SIN-CHN SCIENTIFIC PRESS PTE.LTD.

*Molecular & Cellular Biomechanics* 2025, 22(3), 1032.
https://doi.org/10.62617/mcb1032

Article

# Research on the application of biosensor technology in the detection and prevention of sports injury in college sports training

**Yan Zhang**

School of Economics and Management, Karamay Vocational and Technical College, Karamay 834000, China; zhangyan123_321@126.com

**CITATION**

Zhang Y. Research on the application of biosensor technology in the detection and prevention of sports injury in college sports training. Molecular & Cellular Biomechanics. 2025; 22(3): 1032. https://doi.org/10.62617/mcb1032

**ARTICLE INFO**

Received: 6 December 2024
Accepted: 19 December 2024
Available online: 13 February 2025



**Abstract:** Injury detection plays a critical role in minimizing athlete downtime, ensuring safety, optimizing performance, and preventing long-term physical or mental consequences. In college sports, effective injury prevention and detection strategies enhance athlete safety, support peak performance, reduce healthcare costs, and contribute to sustainable athletic development programs. This research evaluates the application of biosensor technology in identifying injury risks, monitoring physiological metrics, and enhancing preventive strategies in college sports training to improve athlete performance and safety. A novel model, Egret Swarm Search-driven Scalable Deep Convolutional Neural Network (ESS-SDCNN), addresses the limitations of traditional approaches by combining SDCNNs with ESS algorithm for optimized feature selection, hyper parameter tuning, and real-time adaptability. Suitable data for injury detection and prevention include real-time physiological readings, motion sensor data, activity patterns, and injury records, with a focus on wearable technology. The Z-score normalization ensures consistent feature scaling. Independent Component Analysis (ICA) is used to extract hidden components from sensor data for improved feature representation. The SDCNN efficiently processes high-dimensional biosensor data, extracting spatial-temporal patterns related to injuries. The ESS algorithm further optimizes feature selection and hyper parameters, enhancing model accuracy, robustness, and adaptability for real-time applications. Results demonstrate that the hybrid ESS-SDCNN model significantly improves injury detection accuracy, enables faster convergence, and provides real-time monitoring and prevention insights. This approach enhances athlete safety, supports injury prevention, and fosters better performance outcomes in college sports training programs.

**Keywords:** sports injury detection; injury prevention; wearable biosensors; real-time monitoring; Egret Swarm Search-driven Scalable DCNN (ESS-SDCNN)

## 1. Introduction

Sports injuries are a serious issue in sports conditions and are more common in college settings because young players engage in intensive competition and training exercises.

The desire for sports achievement, frequent participation in competitions, and rigorous training are fundamental aspects of college sports teams, all of which elevate the risk of injury [1]. Sports-related injuries persist despite the progressing knowledge in the field of sports, affecting the athletes and the prospects. Athletes' optimal performance and lifetime value can be protected by promoting the accurate identification of sports injuries and implementing strategies for sports injury prevention [2].

The factors include overuse, improper technique, inadequate warm-ups, poor fitness levels, and environmental conditions in college sports training. In college athletes, the most common acute injuries seen are sprains, strains, fractures, and injuries and overuse problems such as stress fractures and muscle strains [3]. These injuries also pose the risk of developing complications that will affect the long-term health of an athlete; complications like persistent pain and restricted flexibility will lower an athlete's performance. To reduce such risks, a better model for analyzing and preventing injuries must be developed [4].

Early identification of these injuries in a sporting activity is very important. An early diagnosis means that the injury can be prevented from worsening, or further problems can be prevented [5]. By incorporating wearables, biomechanical evaluations, and video analysis to assess athletes' movements and discover sources of stress or inefficient biomechanics that may compromise the protective shield of athletes against injuries. Using these techniques, coaches, trainers, and other medical officials receive valuable information to assess the condition of an athlete and track variability in performances, which alerts for the occurrence of certain conditions [6].

However, there are several causes of sports-related injuries, and apart from improvement in technology, there is a need to improve efficient preventative measures. Preventive methods are focused on enhancing and extending athletes' physiological fitness and, at the same time, ensuring that technique is correct and that the recovery period is adequate [7]. There is always a possibility to identify certain athletes who may engage in certain activities that make them more susceptible to certain types of injuries with the help of pre-participation physical examinations and screening for magnitudes of risk factors, and therefore modify their training to accommodate the risks involved. Additionally, it also helps athletes decrease the chance of getting severe injuries through learning the correct warm-up, injury prevention, and recovery techniques [8].

Psychological risks are also factors that determine the experience of injury to athletes in college athletics. This means that athletes are more susceptible to making mistakes that can lead to harm if psychologically drained, stressed, or have performance anxiety [9]. To adopt a holistic approach to preventing any form of injury, there is a need to incorporate physical fitness as well as mental health. There is a need for such a platform to be created where the athletes interact with the coaches and trainers on how to deal with stress or mental burnout [10].

A unique ESS-SDCNN technique was presented to improve tracking of college sports during training for performance evaluation and to better identify injuries.

## Contributions

- The college sports injury data was obtained from the Kaggle platform.
- Z-score normalization is applied to continuously scale features, which helps in normalizing inputs for the system and increasing its ability to predict injury risks.
- Employs ICA to improve feature representation and the model's ability to learn complicated and complex patterns potentially indicative of risk of injury by eliminating the background noise in the sensor data.

- To enhance tracking of college sports during training for performance evaluation and for better identification of injuries, a novel ESS-SDCNN approach was introduced.

The remaining portions are structured in the following manner: Section 2—Related works, Section 3—Methodology, Section 4—Results and Discussion, and Section 5—Conclusion.

## 2. Related works

A methodology powered by AI was presented in research [11] that improved injury management by strategically allocating rest hours throughout athletes' recovery stages. The results demonstrated the significance of AI in producing useful insights, which facilitate better decision-making focused on athletes' welfare.

To evaluate the risk of injury in young football players competing at the highest levels, using anthropometric, motor control, and athletic performance metrics in conjunction with an ML technique was presented in research [12]. The holdout evaluation sample's injured players were identified by the ML system with accuracy (85%), recall (85%), and precision (85%).

The additional diagnosis of the ML-based knee joint sports injury identification method was examined and discussed in research [13]. The findings demonstrated a substantial change in time spent on the equilibrium pad before and following the table tennis players' practice following three months of an ML-based recovery program.

A new DFFNN model for estimating athlete injuries was suggested in research [14]. The findings of the experiment demonstrated that the suggested model successfully detected injuries in athletes and provided more precise predictions, achieving high categorization accuracy.

