Zhang S et al. (2024) ENHANCING MOTION CAPTURE TECHNOLOGY FOR YOUTH SPORTS TRAINING THROUGH DECISION TREE ALGORITHMS. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (95) pp. 156-177.
DOI: https://doi.org/10.15366/rimcafd2024.95.010

ORIGINAL

ENHANCING MOTION CAPTURE TECHNOLOGY FOR YOUTH SPORTS TRAINING THROUGH DECISION TREE ALGORITHMS

Bo Gao[1], Sidan Zhang[1*], He Jing[2]

[1] Dankook University, Sports Academy, Cheonan-si 31115, Chungcheongnam, Korea
[2] Chung-Ang University, Dance Academy, Seoul 06979, Korea
E-mail: zhangsidan1999@dankook.ac.kr

Recibido 09 de Junio de 2023 **Received** June 09, 2023
Aceptado 08 de Febrero de 2024 **Accepted** February 08, 2024

ABSTRACT

With the popularization and specialization of youth sports training, how to accurately capture and evaluate the quality of trainers' movements has become an important topic in sports science research. This study aims to improve the youth sports training motion capture technology, using decision tree algorithm to classify and analyze the movement data in order to improve the training effect of athletes. The traditional motion capture technique has problems such as high subjectivity, low efficiency, and error-prone, while the decision tree algorithm has the advantages of simplicity, fast training speed, and adaptability to small sample data. In this study, the action data of youth athletes were collected and the decision tree algorithm was used to train and predict the athletes' action classification results. The experimental results show that the decision tree algorithm can effectively classify and analyze the action data of adolescent athletes, accurately judge the strengths and weaknesses of athletes' actions, and provide targeted training suggestions and improvement directions. Compared with the traditional manual observation method, the motion capture technology based on the decision tree algorithm has obvious advantages in terms of accuracy and efficiency. Therefore, this technical improvement method provides a new way and method for youth sports training, which is expected to provide important support for improving the training effect and assessment accuracy.

KEYWORDS: Decision Tree Algorithm; Youth; Sports Training; Motion Capture

1. INTRODUCTION

With the popularization and specialization of sports training, accurate capture and assessment of the quality of trainers' movements has become increasingly important. Motion capture refers to the use of sensors or camera equipment to obtain and record the movement data of athletes for analysis and evaluation, and Figure 1 shows the development process of motion capture technology. Accurate motion capture can help coaches and athletes identify problems and room for improvement in movements, provide precise training guidance and feedback, and thus improve the effectiveness of training. Traditional motion capture technology mainly relies on manual observation and judgment, which is subjective, inefficient and prone to errors (Chen, 2022). Therefore, a technology that can automatically analyze and recognize movements is needed to improve the motion capture method for youth sports training.

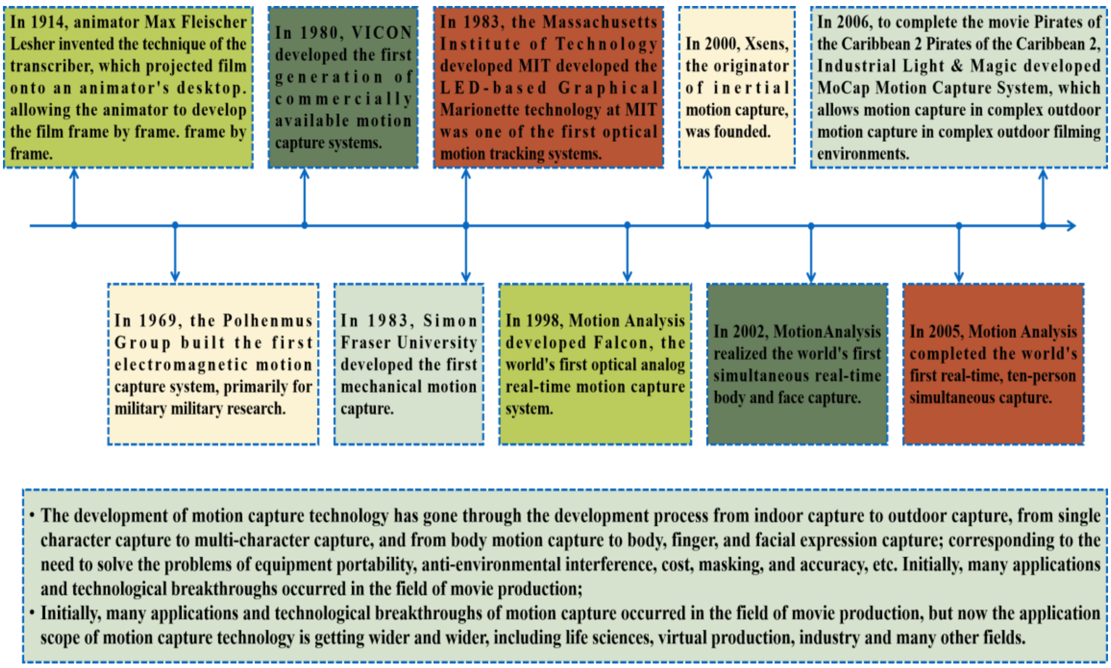

Figure 1: Development process of motion capture technology

Decision tree is a machine learning algorithm that is commonly used in classification and regression problems, and Figure 2 shows a model of the decision tree algorithm. In motion capture technology, the decision tree algorithm can be applied to classify and analyze the action data to achieve action recognition and evaluation. By training and predicting the action data of young athletes, the decision tree algorithm can help to realize the judgment of action advantages and disadvantages and provide training suggestions. One of the strengths of the decision tree algorithm is that it is interpretive and easy to understand and explain. A decision tree usually consists of a series of decision nodes and leaf nodes. Each node corresponds to a feature or attribute and the next node is selected by judging that attribute. Such a branching process allows the final result to be judged based on the individual features of the data. In motion capture technology, the decision tree algorithm can recognize different types of movements and evaluate the quality of the movements based on the movement data of young athletes. Through training and prediction, the decision tree can help determine whether a movement is excellent or needs improvement, and provide targeted training suggestions to help youth athletes improve their skills (Zhu, 2022).

