

# Hybrid Support Vector Machine Integrated Convolution Neural Network Algorithms for Analyzing Body Motion Data for Predicting Leg Injuries

Asha Prashant Sathe<sup>1</sup>, Dr P.Satyanarayana Raju<sup>2</sup>, Dr.M.V.S.S.N.Murthy<sup>3</sup>, Chandra Sekhar Koppireddy<sup>4</sup>, Vijay Kumar Gubbala<sup>5</sup>, Pamula Udayaraju<sup>6</sup>

<sup>1</sup>Department of Computer Engineering, Army Institute of Technology, Pune.

<sup>2</sup>Department of Physical Education, SRKR Engineering College, Bhimavaram, India.

<sup>3</sup>Department Of Physical Education, Adikavi Nannaya University, Rajahmundry, India.

<sup>4</sup>Department of CSE, Pragati Engineering College, Surampalem, India.

<sup>5</sup>Department of CSE, Pragati Engineering College, surampalem, India.

<sup>6</sup>Department of Coputer Science and Engineering , SRM University-AP, Amaravathi, Andhrapradesh, India.

## Abstract

Computer vision is also a key feature for the complete automation of applications, and various research is made to improve the performance of the prediction models to use them in better applications. Computer vision is used to summarize videos and also to analyze the activities of the subjects in the video. Accurate prediction of the subjects helps evaluate the action and provide suggestions through AI. Physical education is an essential application of computer vision, where the players' activities need to be monitored, and real-time recommendations need to be given on their practices. It helps in reducing the sports injuries of the players. This research proposes a hybrid deep learning model, which uses a CNN algorithm to predict the activities of the players and SVM to provide real-time suggestions to the players to avoid sports injuries. A model dataset is considered and used to train the model. Important physical education activities are identified, and relevant data is used to train the CNN algorithm. Based on the prediction, the SVM algorithm classifies the players' activities and provides suggestions for physical education. The model's performance is compared with earlier ones in terms of prediction accuracy. The proposed model offers 89.4 % accuracy in the prediction process for predicting different types of activities that influence physical education.

**Keywords:** Physical Education, Sports Injury, CNN, SVM, Computer Vision, AI.

## Introduction:

In the sports field, computer technology has proved to have great potential. The computer uses the sports posture corrections vision-based virtual. This vision-driven system evaluation was adopted from the decision-making in their training period. Moreover, it can be used for object detection and sports analysis. It helps to improve the performance of athletes. Record raw data is a key position for an athlete. By using a classical method like placing the sensor. The guidelines for the training purpose provided data-driven by using the data science method to analyze. Moreover, the performance of the athletes is noted, and increase the cost of additional sensors for the efficiency and records of the impaired athletes. However, the person cannot be asked by their competitors to analyze the strengths and weaknesses of the competitors.

In the process of improving, athletes are more popular for video analysis and computer vision technology for their efficiency and to prevent injuries. The rapid feedback in the training for effective and minimized accidents for the athletes can be enabled by contactless technology. It also helps the coaches and athletes analyze the strengths and weaknesses of opponent's teams from their past matches videos and the strategies designed in future competitions.

The detection players are analyzed with their step count. With this technology, the development of the convolutional neural network is based on the Football video algorithm. The input in consecutive convolution layers is transformed into sequential images. The active function of these layers is normalized and leaky ReLU to extract different feature levels of the special resolution. The

detection process will proceed with all levels of feature maps around the residual connection of the first three convolutional blocks followed for the normalization. Combining the up-sampling feature maps with feature maps of lower levels results in a player confidence map and a player bounding box [1].

To obtain these features and algorithms for the image processing and machine learning is accomplished by computer vision. To manipulate the image processing into the low-level feature and normalize their data for data augmentation and removal of digital noise. The automation image task for the content and drawing in the computer vision task is the most important in the subfield of deep learning.

Computer vision tasks mostly detect object detection, tracking, and action recognition. It mainly focused on supervised learning tasks. This relies on the learning algorithm for mapping the raw data and the desired label in the training stages. Supervised learning is the preparation of data collection and labeling, feature extraction, and preprocessing. The learning algorithm is followed by the splitting of data into a set of testing and training sets. Performance growth is tested and obtained from the testing and validating models. The performance represented with different appropriate metrics is evaluated for the specific task.

In the sports domain, the image and video detection analysis are analyzed using the computer vision task.

To proceed with computer vision, machine learning, and image processing take a major role. Image processing will normalize the photometric on the low-level features to manipulate the image data. These normalized properties are visual data, data argument, and removing digital noise. This concept is not concerned with the content of visual data understanding. Moreover, the automation task in computer vision can interpret the drawing and content conclusion of the image in machine learning and subfield deep learning (DL).

This computer vision task with little automation involved the required coding and manual effort for extracting images. When the introduction of the DL method as a convolutional neural network extracting the required codes automatically without

manual effect in the process. Even though This CNN framework can process automatically, it still needs manual features for both argumentation and task performance for temporary video detection of active players [2].

### **Contributions**

Various research works are proposed to detect body motion and its actions. It can be seen that various research works have proposed deep learning-based model to analyze the motion of objects in the video and are used to summarize their activities in the videos. This paper proposed a hybrid CNN model that learns the movements of the objects. In the case of application, the movement of players in the videos and their actions are recorded, their unhealthy physical exercise practices are predicted, and suggestions are provided for healthy practices to the players. This paper makes the following contributions,

1. A CNN algorithm is proposed to process the video frames for predicting the players' actions.
2. The actions are recognized and are verified for healthy and unhealthy practices.
3. The proposed model suggests countermeasures to the players based on their physical activities.
4. The effectiveness of the model is verified by comparing it with similar works.

