International Journal of Data Warehousing and Mining
Volume 19 • Issue 4

Dynamic Research on Youth Thought, Behavior, and Growth Law Based on Deep Learning Algorithm

Qi Fu, Zhejiang University of Water Resources and Electric Power, China*

ABSTRACT

The growth and development of youth is related to the destiny of the country and the nation, and the cultivation of young people in the new era is of great significance. Youth is a special stage in life, and the ideological and moral concepts of young people during this period are still very malleable. In the process of educating them, it is necessary to grasp the characteristics and laws of their ideological and moral growth to be targeted. The characteristics of young people's ideological and moral growth are variability, subjectivity, practicality, and hierarchy; the laws of their ideological and moral growth include the law of guiding transcendence, the law of edification and internalization, the law of mutual promotion of knowledge, will and action, and the law of gradual progression. With the intensification of economic globalization, the Internet, and big data, the thoughts, behaviors, and psychology of young people are constantly changing, which poses a huge challenge to youth ideological and political education.

KEYWORDS

deep learning algorithm, education, youth thought

DYNAMIC RESEARCH ON YOUTH THOUGHT, BEHAVIOR, AND GROWTH LAW BASED ON DEEP LEARNING ALGORITHM

In the overall quality of people, ideological and moral qualities play a leading role. The ideological and moral qualities, such as world outlook, outlook on life, values and moral sentiment, which are in line with the progress of the times, dominate the purpose and direction of people's behavior choices (Ozaki et al., 2022). As Einstein said, "It is not enough to educate a man with professional knowledge (Danilov et al., 2022). Through professional education, he can become a useful machine, but he cannot become a harmoniously developed man (Koo et al., 2022). It is necessary for students to have an understanding of values and to generate passionate affection; that is essential (Lee et al., 2022). He must acquire a sense of beauty and morality. Otherwise, he, with his expertise, is more like a well-trained dog than a harmoniously developed man. Therefore, attaching importance to the cultivation of youth's ideological and moral quality has strategic significance for the healthy growth of youth and the sustainable development of national talents (Mondal et al., 2022). In influencing

DOI: 10.4018/IJDWM.333518 *Corresponding Author

young people's beliefs and behaviors through family, education, peer, and cultural influences, the community plays a critical role in the moral development of these young people. They are guided in growing empathy, compassion, and ethical behavior by their parents, teachers, mentors, and community role models. Judging from the development process of individual human life, infancy and childhood are considered as the naive periods of life development; adulthood is considered as the mature period of life development; and adolescence is the transition from child to adult, from immature to mature period (Zhang et al., 2022). My country's psychology circles generally define adolescence as the age of 11 or 12-17 or 18, which is relative to the junior and senior high school education stages in the United States (Li et al., 2020). Western psychologists generally believe that adolescence refers to the period from puberty to completion of university studies (Chang & Wang, 2020). This stage of development is about 11-21 years old (Behnamnia et al., 2020). The above classification methods of Chinese and foreign researchers have certain reference value (Kahana et al., 2020). In this paper, youth is proposed to be defined as the stage of life development that has entered adolescence (i.e., 11-12 years old) but has not yet reached maturity (Mohammed et al., 2022). In view of the diversity of connotations of the concept of maturity, the upper age limit of youth is temporarily treated as flexible (about 18 years old) to define the youth group (Kim et al., 2021).

The growth of youth generally includes the development and maturity of three aspects: physical, intellectual, and values (Ketonen & Malik, 2020). The good ideological and moral development in this period is an important growth foundation for the youth to transition to the mature stage of life (Tian et al., 2022). The so-called ideological and moral growth stage of young people guides them to establish correct values, form a high-level understanding of society and life, cultivate optimistic and confident healthy psychology and correct behavior, and gradually form a harmonious, noble, firm, optimistic, and enterprising healthy personality, so as to ensure the healthy development of the individual and the realization of its social value (Xue & Wu, 2022). Firm principles and convictions provide a strong basis for developing youth socialist core values by offering ideological clarity, moral guidance, resilience, commitment, critical thinking, collective identity, solidarity, and long-term vision. They shape the beliefs and actions of young individuals, empowering them to uphold socialist ideals and work towards a more equitable society.

With the advent of the big data era and increasingly cheaper computer resources, the level of intelligence and continuous improvement of high-definition imaging technology obtaining useful information from video images plays a vital role in many related industries such as national defense security, transportation construction, banking, finance, engineering construction, etc. (Ince, 2022). According to one survey, there are more than 21 million video surveillance cameras across the country (Ma & Yang, 2021). While providing the benefit of having a large number of video resources, data-based video surveillance also faces some problems (Gu et al., 2022). For example, these video surveillances are mainly based on manual monitoring (Yang, 2022). When abnormal behavior occurs, security workers usually check the abnormal events in the video surveillance and take emergency measures after observing and replaying the video footage (Brewster et al., 2018). This method is not only inefficient and produces a large number of useless junk videos, in abnormal monitoring, it will bring huge pressure to the abnormal monitoring system itself. Therefore, understanding massive video data and video semantic information intelligently, quickly, and accurately has become a hot research topic in recent years (Xing et al., 2020).

With the rapid development of network information technology and computers, human behavior recognition technology, as an important part of video semantic understanding, is related to daily life, life and property, social harmony, etc., and has become a key point in the field of computer vision and artificial intelligence. Aiming at a series of problems existing in human target behavior recognition in complex video environments at present, this paper introduces the deep learning algorithm and principal component analysis algorithm commonly used in the field of computer vision into the youth human behavior recognition technology and gives a human action recognition algorithm based on deep learning-assisted principal component analysis network. Furthermore, the

effectiveness and feasibility of the algorithm are verified with the help of video image clips. This study offers important insights into young people's thoughts, behaviors, and developments with methodological rigor, ethical considerations, and practical applicability. It does so by addressing data accessibility, ethical considerations, algorithm selection, model generalization, interpretability, bias, and fairness. Compared with the traditional human behavior recognition technology, the algorithm cannot only quickly extract human targets from complex environments but also avoid a series of problems caused by differences in space or time sequence in the traditional behavior recognition process (Balaji et al., 2021).

MATERIALS AND METHODS

Characteristics of Youth Ideological and Moral Behavior

Under the requirements and influence of the social environment, young people complete their own ideological and moral growth through the process of independent selection combined with their own specific circumstances. Individual ideological and moral growth is influenced by self-reflection, a variety of viewpoints, ongoing learning, conversation, ethical decision-making, ethical frameworks, and significant experiences, with each person's journey being different. The social environment plays a crucial role in shaping the ideological and moral growth of young people. Key factors include the influence of family, peers, school and education, community and cultural context, media and technology, socioeconomic factors, and the political and societal context. The immaturity of young people's physical and mental development and the complexity of the social environment determine the particularity of their ideological and moral formation process. Consequently, the main behavioral characteristics of young people are as follows:

Variability

1. According to historical materialism, social existence determines social consciousness, social conditions, family environment, and mass media and other social factors, which will inevitably play a subtle guiding role in the ideology and morality of young people. The transition of our society from a planned economy to a market economy has made people's moral psychology and behavior gradually change from "false" to "true," from "virtual" to "real," from closed to open, from single to plural, from "dependence and obedience." change from "independent" to "independent." Transitioning from a command to a market economy can occur due to various reasons including inefficiency and stagnation in the command system, globalization and trade liberalization, economic crises or debt burdens, the desire for economic reform and modernization, political changes and regime transitions, and pressure from international financial institutions. Students demonstrate obedience to the overall situation by following rules and instructions, respecting authority, conforming to norms and expectations, participating in group activities, submitting assignments on time, accepting feedback and criticism, and practicing self-discipline and self-control. However, due to the imperfection of the institutional system supporting the market economy, in real society, to varying degrees, violations of social morality have appeared such as "money power transaction," "red envelope phenomenon," "paid saving people," "spending public funds," and others. In the family, the inappropriate words and deeds of parents and elders will also affect the ideological cognition of young people; mass media such as the Internet is intuitive with wide coverage and fast information flow. Mass media has an important role in guiding public opinion for the ideological and moral growth of young people.
2. Young people are active and enterprising; most students are confident and sunny; a few students have poor psychological quality. Poor psychological quality in students can have various causes including academic pressure, peer pressure, challenging family environments,

personal factors, lack of support systems, external stressors, cultural and societal factors, and lack of coping skills. Addressing these factors requires a comprehensive approach involving educational institutions, families, mental health professionals, and communities to provide resources and support for the psychological well-being of students. A supportive learning environment must be created by identifying and addressing issues with students' psychological well-being. In order to serve students effectively, it is crucial to promote diversity, provide access to mental health resources, and foster open communication. Creating a supportive learning environment involves establishing a positive classroom culture, building strong teacher-student relationships, encouraging collaboration, providing a safe and comfortable physical environment, differentiating instruction, fostering a growth mindset, supporting student well-being, engaging families and the community, encouraging student autonomy and voice, and continuously reflecting and improving teaching practices. Good psychological quality is of great significance to the growth of young people. Good mental health is important for young people as it promotes emotional well-being, supports cognitive development and academic performance, fosters healthy relationships, influences physical health, enhances future prospects, prevents mental health disorders, and has a positive societal impact. Most young people are confident and sunny, actively deal with difficulties and setbacks in study and life, have good interpersonal relationships, and can reasonably vent and control their emotions. On the contrary, some young people have some psychological problems. Eating disorders, anxiety disorders, depression, alcoholism, self-injury, thoughts of suicide, stress and strain, autism spectrum disorders, body dysmorphic disorder, post-traumatic stress disorder (PTSD), and attention deficit hyperactivity disorder are a few typical difficulties among young people, as are youth adjustment issues. This problem mostly exists in freshmen. Freshmen often face adjustment issues when transitioning to a new learning environment. Common challenges include academic demands, social adjustment, independence and self-management, homesickness, campus culture and resources shock, financial and practical considerations, academic and career exploration, and mental health and well-being. Facing a new environment, new faces, and new learning methods, students who cannot adapt to the new learning style choose to escape. Difficulty in adapting to a new learning style can lead to elevated levels of learning anxiety and a sense of emptiness in the mind. Uncertainty, fear of the unknown, a struggle with new methods and expectations, a mismatch between learning preferences and style, lack of support and guidance, and loss of confidence can contribute to these challenges. Learning anxiety and mental emptiness are interpersonal problems. Conflicts with roommates and classmates caused by self-centeredness, academic difficulties, living in poverty, relationship issues, and other reasons can develop. Academic difficulties might arise from living with or studying with others, which can affect focus, productivity, and performance. To get beyond these obstacles and identify solutions, effective communication, unambiguous expectations, and open discussions are essential. The ability to control emotional problems in love is low. This can significantly impact an individual's overall well-being and relationships. It can lead to poor emotional well-being, an increased risk of mental health issues, self-esteem and self-confidence issues, interpersonal relationship strains, social functioning hinderances, and physical health impacts and can result in maladaptive coping mechanisms. Once emotionally frustrated, they will become mad, hopeless, and unable to control themselves. If these problems cannot be resolved in due time, extreme events may occur.

3. Young people have active thinking, most students have firm ideals and beliefs, and a few students have distorted values. Ideals and beliefs play significant roles in the personal development of students. They shape values and identity, guide behavior and decision-making, provide motivation and purpose, develop critical thinking and perspective, cultivate empathy and social responsibility, encourage personal growth and resilience, and foster a sense of belonging and connection. The youth of the new era not only have traditional common characteristics

of young people, but also have distinctive characteristics of modern times. They are active in thinking and quick to accept new ideas, and at the same time, they have strong subjective consciousnesses and have their own opinions.

A questionnaire survey was conducted on the students in the class, and information was obtained through interviews and data review. Various methods were used to collect information through interviews and data review. These methods included semi-structured interviews, structured interviews, focus group interviews, in-depth interviews, document reviews, literature reviews, case study analyses, content analyses, and observational research.

Each method offers unique benefits and can be utilized individually or in combination to gather data for research purposes. The results show that most students have firm ideals and beliefs, they agree with the party's line, principles and policies, are proud of the achievements China has made in the new era, and feel they integrate into the great Chinese dream of building your own ideals in life. Firm ideals and beliefs have laid a solid foundation for the formation of youth socialist core values. Socialism's central tenet is the idea that everyone should have access to resources and means of production, which are owned and managed by the community or the state. Addressing social injustices, fostering group welfare, and ensuring equitable resource and wealth distribution are socialism's main objectives. Most of the students love the country and the party, show courtesy and integrity, obey the overall situation, and consciously resist the influence of unhealthy social customs, but a few young people lack belief, have no clear goal planning, and are easily induced by Western values. The assessment of cultural worth can be divided into five categories: aesthetic, social, symbolic, spiritual, and educational value. Western values breed individualism, hedonism, and money worship and result in irrational consumption, exam cheating, and excessive entertainment, which makes the youth life lose its due color.

Advice on the Thinking and Behavior of Contemporary Youth

Infiltrate Correct Values on the Basis of Diverse Ideas 1. This includes fostering an inclusive environment, encouraging open dialogue and active listening, setting ground rules for respectful communication, incorporating diverse perspectives in the curriculum, assigning group projects, challenging biases and stereotypes, cultivating a culture of respect and empathy, providing reflection and self-assessment opportunities, inviting guest speakers from diverse backgrounds, and addressing conflicts constructively. Everyone is a different individual, and college students already have more mature ideas. Academic skills are significantly influenced by a variety of factors, but their effectiveness varies depending on student needs and the educational environment. These factors include a supportive environment, resources, student motivation, parental participation, and individualized assistance. As college educators, we can only guide mainstream positive and healthy values on the basis of the diverse ideas of young people and gradually penetrate these values into the daily life of students. If you still follow the compulsory education method in the middle school era, it will have the opposite effect. College educators guide students towards mainstream positive and healthy values by serving as role models, promoting critical thinking, facilitating ethical discussions, encouraging diversity and inclusion, incorporating values-based education, providing mentorship and guidance, promoting wellness and well-being, collaborating with support services, and building community partnerships. It is through compulsory education that equitable access to education is guaranteed, universal education is promoted, social and economic growth is sped up, educational standards are established, educational attainment is raised, social cohesion is strengthened, and a culture of lifelong learning is fostered. For example, schools can hold some lectures on ideological and political behavior norms for students such as, "How to Use the Internet Correctly," "How to Get Along With Others," "How to Spend Time Better in Youth Life," etc., and use vivid and interesting language to guide students to participate. Starting from the cases in life, let the students immerse themselves in the whole lecture, and then purify the wrong thoughts

of the students. In addition to some lectures organized by the school as a whole, each department and class can carry out these themes in class meetings, requiring each student to speak, linking their daily life and actual performance, so that, in a subtle way, it will have a certain impact on the students' thinking. A student's maturity level influences their learning by enhancing self-discipline, self-regulation, and motivation, which are all vital elements of effective learning. Maturity enables better time management, goal setting, and perseverance. The relevant party courses of the party activists must be carefully and strictly assessed, so that students' theoretical knowledge can be truly rooted in their hearts, and students' theoretical foundations can be further accumulated. Only in this way can students' thinking be sublimated.