Recurrent neural models were employed in the research [15] to forecast the possibility of injuries and conduct continual tracking of the mechanisms behind sports injuries. The findings demonstrated that the suggested approach outperformed the traditional approach in terms of accuracy.

A DLS was introduced in the research [16] to anticipate injuries in sports. Employing several measures, the efficacy of the suggested approach was examined and contrasted with that of the traditional models. The suggested approach outperformed the traditional models in terms of performance.

The utilization of ML algorithms, namely SVM, to forecast injuries in professional athletics and employ BDA approaches to offer valuable player perspectives was examined in research [17]. The outcomes demonstrated the efficacy of the strategy. with the suggested SVM approach achieving an estimation rate of 87.5% and an accuracy of 92.3%.

An athlete injury identification approach based on CNN and sensors was described, and a mobile intelligent healthcare system was developed in research [18]. According to the findings, the CNN-based athlete injury identification system outperformed the conventional approach by 6.73% in terms of detection accuracy.

Ankle joint sports injuries could be more accurately diagnosed by employing ML techniques and image processing approaches to produce objective and reliable diagnostic outcomes that were presented in research [19]. The results demonstrated that the enhanced ResNet algorithm could reliably categorize various ankle joint sports injuries while retaining excellent stability over a range of sample conditions.

A deep learning approach based on ResNet-18 was created by research [20] to identify ACL situations. The findings indicated that their suggested framework and two orthopedic physicians and radiologists do not significantly differ in their ability to diagnose ACL disorders.

A multi-sensor fusion-based sports injury identification approach was presented in research [21]. Initially, calculated the cumulative variance between frames to identify the human body's movement region. The findings of the research demonstrated their suggested method's efficacy.

The development of ML techniques created new opportunities for sports injury monitoring, as presented in research [22]. The proposed system architecture was simple to execute and cost-effective since it determined the possibility of injuries in athletes, offered preventative suggestions based on displacement curves, and had low overall expenses and high testing accuracy.

To comprehend and ascertain the possibility of injuries, research [23] employed big data to construct a psychological evaluation model for risk estimation of sports injuries. The findings demonstrated that their approach outperformed other models in terms of accuracy and assessment effects.

To guarantee that athletes could follow and perform safely and healthily, research [24] assisted athletes preventing and minimizing their chance of injury during training more efficiently by building a sports training injury risk evaluation method. The developed risk evaluation system could meet the requirements of the sports activity injury risk evaluation system due to its high sensitivity and exceptional power consumption capabilities.

The use of predictive ML techniques for data-driven sports injury risk variable detection was described in research [25]. The findings demonstrated the challenge of forecasting future injuries but additionally indicated that some predictive injury risk variables could be reliably identified even with models that exhibited poor predictive power.

## 3. Methodology

The college sports injury detection data was gathered from the Kaggle source. The data was pre-processed using the Z-score normalization. The ICA approach was employed to extract the attributes from the information. An innovative ESS-SDCNN approach was introduced for the detection and prevention of sports injury in college sports training. **Figure 1** display the overview of ESS-SDCNN.

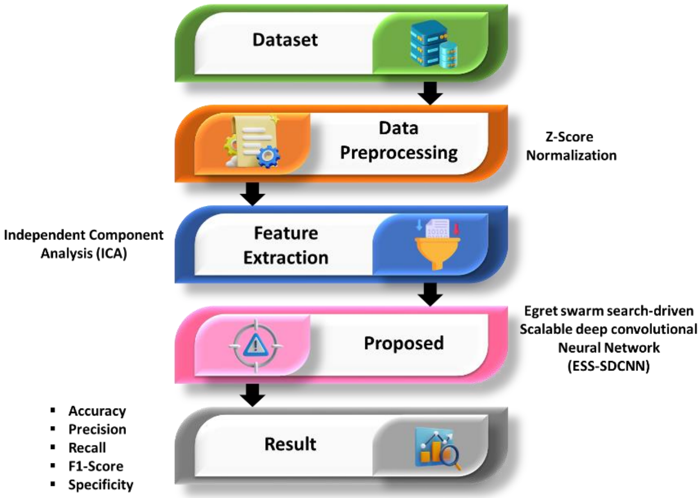


**Figure 1.** Outline of research framework.

### 3.1. Data collection

The data was gathered from Kaggle source [26], includes various physiological measurements including impact force, blood oxygen levels, skin temperature, and heart rates. It also comprises athlete-specific parameters like session duration, exhaustion levels, and activity data that would be helpful for knowing injury risk in the process of training sports. Data was collected from different athletes in various sports, and training conditions, which ensured the data's relevance in injury prediction.

### 3.2. Z-score normalization

Z-score normalization is the most widely used normalizing technique that transforms all input values into a single measure with a SD of 1 and an average of 0. Every feature is computed with its SD and mean. The calculated SD and mean are used to normalize every value of an attribute$W$.

$$Z = \frac{(w - mean\ (W))}{std\ (W)} \tag{1}$$

where $std\ (W)$–SD of feature $W$, and $mean\ (W)$ Mean of feature $W$.

### 3.3. ICA

ICA can be used in sports for identifying early signs of potential sports injuries in sensor data such as muscle strain or improper movement and then apply customized preventive measures. To identify hidden components from a set of observations or observed data, ICA is a relatively novel statistical and mathematical approach that ensures the sources are as independent as possible. The initial and fully independent source $t(s) = t_1(s), t_2(s), \ldots, t_m(s)$at the period $s$is linearly combined with the measured variables $w(s) = w_1(s), w_2(s), \ldots, w_m(s)$ on the computational level such that it may be stated as in Equation (2),

$$w(s) = Bt(s) \tag{2}$$

where $B$ is a full-rank combining matrix. In ICA standards, Equation (2) is frequently expressed as Equation (3),

$$z = Xw \tag{3}$$

where the independent component is indicated by $z = z_1, z_2, \ldots, z_m$, and the demixing matrix (DM) is $X = B^{-1}$.Numerous ICA techniques may be employed to calculate the DM and individual elements only from the mixed data.