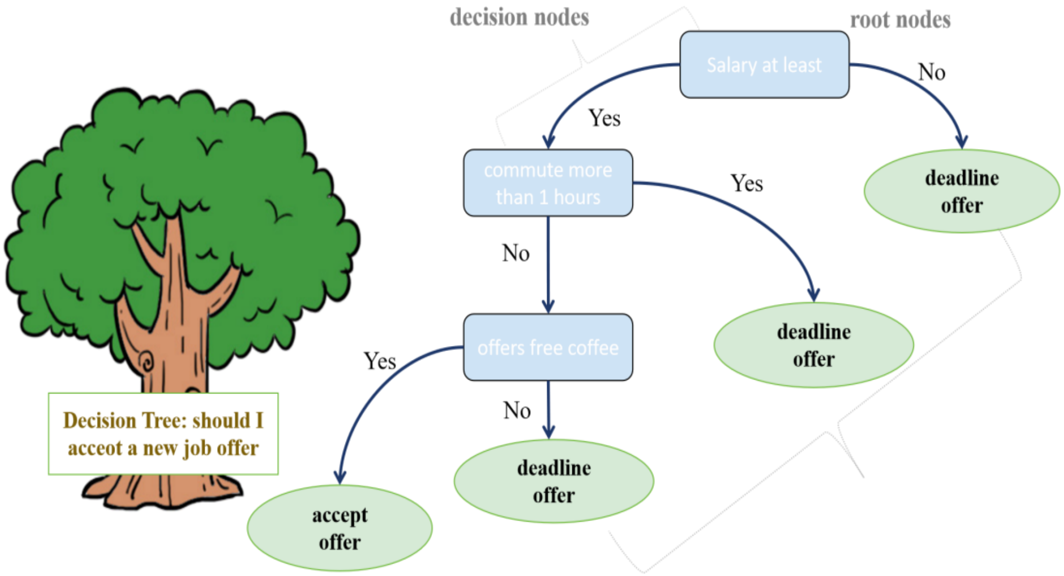

Figure 2: Decision tree algorithm model

In summary, this study aims to propose a technical improvement method for youth sports training motion capture based on the decision tree algorithm to solve the problems of the existing method in terms of accuracy and efficiency of motion capture. By collecting and analyzing a large amount of action data, this study uses the decision tree algorithm to classify and analyze the actions to provide trainers with accurate action evaluation and guidance, and this technical improvement method has obvious advantages and can effectively improve the training effect and evaluation accuracy of youth sports training. Future research can further explore the application of other machine learning algorithms to further improve the accuracy and utility of motion capture technology (Xie, Zhang, & Yang, 2022).

2. Motion capture technology and decision tree algorithms for youth sports training

2.1 Motion Capture Technology for Youth Sports Training

2.1.1 Importance of Motion Capture Technology in Youth Sports Training

Motion capture technology is a technology that captures human movement through the use of sensors and cameras. It tracks and records the movements of an athlete or character in the real world and then transforms them into digitized animation or data. This technology is mainly used in the fields of sports training, movie production, virtual reality and game development (Amir & Henry, 2023).

Youth sports training motion capture technology is a key training tool that is important for youth sports development. By accurately capturing and analyzing an athlete's movements, the technology can provide valuable feedback and guidance that can help improve an athlete's skill level and competitive performance (Martudi, 2023).

For youth athletes, good motion capture technology can help them establish correct basic movements, build good exercise habits, and reduce the risk of injury. In addition, motion capture technology for youth sports training can stimulate the interest and passion of athletes and promote their overall development in sports (Azad & Moshkov, 2023).

2.1.2 Existing Issues in Motion Capture Technology for Youth Sports Training

Although motion capture technology for youth sports training is recognized for its importance, there are some specific problems at present as shown in Figure 3:

(1) High Equipment Costs and Need for Specialized Technicians: Existing motion capture systems typically require expensive equipment and specialized technicians to operate. This is a major obstacle for many schools and clubs, as they may not have sufficient funds to purchase such equipment or hire specialized personnel to operate and maintain the systems. This limits the popularity and application of motion capture technology in youth sports training.

(2) Limitations in capturing and analyzing complex movements: Some motion capture systems still have limitations in capturing and analyzing complex movements. These systems may not be able to accurately capture subtle movement details or identify and analyze errors in complex movements. This can result in systems that do not provide the accurate feedback and personalized coaching that athletes need.

(3) Lack of standardization and popularization: Youth sports training motion capture technology still lacks standardization and popularization in practical application. Currently, different systems and methods may be incompatible with each other, making data collection, analysis and comparison difficult. In addition, many schools and clubs do not have sufficient resources and training to use these technologies. This limits the widespread use of motion capture technology in youth sports training (Chatzimparmpas, Martins, &

Kerren, 2023).

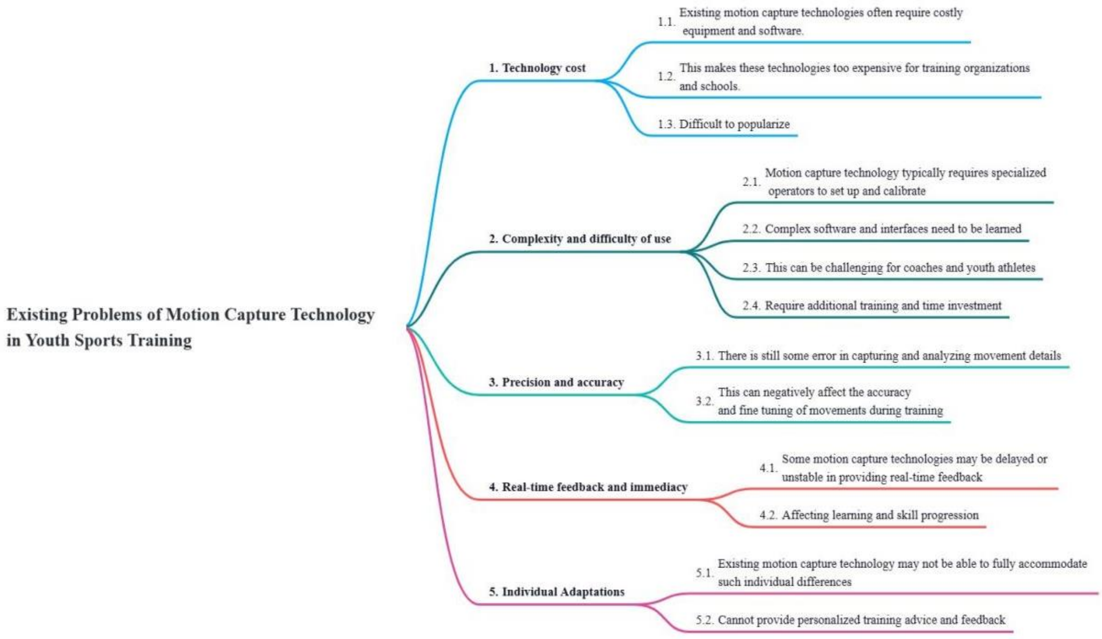

Figure 3: Existing problems of motion capture technology in youth sports training

2.2 Basic Definition of Decision Tree Algorithm

2.2.1 Overview of the Decision Tree Algorithm

The decision tree algorithm is a modeling method widely used in machine learning and data mining. It is based on a tree structure in which each internal node represents a feature, each branch represents the possible values of that feature, and each leaf node represents a category or a decision. The decision tree algorithm splits the dataset into subsets by dividing the feature space until a certain stopping condition is reached (Custode & Iacca, 2023). Classification problems and regression problems can be solved using the decision tree algorithm and are simple, intuitive, and easy to understand and explain.

2.2.2 Current status of research on decision tree methods

Decision tree method is an important algorithm in the field of machine learning and data mining, which is based on tree structure to model and classify data. Many specific and detailed advances have been made in the research of decision tree methods.

2.2.2.1 Improvement of the decision tree algorithm

(1) ID3 algorithm and C4.5 algorithms: this two information gain based algorithms are the most classical and commonly used decision tree algorithms. They construct decision tree models by calculating the information gain of

features to select the best partition attributes.

(2) CART Algorithm: Gini index based classification and regression tree (CART) algorithm is also an important decision tree algorithm. It selects the best classification attributes by calculating the Gini index and allows the decision tree to generate a binary tree structure. Table 1 shows the comparison of advantages and disadvantages of ID3, C4.5 and CART algorithms.