### **Literature survey**

L. Zhang (2022), this analysis concentrates on enhancing Human Motion Recognition (HMR) in sports utilizing the DL approach. They submitted their experience and prior position on HMR. Then, they designed a new HMR technique named KELM-MFF, which performed well in previous systems. The outcomes demonstrated that KELM-MFF had the most elevated recognition speed reached by other algorithms at ten diverse points in a video. It was better than methods like SVM-MFF, CNN-T, IDT, C3D, and CNN. This suggested algorithm is speedy, with 30 seconds for training and testing for 15 seconds. It can manage large datasets and helps capture and recognize minor signs in sports. Finally, this analysis can allow the use of the extreme ML approach in sports competitions.

J. Luo et al. (2022) designed a new approach for following the foot moves of badminton players. Hence, they employed a DL to track the 2D

functions of their shoes and binocular vision to estimate the 3D coordinates. Our technique reaches 74.7% accuracy. Utilizing this technique, they analyzed how participants adjust their footwork during competitions. They discovered that players changed their length and jump height to intercept the shuttlecock. This information delivers functional understandings for training programs, permitting coaches to tailor training to particular players' requirements and enhance their badminton footwork. This process shows a more quantitative and precise analysis than traditional visual observations by biomechanics and motor managing professionals. S. Mekruksavanich & A. Jitpattanakul (2022) Wearable sensors are suitable for identifying primary activities like walking or sitting but struggle with challenging exercises like sports. To handle this, a suggested technique, wearable sensors in distinct body positions, was designed for identifying sports-related actions. They tested this method employing a UCI-DSADS dataset and reached five DL methods. Finally, the outcomes indicate that the proposed model performed far better than the other methods, with an accuracy rate of 99.62%. This advance is essential for real-world applications like remote rehabilitation monitoring and tracking sports implementation. M. Fu et al. (2022) This investigation uses the DL approach and a clustering extraction process to create a sports activity recognition method. It identifies athletes' actions from video frames utilizing DL, incorporates their sports actions, and improves learning with training data. The neural network then specifies non-athlete pictures and enhances the standard iteratively based on false positives using clustering. Simulation investigations indicate that this suggested method achieved and performed well in recognition accuracy, false alarms, and speed. The objective is to examine athletes' activities to support coaches in providing more experienced training and enhance sports action recognition. In this work, S. Pan (2022) focused on assisting trainers in instructing athletes more effectively by recognizing basic training poses. They present a technique named RoI\_KP, which utilizes DL to categorize regions of interest in sports videos. They fine-tune a CNN to classify weightlifting actions in these regions. Then, they use the classification outcomes to remove the

required poses. Finally, the suggested approach demonstrates to be highly competitive in its implementation. J. Fan et al. (2022) presented a new approach known as SCGA, which is trained to identify distinct sensor-fusion basketball shooting positions. It includes SqueezeNet for spatial features, gated recurrent divisions for time-related features, and an attention mechanism to consider the significance of features. They have concluded that the suggested approach achieves high accuracy rates of around 98.79% for intra-testing and 94.06% for inter-testing. I. Dirgová Luptáková et al. (2022) suggested the transformer model, initially developed for language and vision tasks, to analyze motion movements over time. The transformer's self-attention characteristic catches associations between data matters in a time series, and a developed CNN with LSTM is performed. The adjusted transformer reached an outstanding 99.2% accuracy in specifying actions and got 89.67% with previous ML methods. They have concluded that the suggested approach has a hopeful future in human activity recognition. F. Kulsoom et al. (2022) examine Human Activity Recognition (HAR) to determine the numerous sufficient ML approaches. It suggests challenging which algorithms are well-suited for specific applications within HAR. Also, it analyzes the general usage of vision-based and non-vision-based data collection devices in existing analysis, evaluating their usefulness for thorough HAR tasks. Y. Sun and Y. Li (2022) present a process for investigating motion training attitudes employing a multiscale spatiotemporal graph convolution network. It begins by making a spatiotemporal skeleton picture and then involving convolution functions. The outcomes of these convolutions are integrated to capture the features of activities with contrasting time durations. This approach also manages substantial data loss and improves dataset diversity. Investigations show that this system adapts to diverse behaviors, substantially enhancing model implementation and recognition accuracy. K. Host and M. Ivašić-Kos (2022) The HAR analysis initially centered on day-to-day actions employing publicly known datasets. Yet, there is a growing change towards involving HAR in sports, especially within Computer Vision. This work proposes a summary of HAR applications in sports,

emphasizing Computer Vision as a prominent direction, and recognizes generally utilized general datasets in this context.

L. Tong et al. (2022) designed a novel DL model for HAR employing inertial detectors, concentrating on the CATP. The model, known as Bi-GRU-I, featured two Bi-GRU layers and three Inception layers and gained enhanced precision with more periodic parameters. Comparative investigations across three datasets, like CATP, showed outstanding implementation and robustness. Also, sensor design optimization demonstrated the possibility of utilizing the way with fewer sensor units.

Using deep learning, C. Qin and S. Huo's (2022) analysis presents a method to identify limb movements during competitive sports. It computes force differences in limb activities during sports to choose them and employs a DL approach to comprehend their features. The investigations revealed that this technique dramatically enhances detection precision and speed in following limb movements during competitive sports training. J. Liu and B. Liang (2022) submit a method for identifying badminton actions utilizing acceleration and angular velocity movements. The investigations indicate that this approach performs recognition precision of 63%, 84%, and 92% on diverse benchmark datasets, making considerable progress in identifying how humans hit in badminton. These findings suggest coming work in improving DL methods for even better outcomes in badminton action recognition. G. Şengül et al. (2022) designed a smartwatch-based fall detection method to determine falls from everyday actions. It can even differentiate between falls from a chair and a standing position. Our mobile app manages sensor data processed in the cloud utilizing DL. They use cubic Hermite interpolation to improve precision and compute 38 statistical components for input. They have concluded that the proposed accuracy is 99.59%. J. Xiao et al. (2022) present a DNN approach for identifying basketball actions utilizing a dynamic residual attention mechanism. It improves the standard C3D by enhancing feature extraction with a dynamic residual convolution network. The above-said mechanism selects important video frames from the removed components. Comparative investigations against the conventional C3D emphasize the strategy's

progress and significance, performing an average posture recognition exactness of over 97%.

