Journal of Computers Vol. 35 No. 2, April 2024, pp. 183-197
doi: 10.53106/199115992024043502012

The Application Strategy of Intelligent Computer Multimedia Technology in Film and Television Post Production

Jiang-Wei Wan*

Tangshan Polytechnic College,
Tangshan City 063299, Hebei Province, China
jiangwei20080106@126.com

Received 27 March 2024; Revised 3 April 2024; Accepted 15 April 2024

Abstract. The development of short videos and big movies requires high-quality visual design. In the context of digital technology development, computer multimedia technology has begun to emerge in the field of video processing, promoting innovative development in China's industry. The post production of film and television works requires the use of computer multimedia technology to expand the creative space and improve the artistic expression effect of film and television works. This article elaborates on the unique role of computer multimedia technology in the field of film and television post production, and then analyzes the various stages of film and television post production. The focus is on using computer artificial intelligence techniques to improve the quality and efficiency of video editing in the video editing stage. Then, in the color rendering stage of some black and white video resources, computer technology is automatically processed, Therefore, computer multimedia technology can play an advantageous role in film and television post production and should be promoted and applied.

Keywords: computer multimedia, video post, video clips, smart rendering

1 Introduction

With the continuous development of self media technology and the rise of the domestic film market, film and short video works are appearing more densely in our daily lives. With the increasing material and cultural demand, people have higher requirements for the quality of film and short video works. Therefore, it is necessary to continuously improve the techniques and production techniques of film production. Post production is crucial throughout the entire film and television production process, directly determining the final effect of the work presented to the audience [1].

In the post production stage of film and television, the application of computer multimedia technology has become a key means to enhance the artistic effect and video viewing value of film and television works. A complete post production process for film and television, generally including video editing, color rendering, audio processing, and other links, is a key process that endows film and television works with vitality. With the continuous progress of computer technology, multimedia technology has also achieved rapid development. From simple editing to complex visual effects production today, the development of computer multimedia technology has not only promoted the innovation of film and television forms, but also brought unprecedented viewing experiences to audiences. This article focuses on the post production process of film and television, and discusses the application methods and strategies of computer multimedia technology in various stages [2].

The role of film and television post production is reflected in the following aspects:

1) Enhance the visual and auditory experience of film and television works. By adjusting details such as lighting, color, and sound effects, the image is made more vivid and transparent, and the sound effects are more realistic, allowing the audience to better immerse themselves in the plot of the film.

2) Optimize the rhythm and structure of film and television works. Possible issues during the filming process, such as actor performances or imperfect camera shots, can be finely adjusted through post production to make the film more logical and captivating with a sense of rhythm.

3) Add special effects and animation effects. Post production can incorporate various special effects and animation effects to enhance the visual impact of the film, and even create scenes that do not actually exist, enhancing the audience's viewing experience.

* Corresponding Author

4) Satisfy the spiritual enjoyment of modern people. Film and television post production uses technological means to showcase past or future events, satisfying the audience's spiritual enjoyment.

5) Promote the integration of art and technology. Post production of film and television not only showcases the charm of modern technology, but also embodies the beauty of art, and the combination of the two injects vitality into the cultural industry.

6) Meet the cultural consumption needs of the audience. With the advent of the 5G era, users have increasingly high demands for the content of their works. Film and television post production showcases better artistic beauty through the artistic recreation of materials, while also meeting people's cultural consumption needs.

2 Analysis of the Role of Computer Multimedia Technology in Post Production of Film and Television

Compared to manual processing of film and television works, computer multimedia technology not only has a faster processing speed for film and television editing and rendering, but more importantly, the processing effect. When manual production is used in post production, the production level of film and television works depends on the individual's technical level and artistic literacy [3]. However, with the support of artificial intelligence, the use of computer multimedia technology will only continuously improve the quality of film and television post production, Through the accumulation of technical experience and the current situation of the industry, the role of computer multimedia technology is reflected in the following aspects:

1) Enhance the artistic expression and infectiousness of film and television works

Compared to traditional post production of film and television works, the addition of computer multimedia technology will provide technical support for the creative space of post production personnel, compensate for the shortcomings in the original shooting images of film and television works, restore the ideas of the creators of film and television works to the greatest extent, and enhance the artistic expression of film and television works. The integration of captured images, audio and other information as a dataset can provide rich learning materials for artificial intelligence technology, thereby optimizing the content expression effect of film and television works and improving their artistic appeal. Meanwhile, computer multimedia technology can add narration, subtitles, and transitions to a single image, making the content of the captured image more dynamic. In short, computer multimedia technology can improve the landscape and character expression of images, optimize visual expression effects, and enhance the artistic value of film and television works based on different themes [4].