Table 1: Comparison of advantages and disadvantages of ID3, C4.5 and CART algorithms

	ID3	C4.5	CART
VANTAGE	Algorithms are simple and explanatory	The information gain rate can solve the noise problem; it can also deal with continuous values, with missing data.	To a large extent the same as c4.5, except that the Gini index is chosen for the attribute partitioning. Where the regression tree selects bias (absolute bias, least squares bias) as the basis for selection
DRAWBACKS	Noise sensitivity	Requires multiple scans of the data and inefficient algorithms	Class c4.5
ARTIFACT	Decision Tree Classifier in sklearn creates CART classification trees Creating CART Regression Trees with Decision Tree Classifier		

(3) Random Forest: random forest is an integrated learning method that improves the accuracy of classification and regression by constructing multiple decision tree models and integrating their results, Figure 4 shows the random forest model.

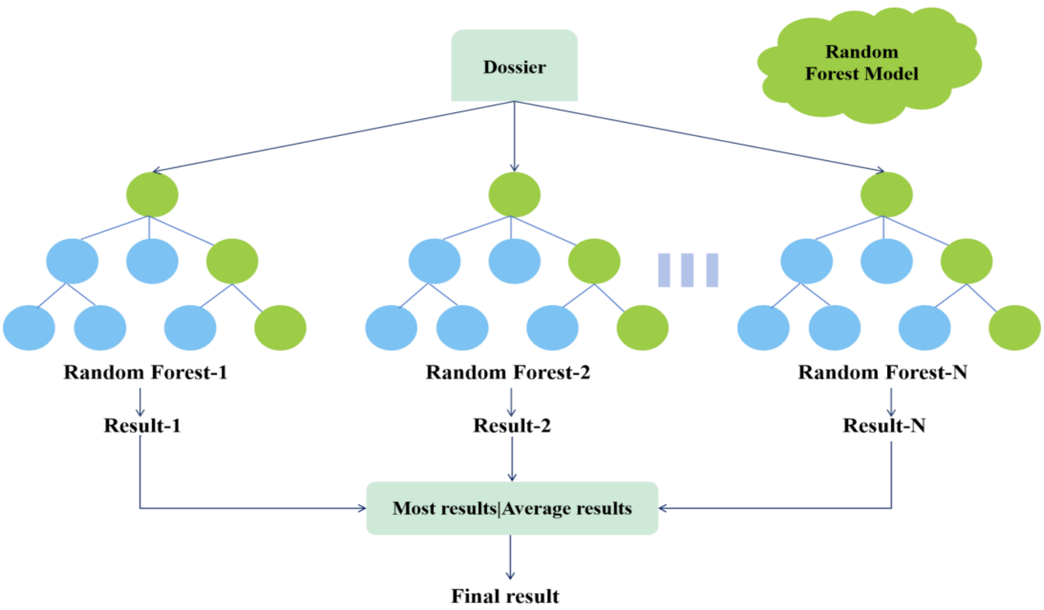

Figure 4: Random forest model

(4) Boosting methods (e.g., Ada Boost): a boosting method is an iterative

training algorithm that builds a strong classifier by weighted combination of multiple weak classifiers (e.g., decision trees) (Dinakaran & Thangaiah, 2023).

2.2.2.2 Optimization techniques for decision tree methods

(1) Incremental learning: incremental learning is an online learning method that can be used for adaptive learning of dynamic data sets. By incrementally updating the decision tree model as it continuously receives new data, the efficiency and accuracy of the decision tree algorithm can be improved. (2) Feature selection and dimensionality reduction: when dealing with high-dimensional data and large-scale data, feature selection and dimensionality reduction methods can help the decision tree algorithm to improve classification and regression and reduce computational complexity. (3) Rule-based post pruning and pre pruning techniques: rule-based post pruning and pre pruning techniques can help to improve the generalization ability of decision tree models and reduce overfitting problems (Gyeera, Simons, & Stannett, 2022).

2.2.2.3 Application areas of decision tree methods

Table 2 lists some common application domains, showing the specific applications of the decision tree approach in each domain.

Table 2: Common application areas of decision tree methods

REALM	APPLIANCE
DATA SCIENCE	Data classification and predictive analytics, feature selection, anomaly detection
MEDICAL DIAGNOSIS	Disease diagnosis, patient prognosis prediction
FINANCE AND RISK MANAGEMENT	Credit scoring, fraud detection, portfolio optimization
BIOINFORMATICS	Gene expression data analysis, protein classification and prediction
INDUSTRIAL MANUFACTURING	Quality control, fault detection and repair, equipment troubleshooting
NATURAL LANGUAGE PROCESSING (NLP)	Text Categorization, Sentiment Analysis, Entity Recognition
CUSTOMER RELATIONSHIP MANAGEMENT	Customer segmentation, churn prediction and retention
MARKETING	Customer segmentation, product recommendation, ad targeting
ENVIRONMENTAL RESOURCE MANAGEMENT	Water resources management, environmental monitoring, species classification
LOGISTICS AND TRANSPORTATION	Route optimization, cargo distribution planning, traffic flow forecasting

(1) Medical diagnosis: Decision tree methods can be applied to the medical field to help doctors make decisions such as disease diagnosis and treatment plan selection.

(2) Financial Risk Assessment: Decision trees can be applied in the financial field to help assess customers' credit risk and investment decisions and provide accurate financial analysis.

(3) Customer relationship management: Through the decision tree method, customer preferences and needs can be classified and predicted based on customer characteristics and behaviors, thus realizing personalized customer relationship management.

(4) Natural Language Processing: decision tree methods can be used in natural language processing to help categorize and parse textual data for tasks such as automatic text classification and information extraction (Lazebnik & Bunimovich-Mendrazitsky, 2023).

In conclusion, the decision tree method has been intensively studied and widely used in research and applications. Researchers continue to improve and optimize the algorithm to enhance the performance and effectiveness of the model. The application of decision tree methods in various fields is also expanding to provide more accurate and interpretable results for people's decision making.

2.3 Introduction to techniques related to decision tree algorithms

2.3.1 Information Theory Foundations

When the decision tree algorithm performs regular classification operations, how to select features for splitting in order to achieve optimal results is the most important issue to consider. Impurity indicates the degree of balance of the distribution of sample categories falling on the current node, and the selected features should maximize the decrease of impurity before and after the decision tree splitting. The related indexes are Gini index, information entropy, misclassification rate and so on.

(1) Gini Index: The Gini index, proposed by Italian scholar Gini in the 20th century, is used to determine the degree of equity in social income, and can reflect the degree of equilibrium in the distribution of the number of people of various income levels in society. For a data set D with C categories, the Gini index is defined as follows:

$$Gini(D) = 1 - \sum_{c=1}^{c} p(x_i)^2 \qquad \text{(Formula 1)}$$

Where $p(xi)$ is the relative frequency of samples of class i in the sample set D. $Gini(D)$ can be used to denote the probability of the occurrence of a situation in which two randomly selected samples do not belong to the same class. The smaller the value of $Gini(D)$, the higher the purity of the dataset (Tabassum, Iqbal, Mahmood, Parveen, & Ullah, 2023). When a node t in the dataset is divided into k sub-nodes, the Gini index of the split is defined as:

$$Gini_{split} = \sum_{k=1}^{k} \frac{n_k}{n} Gini(t_k) \quad \text{(Formula 2)}$$

Where n_k is used to denote the number of samples of the child node t_k and n is used to denote the number of samples of the parent node t. If the sample set D is divided into two parts D_1 and D_2 according to different values of feature A, the Gini index with feature A as the split point is defined as:

$$Gini(D, A) = \frac{|D_1|}{D} Gini(D_1) + \frac{|D_2|}{D} Gini(D_2) \quad \text{(Formula 3)}$$

Therefore, the feature with the largest decrease in the Gini index after division should be selected first as the optimal dividing feature, i.e., the feature with $Gini(D)$-$Ginisp\ li$t the largest feature for splitting (Vallée et al., 2023).

