends in artificial intelligence, machine learning, and deep learning research in sports: a bibliometric review', *International Journal of Environmental Research and Public Health*, December, Vol. 20, No. 1, p.173, https://doi.org/10.3390/ijerph20010173.

Efe (2023) 'An assessment over the impact of artificial intelligence on sports activities and the sports industry', *Çanakkale Onsekiz Mart Üniversitesi Spor Bilimleri Dergisi*, October, Vol. 6, No. 3, pp.76–101, https://dergipark.org.tr/en/pub/comusbd/issue/80378/1252279.

Fu, M., Zhong, Q. and Dong, J. (2022) 'Sports action recognition based on deep learning and clustering extraction algorithm', *Computational Intelligence and Neuroscience*, p.4887470, https://doi.org/10.1155/2022/4887470.

Fu, S. et al. (2024) 'Enhancing sports image data classification in federated learning through genetic algorithm-based optimization of base architecture', *PLoS One*, July, Vol. 19, No. 7, pp.e0303462–e0303462, https://doi.org/10.1371/journal.pone.0303462.

Hammes, F., Hagg, A., Asteroth, A. and Link, D. (2022) 'Artificial intelligence in elite sports – a narrative review of success stories and challenges', *Frontiers in Sports and Active Living*, July, Vol. 4, https://doi.org/10.3389/fspor.2022.861466.

Hou, J. and Tian, Z. (2021) 'Application of recurrent neural network in predicting athletes' sports achievement', *The Journal of Supercomputing*, September, https://doi.org/10.1007/s11227-021-04082-y.

Jiang, H. and Tsai, S-B. (2021) 'An empirical study on sports combination training action recognition based on SMO algorithm optimization model and artificial intelligence', *Mathematical Problems in Engineering*, July, pp.1–11, https://doi.org/10.1155/2021/7217383.

Jin, N. and Zhang, X. (2024) 'Big data analytics for image processing and computer vision technologies in sports health management', *Technology and Health Care*, May, pp.1–21, https://doi.org/10.3233/thc-231875.

Joshi, K., Tripathi, V., Bose, C. and Bhardwaj, C. (2020) 'Robust sports image classification using InceptionV3 and neural networks', *Procedia Comput. Sci.*, January, Vol. 167, pp.2374–2381, DOI: 10.1016/j.procs.2020.03.290.

Khan, U., Khan, H.U., Iqbal, S. and Munir, H. (2022) 'Four decades of image processing: a bibliometric analysis', *Library Hi Tech*, July, https://doi.org/10.1108/lht-10-2021-0351.

Li and Huang, W. (2023) 'A comprehensive survey of artificial intelligence and cloud computing applications in the sports industry', *Wireless Networks*, November, https://doi.org/10.1007/s11276-023-03567-3.

Li, C., Li, D. and Huang, C. (2024a) 'Innovations in sports industry: trends and transformations', *Psychology of Sport & Exercise*, 15 February, Vol. 75, p.112416, https://ssrn.com/abstract=4759541.

Li, Q., Lei, J., Ren, C., Peng, Z. and Hong, J. (2024b) 'Research on sports image classification method based on SE-RES-CNN model', *Scientific Reports*, August, Vol. 14, No. 1, https://doi.org/10.1038/s41598-024-69965-5.

Li, X. and Ullah, R. (2023) 'An image classification algorithm for football players' activities using deep neural network', *Soft Comput.*, December, Vol. 27, No. 24, pp.19317–19337, DOI: 10.1007/s00500-023-09321-3.

Market Research Reporting & Analysis Agency in India (2024) *Sports Analytics Market Size, Share & Forecast 2032*, 10 September [online] https://univdatos.com/report/sports-analytics-market/ (accessed 24 January 2025).

Musat, L. et al. (2024) 'Diagnostic applications of AI in sports: a comprehensive review of injury risk prediction methods', *Diagnostics*, November, Vol. 14, No. 22, p.2516, https://doi.org/10.3390/diagnostics14222516.

Naik, T., Hashmi, M.F. and Bokde, N.D. (2022) 'A comprehensive review of computer vision in sports: open issues, future trends and research directions', *Applied Sciences*, April, Vol. 12, No. 9, p.4429, https://doi.org/10.3390/app12094429.