The retrieved components are independent and non-gaussian according to the ICA assessment criteria. Non-gaussianity could be measured using kurtosis($\beta_1$). Kurtosis values for the Gaussian ICs are equal to zero, sub-gaussian, $\beta_1 \leq 0$, and super-gaussian, $\beta_1 \geq 0$.The standard description of kurtosis is shown in Equation (4),

$$\beta_1 = \frac{E(w-\mu)^4}{(E(w-\mu)^2)^2} - 3 = \frac{\mu_4}{\sigma_4} - 3 \tag{4}$$

The standard kurtosis methods are also susceptible to outliers because, mostly dependent on sample averages. Furthermore, recognizing that outliers are increased to the third and fourth powers in standard kurtosis assessments significantly increases their impact. An attempt is made to employ an accurate measure of kurtosis in ICA to solve the provided issue. Moors suggested a quantile kurtosis substitute for $\beta_1$. The amount of Moors kurtosis is shown in Equation (5),

$$Kurtosis = \frac{(F_7 - F_5) + (F_3 - F_1)}{(F_6 - F_2)} \tag{5}$$

$F_j$is the $j^{th}$ octile, therefore $F_j = E^{-1}\left(\frac{j}{8}\right)$. For independent Gaussian elements, Moor's quantile kurtosis equals 1.23. One benefit of using quantile measurements of kurtosis is independence from the first and second moments. It is more reliable than the standard kurtosis measure.

### 3.4. ESS-SDCNN

To prevent sports afflictions during college athletic training, an advanced model known as the ESS-SDCNN was developed. It employs ESS, an optimization technique inspired by the cooperative behavior of egrets, to fine-tune an SDCNN's hyper parameters for accurate injury prediction and detection. The model analyses the risk of injury and predicts based on sensor information such as movement, speed, and biomechanics during a sporting event. The ESS method optimizes the network structure and parameters of the learning process in the SDCNN. The suggested approach offers a highly effective solution to the prevention of injuries with the help of ESS, which adapts to different conditions in sports as well as individual training patterns.

Using real-time analysis, the ESS-SDCNN may identify risks for injury such as overexertion, poor form, or overstraining to enable precautionary measures by trainers or medical personnel. This lowers the rate of occurrences of injuries during practice sessions and training in constructing a unique practice regimen of the athlete under consideration guided by past performance and health status to boost safety and athletic performance in college sports activities. Algorithm 1 pseudo code for ESS-SDCNN.

**Algorithm 1** ESS-SDCNN

```
import numpy as np
import tensorflow as tf
from ESS import ESS_optimizer
from sklearn.preprocessing import StandardScaler
def preprocess_data(sensor_data):
scaler = StandardScaler()
normalized_data = scaler.fit_transform(sensor_data)
return normalized_data
def build_sdcnn_model(input_shape):
model = tf.keras.Sequential()
model.add(tf.keras.layers.Conv2D(filters, kernel_size, activation = 'relu', input_shape = input_shape))
model.add(tf.keras.layers.MaxPooling2D(pool_size))
model.add(tf.keras.layers.Conv2D(filters, kernel_size, activation = 'relu'))
model.add(tf.keras.layers.MaxPooling2D(pool_size))
model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dense(units, activation = 'relu'))
model.add(tf.keras.layers.Dense(1, activation = 'sigmoid'))
model.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])
return model
def apply_ESS_optimizer(model, train_data, train_labels):
ESS = ESS_optimizer(model)
optimized_model = ESS.optimize(train_data, train_labels)
return optimized_model
def predict_injury(model, test_data):
injury_risk = model.predict(test_data)
return injury_risk
def train_injury_detection_model(sensor_data, labels):
processed_data = preprocess_data(sensor_data)
train_data, test_data, train_labels, test_labels = train_test_split(processed_data, labels, test_size = 0.2)
model = build_sdcnn_model(input_shape = train_data.shape[1:])
optimized_model = apply_ESS_optimizer(model, train_data, train_labels)
optimized_model.fit(train_data, train_labels, epochs = epochs, batch_size = batch_size)
loss, accuracy = optimized_model.evaluate(test_data, test_labels)
print(f"Test Accuracy: {accuracy}")
return optimized_model
# Example usage
if __name__ == "__main__":
sensor_data = load_sensor_data()
labels = load_labels()
trained_model = train_injury_detection_model(sensor_data, labels)
new_sensor_data = preprocess_data(new_data)
injury_risk = predict_injury(trained_model, new_sensor_data)
print(f"Injury Risk Prediction: {injury_risk}")
```

### 3.4.1. Scalable deep convolutional neural network

An SDCNN depends on the sensor data and performs complex analysis for detecting and preventing sports injuries during college athletic training. It strengthens athlete safety and optimizes training by analyzing movement to estimate risk factors, in real time, to predict and prevent potential injuries. A DCNN is a type of multilayered neural network that is primarily composed of the IL, PL, OL, and CL.CL and PL are two of these hidden layers as. The original input provides values to the IL initially in a DCNN. The CL then uses a convolution kernel to extract attributes from the data. The PL then uses the local relationship principle to reduce the data that has to be examined. Subsequently, the feature is mapped to the labels through the OL.

In a DCNN, the OL of the top layer is the IL of the present layer since the mapping is a forward propagation procedure that explains the data flow from the neural network's IL to its OL. Neurons with nonlinear activation functions must be introduced into every layer in the forward procedure to overcome the errors in the linear model. The second layer's nonlinear activation functions are used, and the $k^{th}$layer's output may be represented as in Equation (6),

$$\left.\begin{array}{c} y^k = X^k * w^{k-1} + a^k \\ b^k = \sigma(y^k) \end{array}\right\} \tag{6}$$

where, $*$—Convolution operation, and $k$—$k^{th}$ Layer, and.

The feature map matrix retrieved from the $(1-2)^{th}$ layer is $w^{k-1} = b^{k-1} = \sigma(y^{k-2})$, and $k > 2, w^{k-1},$ the matrix whose components are values is $k = 2, w^{2-1} = w^1$. The weighted input of the$k$th layer is represented by $X^k, a^k, and\ y^k$, while the nonlinear activation function is denoted by $\sigma$. A $b^K$ will represent the final output vector, while $K$ is the OL.

The parameters of $X^k\ and\ a^k$ are often updated repeatedly using the BP approach, a supervised learning technique. It creates a CF by initially integrating the actual output and the desired values, and then it adjusts the parameters by applying GD across the CF's negative gradient orientation. The specific procedure is as follows.

*Selection of CF*

When choosing an error cost function, the quadratic function is typically used. However, it would require a lot of time if the neurons made a clear error when DCNN was being trained. Therefore, it uses Cross Entropy ($F_0^K$) as the error cost function instead of the quadratic function. The Cross-Entropy may be determined using the forward propagation approach using Equation (7).

$$F_0^K = \frac{1}{m}\sum\nolimits_{j=1}^{m}\sum\nolimits_{l=1}^{M}\left[s_l^K Inb_l^K + \left(1 - s_l^K\right)In(1 - b_l^K)\right] \tag{7}$$

DCNN is effectively separated into $M$ classes, where $M$ is the number of neurons in the OL and $m$ is the overall number of training sets.$s_l^K$is the desired value that corresponds to the OL's $l^{th}$ neuron, while $b_l^K$ is the neuron's actual output value.