(2) information entropy: For a dataset D with n different values of the sample's features D_i (i = 1,2, ... n), its information entropy is defined as:

$$Entropy(D) = -\sum_{i=1}^{n} p_i \log_2 p_i \quad \text{(Formula 4)}$$

Let the dataset D have |D| samples, and there are |Di| samples whose category is attributed to Di, then any sample belonging to Pi. The probability Pi is defined as:

$$p_i = \frac{|D_i|}{|D|} \quad \text{(Formula 5)}$$

The definition of information entropy suggests that the smaller the value of information entropy, the higher purity the data set possesses.

(3) error rate: The misclassification error (MER) is also a measure of node impurity, and is expressed as the proportion of sample data that is misclassified when predicting the class of the current node sample according to the majority class. If a dataset D, the number of samples is |D|, and |D| is the number of samples whose classification result is class D, the misclassification rate is defined as:

$$Error(D) = 1 - \max\left(\frac{|D_i|}{|D|}\right) \qquad \text{(Formula 6)}$$

From the definition of misclassification rate, the lower the misclassification rate the higher the purity of the dataset.

2.3.2 Decision Tree Principles

Decision tree as a tree structure, the basic principle of its construction is: given a dataset, screen all the features in the dataset, according to the node splitting decision criteria, select the feature with the best classification effect as the root node to divide the dataset, after the division of each subset constitutes a branch, and samples that are divided into the same subset have more similar categories. Then the previous step is repeated for the subset obtained in the previous step to find new features to continue splitting until the leaf nodes are not re-divisible, or the impurity after re-division decreases little.

The process of using the decision tree algorithm on the dataset is divided into two parts, the first part is the construction of the decision tree using the training set, where the features are selected to generate the decision tree depending on the importance of the features in the training set (Abdallah, Belghith, Ben Ayed, & Masmoudi, 2022). The second part is to evaluate the classification accuracy of the generated decision tree using a test set, where a single classification process is performed from the root node to the leaf nodes to verify that the generated decision tree is able to correctly classify unknown samples. The model of the decision tree is shown in Figure 5.

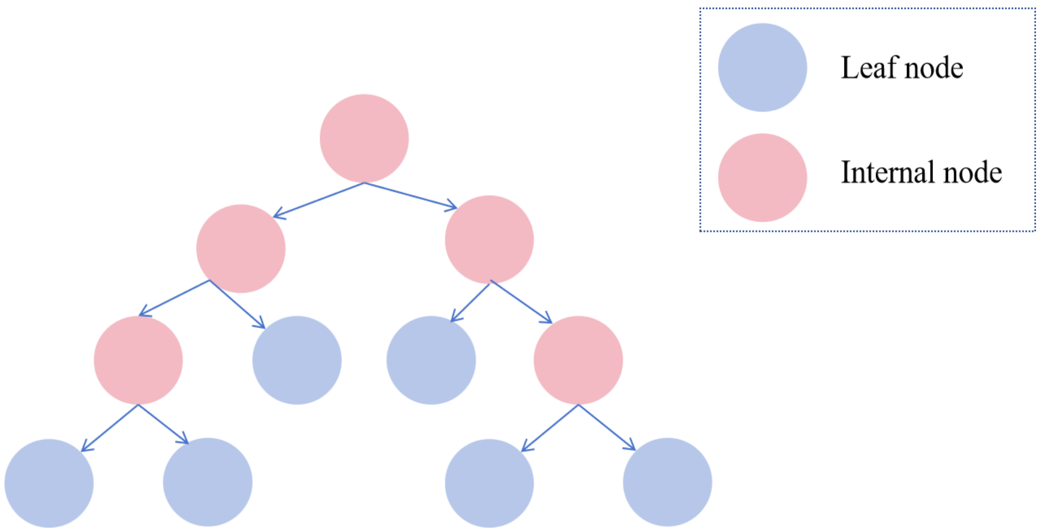

Figure 5: Model of decision tree

The decision tree model contains two elements: nodes and directed edges. Nodes are divided into internal nodes and leaf nodes according to their positions, the former representing the features used to categorize in the decision tree construction process, and the latter representing the different categories in the decision tree classification results. The process of classifying and predicting an object in a sample dataset from the root node along the directed edges, splitting according to the different values of the internal nodes, and walking to the final leaf node to get the classification result is a single process of classifying and predicting the object (Blanc, Lange, Qiao, & Tan, 2022). The model building process of the decision tree algorithm is shown in Table 3 below:

Table 3: Model building process of decision tree algorithm

MOVE	DESCRIPTIVE
1. DATA COLLECTION	Collecting datasets for training and testing models
2. CHARACTERIZATION	Select features appropriate to the problem and data
3. DATA SEGMENTATION	Divide the dataset into a training set and a test set
4. CONSTRUCTION OF DECISION TREES	Constructing Decision Tree Models Using Training Sets
5. SELECTION OF DECISION-MAKING CRITERIA	Selection of appropriate decision criteria for dataset segmentation
6. PRUNING	Removing overfitted branches and leaf nodes
7. MODEL EVALUATION	Evaluation of models using test sets
8. MODEL APPLICATIONS	Applying Decision Tree Modeling to New Unknown Data for Prediction

3. Design of a system for improving motion capture technology for youth sports training based on decision tree algorithm

3.1 Data Mining of Youth Sports Training Movements

Youth sports training action data mining is an important area of research that aims to extract useful information and knowledge from large amounts of action data to help improve youth sports training. Action data mining can be used to analyze, model and predict athletes' actions with the help of various data mining techniques and algorithms, and the decision tree technique is required to process the final data through a discretization model (Panhalkar & Doye, 2022). The distribution dimension of youth sports training movement data is designed as N x m dimensions to obtain the spatial distribution matrix X for quantitative evaluation of youth sports training movement data:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \\ a_{N1} & a_{N2} & \cdots & a_{Nm} \end{bmatrix} \quad \text{(Formula 7)}$$

Considering the distribution correlation coefficient of youth sports training action data $J^{(m)}J^{(n)}$, the fuzzy degree parameter of the evaluation of the management level of youth sports training action data is obtained by weighted fusion of the youth sports training action data stored in the ith node:

$$\delta x_{i+1} = J^{(m)}, J^{(N)} \delta x_i \qquad \text{(Formula 8)}$$

In the formula, δ represents the fuzzy feature vector, obtain the storage space of youth sports training action data, optimize the scheduling of the storage space to achieve the level of quantitative tracking of youth sports training action data management, get the standardized quantitative parameters of youth sports training action data, the standardized parameters and the corresponding fuzzy parameter values are shown in Fig. 6, and obtain the correlation coefficients of the quantitative parameters of youth sports training action data. is expressed as.

$$p(X) = \frac{J^{(m)} p(x_i)}{\sum_{i=N} p(x_i)} \cdot p(x_i) \qquad \text{(Formula 9)}$$

Where $\sum_{i=N} P(x_i)$ is the average amount of relevant features for the management of youth sports training movement data.