Pay Attention to the Construction of Campus Culture 2. The campus cultural environment plays an important role in the formation of a good school spirit and study style and affects the quality of student training. Educational institutions can establish a setting that supports student learning, engagement, and success by addressing cultural variables like diversity, support, ethics, collaboration, learning styles, well-being, and global awareness. In addition to the education and teaching in the first classroom, we should make full use of the second classroom to guide and nurture students. By guiding students to participate in activities, enhance their teamwork spirit and collective sense of honor, constantly improve themselves in activities, and establish good interpersonal relationships. According to the needs of students, practice activities are carried out in combination with majors. Through the development of the second classroom activities, students' spare time life is enriched, their comprehensive quality is improved, and their noble character of loving the motherland and being grateful to the society is cultivated.

RESULTS AND DISCUSSION

Human Behavior Recognition Methods

In recent years, with the large-scale growth of video image data sets, there have been many factors such as complex background, changing perspective, multiple targets, and small differences between targets, which have brought great impact on the accuracy and computational efficiency of human action recognition influences. The large-scale growth of video image datasets has a significant impact on human action recognition. It improves accuracy by providing diverse and comprehensive training data, enhancing generalization, and adding robustness to variability. However, it also poses computational challenges, requiring efficient algorithms and resources. Therefore, the efficient acquisition of feature information of video images is a crucial part of human target behavior recognition technology (Hong et al., 2020). The analysis of motion patterns, gestures, and interactions, which are all essential for a precise knowledge of human behavior, is made possible by video pictures, which are essential for human target behavior recognition technology.

Generally, a good feature extraction method has better robustness for factors such as illumination intensity, variable viewing angle, and complex background. Feature extraction involves preprocessing the data to remove noise, normalize or standardize it, and handle missing values. Relevant features are then selected or extracted using statistical, mathematical, or domain-specific techniques, capturing important patterns for the task. In view of this, this paper analyzes the advantages and disadvantages of the behavior recognition algorithm based on artificial feature extraction and the behavior recognition algorithm based on deep learning feature extraction (Berggren et al., 2020). Artificial feature extraction-based behavior recognition algorithms offer interpretable features, flexible feature selection, and efficient computing. The methods, tactics, and context all affect these advantages, though. Manual feature engineering, little generalization, feature sensitivity, and redundancy are drawbacks. Among them, the advantage of the method based on manual feature extraction is that it avoids some complex object labeling and segmentation steps and directly calibrates human behavior and actions through interest points such as space-time

and dense trajectories. In addition to increasing efficiency, robustness, high-level understanding, generalization, and enabling real-time analysis of actions or behaviors in movies, behavior recognition systems do away with the requirement for complicated object labelling and segmentation. It includes the main following two representative methods:

Behavior Recognition Based on Spatiotemporal Interest Points 1. This method describes and locates the spatiotemporal regions in the video image sequence that have obvious changes in both the spatial and temporal dimensions, so as to extract the spatiotemporal interest points in the sequence. For example, Equation 1 defines the model of the second-order spatiotemporal matrix in the video image sequence:

$$\mu = g\left(.;\sigma_i^2,\tau_i^2\right) * \begin{bmatrix} L_x^2 & L_xL_y & L_xL_t \\ L_xL_y & L_y^2 & L_yL_t \\ L_xL_t & L_yL_t & L_t^2 \end{bmatrix} \tag{1}$$

where the first term on the right side of the equation is a Gaussian weighting function, and σ and τ represent parameters that control the scale of the weight, respectively. The mean and standard deviation of the Gaussian distribution are used to calculate the Gaussian weight. It rises for values near the mean and falls as the deviation from the mean rises. It evaluates the strength or saliency of a potential interest point considering factors such as uniqueness, stability, corner or edge characteristics, blob-like shapes, or motion consistency. Different techniques like scale-space extrema detection, corner or edge detection, blob detection, or motion analysis can be used to design the response function. When the eigenvalue corresponding to a certain pixel in the image sequence is large, the point is the detected spatiotemporal interest point, and the relevant position of the spatiotemporal interest point is determined by the response function such as Equation 2.

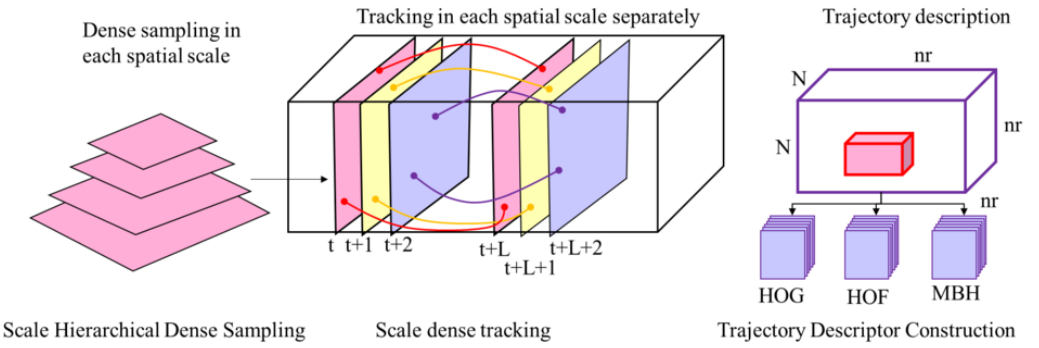
$$H = \det(\mu) - k \cdot trace^3(\mu) = \lambda_1\lambda_2\lambda_3 - k\left(\lambda_1 + \lambda_2 + \lambda_3\right)^3 \tag{2}$$

Among them, the first item and the second item represent the trace of the second-order matrix and the corresponding determinant, respectively, and k represents the possible influence factor, which is usually the best when the value is 0.005.

Target Behavior Recognition Based on Dense Trajectories 2. Since the human behavior recognition method based on spatiotemporal interest is not suitable for describing the human body with long and complex movements, it cannot capture and track the characteristic interest points; that is, it is difficult to obtain long-term movements. Due to the complexity of the human body, which includes various body parts and a variety in joints, it can be difficult to recognize behavior. Occlusions, appearance changes, and information loss make it difficult to track interest sites throughout time, restricting spatiotemporal interest-based techniques. Action information in the range. In this method, each frame of the video image is processed in three steps: dense sampling, dense trajectory tracking, and feature descriptor modeling, and the motion boundary histogram is introduced to enhance the robustness of target motion. A picture is fed into a feature descriptor algorithm, which outputs feature descriptors or feature vectors. For the purpose of detecting features, the Scale-Invariant Feature Transform (SIFT) algorithm is a well-liked computer vision method. As shown in Figure 1, the process of feature extraction from dense trajectories consists of three parts: .