*Error vector assessment*

Every layer has a predetermined error vector, and the OL's $l^{th}$ neuron's error vector is described as shown in Equation (8),

$$\delta^K = \frac{\partial F_0^K}{\partial y^k} \tag{8}$$

In the BP method, $\delta^K$ may be utilized to determine $\delta^K - 1$ in reverse. Additionally, assume that $\delta^K\ and\ \delta(k+1)$represent the $k^{th}\ and\ (k+1)^{th}$ layers' corresponding error vectors. Then, $\delta^K$ is expressed as follows in Equation (9) by Equations (6) and (8), and the Chain Rule,

$$\delta^k = X^{k+1}\delta^{k+1} \odot \sigma'(y^k) \tag{9}$$

where the Hadamard product, also known as the Schur product, is represented by the $\odot$, signifying the element wise result of the two vectors.

*Weight updates*

The gradients $\frac{\partial F_0^K}{\partial X^k}$ and $\frac{\partial F_0^K}{\partial a^k}$represent the values of $X^k$and $a^k$. The PD operation is denoted by the $\partial(.)$. Equations (6) and (8) could be used to compute the PD $F_0^K\ to\ X^k\ and\ a^k$,

$$\left.\begin{array}{c}\frac{\partial F_0^K}{\partial X^k} = \frac{\partial F_0^K}{\partial b^k} \odot \frac{\partial b^k}{\partial X^k} = \delta^k \odot w^{k-1} \\ \frac{\partial F_0^K}{\partial a^k} = \frac{\partial F_0^K}{\partial b^k} \odot \frac{\partial b^k}{\partial a^k} = \delta^k\end{array}\right\} \tag{10}$$

The modified values of $X^k\ and\ a^k$ are computed by minimizing the cost function, and are denoted by $\Delta X^k\ and\ \Delta a^k$,

$$\left.\begin{array}{c}\Delta X^k = -\eta \frac{\partial F_0^K}{\partial X^k} \\ \Delta a^k = -\eta \frac{\partial F_0^K}{\partial a^k}\end{array}\right\} \tag{11}$$

where the learning rate is represented by$\eta$.

*Scalable DCNN*

The class separability data is contributed to the cross-entropy CF as a term for regularization to train the DCNN approach, improving the class separability of the attributes retrieved by the algorithm. ICD and intra-class compactness, denoted by the$F_1\ and\ F_2$, comprise the class separability data is shown in Equations (12) and (13),

$$F_1 = \frac{1}{2}\sum \|z_d^m - N_d\|_2^2 \tag{12}$$

$$F_2 = \frac{1}{2}\sum \|N_d - N_{d'}\|_2^2 \tag{13}$$

where $z_d^m$ is the $m^{th}$ training sample's actual output value, which is a member of the $m^{th}$ class. $N_d\ and\ N_{d'}$, Represent the training sample output average values for $d^{th}\ and\ (d')^{th}$ classes. In $(d')^{th}$classes, $C$ is the number of training samples. The output characteristics' intra-class and ICDs are indicated by $F_1\ and\ F_2$.To improve the separability of the output characteristics, the intra-class distance must be shortened and the ICD must be increased in each iteration. The CF is altered as follows when it is included as a regularization factor is shown in Equation (14).

$$F = F_0^K + \alpha F_1 - \beta F_2 \tag{14}$$

where, $\alpha\ and\ \beta$—Weight parameters.

The updated CF's error vector is crucial for adjusting $X^k$ $and$ $a^k$ to make the system improve in favor of categorization.

The error vector for $F_1$ in the OL is,

$$\begin{aligned} \delta_1^K &= \frac{\partial F_1}{\partial_{y^K}} \\ &= \frac{\partial}{\partial_{y^K}} \frac{1}{2} \| z_d^m - N_d \|_2^2 \\ &= \left(1 - \frac{1}{m^d}\right) \sigma'(y^K) \odot (z_d^m - N_d) \end{aligned} \tag{15}$$

where $m^d$ is the number of samples in class $d$.

The error vector for $F_2$ in the OL $K$ is,

$$\begin{aligned} \delta_2^K &= \frac{\partial F_2}{\partial_{y^K}} \\ &= \frac{\partial}{\partial_{y^K}} \frac{1}{2} \| M_d - N_{d'} \|_2^2 \\ &= \frac{1}{m^d} \sigma'(y^K) \odot \sum\nolimits_{d'}^{A} (M_d - N_{d'}) \end{aligned} \tag{16}$$

The updated error vector of $F$ in the OL $K$ is as follows, using Equation (8), and Equations (14)–(16),

$$\delta' = \delta^K + \alpha\delta_1^K - \beta\delta_2^K = \sigma'(Y^K) \odot (z^m - s^m) + \alpha\left(1 - \frac{1}{m^d}\right)\sigma'(Y^K) \odot (z_d^m - N_d) - \beta\frac{1}{m^d}\sigma'(Y^K) \odot (M_d - N_{d'}).$$

Once the output layer's error vector has been obtained, it may use Equation (9) to repeatedly estimate the error vector in every layer. Then, using Equations (10) and (11), it could determine the updated $X^k$ $and$ $a^k$ parameters for every layer.

### 3.4.2. Egret swarm search (ESS) optimization

In college sports training, a sophisticated approach known as ESS optimization is used to identify and prevent sports injuries. ESSO enhances the forecast of the possibility of injuries by imitating the egrets' behaviors, which enhances the real-time monitoring of the movements and biomechanics of athletes. This prevents possible harm during training since it can identify early signs of injury.

*Inspiration*

The Great Egret, Middle Egret, Little Egret, and Yellow-billed Egret are the four bird species that are collectively referred to as egrets. Distinguished by their beautiful white feathers. Most egrets live on coastal islands, along estuaries, and coastlines, and near rice paddies, wetlands, lakes, ponds, and rivers. Maccarone discovered that Great Egrets balance their motions and energy consumption when hunting, flying at a standard speed of 9.2 m/s. Since flying requires a lot of energy, choosing to prey usually requires a careful analysis of the direction to ensure that the position of food will provide more energy than the energy used for flight. Snowy egrets frequently visit more locations than Great Egrets, observing and selecting food sources where other birds have previously found it. Snowy Egrets frequently use a SAW technique, which is a method of hunting with the least amount of energy consumption that entails watching prey behavior for a while and then predicting their next move. According to Maccarone, Snowy Egrets who use the technique are 50% more successful at capturing prey than other egrets, in addition to using less energy. The active searching approach of Great Egrets balances high energy consumption for possibly higher returns, while the SAW technique of Snowy Egrets balances lower energy usage for smaller but more consistent benefits.