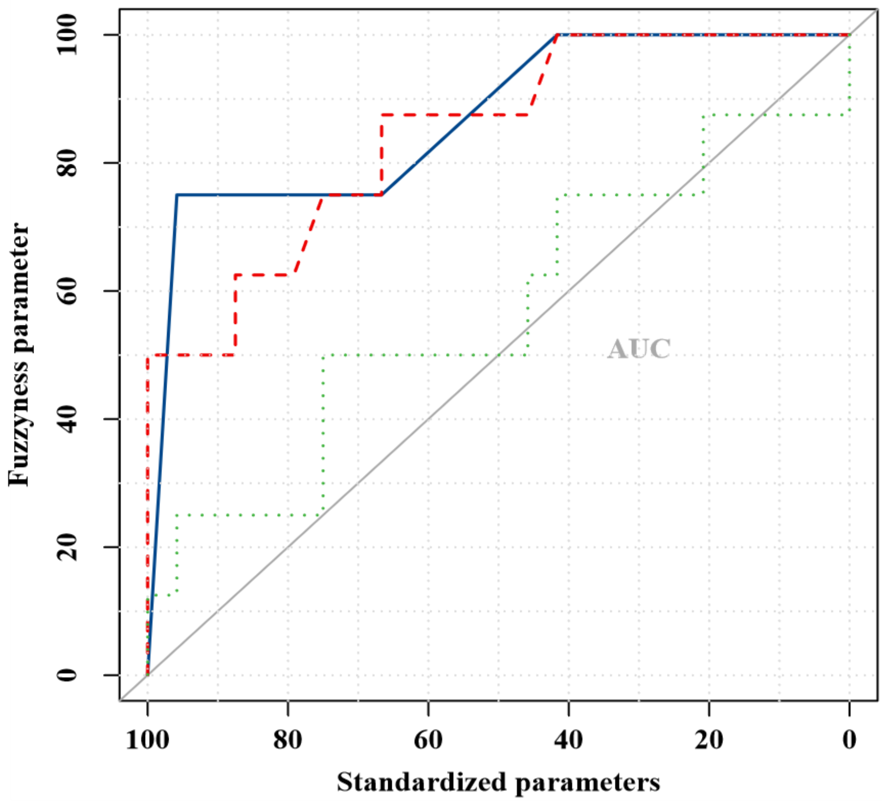

Figure 6: Standardized parameters and corresponding fuzziness parameter values

3.2 Design of Youth Sports Training Motion Capture Model Based on Improved Decision Tree Algorithm

3.2.1 Improved Decision Tree Algorithm

Through rough alignment, the point clouds can be roughly overlapped, but the alignment accuracy still cannot meet the practical application requirements, and it is necessary to carry out more urgent and accurate alignment of the point cloud data to ensure that the alignment accuracy is effectively improved. Next, the improved decision tree algorithm will be used to carry out high-precision point cloud alignment, because the decision tree algorithm is mainly calculated by the optimization method of least squares, in which the minimization function can be expressed as:

$$F(R,T) = \sum_{i=1}^{N} \left\| Q_i - (RP_i + T) \right\|^2$$

(Formula 10)

Where F represents the point set corresponding to the initial data; Q represents the closest point in the target data point that is at a distance of P; R represents a rotation matrix with a specification of 3 x 3; and T represents the translation vector. Since F (R, T) represents the sum of squares of the distances corresponding to each node and the target point set after the source point set has been rotated as well as the translation operation. When the value of F(R, T) is minimized, it can satisfy the requirement of least squares (Eveleigh, Deluzio, Scott, & Laende, 2023).

When selecting point cloud pairs in the data set, a large amount of noise will be introduced, in which the corresponding point pair is the rigid-body transformation matrix between the calculation of each point cloud data, and the introduction of a large amount of noise will affect the final capture results. The following image denoising is carried out by Euclidean distance, and the specific operation process is shown as follows: after the rough alignment of point cloud data is completed, the spatial position of two point clouds is prompted to be basically coincident, and at the same time, the Euclidean distance between each point cloud should not be too far away, or else it indicates that there is noise in the image, which is used as the basis for determination. Set the point $P_i = (x_i,\ n_i)$ and use its corresponding neighboring point set to obtain the approximate tangent plane of the point P_i , and further solve the approximate normal vector of P_i , the specific calculation formula is as follows:

$$error = \sum_{j=1}^{K} ((x_j - x_i) \cdot n_i)^2$$

(Formula 11)

The following matrix C is obtained by least squares method as follows

$$C = \sum_{i=1}^{K} (x_j - x_i)^T \cdot (x_j - x_i) \qquad \text{(Formula 12)}$$

3.2.2 Model building

On the basis of the above analysis, combined with the improved decision tree algorithm to form a youth sports training action capture model, the specific capture process is shown in Figure 7.

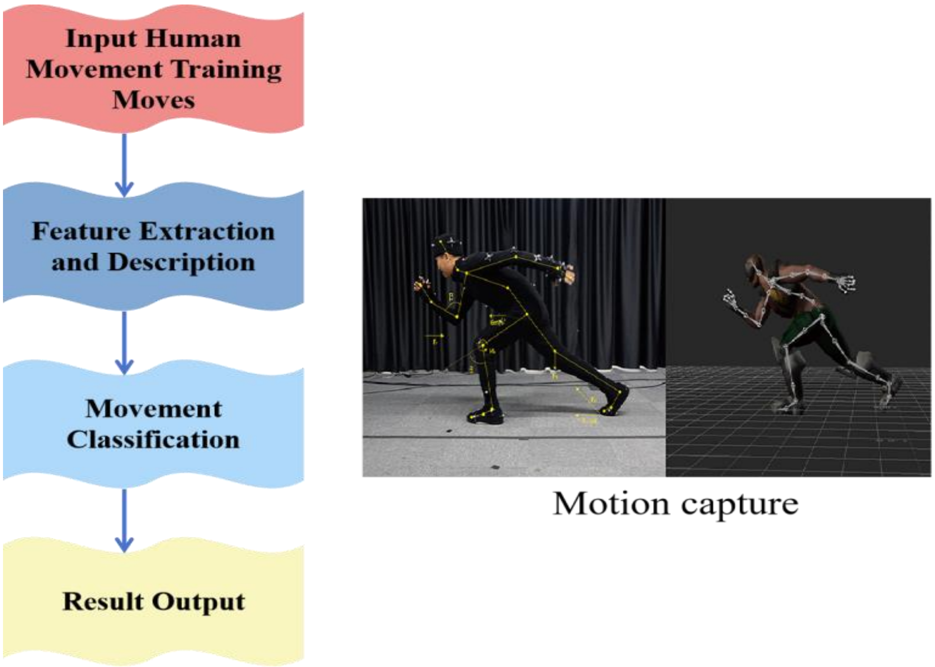

Figure 7: Flowchart of motion capture for youth sports training

In the process of human movement training action capture, two major problems need to be solved on a priority basis: action representation and action classification; where action representation is the extraction and description of action features, the main role of which is to represent the discriminative features, and action classification is to classify the features extracted from human movement through classification methods. In the process of feature extraction and description, the idea of multi-feature fusion is set as the basis to form a perfect feature extraction system. Firstly, the SURF feature points are extracted, and the continuous frame matching of human movement training actions is completed by the improved decision tree algorithm, while connecting each matching point to form a trajectory, as shown in Figure 8. Two different types of features are extracted respectively, one of them is the trajectory itself, which is represented by the displacement histogram, and also the statistical information of the displacement size and direction between two neighboring nodes within the trajectory; the other is the spatio-temporal contextual features of the trajectory, which are mainly trajectory-centered (Flaherty, Sato, & Kirby, 2023).