With great progress made by deep learning in the field of computer vision, the advantage of deep learning-based human action recognition compared with traditional artificial feature extraction methods is that convolutional neural networks can spontaneously obtain feature representations

Figure 1. Schematic diagram of feature extraction for dense trajectory detection



with content from video images, and multi-level nonlinear changes of parameters can have stronger generalization ability and semantic understanding of video image feature information.

Theories Related to Convolutional Neural Networks

The convolutional layer is the core part of the convolutional neural network, and this layer is mainly based on convolutional operations. When the input is a single image, the corresponding feature map will be obtained after the convolutional layer completes the convolution of the image. The calculation formula is shown in Equation 3:

$$v_{ij}^{xy} = \tanh\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)}\right) \tag{3}$$

Activation functions play an important role in neural networks, usually after convolutional and batch normalization (BN) layers. When choosing activation functions to enhance the performance of neural networks in various tasks, researchers and practitioners can take into account interpretability, saturation handling, sparsity, computing efficiency, activation range, gradient propagation, and nonlinearity. With the deepening of the network layers, the activation function can improve the nonlinear expression ability of the network and better fit the input data. At present, researchers prefer to use the ReLu function and its variants such as ELU, Swish, P-ReLu, etc.

Sigmoid Function 1. The f(x) expression of the sigmoid activation function is shown in Equation 4. The function is symmetric about f(x)=0.5. When x>5 or x<5, its gradient tends to 0.

$$f(x) = \frac{1}{1+e^{-x}} \tag{4}$$

Tanh Function 2. Tanh activation function is also called double tangent function, and the expression of f(x) is shown in Equation 5. The value range of this function is distributed between [-1, 1] and is symmetrical about f(x)=0, so that the mean of the input is 0, which can effectively train the network parameters.

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5}$$

ReLu Function 3. The ReLu activation function effectively solves the vanishing gradient problem. ReLU activation function effectively addresses the vanishing gradient issue, promoting gradient propagation, non-linearity, sparsity, computational efficiency, and avoids saturation problems. This enhances deep neural network performance, improving model expressiveness and training efficiency. The expression of the function f(x) is shown in Equation 6. When x<0, the value range is 0, and when x>0, the value range is x.

$$f(x) = \begin{cases} 0, \text{ if } x \leq 0 \\ x, \text{ if } x > 0 \end{cases} \tag{6}$$

Accuracy 4. In the field of image recognition and video classification, the most important evaluation index is the accuracy. The evaluation index of accuracy is the probability of correct prediction in all samples, and the calculation formula is shown in Equation 7:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

Loss Function 5. The loss function refers to the difference between the predicted value obtained by the network training sample and the real value. In the field of image recognition and video classification, the cross-entropy function is usually used as the loss function of the convolutional neural network, and the expression is as shown in Equation 8. Due to its compliance with probabilities, gradient efficiency, discriminative power, logarithmic transformation, and compatibility with softmax activation, cross-entropy loss is recommended in CNNs and allows for efficient classification training and optimisation. Cross-entropy can measure the difference between the real probability distribution and the predicted probability distribution. The smaller the value, the better the prediction effect.

$$L(p,q) = -\sum_{i=1}^{n} p\left(x_i\right) \log\left(q\left(x_i\right)\right) \tag{8}$$

Convolutional Layer 6. Although two-dimensional convolution can process spatial feature information, it lacks time series modeling ability and cannot well complete tasks including time domain information such as video analysis and video classification. Multiple fields, including image classification, object recognition, semantic segmentation, picture super-resolution, style transfer, and image generative models, use 2D convolution. Its adaptability and efficiency have sped up picture processing and computer vision. The time dimension of three-dimensional convolution kernels is the same color to represent weight sharing. The convolution calculation process is shown in Equation 9:

$$v_{ij}^{xyz} = \tanh\left(b_{ij} + \sum_{m} \sum_{p=0}^{P_i - 1} \sum_{q=0}^{Q_i - 1} \sum_{r=0}^{R_i - 1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+\boldsymbol{p})(y+q)(=+r)} \right) \tag{9}$$

Space-time Separation Convolution 7. Compared with two-dimensional convolution, three-dimensional convolution can handle video tasks better, but the amount of parameters and computation is large. When it comes to video applications, 3D convolution outperforms 2D convolution because

it can capture temporal dependencies, extract spatial-temporal features, optimize parameters, capture localized motion patterns, and support hierarchical representation learning. Suppose N is the input channel, M is the output channel, DhxDw is the size of the two-dimensional convolution kernel, DtxDnxDw is the size of the three-dimensional convolution kernel, and FnxFw is the output feature map size. In the case of ignoring the bias, a convolutional layer is used as the benchmark; the parameters and computations of two-dimensional convolution and three-dimensional convolution are shown in Equations 10 and 11:

$$\begin{cases} 2\mathrm{D}_{\mathrm{params}} = D_h \times D_w \times M \times N \\ 3\mathrm{D}_{\_\mathrm{params}} = D_t \times D_h \times D_w \times M \times N \end{cases} \tag{10}$$

$$\begin{cases} 2\mathrm{D}\_\mathrm{FLOPs} = D_h \times D_w \times M \times N \times F_h \times F_w \\ 3\mathrm{D}\_\mathrm{FLOPs} = D_t \times D_h \times D_w \times M \times N \times F_h \times F_w \end{cases} \tag{11}$$

It can be seen from the above that the parameter amount and calculation amount of the three-dimensional convolution are Dt times that of the two-dimensional convolution, respectively. The parameters and calculation formulas of the space-time separation convolution are shown in Equations 12 and 13. If and only when Dt=Dh=Dw=3, the parameters and calculations account for 44.4% of the three-dimensional convolution, which can alleviate network overfitting and increase the amount of training data, feature selection, regularization, cross-validation, early halting, dropout, assembling, model complexity reduction, and data augmentation. By implementing these methods, overfitting can be reduced and the model's capacity to perform effectively on untested data can be enhanced. At the same time, the separated convolution layer adds an additional activation function layer, which improves the nonlinear expression ability of the network. The network's expressive capability is increased by introducing non-linearity by the addition of an activation function layer following a separated convolution. Separated convolutions lessen the complexity of the computation but are unable to capture complex relationships. The activation function makes it possible to model complex patterns and learn sophisticated aspects. Therefore, in the subsequent design of the network, the space-time separation technology is used to reduce the amount of parameters and calculation of the three-dimensional convolution. By strengthening computational and memory efficiency, reducing overfitting, promoting transfer learning, and improving interpretability, parameter reduction through space-time separation improves video analysis and results in high-performance models for video-related tasks.

$$\mathrm{Separa}\_\mathrm{params} = N \times M \times \left(D_h \times D_w + D_t\right) \tag{12}$$

$$\mathrm{Separa}\_\mathrm{FLOPs} = N \times M \times \left(D_h \times D_w + D_t\right) \times F_h \times F_w \tag{13}$$

Behavior Recognition Based on Two-Stream Convolutional Network Feature Extraction