*Algorithms and mathematical models*

ESS created an equation to quantify behaviors after being influenced by the aggressive strategy of the Great Egret and the Snowy Egret's SAW method. The SAW approach, the aggressive approach, and the discriminant criterion are the three key elements of ESS.A group of three egrets, consisting of Egret A, B, and C, employs a directing forward process, random walk, and encircling methods.

Egret B provides worldwide random travel, Egret C specifically examines superior egret positions, and Egret A estimates a descending plane and searches based on plane gradient. Exploration and exploitation will be more evenly distributed in ESS, which will also be able to conduct quick searches for feasible responses. ESS is less probable to enter the saddle area of the optimization issue than GD because it incorporates randomness and previous data into the gradient calculation. Additionally, ESS is different from previous meta-heuristic approaches in that it assess the optimization issue's tangent plane, permitting for a rapid fall to the present ideal position.

*SAW technique*

Considering that the $j^{th}$ egret group's location is $w_j \in \mathbb{R}^m$, $m$ is the difficulty's size, and $A(*)$is the Snowy Egret's estimation of the possible incidence of prey at its present location. $\hat{z}$is the prey's estimated current position.

$$\hat{z}_j = A(w_j) \tag{17}$$

The estimation algorithm might then be modified as follows,

$$\hat{z}_j = w_j \times x_j \tag{18}$$

where the weight of the estimation technique is denoted by $x_j \in \mathbb{R}^m$. Error $f_j$might be explained as

$$f_j = \left\|\hat{z}_j - z_j\right\|^2/2 \tag{19}$$

In addition, by calculating the PD of $x_j$ for the error in Equation (19), $\hat{h}_j \in \mathbb{R}^m$, the actual gradient of $\omega_j$, may be obtained, with $\hat{c}_j$ as its direction.

$$\begin{aligned}\hat{h}_j &= \frac{\partial \hat{f}_j}{\partial x_j}\\ &= \frac{\left(\partial\left\|\hat{z}_j - z_j\right\|^2\right)/2}{\partial x_j}\\ &= \left(\hat{z}_j - z_j\right). w_j,\end{aligned}$$

$$\hat{c}_j = \hat{h}_j/\left|\hat{h}_j\right| \tag{20}$$

where Egrets use their knowledge of predicting prey behavior and integration of their ideas to indicate superior Egrets during preying. While $c_{h,j} \in \mathbb{R}^m$ is the directional modification of the ideal position of the entire group, $c_{g,j} \in \mathbb{R}^m$ is the directional modification of the group's best position.

$$c_{g,j} = \frac{w_{jbest} - w_j}{|w_{jbest} - w_j|} \times \frac{e_{jbest} - e_j}{|w_{jbest} - w_j|} + c_{jbest} \quad (21)$$

$$c_{h,j} = \frac{w_{hbest} - w_j}{|w_{hbest} - w_j|} \times \frac{e_{hbest} - e_j}{|w_{hbest} - w_j|} + c_{hbest} \quad (22)$$

The combined gradient $h_j \in \mathbb{R}^m$ may be shown as follows, where $q_g \in [0,0.5), q_h \in [0,0.5)$,

$$h_j = (1 - q_g - q_h) \times \hat{c}_j + q_g \times c_{g,j} + q_h \times c_{h,j} \quad (23)$$

An adaptive weight updating approach is employed, $\beta_1$ is 0.9 and $\beta_2$ is 0.99,

$$n_j = \beta_1 \times n_j + (1 - \beta_1) \times h_j,$$

$$u_j = \beta_1 \times u_j + (1 - \beta_1) \times h_j^2, x_j = x_j - n_j/\sqrt{u_j} \quad (24)$$

Egret A's assessment of the present situation indicates that the subsequent sample site, $w_{b,j}$, may be characterized as

$$w_{b,j} = w_j + step_b \times exp(-s/(0.1 \times s_{\max})) \times hop \times h_j \quad (25)$$

$$z_{b,j} = e(w_{b,j}) \quad (26)$$

where,

$hop$—Difference between the solution space's low and upper bounds,

$s \ and \ s_{max}$—Iteration time at the moment and the maximum iteration time,

$step_b \in (0,1]$—Egret A's SS factor, and

$z_{b,j}$—The fitness of $w_{b,j}$.

*Aggressive approach*

Egret B's habit of searching for food at random may be illustrated as follows,

$$w_{b,j} = w_j + step_a \times \tan(q_{a,j}) \times hop/(1 + s) \quad (27)$$

$$z_{a,j} = e(x_{a,j}) \quad (28)$$

where,

$w_{a,j}$—Anticipated subsequent position of Egret B,

$q_{a,j}$—Arbitrary number in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and

$z_{a,j}$—Fitness.

Due to its preference for aggressive prey chase, Egret C uses the encircling process to modify its location,

$$C_g = w_{jbest} - w_j.$$

$$C_h = w_{hbest} - w_j w_{d,j} = \left(1 - q_j - q_h\right) \times w_j + q_g \times C_g + q_h \times C_h \tag{29}$$

$$z_{d,j} = e(x_{d,j}) \tag{30}$$

where as $C_h$ associates with the ideal position of all Egret groups, $C_g$ is the gap matrix between the present position and the ideal location of the particular group.

Egret C is predicted to be located at $w_{d,j}$. Egret B's SS factor is $step_a \in (0,1]$. In $[0,0.5], q_g\ and\ q_h$ are arbitrary numbers.

*Discriminant phase*

After each member has decided on their approach, the Egret group decides on the most effective path of action and works as a cohesive one. The $j^{th}$ Egret group's solution matrix is $w_{t,j}$,

$$w_{t,j} = [w_{b,j} \quad w_{a,j} \quad w_{d,j}] \tag{31}$$

$$z_{t,j} = [z_{b,j} \quad z_{a,j} \quad z_{d,j}] \tag{32}$$

$$d_j = arg\ \min\ (z_{t,j}) \tag{33}$$

$$w_j = \begin{cases} w_{t,j|d_s} & if\ \ z_{t,j|d_s}\ \ pr\ q < 0.3 \\ w_j & else \end{cases} \tag{34}$$

The Egret group supports the decision if the minimum value of $z_{t,j}$ is greater than the present fitness$z_j$. Additionally, if an arbitrary number $q \in (0,1)$is smaller than 0.3, there is a 30% chance that a less effective plan will be accepted.