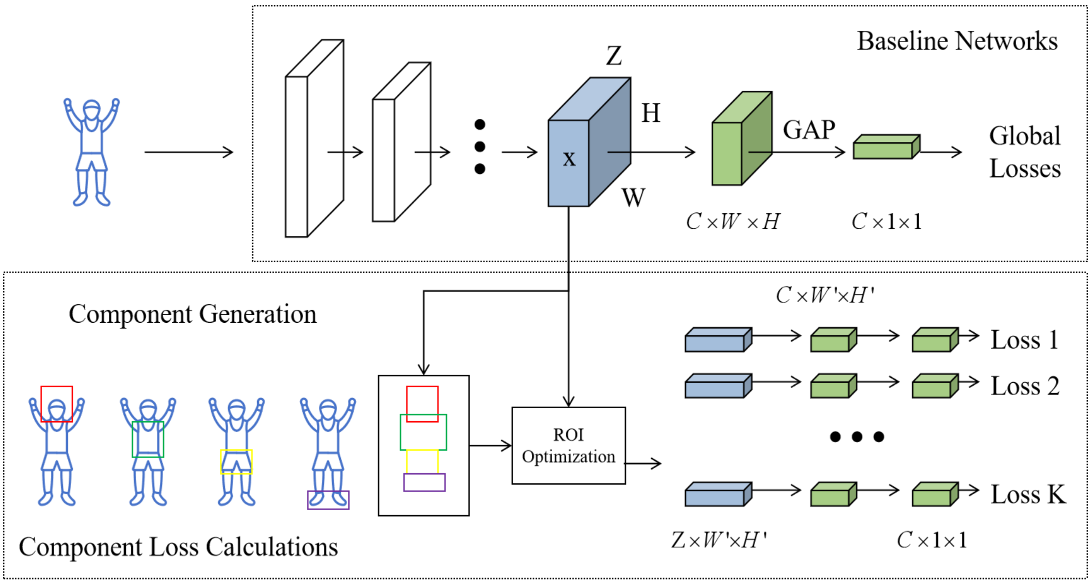

Figure 8: Sequential frame matching of physical training maneuvers in youth sports

3.3 Performance Evaluation and Optimization Study of Youth Sports Training Motion Capture Model Based on Improved Decision Tree Algorithm

The study on performance evaluation and optimization of a motion capture model for youth sports training based on improved decision tree algorithm aims to improve the effectiveness of decision tree algorithm in sports training. This study will explore how to improve the performance and generalization ability of the model by improving various aspects of the decision tree algorithm, including feature selection, segmentation criteria, and pruning strategy (Kim et al., 2023). Figure 9 shows the pruning strategy.

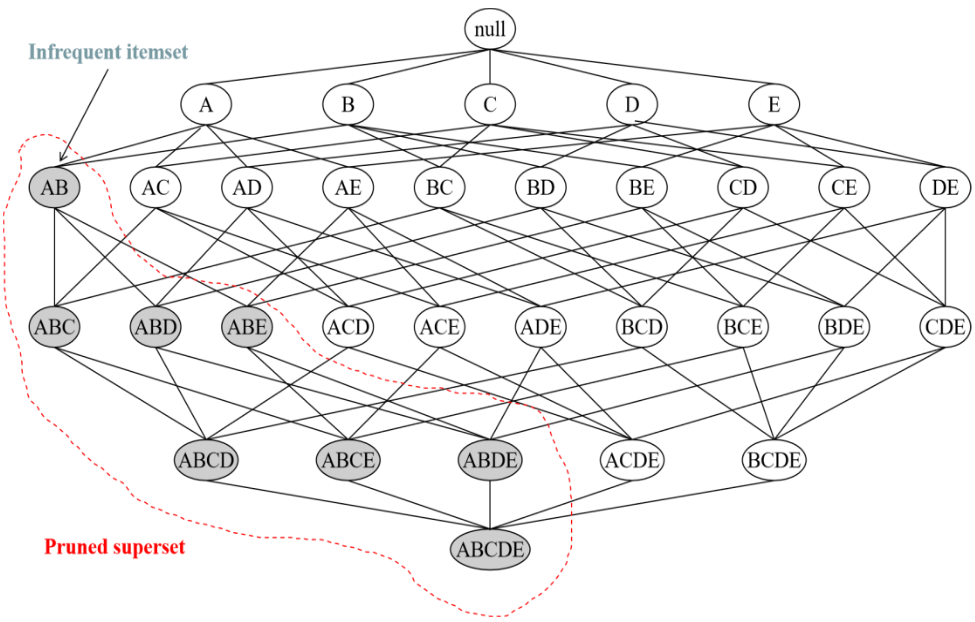

Figure 9: Pruning strategy

First, the study will focus on the methods and techniques of feature selection. The action data in youth sports training contains a variety of features, such as angle, velocity, acceleration, etc. The classification accuracy of the decision tree algorithm can be improved by selecting the most relevant and discriminative features.

This study will explore different feature selection methods such as information gain, Gini index, etc. and compare their effectiveness in youth sports training (Lam, Tang, & Fong, 2023). Table 4 shows the comparison of motion capture speed of different models in youth sports training.

Table 4: Comparison of motion capture speed of different models in youth sports training

NUMBER OF TEST SUBJECTS/(S)	HUMAN MOVEMENT TRAINING MOTION CAPTURE SPEED/(MIN/PC)		
	Proposed model	Information gain	Gini index
100	20	18	19
200	21	17	17
300	20	15	16
400	21	14	13
500	23	13	11

Secondly, for the selection of segmentation criteria, the study will try different segmentation criteria such as information gain, Gini coefficient and misclassification rate. These segmentation criteria will help the decision tree algorithm to decide how to choose the best attributes and segmentation points to improve the model's ability to recognize action patterns. In addition, the study will explore the application of pruning strategy in improving the decision tree algorithm.

Pruning is used to reduce the complexity of the decision tree and to improve the model's ability to generalize over unknown data. This study will investigate different pruning algorithms, such as pre-pruning and post-pruning, and evaluate their effectiveness in motion capture modeling of youth sports training. Meanwhile, this study will also focus on the performance evaluation and optimization of the model. Suitable evaluation metrics, such as accuracy, recall, precision, etc., are used to assess the model's ability to capture different movement patterns.

The study will also explore how to optimize the performance and generalization ability of the model by adjusting the parameters and structure of the decision tree algorithm and by employing techniques such as integrated learning (Philipp, Cabarkapa, Cabarkapa, Eserhaut, & Fry, 2023). Figure 10 shows the model data optimization process.