#### **Limitations and Motivation**

With the development of computer vision models, object detection has considerably improved over the past decade, with fast-tracking of the subject in a video with better accuracy. Most of the research works have considered LSTM models for predicting objects in videos. Detecting the motion of the objects in a network is difficult for the CNN algorithms, as they are mostly employed in the image recognition process and are incapable of processing continuous time-series data and comparing them. Hence, LSTM models are widely preferred by researchers. The reinforcement learning-based modules are widely considered for video processing tasks for better processing of the time series data. Along with that, in computer vision applications, understanding the actions of the objects is also required for summarizing the context of the video. Various research works have used these summaries to predict and analyze the subject's activities. However, based on the actions of the individuals, a real-time analysis is carried out, and suggestions are given based on their activities. Though it is only in the development stage, it needs to be adopted for several applications like sports, education, and law.

#### **Proposed approach**

In this study, a deep learning-based CNN model is used to track and detect the body motion of athletes to prevent them from injuries. The spatial information of the input body motion samples is detected based on the movements in the muscles and surface of the skeletons. To extract this information from the body motion, the proposed CNN model includes five layers: input, output, convolutional, pooling, and fully connected. The input layer is used to gather the input signal, the convolutional and pooling layer is used to process the input signal, the fully connected layer is used to classify the input data, and the output layer is used to produce the output. To extract the spatial information from the input signal, a 3D CNN model is proposed. The basic structure of the CNN model is depicted in figure-1.

### Convolutional Neural Network (CNN)

CNNs define a technical structure of the DL framework that combines features from ANN and DL networks. CNNs excel in identifying two-dimensional images and submit two further layers reached to previous neural networks: the convolutional and the pooling layers. This neural network paradigm discovers comprehensive applications in image data processing, showing remarkable proficiency in removing valuable components from images. Also, CNNs have achieved traction in various domains of further image research, including but have yet to define NLP and numerous regions within the biological sciences.

Before supervised training, the CNN request values are initialized utilizing diverse small random numbers. Training a folded neural network applies a two-stage method.

During the pre-propagation stage, they remove models from the dataset. The network carries input A, uses weight Q, and involves mapping function F to transfer data from the input layer to the output layer via a one-step transformation. After that, they sum the precise corresponding outcomes.

$$o_p = F_n \quad (1)$$

At that time phase of backtracking, the contrast between the actual output and the outcome of an ideal is calculated.

$$E_p = \frac{1}{2} \sum_j (B_{pj} - o_{pj})^2 \quad (2)$$

The weight matrix is then modified in a method that reduces error.

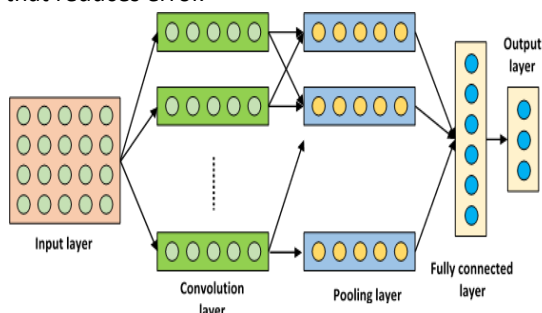


Figure-1 Basic structure of the CNN model

**Convolutional layer:** It is the major building block in the CNN model, mainly used to perform feature extraction. The features of the input data are extracted using the convolutional kernels or filters. The process of the convolutional layer is performed by multiplying the raw inputs with the extracted data and summing them to produce the final

output. The structural process of the convolutional layer is clearly shown in figure-2. At first, the calculation starts with the upper left corner and ends with the lower right corner, and then the output is generated. After extracting the features from the input data, the feature-extracted data are transferred to the pooling layer to perform further tasks.

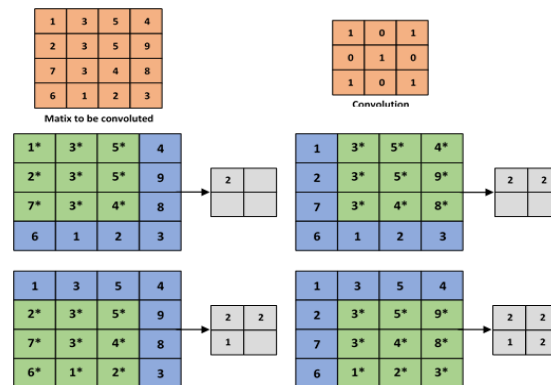


Figure-2 Process Of Convolutional Layer

**Pooling layer:** In CNN architecture, the pooling layer is mainly used to reduce the dimensionality of the feature-extracted input data. The major process of the convolutional layer is to extract the essential features from the input samples. Likewise, the pooling layer is applied for two major benefits: eliminating the irrelevant data and reducing the computational complexity by minimizing the data dimensionality. This will also improve the processing speed of the model. Pooling layers are classified into average, maximum, and mean. Of these, max-pooling is one of the most common types of methods widely used in many applications. The example  $2 \times 2$  max\_ pooling sliding window is shown in Figure 3. Figure-3 illustrates how the pooling layer reduces the dimensionality of the input data. The result shows a step-size window extracted from the input data.