## 4. Result

The ESS-SDCNN method was implemented in Python 3.11 and tested on a Windows 11 laptop equipped with an Intel i7 5th Gen processor and 16 GB of RAM. In this research, the performance of the proposed ESS-SDCNN was compared against conventional algorithms, including AdaBoost, Random Forest (Ada-RF) [27], and k-Nearest Neighbor (k-NN) [28], using a dataset of labeled injury-prone behaviors acquired through sensor-based activity tracking in training environments. The data collection was intended to capture both injury-prone and safe behaviors across several training sessions.

### 4.1. Wearable device data insights

The heart rate is 60% of the normal range (71–120 beats per minute) considered to be healthy. It covers 40% of the total and remains average and acceptable with a sleep length of 6–7 h. There is continuous exercise shown by a moderate exercise intensity rating of 40%. The moderate stress level is measured at 40% proving that the tension is still well contained. Finally, 45% have provided a reasonable blood oxygen level indicating that their oxygen saturation could be improved potentially. **Figure 2** displays many health-related measurements, all of which can be described using a percentage that corresponds to its normal/abnormal state.

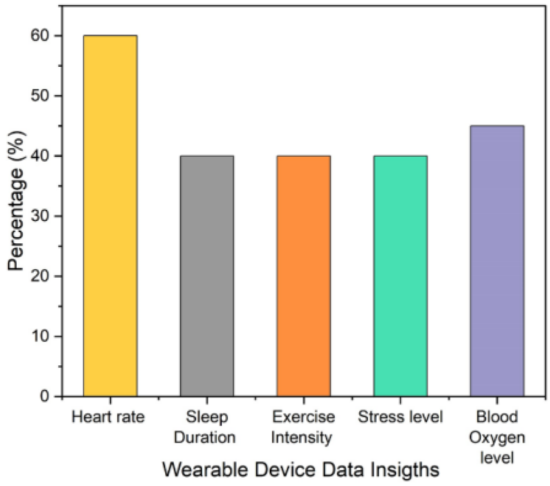


**Figure 2.** Outcome of health data insight from wearable devices**.**

## 4.2. Accuracy

The accuracy evaluates the effectiveness in recognizing injury-prone behaviors, training errors, and other risks. It determines the percentage of false negative and true positive injury identification in different training sessions. **Figure 3** and **Table 1** illustrate the comparison between traditional Ada-RF and the proposed approach in terms of accuracy for injury prone behavior recognition. Compared to the traditional Ada-RF approach, which has an accuracy of 86.9%, the proposed ESS-SDCNN approach demonstrate superior performance with accuracy of 97.1%.

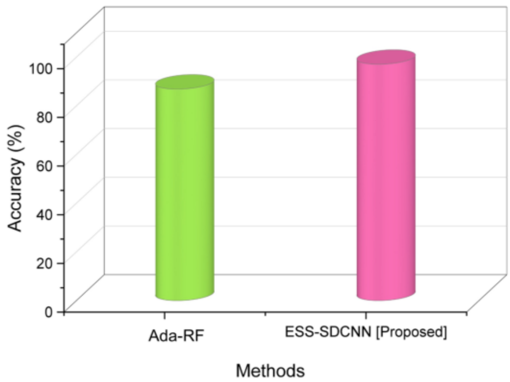


**Figure 3.** Accuracy result of injury identification.

## 4.3. Specificity

Specificity is the ability of a system or a model to correctly predict no-injury athletes. It evaluates the extent to which the system reduces the rate of false negatives, which prevents healthy athletes from being incorrectly identified as injured. The comparison between the suggested method and conventional Ada-RF in terms of specificity for injury-prone behavior identification is shown in **Figure 4** and **Table 1**. The proposed ESS-SDCNN technique with a specificity of 95.4% outperforms better than the conventional Ada-RF strategy, which has a specificity of 75.5%.

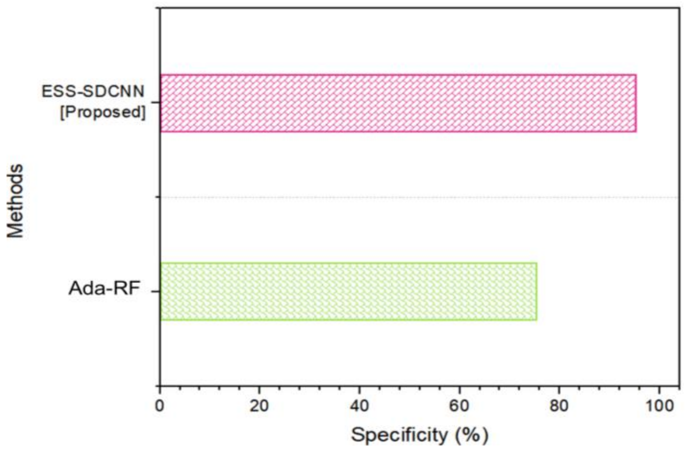


**Figure 4.** Specificity result of injury identification.

**Table 1.** Result of accuracy and specificity.

| Methods | Accuracy (%) | Specificity (%) |
|---|---|---|
| Ada-RF | 86.9% | 75.5% |
| ESS-SDCNN [Proposed] | 97.1% | 95.4% |

### 4.4. Recall

Recall measures the capability of a system or a model to classify every sports injury instance effectively. It is calculated by identifying the average of the total number of injuries that were identified to the number of injuries that were not identified. **Figure 5** and **Table 2** show the comparison between the proposed ESS-SDCNN approach and convolutional Ada-RF and k-NN in terms of recall injury identification. When compared to the traditional approaches such as Ada-RF (88.5%), and k-NN (66%) the suggested ESS-SDCNN approach has a superior recall value of 94.3%.

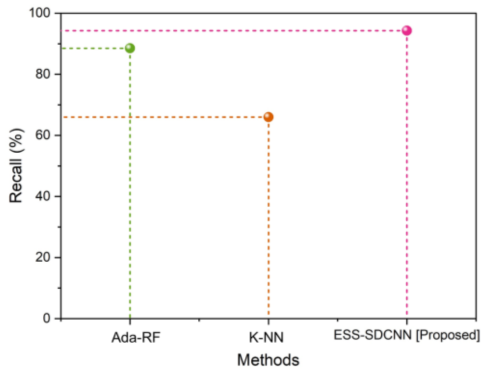


**Figure 5.** Recall result of injury identification.

### 4.5. Precision

Precision is the measure of the number of accurately forecasted injuries about all well-forecasted injuries. It measures how well the system is capable of recognizing actual injuries without incorrectly identifying healthy athletes as injured. The suggested ESS-SDCNN strategy has a precision value of 95.1% when compared to the conventional methods such as Ada-RF (75.4%) and k-NN (67%) approaches have low precision values. **Table 2** and **Figure 6** demonstrate the precision result of injury identification.