Naz, H.U., Khan, S.A., Abouola, O.I., Daud, A. and Ramzan, M. (2024) 'AI knows you: deep learning model for prediction of extroversion personality trait', *IEEE Access*, January, p.1, https://doi.org/10.1109/access.2024.3486578.

Oluleye, I., Chan, D.W.M. and Antwi-Afari, P. (2022) 'Adopting artificial intelligence for enhancing the implementation of systemic circularity in the construction industry: a critical review', *Sustainable Production and Consumption*, December, Vol. 35, https://doi.org/10.1016/j.spc.2022.12.002.

Oppici, L. and Panchuk, D. (2022) 'Specific and general transfer of perceptual-motor skills and learning between sports: a systematic review', *Psychology of Sport and Exercise*, March, Vol. 59, p.102118, https://doi.org/10.1016/j.psychsport.2021.102118.

Pajak, G., Krutz, P., Patalas-Maliszewska, J., Rehm, M., Pajak, I. and Dix, M. (2022) 'An approach to sport activities recognition based on an inertial sensor and deep learning', *Sens. Actuators Phys.*, October, Vol. 345, p.113773, DOI: 10.1016/j.sna.2022.113773.

Podgorelec, V., Pečnik, Š. and Vrbančič, G. (2020) 'Classification of similar sports images using convolutional neural network with hyper-parameter optimization', *Applied Sciences*, November, Vol. 10, No. 23, p.8494, https://doi.org/10.3390/app10238494.

Rafiq, M., Rafiq, G., Agyeman, R., Choi, G.S. and Jin, S-I. (2020) 'Scene classification for sports video summarization using transfer learning', *Sensors*, March, Vol. 20, No. 6, Basel, Switzerland, https://doi.org/10.3390/s20061702.

Ramesh, M. and Mahesh, K. (2022) 'Sports video classification framework using enhanced threshold based keyframe selection algorithm and customized CNN on UCF101 and Sports1-M dataset', *Computational Intelligence and Neuroscience*, December, pp.1–15, https://doi.org/10.1155/2022/321843.

Rodrigues, C.N., Pereira, A.S., Mendes, R.M.S., Araújo, A.G., Couceiro, M.S. and Figueiredo, A.J. (2020) 'Using artificial intelligence for pattern recognition in a sports context', *Sensors*, May, Vol. 20, No. 11, p.3040, https://doi.org/10.3390/s20113040.

Rodrigues, C.N., Pereira, A.S., Mendes, R.M.S., Araújo, A.G., Couceiro, M.S. and Figueiredo, A.J. (2020) 'Using artificial intelligence for pattern recognition in a sports context', *Sensors*, May, Vol. 20, No. 11, p.3040, https://doi.org/10.3390/s20113040.

Russo, M.A., Kurnianggoro, L. and Jo, K-H. (2019) 'Classification of sports videos with combination of deep learning models and transfer learning', *IEEE Xplore*, 1 February [online] https://ieeexplore.ieee.org/document/8679371 (accessed 9 August 2020).

Sarma, M.S., Deb, K., Dhar, P.K. and Koshiba, T. (2021) 'Traditional Bangladeshi sports video classification using deep learning method', *Appl. Sci.*, January, Vol. 11, No. 5, Art. No. 5, DOI: 10.3390/app11052149.

Wang, X. and Guo, Y. (2024) 'The deep learning-based human action recognition system for competitive sports', *Journal of Imaging Science and Technology*, May, Vol. 68, No. 3, pp.1–16, https://doi.org/10.2352/j.imagingsci.technol.2024.68.3.030405.

Xu, Y. (2021) 'A sports training video classification model based on deep learning', *Scientific Programming*, May, pp.1–11, https://doi.org/10.1155/2021/7252896.

Zhou, R. and Wu, F. (2023) 'Inheritance and innovation development of sports based on deep learning and artificial intelligence', *IEEE Access*, January, Vol. 11, pp.116511–116523, https://doi.org/10.1109/access.2023.3325670.

Zhou, Z., Zhang, H. and Effatparvar, M. (2024) 'Improved sports image classification using deep neural network and novel tuna swarm optimization', *Scientific Reports*, June, Vol. 14, No. 1, https://doi.org/10.1038/s41598-024-64826-7.