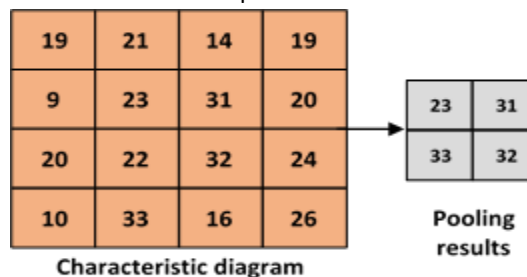


Figure-3 Process of pooling layer

### Activation Layer

CNNs depend on linear functions during convolution. When multiple matrix expansions are

executed successively, the prevalent modification remains linear. CNNs contain an important non-linear element called an activation function to handle the challenge of non-linear classification tasks. Typically employed activation functions have Sigmoid and ReLU, among others. This activation process presents nonlinearity to the network, allowing it to manage non-linear patterns and solve complex classification problems effectively.

The sigmoid-type activation operation emulates how biological neurons perform. It triggers when the input motion gets a specific threshold. This process compresses the output value into the field between the values (0, 1), as Figure 5(a) illustrates the sigmoid function diagram. The subsequent formula represents its mathematical term:

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}}. \quad (3)$$

However, there's a typical problem of vanishing gradients when reversing solutions via the network. Correcting the network's parameters effectively utilizing gradient descent becomes problematic when this occurs. To handle this issue and to reduce gradient flow at the last points, the sigmoid function is substituted with the ReLU method, as depicted in Figure 5(b). The essential benefit of ReLU is that its result is  $1 < 0$  and  $1 > 0$ , which mitigates the issue of vanishing gradients.

#### Fully Connected Layer

The fully connected layer arrives after the convolutional layer, which strives to obtain the local components extracted through multiple convolutional functions. These local elements are reduced via densely associated neurons, including a weight matrix to incorporate them into a complete feature model. The characteristic technique utilizes the softmax function as the final output layer for classification tasks. This method authorizes the developing probabilities for each classification, finally achieving the input data classification. The basic structure of the fully connected layer is shown in figure-4.

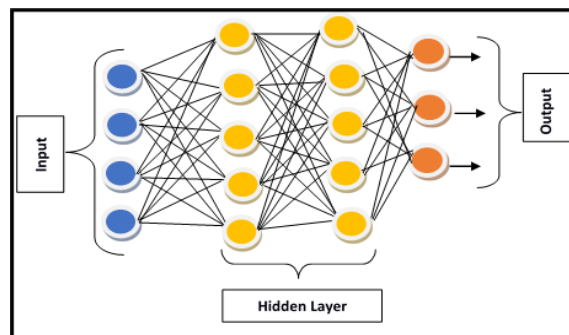


Figure-4 Structure Of Fully Connected Layer

The following steps, such as data preprocessing, data splitting, feature extraction, and data classification, are followed to classify the body motion of the players. Figure-5 depicts the overall function of the proposed CNN-based injury prevention model.

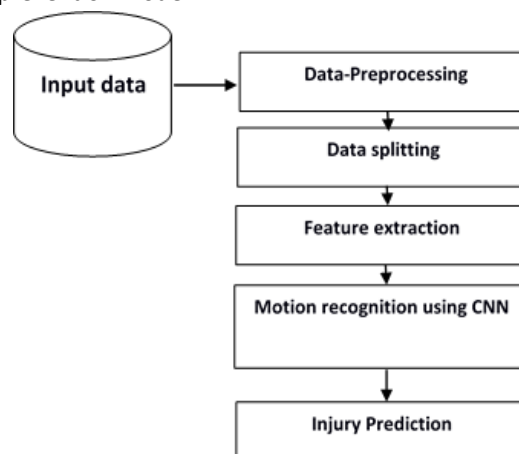


Figure-5 Overall Workflow Of The Proposed Work  
Data preprocessing

In injury prediction models, data aggregation arises over a sliding time window. This process is selected because muscle injury from over-training usually doesn't originate only from an escalation in the training workload over a session sequence. It can happen independently. Therefore, this perfection utilizes data aggregation to identify discrepancies in injury-related data. This data is collected from diverse training sessions, making recognizing practices associated with injury risk more painless. This step is specially performed to reduce the noise in the input samples using the kernel filters and using the following equation-4, and the data preprocessing process is performed.

$$X_T = \frac{(X_{T-2}+X_{T-1}+X_T+X_{T+2})}{n} \quad (4)$$

In equation-1,  $X$ ,  $T$ , and  $n$ , represent the input samples, time, and total number of samples. After denoising the input data, the samples are split into

two phases. The following section discusses the steps involved in the proposed model for detecting athlete body motion.

**Split Training and testing data**

Maintaining the documented hierarchy of activity sequence data is essential for sports prediction studies since game results are indicated based on past games. When post-test series and test examples apply mixed data during training, this can show an inappropriate partition, commonly resulting in problems with the precision of motion forecast results. Utilizing an outer split between the training and test datasets is more appropriate, maintaining the order of models.

**Feature Selection**

Feature selection is an essential phase in sports player injury prediction. It concerns the automated or manual designation of the reading's multiple practical aspects or result processes. By eliminating redundant data, they need to be aware of the possibility of creating conclusions affected by noise. Also, lowering the number of data points reduces algorithm complexity and faster activity speeds. Generally, the features are extracted from the denoised input data based on reference points like the player's shoulder, hip, foot, and muscle movements. The following equation-(5) is expressed to extract the features from the input motion data.

$$v_{TK} \triangleq x_T - R_K \quad (5)$$

The above equation is performed based on the reference points evaluated from the input samples. As mentioned above, T, X, R, V, and K represent the time, input point, coordinate value of the reference point, extracted features, and Kth reference point in the input samples, respectively. After extracting the features from the input samples, the proposed DL-based CNN model is applied to detect athlete body motion.