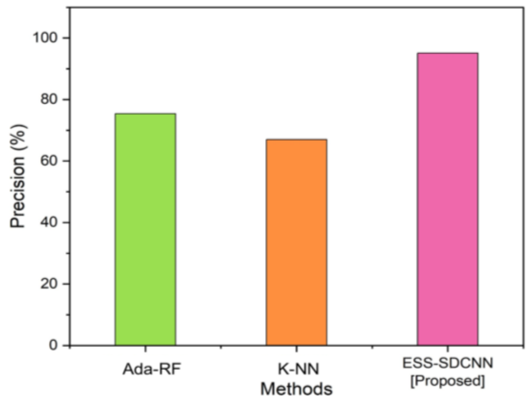


**Figure 6.** Precision result of injury identification.

### 4.6. *F*1-score

The *F*1-score measure is suitable for evaluating the efficiency of sports injury identification and prevention because it balances precision and recall. It offers a comprehensive assessment of the performance of the developed model in identifying injuries and the effectiveness of the prevention by proving the model's ability to identify the correct injuries and minimize both false positives and false negatives. The proposed ESS-SDCNN approach significantly better *F*1-score of 96.2%, outperforming traditional approaches such as Ada-RF (68.3%) and k-NN (66%). **Figure 7** and **Table 2** shows the comparison of traditional Ada-RF and k-NN method for injury-prone behavior identification.

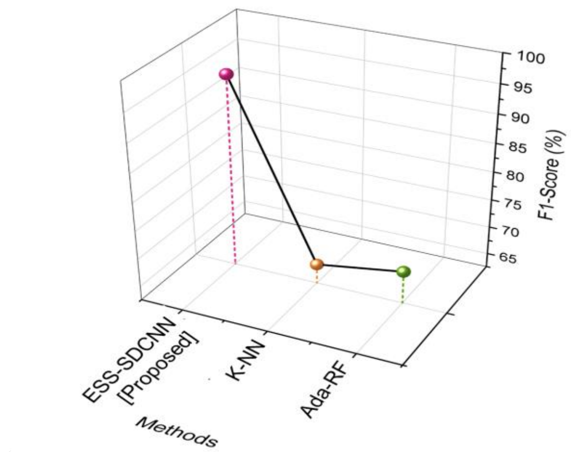


**Figure 7.** *F*1-score result of injury identification.

**Table 2.** Result of *F*1-score, recall, and precision.

| Methods | *F*1-score (%) | Recall (%) | Precision (%) |
|---|---|---|---|
| Ada-RF | 68.3% | 88.5% | 75.4% |
| k-NN | 66% | 66% | 67% |
| ESS-SDCNN [Proposed] | 96.2% | 94.3% | 95.1% |

### 4.7. Discussion

There are several disadvantages to using the Ada-RF [27] method, towards the identification and prevention of sports injuries during college athletic training. Its performance heavily depends on the quality of the training data which might not be distinctive or might contain a lot of noise in the case of actual sports injuries datasets. Moreover, when the number of decision trees is very large, the method may be highly influenced by noise in the data set. In addition, it could fail to accurately determine, highly, complex, non-linear patterns of the athletes' movement which could severely limit its application in many types of training conditions.

K-NN [28] faces challenges when it comes to sports injury identification particularly when applied to high–dimensional data within which it may have problems due to the curse of dimensionality. A disadvantage of K-NN is the long processing times because the method depends on the distances between data points, a drawback in sports training, which may need real-time forecasts. Preventing and detecting injuries more concretely may be a challenge due to the sensitivity of k-NN to noise or irrelevant attributes in data. In addition, it performs poorly when given an unbalanced data set, which leads to the development of inaccurate results.

To overcome these issues, a novel ESS-SDCNN method is superior to the other traditional methods, including Ada-RF and k-NN, as it integrates Egret Swarm Search (ESS) with Scalable Deep Convolutional Neural Networks (SDCNNs) to optimize feature selection, hyper parameter tuning, and real-time adaptability. The hybrid approach, therefore, results in higher accuracy for injury detection, model robustness, and faster convergence rates, which make it more effective for real-time sports injury prevention. Appendix **Table A1** shows the list of abbreviations.

## 5. Conclusion

Athlete injury identification is important to reduce athlete downtime, improve safety measures, enhance performance, and eliminate future consequences. The college sports injury data was gathered using the Kaggle platform. A novel ESS-SDCNN method was introduced to enhance college sports monitoring during training for performance assessment and to accurately detect injuries. The ESS-SDCNN approach was evaluated to be highly effective in the detection and prevention of sports injuries in college sports training, with high performance in key metrics in terms of precision (95.1%), specificity (95.4%), recall (94.3%), accuracy (97.1%), and $F1$-score (96.2). This indicates biosensor technology's potential role in enhancing injury prevention and improving the safety standards of training. Several conditions include poor access to advanced technology, inadequate protocols, inadequate staff education, athletes' compliance issues, and challenges relating to real-time assessment and treatment of injuries in college athletic training. Future work also entails the advance of advanced technologies for the estimate and diagnosis of injuries in college sports training, establishing standard operating procedures, enhancing staff development, increasing staff engagement, and enhancing access to advanced technologies.