The static information of the spatial dimension of the video image and the dynamic information of the temporal dimension are the keys to determine the category judgment in the behavior recognition. In behavior recognition, both static information (visual cues and patterns in a single frame) and dynamic information (changes and motion patterns across consecutive frames) of video images are essential for determining category judgments. Static information focuses on appearance and spatial relationships while dynamic information captures motion and temporal relationships. The spatial behavior information includes the information of the specific scene and the description object, and the temporal behavior information includes the dynamic video frame optical flow motion. Optical flow is a method used in computer vision to estimate the motion of objects within an image or video

sequence. It provides information about the apparent motion of pixels between consecutive frames, allowing us to analyze and understand the movement of objects in a scene. Among them, the spatial stream convolution is to classify the specific object behavior of a single video frame. By pre-training the deep convolutional neural network on the large-scale dataset ImageNet, the video image can be recognized for human behavior. Temporal video stream convolution is to perform convolution network calculation on the optical flow displacement field of stacked video frames, which can be accurately represented as behavioral features between video frames. In contrast to temporal stream convolution, which collects temporal patterns and motion data over consecutive frames, spatial stream convolution concentrates on spatial patterns inside individual frames. Performance of video analysis is enhanced when both streams are combined. The input model of temporal convolution network is exhibited in Equations 14 and 15:

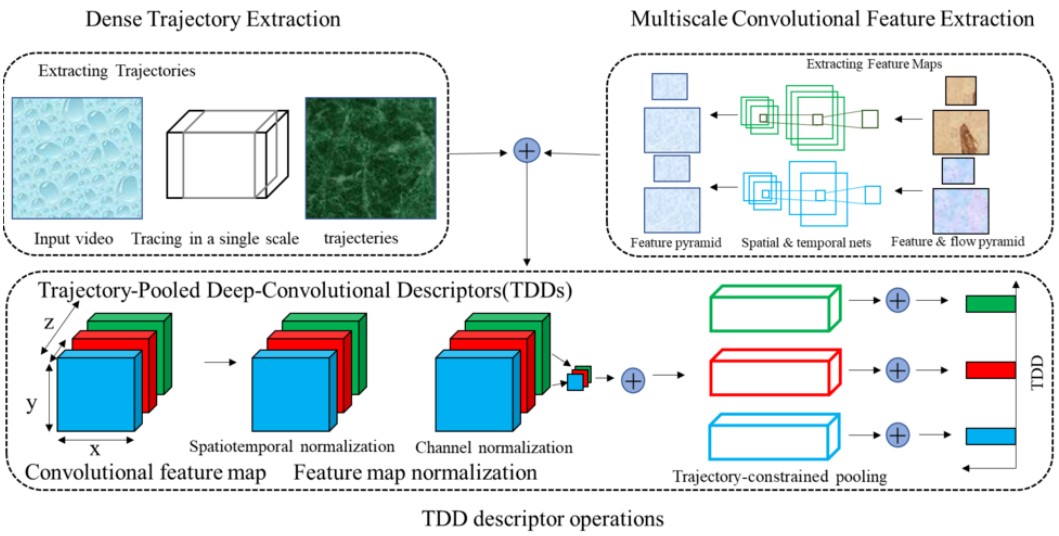
$$I_t(\mu, v, 2k-1) = d^x_{t+k-1}(\mu, v) \tag{14}$$

$$I_t(\mu, v, 2k) = d^y_{t+k-1}(\mu, v), \mu = [1; w], v = [1; h], k = [1; L] \tag{15}$$

Among them, a set of displacement vector field quantities between consecutive video image frames t to t+1 is represented by the dt symbol. Considering the difference between the temporal dimension and the spatial dimension in the video image sequence, Wang L et al. A method for trajectory pooling convolutional description based on artificial feature extraction and deep learning. The method first extracts dense trajectories, then uses a multi-scale convolutional neural network to perform deep mapping of features, calculates the trajectory depth convolution descriptor (TDD), and calculates

The descriptor is encoded by Fisher vector, and finally the encoded vector is put into the classifier to identify and judge the behavior of the target human body. The feature extraction process of the trajectory depth convolution descriptor is shown in Figure 2. The TDD descriptor operation is to perform local trajectory alignment around dense trajectories by extracting convolutional feature map normalization and trajectory motion pooling. The steps involved in local trajectory alignment using TDD are dense trajectory extraction, descriptor calculation, trajectory clustering, local alignment, reference trajectory selection, alignment transformation estimation, trajectory alignment, and descriptor update.

Figure 2. Schematic diagram of TDD descriptor feature extraction



RESULT ANALYSIS AND DISCUSSION

Experimental Samples

In this section, relevant experiments are carried out on the human action recognition data set UCF10l. A frequently used benchmark dataset in the fields of action detection and computer vision research is the UCF101 dataset. It comprises of a selection of video clips covering 101 different action genres. The University of Central Florida, where the dataset was first assembled, is referenced by the initials "UCF." In the data preprocessing stage, the UCF101 data set is first decompressed, and the data set is divided according to spilt1. Then, using Python libraries such as OpenCV, OS, and Math, the image frames of the video are extracted. Finally a random horizontal flip is carried out and the image is randomly cropped to a 112x112 size (Inapakurthi et al., 2021).

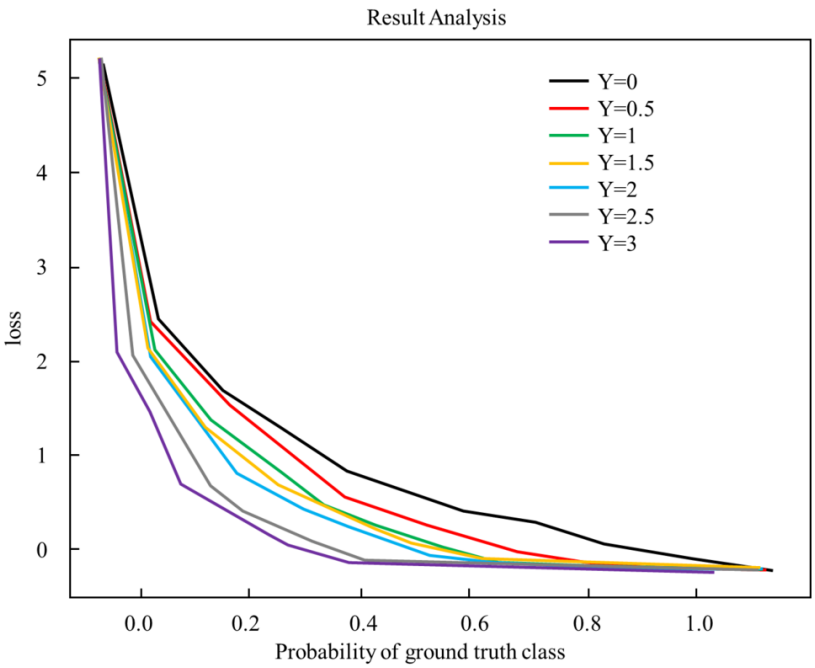
After the data preprocessing stage, the input size of the network is 16x3x16x112x112. During training, a total of 100 training epochs are used from scratch, the initial learning rate is set to 0.01, and an adaptive learning rate decay strategy is used to update the current learning rate of the network. The network optimizer is SGD, the momentum coefficient is 0.9, the weight decay coefficient is 0.0005, and the scale of the dropout layer is 0.5.The grid used in this paper is divided into 7X7 small grids to divide the entire image, the prediction box B of each grid is 2, and the category c is 20 categories. In order to simplify the calculation and improve the training rate, the feature extraction network structure, GoogLeNet in YOLO, uses 1X1 and 3X3 convolutional layers to replace the 1X1, 3X3, 5X5 convolution kernels, and 3X3 maximum pooling in the inception layer in the original structure. The YOLO framework is highly important due to its fast and accurate object detection capabilities, making it suitable for surveillance, autonomous driving, and video analysis. It achieves real-time performance by processing whole images and predicting bounding boxes directly, enabling efficient detection in dynamic environments. Among them, the image specification of the first layer of the input is a two-dimensional single-frame image with a resolution dimension of 448x448, and each convolutional layer will be followed by a nonlinear activation function leakyReLu. Among them, in the first-order target detection model, the selection of this function effectively overcomes the possibility that the ReLu piecewise function is always 0. That is, by introducing the rectification linear parameter, the data with the input value $x \leq 0$ will not cause data when it is activated. The possibilities of ReLU include dead neurons, saturation, limited output range, gradient explosion, and biased outputs. Alternative activation functions have been developed to address these limitations and enhance neural network performance. The loss of, which is consistent with the activation value when the input value is positive, ensures the diversity of data feature information. Usually, λ is between 0 and 1, and h=0.1 in the YOLO model.