**CNN-based athlete body motion detection**

The proposed CNN-based human body motion detection technique incorporates the features of the pyramid network to enhance the model's performance, and based on the mapped features, the sequential network is used. Then, the interested region of the mapped features is transmitted to the fully connected to detect the changes in the player's body. The proposed model uses the backpropagation model to learn the input data in

this research work. to learn the input data, initially using filters, the input samples are preprocessed. Then, the preprocessed data are provided as input to the model, and forward calculation is performed to evaluate the value of the activation function. The activation function value of the proposed model is calculated using the equation (6). Then, using the reverse error method, the weight and bias value of the gradient is evaluated.

$$y_i^n(m) = f \quad (6)$$

Here, n represents the number of layers, i and m represent the feature graph of the layer and neurons, respectively,  $x_i^n(m)$  and  $y_i^n(m)$  denotes the input and output data. Using the following equation (7) the pooling and convolutional layer activation function is expressed.

$$f(x) = \text{atanh}(bx) \quad (7)$$

In equation-8 a and b are the constant elements. In the fully connected and output layer in the proposed CNN model used the sigmoid function as the activation function, which is expressed as:

$$f(x) = \frac{1}{1+e^{-x}} \quad (8)$$

In the proposed CNN model, the relationship between each layer is data transfer. Based on the following step, the relationship between the layer is emphasis: at first, the data in the input layer is sampled and convoluted using the filters in the convolutional layer, it is expressed as:

$$y_i^2(m) = f \left( \sum_1^l Y_{l,m} X k_i^2 + b_i^2(m) \right) \quad (9)$$

Where,  $Y_{l,m}$  indicates the sampled data,  $k_i^2$  indicates the convolutional core,  $b_i^2(m)$  denotes the network offset value. after per-processed the input sampled data in the convolutional layer, it is transferred to the pooling layer. The data transmission from the convolutional to pooling layer is expressed as equation (10)

$$y_i^3(m) = f \left( \sum_1^n y_i^2 \left( ((m-1) \times 10 + n) \times k_i^3 + b_i^3(m) \right) \right) \quad (10)$$

In the equation-5,  $k_i^3$  indicates the convolutional kernel. The overall neurons in the proposed detection model is connects using the connection and pooling layer, which is performed through following equation-11.

$$y^4(m) = f \left( \sum_1^l \sum_1^q y_i^3(q) \omega_i^4(q) + b^4(m) \right) \quad (11)$$

$\omega_l^4(q)$  and  $b^4(m)$  denotes the wight and bias value of the neurons in the pooling layer. then to establish the connection between the fully connected and output layer to produce the player motion classification result, equation-12 is used.

$$y^5(m) = f\left(\sum_1^l y_l^4(l)\omega_l^5(l) + b^5(m)\right) \quad (12)$$

To set the actual weight and bias value into the proposed trained CNN model, the learning rate of the functional convolutional and pooling layers is fixed as:

And to produce the final classification result the learning rate of the end layers such as: fully connected and output layer is set as:

**Prediction**

After finding the average learning rate, weight, and bias value of the convolutional, pooling, fully connected, and output layer, the joint and end point of the body motion is detected from the grey scale images extracted from the input sports video. It is expressed using the following formula-13.

$$TF = K - P \quad (13)$$

At last, based on the reference point of the all extracted images, the human body motion is recognised. Using the equation-, the information regarding the body motion to prevent athletes from injuries is expressed as equation-(14).

$$M_T = x_{T+1} - X_T \quad (14)$$

In equation-  $M_T$  and  $X_T$  represent the body motion information and classification result obtained after preprocessing. Using these set of value, the final healthy and unhealthy body condition of the athlete is detected.

**Multi-class SVM for prediction**

SVM is one of the widely considered model for decision making process, as it provides better classification of data with different classes. It maps the values of different classes in a high dimensional feature space. A hyperplane is drawn between different classes. A linear separator is constructed to evaluate the distance between different classes. It can be seen that, the predictions made by the CNN algorithm need to be classified, for which the multiclass SVM is used. The SVM model is used to label the data as positive or negative based on the objective of the application. It is capable of processing the higher dimensional space. Each row

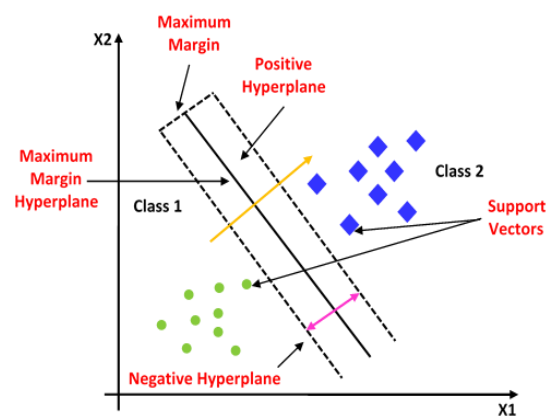
in the input space is mapped in the high-dimensional space, and based on the number of attributes the dimensionality is defined. It can solve both linear and non-linear problems effectively. Most of the attributes considered for the evaluation of physical education are binary in nature, and let us consider N number of training samples. Each sample is represented as  $(x_i, y_i)$  where,  $(i = 1, 2, \dots, N)$ , and the  $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$  represent the attribute set up to the  $i$ th sample. It is general that  $y_i \in \{-1, 1\}$ , where the  $y$  represents the class label allotted to the respective group of data. The hyperplane that classifies the different classes through the linear classifier is written as,

$$w^T x + b = 0,$$

The weight of the class is  $w$  and the bias term  $b$ . Similar to this, different types of linear separators are carried out based on the application. The decision boundary drawn between two different classes is represented as the classifier margin. The data points that are mapped near the classifier margins are called the support vectors. These vectors helps to classify the data. The figure-1, shows the support vector, where the classification process between two classes is given. The prediction obtained from the CNN algorithm are linearly separable, and hence, the following pair of data  $(w, b)$  is written in the following manner as,

$$w^T x_i + b \geq 1 \text{ if } y_i = 1,$$

$$w^T x_i + b \leq -1 \text{ if } y_i = -1$$



**Figure-6 SVM algorithm for classification**

The linear classification process can be formulated in the following manner, as,

$$f(x) = \text{sign}(w^T x + b)$$

where, the functional margin for the  $i$ th sample with respect to  $x_i$  for the corresponding hyperplane is  $(w, b)$  which can be formulated as,

$$\gamma_i = y_i(w^T x_i + b)$$

The shortest distance separating the data point from the hyperplane is perpendicular to it, and parallel to the weight  $w$ . The direction of the unit vector is  $w/\|w\|$ . The maximum separation between the support vectors of two different classes is defined as the geometric margin  $\rho$ , which is also shown in figure-1. The distance between the  $x_i$  and the separator is found using,