2) Improve the efficiency of film and television post production

On the one hand, the use of computer multimedia technology in the post production process of film and television works can make the working methods of relevant post production personnel more flexible, breaking through the limitations of time and space. The traditional linear production method takes a long time, and the progress of production depends on the energy of the staff, resulting in more errors. However, computer multimedia technology can simplify the post production process, reduce work difficulty, and automatically correct errors made by staff, ensuring the efficiency and accuracy of film and post production. In addition, computer multimedia technology has strong replicability for video processing, making it more suitable for batch processing of similar videos. On the other hand, it is reflected in the storage of media materials in computer multimedia technology. Using computers as storage media makes storage more flexible and has stronger storage capabilities. Moreover, computer multimedia technology can quickly retrieve corresponding film and television data information during the post production process. The retrieved data will be quickly and completely transmitted to the operating platform of the staff, reducing the storage cost of film and television data while improving post production efficiency [5].

3) Improve the economic benefits of film and television companies

Computer multimedia technology can compensate for the shortcomings in the pre filming process of videos, save pre filming time, shorten the production cycle of film and television works, and improve the economic benefits of film and television companies. At present, some film and television companies in China only need a few computers and staff to complete their post production work.

3 Implementing Video Editing with Computational Multimedia Technology

During the editing process, the computer, as the first person perspective, needs to complete tasks such as recognizing objects, analyzing behavior, understanding images, understanding video content, and even constructing video scenes and rendering colors. The camera is the basic unit of editing, and this section mainly discusses the application of artificial intelligence techniques in automatic processing of camera operation. Using this as a technical foundation to achieve automatic video editing in film and television post production.

3.1 Intelligent Mirror Operation Classification Model

The general steps of mirror classification include: motion information extraction, noise removal, motion parameter model fitting, and sequence smoothing. If the representation method of a point in the video is (x, y), then the motion vector is represented as $(\Delta x, \Delta y)$ and the displacement amplitude is represented as $\sqrt{(\Delta x)^2 + (\Delta y)^2}$. Then, using the deep learning network Lightweight MobileNet [6], multi-scale motion direction templates are further learned from the one-dimensional block angle histogram sequence to obtain template matching information. The deep learning network is shown in Fig. 1.

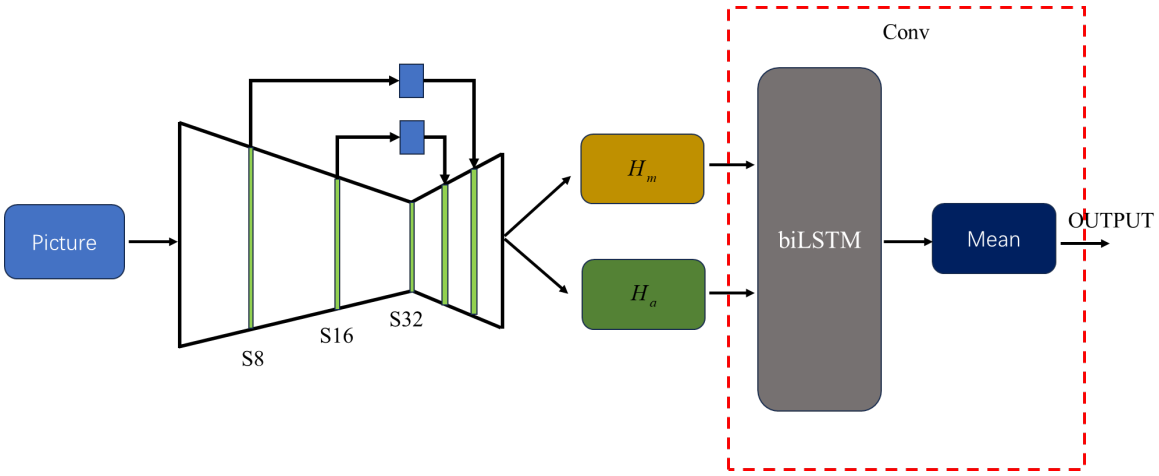

Fig. 1. Network model structure

Then, the frame is divided into multiple non overlapping macroblocks. For each block in the current frame, the most similar matching block is found within the search range of the previous frame. The relative displacement between the current macroblock and the matching block is the motion vector. Using a segmentation model to remove noise data from the foreground area. Use an improved convolutional neural network to remove noisy foreground.

In the above network, the first step is to perform sparse sampling on the input video shots, randomly selecting intervals of two frames in each segment. Then, input each pair of frame samples into the instantaneous mirror network for instantaneous motion feature extraction. Next, sequentially input the motion features of all segments into a segment consensus module. Because the classification of camera operations at the lens level is related to the motion characteristics and order of each sample point, a network that can be used for temporal modeling is needed. Therefore, a single-layer biLSTM [7] is added to the network to obtain consensus on all features.