Analysis of Experimental Results

Since the parameter Y of the optimal theory of focal loss is a control factor for predicting the target probability, the experimental range of the setting value of this factor is 0~3, and the adjustment step size is 0.5. Through the training analysis of the target data set, the experimental results shown in Figure 3 can be obtained. As shown in the figure, the convergence speed of the loss function loss increases as the value of Y increases. When γ=0, the curve represents the loss speed of the original first-order model; when y=2, the convergence speed and target accuracy are preferably, the loss function is made to pay more attention to the loss contribution of indistinguishable objects and suppresses the loss contribution of easy-to-classify objects. In addition, it can be seen from the figure that it is not that the larger the parameter Y, the better the effect. When the parameter Y is large, although the loss converges quickly, the target recognition rate drops relatively low.

First, verify that the DenseNet (default DenseNet_ 121(k=32)) model is better than other convolutional neural networks, and conduct comparative experiments with VGGNet-16, VGGNet-19, GoogLeNet, ResNet50, and ResNet101 on the dataset, and compare them to each other. The parameter complexity and error rate values are evaluated. As shown in Figure 4, GoogLeNet in the original

Figure 3. Experimental diagram of FocalLoss parameter determination



first-order target detection model has a lower classification error rate than VGGNet as a whole, but due to the limitation of the number of layers, the recognition accuracy of these two feature extraction networks is generally low, so, in order to reduce the influence of the number of network layers on the recognition accuracy, the deeper network models of ResNet and DenseNet are used for experiments.

Secondly, in order to better verify the effectiveness of the deepening of neural network layers and the increase of learning width on feature extraction, this paper uses DenseNet-121 (k=32) model, DenseNet-161 (k=32) model, and DenseNet- 161 (k=64) models for network training. Figure 5 (a) shows the relationship between the number of iterations and the loss value in the training process. Through experimental comparison, it is found that the loss degree curves of the three network models with the number of iterations before 12,000 have a high degree of overlap and similarity. The loss value of the DenseNet121 network model is smaller than that of the DenseNet161 network model when the number of iterations is between 23,000 and 30,000. Figure 5 (b) describes the relationship between the training parameter complexity and the network error rate value. As the parameter complexity gradually increases, the error rate of the feature extraction network model with more layers is lower, and at the same time, the wider the learning width. The error rate of the network model then decreases more significantly.

Because the parameter filter size k in the principal component analysis network used in this experiment, the number of filters (L1, L2, L3) in each layer of the transfer network directly affects the effect of the behavior recognition and classification of human targets in the principal component analysis network model. During the experiment, since the parameter k of each layer and the number of filters is independent of each other, the experiment strategy of independent combination is adopted. That is, the parameter selection of the current network layer is carried out on the basis of selecting the optimal parameters of the previous layer. The filter size k value ranges from 6-10, and the number of filters ranges from 3-13. The experimental results are shown in Figure 6.

Figure 4. Error rate comparison experiment

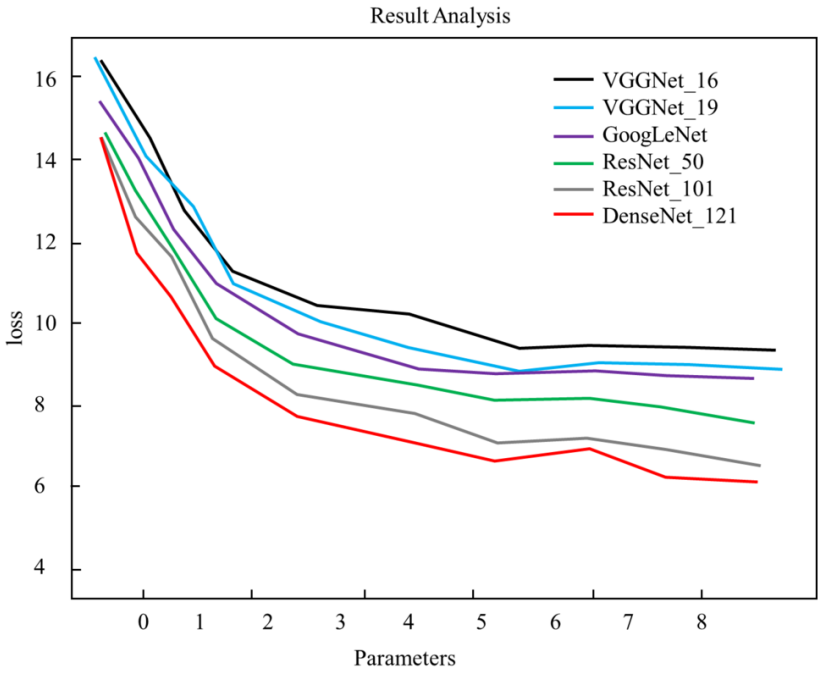


Figure 5. Experimental comparison diagram

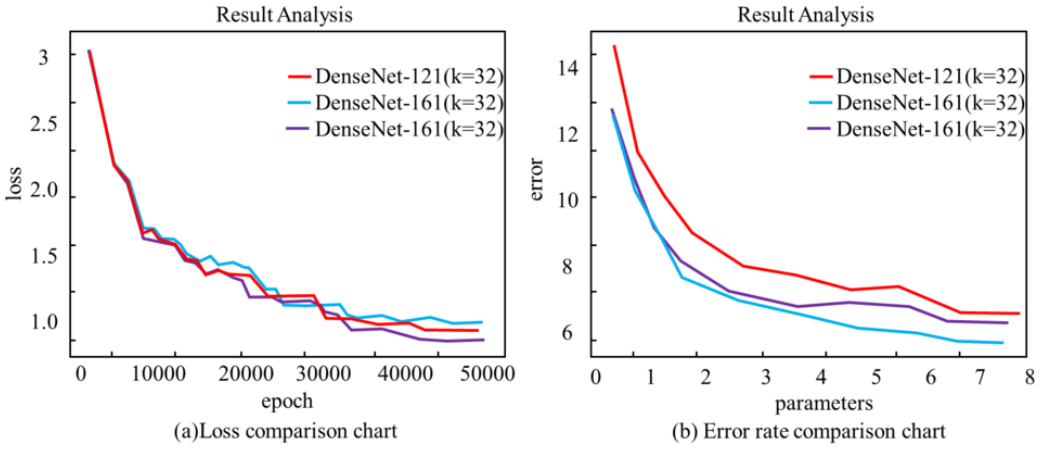


Figure 7 shows the test set accuracy rates of different networks on the UCF101 dataset. It can be seen that 0~17epoch, the SR3D network, has the highest accuracy, and the VT3D network has the lowest accuracy; at 18~60epoch, the VT3D network has the highest accuracy. The SR3D network has the lowest accuracy and has converged; at 61~100epoch, the VTR3D network has the highest accuracy, and the VT3D network has the second highest accuracy while the SR3D network has the lowest accuracy.