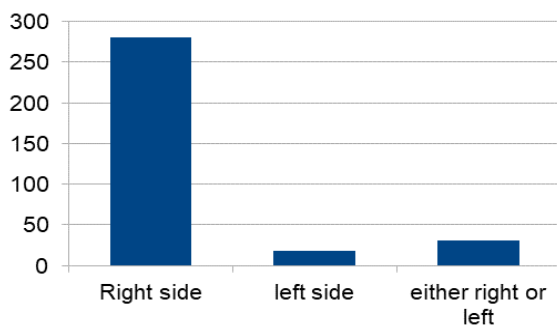
$$\tilde{\gamma} = y_i (w^T x_i + b) / \|w\| = 1 / \|w\|$$

The linear separator is designed in such a manner to maximize the geometric margin, which is formulated as,

$$\rho = 2 / \|w\|$$

where the following conditions is set,  $(x_i, y_i), i = 1, \dots, N: s. t. y_i(w^T x_i + b) \geq 1$ . The accuracy of the prediction model is evaluated. Thus, the physical education of the players are analyzed through CNN and real-time healthy practices and suggestions are provided by the SVM model.

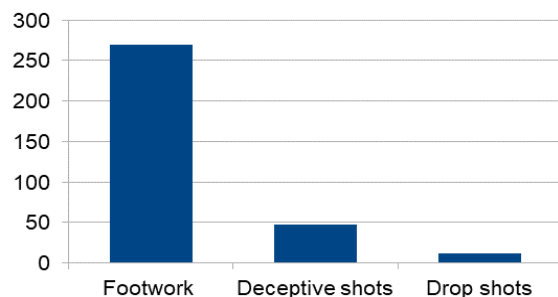
**Results and discussion**



**Figure-7. Players type of footwork**

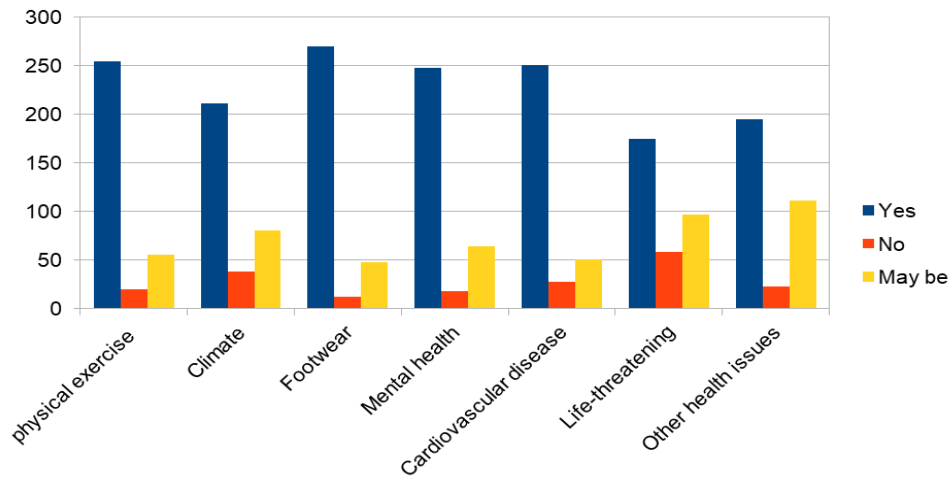
The figure-3 shows the factors and their influence on the players performance. Different types of factors like physical exercise, mental health, climate, footwear, cardiovascular disease, life threatening and other health issues and their

The proposed model is trained with different model datasets, and the predictions obtained from the proposed model are discussed in detail. The model is subjected to 10 fold cross validation. The generalization of the model is also verified by splitting the dataset into training and testing data. A 5-fold cross validation is carried out to verify the performance of the model. Initially, the obtained dataset is evaluated to analyse the performance of the model. A similar dataset with videos of 300 different players and their experiences are obtained and analysed. Their performance and prediction accuracy of the models are verified in and the predictions made by the models are also summarized. Different types of categories are considered for the prediction of the physical well being of the player using the video processing. Some of the important aspects considered for evaluating the physical well being are, right or left leg user during the match, skills they are good at like footwork, deceptive shots and drop shots. A basic analysis of the data helps the model to provide suggestions, and also to train the model.