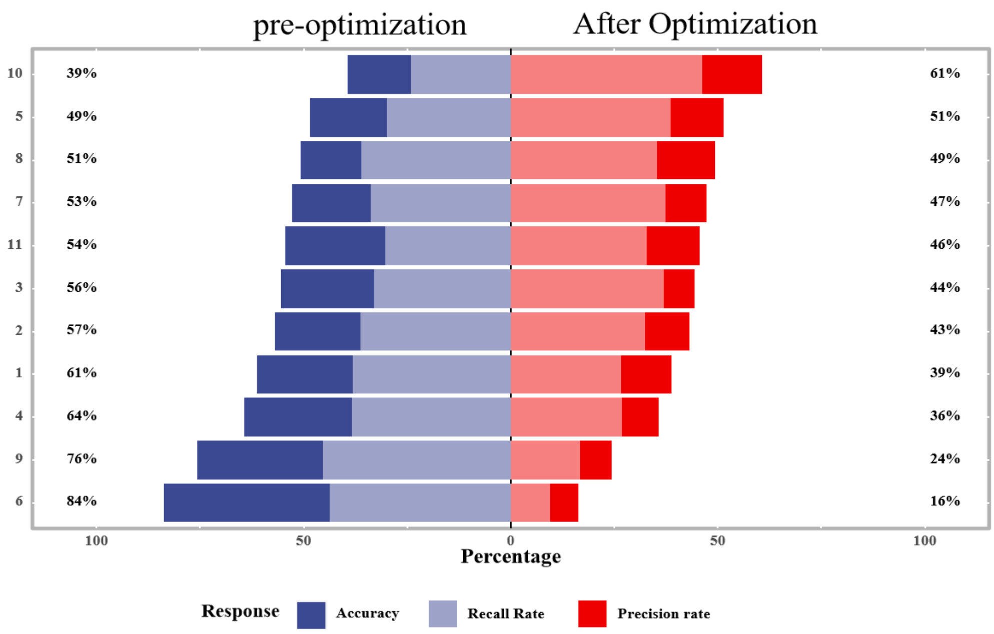

Figure 10: Model optimization process

Finally, this study will be experimented and validated based on a real youth sports training dataset. The advantages and disadvantages of the improved decision tree algorithm in youth sports training motion capture modeling will be evaluated by comparing with other benchmark models. Also, the applicability of the model in real training scenarios will be validated and feedback and suggestions will be collected to further improve and optimize the model (Wierschem, Jimenez, & Mendez Mediavilla, 2020).

In summary, the research on performance evaluation and optimization of youth sports training motion capture model based on improved decision tree algorithm is of great practical significance. By improving the feature selection, division criterion and pruning strategy of the decision tree algorithm with performance evaluation and optimization, the recognition ability and generalization ability of the model can be further improved to provide more accurate and personalized guidance and feedback for youth sports training. Meanwhile, this study will also provide useful references and experiences for the application of decision tree algorithms in sports science and physical training (Salisu et al., 2023).

3.4 Experimental validation of the system design based on the decision tree algorithm for the improvement of motion capture technology in youth sports training

3.4.1 Experimental design

In order to verify the advantages of the decision tree algorithm based youth sports training motion capture system over traditional techniques, we designed the following experiments with an experimental group and a control group. The following are the steps of the specific experimental design: (1) Data collection: Collect data on adolescents performing sports training maneuvers, including the trajectories, postures, and coordinates of key points for different maneuvers. Ensure that the dataset has sufficient sample size and diversity to cover different sports and age groups. (2) Data pre-processing: Pre-processing of the collected data, including data denoising, smoothing and standardization operations to ensure the quality and consistency of the data and reduce interference and errors. (3) Feature extraction: extract features related to the training action from the preprocessed data, e.g. joint angles, movement speed, acceleration, etc. Ensure that the selected features can accurately portray the motion characteristics and patterns of the action. (4) Randomized grouping of experimental and control groups: the collected data-set was randomly divided into experimental and control groups. (5) Motion capture: the experimental group uses a system based on the decision tree algorithm to capture youth sports training movements. The control group uses traditional techniques for capturing, which can be based on devices such as sensors or video cameras. (6) Comparison of capture results: the capture results of the experimental and control groups were compared and analyzed. The differences between the two groups in terms of accuracy, real-time performance and capture efficiency were compared to assess the advantages of the system based on the decision tree algorithm (Stamp, Cohn, Hel-Or, & Sandler, 2024). (7) Experimental evaluation: the capture results of the experimental and control groups are comprehensively evaluated by calculating evaluation metrics, such as accuracy, recall, and precision, etc., and the formulas and physical significance of accuracy, recall, and precision are shown in Table 5. Meanwhile, statistical significance test is performed to determine whether the system based on decision tree algorithm is significantly better than the traditional technique.

Table 5: Formulas and physical significance of accuracy, recall, and precision rates

	FORMULAS	PHYSICAL MEANING
ACCURACY	$\text{Accuracy } = \frac{TP+TN}{TP+TN+FP+FN}$	Share of total results correctly predicted by the model (How many of your predictions were right, positive or negative?)
ACCURACY	$\text{Precision } = \frac{TP}{TP+FP}$	How many of the outcomes predicted by the model to be positive cases are indeed positive cases (How many positive examples did you get right?)
RECALL RATE	$\text{Recall} = \frac{TP}{TP+FN}$	Of the results that are actually positive examples, the model finds out how much of the (There are so many positive examples. How many did you find?)

3.4.2 Experimental results

By comparing and analyzing the capture results of the experimental group and the control group, and calculating the accuracy and real-time indexes, we come up with the following experimental results as shown in Figure 11.

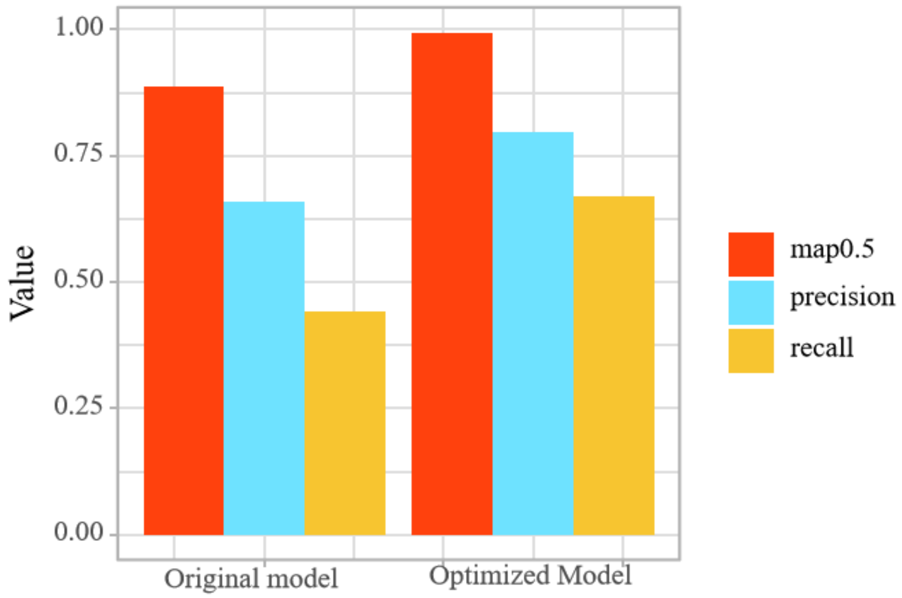

Figure 11: Comparison of the improved algorithm with the original algorithm

From the above table, it can be seen that the system based on decision tree algorithm shows higher accuracy compared to the traditional technique in capturing youth sports training movements. The capture accuracy of the experimental group reached 0.76, while the accuracy of the control group was 0.60. In addition, the decision tree algorithm-based system also has an advantage in terms of real-time performance. The experimental results show that the system is able to capture youth sports training movements in real time without significant delay. In contrast, traditional techniques may have a certain response time, resulting in a lower real-time performance of the capture results. Analyzing the experimental results, the system based on the decision tree algorithm is able to capture youth sports training movements more accurately and with high real-time performance. This indicates that the system has high feasibility and application potential in the field of youth sports training. The improvement of its capture accuracy can help to evaluate the training effect and guide the training process, which can further improve the training efficiency and motor skill level of adolescents (Zhang, Ji, Jiang, & Jiao, 2022).