Figure 6. Principal component analysis network parameters to determine the experimental results

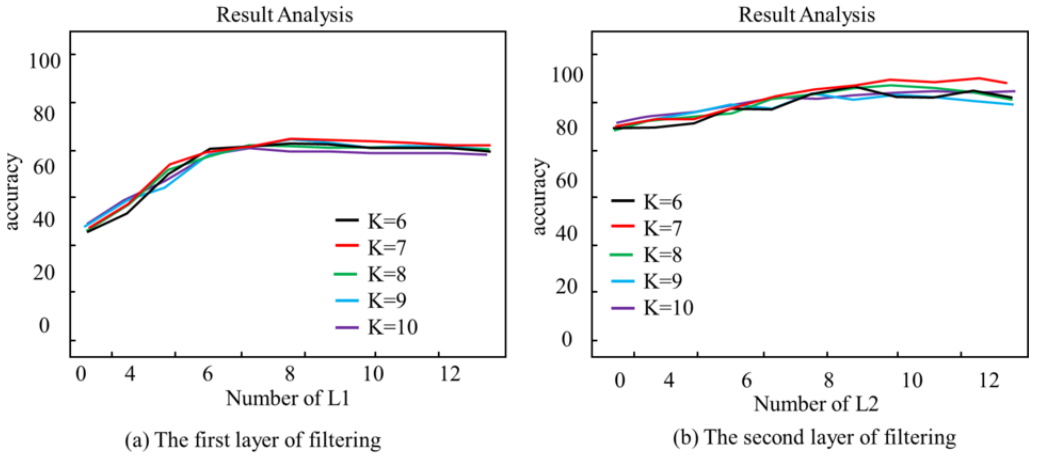
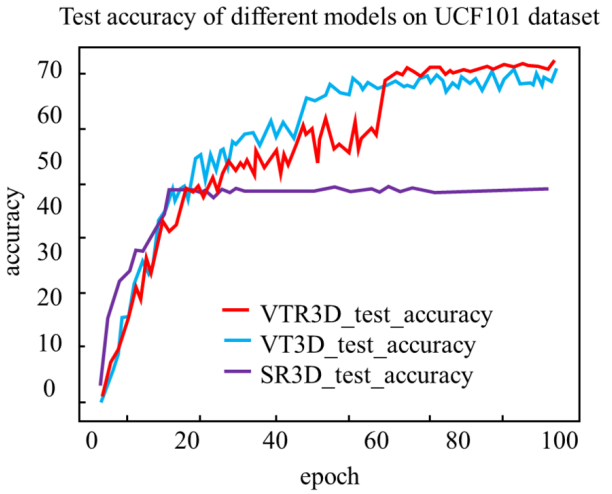


Figure 7. Comparison of the accuracy of different networks on the UCF101 dataset

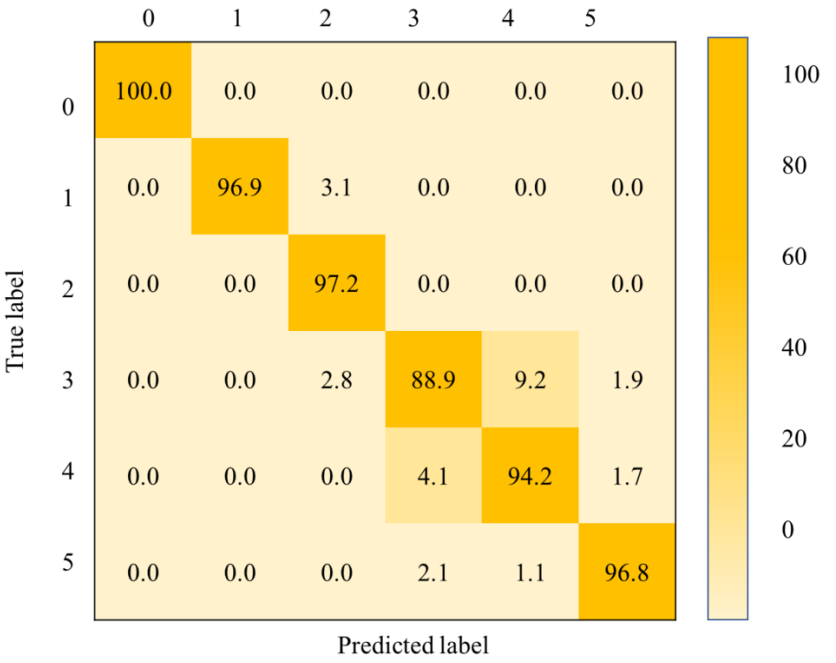


In order to more accurately observe the experimental effect of behavior recognition and classification, this paper uses confusion matrix to conduct recognition accuracy statistics for different data sets. By dividing the total number of predictions by the sum of correct predictions (true positives and true negatives), the accuracy of a confusion matrix can be determined, giving an evaluation of the model's overall correctness and performance. Actual predictions Each row is represented as an actual class ground truth. Figure 8 shows the classification and recognition performance of each action in the three behavior datasets.

CONCLUSION

This article analyzes the characteristics of contemporary youth's thinking and behavioral laws and then puts forward some targeted suggestions. In view of the characteristics of youth's diversity of

Figure 8. Schematic diagram of KTH behavioral data confusion matrix



thinking, schools and teachers should cooperate to provide suggestions for improving students' ideological levels and provide good suggestions. To implement it, they should also pay attention to combining the individual characteristics and learning situation of students to develop better study plans and methods for some students. Youth life is colorful, and young people are inevitably addicted to play. For students with insufficient self-control, some ideological education may be needed to change their living and study habits. In this paper, a human body recognition algorithm based on deep learning-assisted principal component analysis network is proposed. The human body target in complex scenes is detected and extracted through the deep learning model, and then the behavior of the human body target is determined by the method of principal component analysis network. Identify and judge. Establish a human target detection model based on deep learning. Firstly, the first-order target detection model and the second-order target detection model are compared and analyzed, the YOLO framework is selected as the main structure of human target detection, and the optimization strategy of Focal_Loss theory is given to improve the detection rate of human targets. Finally, the feasibility of the two optimization methods is verified by experiments.

CONFLICT OF INTEREST

The author declares that they have no conflicts of interest.

FUNDING STATEMENT

This work was supported by the Zhejiang province social science planning "University ideological and political work" special subject (21GXSZ074YBM); and the Higher Education Research Project

of Zhejiang Province (KT2020143); and the General Project of Department of Education of Zhejiang Province (Y201942954).

ACKNOWLEDGMENT

The author would like to show sincere thanks to those techniques who have contributed to this research.