**Figure-8. The skill of the players**

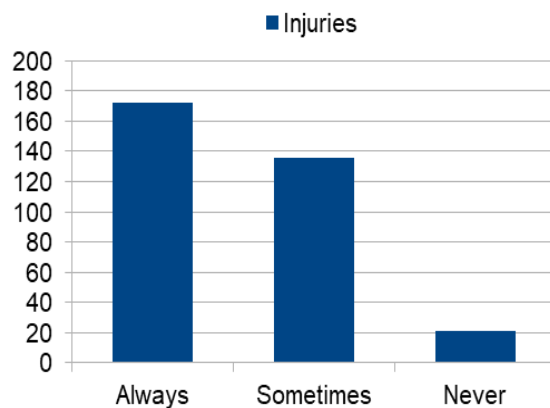
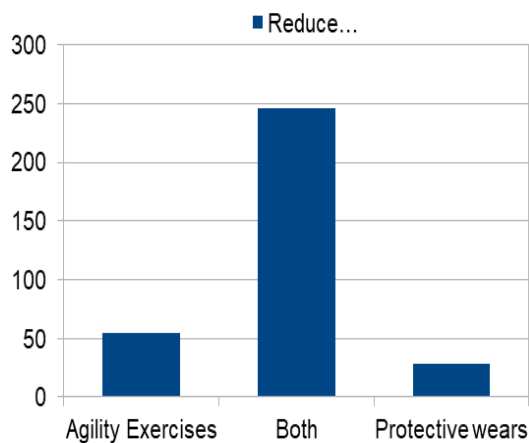
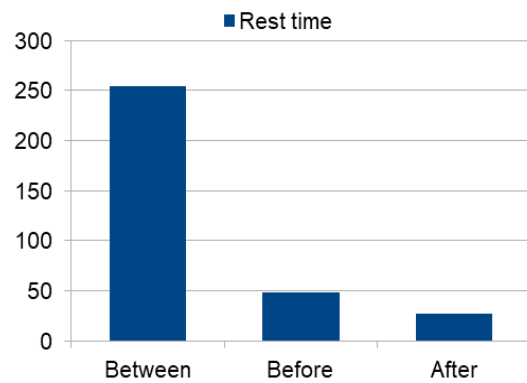
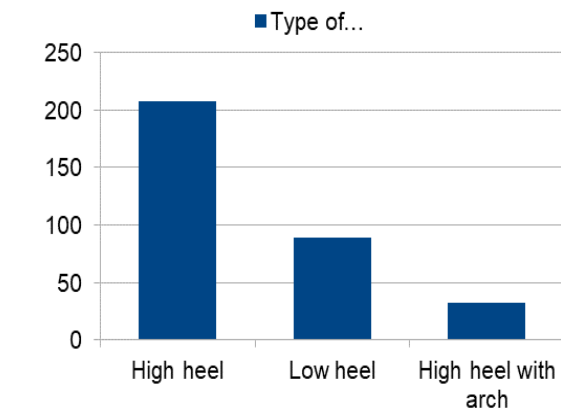
impact on the players are also calculated. These predictions are obtained from the dataset, and this data is used to train the proposed CNN-SVM model to obtain the suggestions based on the activities of the players.



**Figure-9. Factors that affect the players performance**

Different types of parameters are discussed in the figure-4, where the factors that influence the game of the players, are compared and a comparative analysis is given. The preference of the players and

their physical and mental state are evaluated for training the SVM model to give suggestions to the players for physical education.



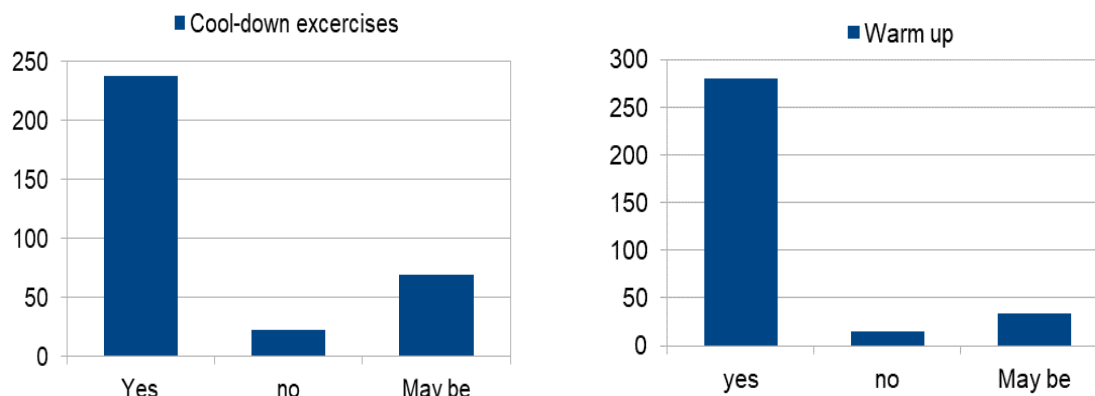


Figure-10. Comparative analysis of data obtained from the players

The CNN algorithm is used to predict the actions of the players, and the dataset is modelled in such a way to predict some of the important aspects of the players activities and appearances, for the SVM algorithm to classify and give suggestions on the

physical education of the player. Some of the important aspects are Right or left foot player (RL), skill (SK), footwear (FW), warm up (WP), cool down (CD), injuries (IJ) and protective wears (PW).

Table-1. Summary of predictions for different actions from the video frames

	Mean	Std	Min	0.25	0.5	0.75	Max
RL	54.8	5.58	53.8	50.8	53.8	57.8	63.8
SK	62.4	6.34	57.4	59.4	60.4	64.4	68.4
FW	72.5	7.35	63.5	68.5	71.5	74.5	81.5
WP	63.7	6.47	57.7	59.7	61.7	68.7	72.7
CD	45.4	4.64	36.4	40.4	44.4	47.4	54.4
IJ	51.3	5.23	44.3	50.3	50.3	54.3	59.3
PW	67.9	6.89	64.9	65.9	66.9	70.9	72.9

Table-2. Confusion matrix for the prediction of actions

	RL	SK	FW	WP	CD	IJ	PW
RL	87.3	7.3	2.2	-	1.4	1.8	-
SK	1.3	83.2	-	3.1	-	0.5	11.9
FW	3.8	0.2	93.2	-	2.8	-	-
WP	3	3.1	-	91.4	-	-	2.1
CD	-	4.6	-	-	95	-	0.4
IJ	4.1	0.4	4.2	0.3	-	91	-
PW	0.5	1.2	0.4	5.2	0.8	6.7	85

The confusion matrix shown in table-2 shows the performance of the CNN algorithm in predicting different actions by the players from the video

frames that are obtained from the dataset. The table-3 shows the comparison of accuracy in the prediction processes with the existing models.