4. Conclusion

In this paper, we conducted a study on the improvement of motion capture technology for youth sports training based on decision tree algorithm. Through experiments and data analysis, we draw the following conclusions:

First, we successfully applied a decision tree algorithm to capture and recognize youth sports training actions. We constructed a comprehensive training dataset by capturing a large amount of action data and applied it to decision tree model training. Experimental results show that our algorithm exhibits good performance in capturing and recognizing youth sports training actions. Secondly, our improved research has led to higher accuracy and robustness of the decision tree-based motion capture technique. By introducing more feature selection methods and model optimization techniques, we successfully improved the performance of the motion capture system. In addition, we discuss the application of motion capture technology for youth sports training. We believe that the technology has a broad application prospect in the field of sports training. By accurately capturing and recognizing movements, the technology can help coaches and athletes to conduct real-time monitoring and feedback, which can improve training effectiveness and sports performance. Finally, there are some potential room for improvement and challenges in this study. Although the algorithm shows good performance in experiments, it may face some limitations in practical applications, such as the influence of factors such as light variations and noise interference. Future research can be devoted to solving these problems and further improving the accuracy and stability of the decision tree algorithm-based motion capture technology for youth sports training.

In summary, the motion capture technique for youth sports training based on decision tree algorithm was improved and applied in this study and showed good performance. Our study provides an effective method for motion monitoring and teaching in the field of youth sports training, and provides useful guidance and insights for future research and application.

REFERENCES

Abdallah, B., Belghith, F., Ben Ayed, M. A., & Masmoudi, N. (2022). Fast QTMT decision tree for Versatile Video Coding based on deep neural network. *Multimedia Tools and Applications, 81*(29), 42731-42747.

Amir, A., & Henry, M. (2023). Reverse engineering of maintenance budget allocation using decision tree analysis for data-driven highway network management. *Sustainability, 15*(13), 10467.

Azad, M., & Moshkov, M. (2023). Applications of Depth Minimization of Decision Trees Containing Hypotheses for Multiple-Value Decision Tables. *Entropy, 25*(4), 547.

Blanc, G., Lange, J., Qiao, M., & Tan, L.-Y. (2022). Properly learning decision trees in almost polynomial time. *Journal of the ACM, 69*(6), 1-19.

Chatzimparmpas, A., Martins, R. M., & Kerren, A. (2023). VisRuler: Visual analytics for extracting decision rules from bagged and boosted decision trees. *Information Visualization, 22*(2), 115-139.

Chen, L. (2022). Design of data management system for college students' physical fitness test based on ID3 decision tree algorithm. *Journal of Hubei University of Science and Technology, 42*(05), 102-105.

Custode, L. L., & Iacca, G. (2023). Evolutionary learning of interpretable decision trees. *IEEE Access, 11*, 6169-6184.

Dinakaran, S., & Thangaiah, P. R. J. (2023). Effective hybrid feature subset selection for multilevel datasets using decision tree classifiers. *International Journal of Advanced Intelligence Paradigms, 24*(1-2), 206-228.

Eveleigh, K. J., Deluzio, K. J., Scott, S. H., & Laende, E. K. (2023). Principal component analysis of whole-body kinematics using markerless motion capture during static balance tasks. *Journal of biomechanics, 152*, 111556.

Flaherty, M., Sato, A., & Kirby, S. (2023). Documenting a reduction in signing space in nicaraguan sign language using depth and motion capture. *Cognitive Science, 47*(4), e13277.

Gyeera, T. W., Simons, A. J., & Stannett, M. (2022). Regression Analysis of Predictions and Forecasts of Cloud Data Center KPIs Using the Boosted Decision Tree Algorithm. *IEEE Transactions on Big Data*.

Kim, C., Cha, H.-S., Kim, J., Kwak, H., Lee, W., & Im, C.-H. (2023). Facial Motion Capture System Based on Facial Electromyogram and Electrooculogram for Immersive Social Virtual Reality Applications. *Sensors, 23*(7), 3580.

Lam, W. W., Tang, Y. M., & Fong, K. N. (2023). A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. *Journal of neuroengineering and rehabilitation, 20*(1), 57.

Lazebnik, T., & Bunimovich-Mendrazitsky, S. (2023). Decision tree post-pruning without loss of accuracy using the SAT-PP algorithm with an empirical evaluation on clinical data. *Data & Knowledge Engineering, 145*, 102173.

Martudi, S. (2023). An empirical study of fish breeding and biotechnology: Evidence from Indonesia. *FishTaxa-Journal of Fish Taxonomy, 28*.

Panhalkar, A. R., & Doye, D. D. (2022). Optimization of decision trees using modified African buffalo algorithm. *Journal of King Saud University-Computer and Information Sciences, 34*(8), 4763-4772.

Philipp, N. M., Cabarkapa, D., Cabarkapa, D. V., Eserhaut, D. A., & Fry, A. C. (2023). Inter-Device Reliability of a Three-Dimensional Markerless Motion Capture System Quantifying Elementary Movement Patterns in Humans. *Journal of Functional Morphology and Kinesiology, 8*(2), 69.

Salisu, S., Ruhaiyem, N. I. R., Eisa, T. A. E., Nasser, M., Saeed, F., & Younis, H. A. (2023). Motion Capture Technologies for Ergonomics: A Systematic Literature Review. *Diagnostics, 13*(15), 2593.

Stamp, R., Cohn, D., Hel-Or, H., & Sandler, W. (2024). Kinect-ing the dots: Using motion-capture technology to distinguish sign language linguistic from gestural expressions. *Language and Speech, 67*(1), 255-276.

Tabassum, H., Iqbal, M. M., Mahmood, Z., Parveen, M., & Ullah, I. (2023). Gender classification from anthropometric measurement by boosting decision tree: A novel machine learning approach. *Journal of the National Medical Association, 115*(3), 273-282.

Vallée, R., Vallée, J.-N., Guillevin, C., Lallouette, A., Thomas, C., Rittano, G., . . . Vallée, A. (2023). Machine learning decision tree models for multiclass classification of common malignant brain tumors using perfusion and spectroscopy MRI data. *Frontiers in oncology, 13*.

Wierschem, D. C., Jimenez, J. A., & Mendez Mediavilla, F. A. (2020). A motion capture system for the study of human manufacturing repetitive motions. *The International Journal of Advanced Manufacturing Technology, 110*, 813-827.

Xie, X., Zhang, X., & Yang, J. (2022). Decision Tree Algorithm Fusing Information Gain and Gini Index. *Comput. Eng. Appl, 58*, 139-144.

Zhang, D., Ji, Z., Jiang, G., & Jiao, W. (2022). Using AI Motion Capture Systems to Capture Race Walking Technology at a Race Scene: A Comparative Experiment. *Applied Sciences, 13*(1), 113.

Zhu, L. (2022). Research on the application of enrollment data mining based on decision tree algorithm. *Modern Information Technology, 6*(17), 109-112.