**Ethical approval:** Not applicable.

**Conflict of interest:** The author declares no conflict of interest.

## References

1. Rodrigues, F., Monteiro, D., Ferraz, R., Branquinho, L. and Forte, P., 2023. The association between training frequency, symptoms of overtraining, and injuries in young men soccer players. International journal of environmental research and public health, 20(8), p.5466.
2. Zadeh, A., Taylor, D., Bertsos, M., Tillman, T., Nosoudi, N. and Bruce, S., 2021. Predicting sports injuries with wearable technology and data analysis. Information Systems Frontiers, 23, pp.1023-1037.
3. Wang, L. and Sun, Q., 2023. Common Injuries and Healthcare in College Basketball Sports. International Journal of Healthcare Information Systems and Informatics (IJHISI), 18(1), pp.1-11.
4. Zhao, F., 2024. The Application of Sports Biomechanics in Sports Injury Prevention and Rehabilitation. Frontiers in Sport Research, 6(3).
5. Xiang, C., 2021, September. Early warning model of track and field sports injury based on RBF neural network algorithm. In Journal of Physics: Conference Series (Vol. 2037, No. 1, p. 012084). IOP Publishing.
6. Wang, X. and Guo, Y., 2021. Attribute Reduction Algorithm Based Early Warning Model of Sports Injury. Informatica, 45(5).
7. Sillero-Quintana, M., Gomez-Carmona, P.M. and Fernández-Cuevas, I., 2021. Infrared thermography as a means of monitoring and preventing sports injuries. In Research Anthology on Business Strategies, Health Factors, and Ethical Implications in Sports and eSports (pp. 832-865). IGI Global.
8. Wenjie, E., Yu, Q. and Guo, H., 2022. Sports medicine image modeling for injury prevention in basketball training. Contrast Media & Molecular Imaging, 2022.
9. Schiepek, G., Schorb, A., Schöller, H. and Aichhorn, W., 2023. Prediction of sports injuries by psychological process monitoring. Sports Psychiatry: Journal of Sports and Exercise Psychiatry.
10. Salemi, S., Hamidi, M., Moghadas, Y., Alizadeh, H. and Karimizadeh, M., 2023. Strategy of Preventing Sports Injuries of Students in Schools with Grounded Theory Approach. Physical Treatments-Specific Physical Therapy Journal, 13(4), pp.287-294.
11. Mishra, N., Habal, B.G.M., Garcia, P.S. and Garcia, M.B., 2024, June. Harnessing an AI-Driven Analytics Model to Optimize Training and Treatment in Physical Education for Sports Injury Prevention. In Proceedings of the 2024 8th International Conference on Education and Multimedia Technology (pp. 309-315).
12. Rommers, N., Rössler, R., Verhagen, E., Vandecasteele, F., Verstockt, S., Vaeyens, R., Lenoir, M., D'Hondt, E. and Witvrouw, E., 2020. A machine learning approach to assess injury risk in elite youth football players. Medicine and science in sports and exercise, 52(8), pp.1745-1751.
13. Liu, H., 2023. Value evaluation of knee joint sports injury detection model-aided diagnosis based on machine learning. Frontiers in Physics, 11, p.1166275.
14. Meng, L. and Qiao, E., 2023. Analysis and design of dual-feature fusion neural network for sports injury estimation model. Neural Computing and Applications, 35(20), pp.14627-14639.
15. Dhanke, J.A., Maurya, R.K., Navaneethan, S., Mavaluru, D., Nuhmani, S., Mishra, N. and Venugopal, E., 2022. Recurrent neural model to analyze the effect of physical training and treatment in relation to sports injuries. Computational Intelligence and Neuroscience, 2022(1), p.1359714.
16. Wu, X., Zhou, J., Zheng, M., Chen, S., Wang, D., Anajemba, J., Zhang, G., Abdelhaq, M., Alsaqour, R. and Uddin, M., 2022. Cloud-based deep learning-assisted system for diagnosis of sports injuries. Journal of Cloud Computing, 11(1), p.82.
17. Li, W., 2024. A Big Data Approach to Forecast Injuries in Professional Sports Using Support Vector Machine. Mobile Networks and Applications, pp.1-17.
18. Zhang, Y., Zhang, Q. and Liu, Y., 2023. Athlete injury detection and emergency treatment in mobile smart medical system. Frontiers in Physics, 11, p.1191485.
19. Han, H. and Li, Y., 2024. Magnetic resonance imaging diagnosis of ankle joint athletic injury based on machine learning algorithms. Molecular & Cellular Biomechanics, 21(3), pp.414-414.

20. Luo, Y., Liang, J., Lin, S., Bai, T., Kong, L., Jin, Y., Zhang, X., Li, B. and Chen, B., 2024. The application of deep learning methods in knee joint sports injury diseases. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11(7), pp.2689-2696.
21. Liu, Z., Han, X., Yang, Y., Wang, W. and Liu, F., 2023. Sports injury detection mechanism based on multi-sensor fusion. International Journal of Computer Applications in Technology, 71(3), pp.229-236.
22. Li, J., 2024. An investigation of an athlete injury likelihood monitoring system using the random forest algorithm and DWT. Technology and Health Care, (Preprint), pp.1-15.
23. Chen, S., Guo, L., Xiao, R., Ran, J., Li, H. and Reynoso, L.C., 2023. Establishing a cognitive evaluation model for injury risk assessment in athletes using RBF neural networks. Soft Computing, 27(17), pp.12637-12652.
24. Liu, F., A Risk Assessment System for Sports Training Injuries Based on Artificial Intelligence and Big Data. Applied Mathematics and Nonlinear Sciences, 9(1).
25. Jauhiainen, S., Kauppi, J.P., Leppänen, M., Pasanen, K., Parkkari, J., Vasankari, T., Kannus, P. and Äyrämö, S., 2021. New machine learning approach for detection of injury risk factors in young team sport athletes. International journal of sports medicine, 42(02), pp.175-182.
26. Kaggle. College Sports Injury Detection. Available online: https://www.kaggle.com/datasets/ziya07/college-sports-injury-detection/data (accessed on 6 December 2024).
27. She, H. Application of Big Data Analysis in Model Construction to Prevent Athlete Injury in Training. Applied Mathematics and Nonlinear Sciences, 2024, Sciendo, vol. 9 no. 1. https://doi.org/10.2478/amns-2024-1723
28. Arciniega-Rocha, R.P., Erazo-Chamorro, V.C., Rosero-Montalvo, P.D. and Szabó, G., 2023. Smart wearable to prevent injuries in amateur athletes in squats exercise by using lightweight machine learning model. Information, 14(7), p.402.

# Appendix

**Table A1.** List of abbreviations.

| **PL** | **Pooling layer** | **BP** | **Back Propagation** |
|---|---|---|---|
| DFFNN | Dual-feature fusion neural network | BDA | Big Data Analytics |
| ACL | Anterior Cruciate Ligament | SD | Standard deviation |
| IL | Input layer | ICD | Inter-class distance |
| GD | Gradient Descent | DLS | Deep learning-assisted system |
| AI | Artificial intelligence | SAW | Sit-and-wait |
| OL | Output layer | CF | Cost function |
| DM | Demixing matrix | CNN | Convolutional Neural Network |
| SVM | Support Vector Machine | CL | Convolutional layer |
| PD | Partial derivative | SS | Step size |
| ResNet | Residual Neural Network | ML | machine learning |