**Table-3. Comparison of the prediction accuracy with the existing ones**

S.No	Methods	Accuracy
1	JLJA [1]	86.3
2	LOP/FTP [2]	76.1
3	BiFenceNet[3]	87.6
4	CNN-SVM	89.4

**Conclusion**

With the development of computer vision models, evaluating the activities of subjects in videos. In this paper, the physical education for the players are evaluated with CNN and SVM algorithms. The videos of the players are preprocessed and are segmented into frames, and these frames are evaluated through CNN algorithm. The CNN algorithm predicts the activities of the players, and these predictions are given to the SVM algorithm that classifies the data obtained from the CNN model and predicts the suggestions to be given to the players. The CNN algorithm proposed provides better processing of the data and predicts the actions of the players. Some of the important activities that correspond to the physical education of the players like Right or left foot player (RL), skill (SK), footwear (FW), warm up (WP), cool down (CD), injuries (IJ) and protective wears (PW) are defined and videos and instances with such activities are used to train and test the CNN algorithm. The CNN algorithm predicts such instances in the videos, and the SVM algorithm predicts the physical activities and practices of the players and provides suggestions accordingly for better physical education. This model is an initiative towards computer vision based recommendation and suggestion system. A model dataset of the players, videos and their views on the physical education is obtained and analysed. The analysis of on the dataset and the prediction made by the CNN-SVM model shows the effectiveness of the prediction for the players on physical education and healthy practices. It is also compared with similar models and the performance of the proposed model is verified. The proposed CNN-SVM model provides better prediction accuracy of 89.4 percent in predicting different activities from the videos for physical education and health suggestions.

**Reference**

1. Malawski, F., & Kwolek, B. (2019). Improving multimodal action representation with joint motion history context. *Journal of Visual Communication and Image Representation*, 61, 198-208.
2. Wang, J., Liu, Z., Wu, Y., & Yuan, J. (2012, June). Mining actionlet ensemble for action recognition with depth cameras. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 1290-1297). IEEE.
3. Zhu, K., Wong, A., & McPhee, J. (2022). Fencenet: Fine-grained footwork recognition in fencing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3589-3598).

**Reference:**

- [1] Tianyi Wang, Tongyan Li, "Deep Learning-Based Football Player Detection in Videos", *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3540642, 8 pages, 2022. <https://doi.org/10.1155/2022/3540642>
- [2] M. Ivasic-Kos, K. Host, and M. Pobar, 'Application of Deep Learning Methods for Detection and Tracking of Players', *Deep Learning Applications*. IntechOpen, Jul. 14, 2021. doi: 10.5772/intechopen.96308.
- [3] Zhang, L. (2022). Applying deep learning-based human motion recognition system in sports competition. *Frontiers in Neurorobotics*, 16, 860981.
- [4] Luo, J., Hu, Y., Davids, K., Zhang, D., Gouin, C., Li, X., & Xu, X. (2022). Vision-based movement recognition reveals badminton player footwork using deep learning and binocular positioning. *Heliyon*, 8(8).
- [5] Mekruksavanich, S., & Jitpattanukul, A. (2022). Multimodal wearable sensing for sport-related activity recognition using deep learning

- networks. *Journal of Advances in Information Technology*
- [6] . Fu, M., Zhong, Q., & Dong, J. (2022). Sports action recognition based on deep learning and clustering extraction algorithm. *Computational Intelligence and Neuroscience*, 2022.
- [7] Pan, S. (2022). A method of key posture detection and motion recognition in sports based on Deep Learning. *Mobile Information Systems*, 2022.
- [8] Fan, J., Bi, S., Xu, R., Wang, L., & Zhang, L. (2022). Hybrid lightweight Deep-learning model for Sensor-fusion basketball Shooting-posture recognition. *Measurement*, 189, 110595.
- [9] Dirgová Luptáková, I., Kubovčík, M., & Pospíchal, J. (2022). Wearable sensor-based human activity recognition with transformer model. *Sensors*, 22(5), 1911.
- [10] Kulsoom, F., Narejo, S., Mehmood, Z., Chaudhry, H. N., Butt, A., & Bashir, A. K. (2022). A review of machine learning-based human activity recognition for diverse applications. *Neural Computing and Applications*, 34(21), 18289-18324.
- [11] Sun, Y., & Li, Y. (2022). A Deep Learning Method for Intelligent Analysis of Sports Training Postures. *Computational Intelligence and Neuroscience*, 2022.
- [12] Host, K., & Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. *Heliyon*.
- [13] Tong, L., Ma, H., Lin, Q., He, J., & Peng, L. (2022). A novel deep learning Bi-GRU-I model for real-time human activity recognition using inertial sensors. *IEEE Sensors Journal*, 22(6), 6164-6174.
- [14] Qin, C., & Huo, S. (2022). Detection method of limb movement in competitive sports training based on deep learning. *Journal of Mathematics*, 2022, 1-8.
- [15] Liu, J., & Liang, B. (2022). An Action Recognition Technology for Badminton Players Using Deep Learning. *Mobile Information Systems*, 2022.
- [16] Şengül, G., Karakaya, M., Misra, S., Abayomi-Alli, O. O., & Damaševičius, R. (2022). Deep learning based fall detection using smartwatches for healthcare applications. *Biomedical Signal Processing and Control*, 71, 103242.
- [17] Xiao, J., Tian, W., & Ding, L. (2022). Basketball Action Recognition Method of Deep Neural Network Based on Dynamic Residual Attention Mechanism. *Information*, 14(1), 13.
- [18] Proposed Base paper
- [19] <https://scihub.ee/10.1016/j.micpro.2021.103900>
- [20] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9377869/>
- [21] <https://www.hindawi.com/journals/cin/2021/7006541/>