REFERENCES

Balaji, T. K., Annavarapu, C. S. R., & Bablani, A. (2021). Machine learning algorithms for social media analysis: A survey. *Computer Science Review*, *40*, 100395. doi:10.1016/j.cosrev.2021.100395

Behnamnia, N., Kamsin, A., Ismail, M. A. B., & Hayati, A. (2020). The effective components of creativity in digital game-based learning among young children: A case study. *Children and Youth Services Review*, *116*, 105227. doi:10.1016/j.childyouth.2020.105227

Berggren, K. K., Xia, Q., Likharev, K. K., Strukov, D. B., Jiang, H., Mikolajick, T., Querlioz, D., Salinga, M., Erickson, J. R., Pi, S., Xiong, F., Lin, P., Li, C., Chen, Y., Xiong, S., Hoskins, B. D., Daniels, M. W., Madhavan, A., Liddle, J., & Raychowdhury, A. (2020). Roadmap on emerging hardware and technology for machine learning. *Nanotechnology*, *32*(1), 012002. doi:10.1088/1361-6528/aba70f PMID:32679577

Brewster, L. R., Dale, J. J., Guttridge, T. L., Gruber, S. H., Hansell, A. C., Elliott, M., Cowx, I. G., Whitney, N. M., & Gleiss, A. C. (2018). Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data. *Marine Biology*, *165*(4), 1–19. doi:10.1007/s00227-018-3318-y PMID:29563648

Chang, C., & Wang, X. (2020, November). Research on Dynamic Political Sentiment Polarity analysis of specific group Twitter based on Deep learning method. *Journal of Physics: Conference Series*, *1651*(1), 012108. doi:10.1088/1742-6596/1651/1/012108

Danilov, V. V., Litmanovich, D., Proutski, A., Kirpich, A., Nefaridze, D., Karpovsky, A., & Gankin, Y. (2022). Automatic scoring of COVID-19 severity in X-ray imaging based on a novel deep learning workflow. *Scientific Reports*, *12*(1), 12791. doi:10.1038/s41598-022-15013-z PMID:35896761

Gu, H., Guo, H., Li, X., Li, F., & Liu, Z. (2022). Nonlinear dynamic analysis and fatigue study of steep wave risers under irregular loads. *Journal of Ocean University of China*, *21*(3), 669–680. doi:10.1007/s11802-022-4850-4

Hong, T., Wang, Z., Luo, X., & Zhang, W. (2020). State-of-the-art on research and applications of machine learning in the building life cycle. *Energy and Building*, *212*, 109831. doi:10.1016/j.enbuild.2020.109831

Inapakurthi, R. K., Miriyala, S. S., & Mitra, K. (2021). Deep learning based dynamic behavior modelling and prediction of particulate matter in air. *Chemical Engineering Journal*, *426*, 131221. doi:10.1016/j.cej.2021.131221

Ince, M. (2022). Automatic and intelligent content visualization system based on deep learning and genetic algorithm. *Neural Computing & Applications*, *34*(3), 2473–2493. doi:10.1007/s00521-022-06887-1 PMID:35068702

Kahana, A., Turkel, E., Dekel, S., & Givoli, D. (2020). Obstacle segmentation based on the wave equation and deep learning. *Journal of Computational Physics*, *413*, 109458. doi:10.1016/j.jcp.2020.109458

Ketonen, V., & Malik, A. N. (2020). Characterizing vaping posts on Instagram by using unsupervised machine learning. *International Journal of Medical Informatics*, *141*, 104223. doi:10.1016/j.ijmedinf.2020.104223 PMID:32623330

Kim, E. Y., Kim, Y. J., Choi, W. J., Lee, G. P., Choi, Y. R., Jin, K. N., & Cho, Y. J. (2021). Performance of a deep-learning algorithm for referable thoracic abnormalities on chest radiographs: A multicenter study of a health screening cohort. *PLoS One*, *16*(2), e0246472. doi:10.1371/journal.pone.0246472 PMID:33606779

Koo, J., Hwang, S., Han, S. H., Lee, J., Lee, H. S., Park, G., Kim, H., Choi, J., & Kim, S. (2022). Deep learning-based tool affects reproducibility of pes planus radiographic assessment. *Scientific Reports*, *12*(1), 12891. Advance online publication. doi:10.1038/s41598-022-16995-6 PMID:35902681

Lee, C., Light, A., Saveliev, E. S., Van Der Schaar, M., & Gnanapragasam, V. (2022). Developing machine learning algorithms for dynamic estimation of progression during active surveillance for prostate cancer. *NPJ Digital Medicine*, *5*(1), 110. doi:10.1038/s41746-022-00659-w PMID:35933478

Li, G., Zhou, W., Chen, W., Sun, F., Fu, Y., Gong, F., & Zhang, H. (2020). Study on the detection of pulmonary nodules in CT images based on deep learning. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 67300–67309. doi:10.1109/ACCESS.2020.2984381

Ma, C., & Yang, P. (2021). Research on classroom teaching behavior analysis and evaluation system based on deep learning face recognition technology. *Journal of Physics: Conference Series*, *1992*(3), 032040. doi:10.1088/1742-6596/1992/3/032040

Mohammed, A. S., Mohammed, A. S., & Kareem, S. W. (2022). Deep Learning and Neural Network-Based Wind Speed Prediction Model. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, *30*(03), 403–425. doi:10.1142/S021848852240013X

Mondal, D., Roy, K., Pal, D., & Kole, D. K. (2022). Deep Learning-Based Approach to detect and classify signs of crop leaf diseases and pest damage. *SN Computer Science*, *3*(6), 433. Advance online publication. doi:10.1007/s42979-022-01332-5

Ozaki, J., Fujioka, T., Yamaga, E., Hayashi, A., Kujiraoka, Y., Imokawa, T., Takahashi, K., Okawa, S., Yashima, Y., Mori, M., Kubota, K., Oda, G., Nakagawa, T., & Tateishi, U. (2022). Deep learning method with a convolutional neural network for image classification of normal and metastatic axillary lymph nodes on breast ultrasonography. *Japanese Journal of Radiology*, *40*(8), 814–822. doi:10.1007/s11604-022-01261-6 PMID:35284996

Tian, X., Chen, C., Chen, C., Yan, Z., Wu, W., Chen, F., Chen, J., & Lv, X. (2022). Application of Raman spectroscopy technology based on deep learning algorithm in the rapid diagnosis of glioma. *Journal of Raman Spectroscopy : JRS*, *53*(4), 735–745. doi:10.1002/jrs.6302

Xing, Y., Lv, C., Wang, H., Cao, D., & Velenis, E. (2020). An ensemble deep learning approach for driver lane change intention inference. *Transportation Research Part C, Emerging Technologies*, *115*, 102615. doi:10.1016/j.trc.2020.102615

Xue, B., & Wu, Z. (2022). Defect Detection and Classification Algorithm of metal Nanomaterials based on Deep learning. *Integrated Ferroelectrics*, *226*(1), 277–292. doi:10.1080/10584587.2022.2065568

Yang, F. (2022). Construction of graduate Behavior Dynamic model based on dynamic Decision Tree Algorithm. *Discrete Dynamics in Nature and Society*, *2022*, 1–10. doi:10.1155/2022/3684176

Zhang, T., Li, Y., Wang, T., Wang, H., Chen, T., Sun, Z., Luo, D., Li, C., & Han, L. (2022). Evaluation of different machine learning models and novel deep learning-based algorithm for landslide susceptibility mapping. *Geoscience Letters*, *9*(1), 26. Advance online publication. doi:10.1186/s40562-022-00236-9

Qi Fu is a professor of Zhejiang University of Water Resources and Electric Power.