In the field classification task, it can be inferred from the definition of field that field classification is a position related image classification task. It is not only related to the body parts present in the image, but also to their vertical position and proportion. That is, the position features and non position features in the image together deter-

mine the field of the image. Almost all existing scene classification algorithms use global pooling to downsample features, which shields vertical positional features. Vertical position features and non-position features play an important role in scene classification, while horizontal position features (the left and right position information of body parts in the picture) have a weaker relationship with the scene. Therefore, the pooling operation method in the algorithm has been changed, and the pooling module design is shown in Fig. 2.

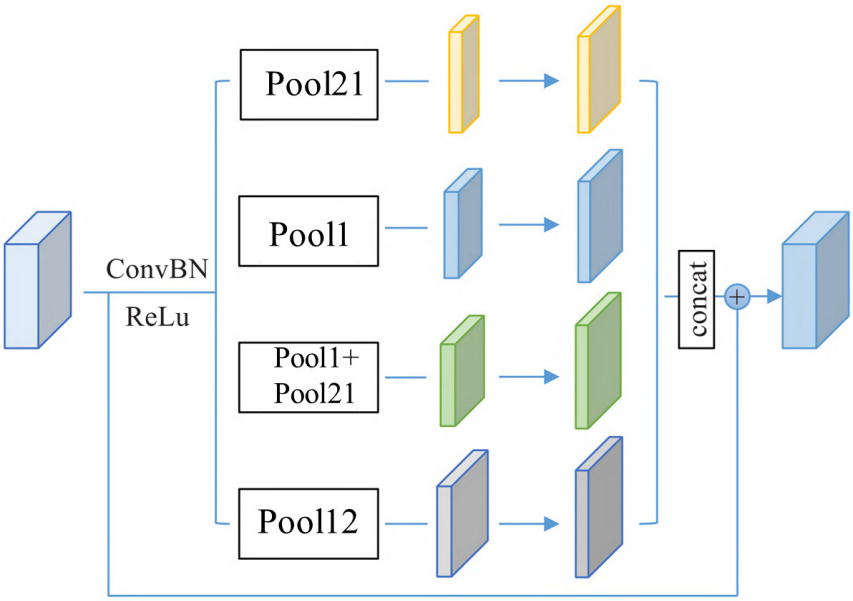

Fig. 2. Improved pooling structure

For these four categories of pooling layers with fixed output sizes, the pooling domain and pooling step size are adaptively calculated based on the input and output feature sizes. The formula is as follows:

$$s = I_{down}\left(\frac{S_{input}}{S_{output}}\right). \tag{1}$$

$$t = S_{input} - s\left(S_{output} - 1\right). \tag{2}$$

S_{input}- Input the feature size of the image;
S_{output}- Output the feature size of the image;
k- Pooled domain;
s- Pooling step size;
I_{down}- Rounding down function.

3.2 Intelligent Mirror Operation Classification Experiment

Using a self-made mirror classification dataset, the data is first preprocessed and shot boundary detection algorithms are used to segment video shots. Each video segment contains between 1 and 20 shots, with each shot lasting from 0.5 seconds to 80 seconds. When there is no smooth motion segment in the lens, the lens is not labeled and becomes an invalid lens. The entire dataset is divided into training and testing sets based on the effective instantaneous motion sample logarithm of approximately 3:1.

The image characteristics of the training and testing sets are shown in Fig. 3.

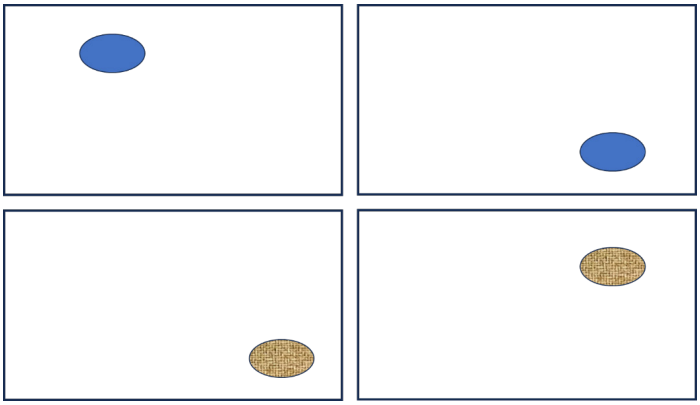

Fig. 3. Image feature diagram

All experiments in this section were conducted under the GPU acceleration environment of RTX4740. The model is developed based on the PyTorch deep learning framework. The convolutional part of the pre trained MobileNet model is used as a feature extractor.

The purpose of the experiment is to test the contribution of vertical position features and texture features to the "up/down" classification of images during the image mirror operation classification process. Among them, texture features are an example of non positional features.

In the training set images, the solid elliptical shape features are located in the upper half of the image, and the striped ellipses are located in the lower half of the image [8]. The four pooling layers in Fig. 2 are respectively used in the convolutional neural network for transfer learning on the training set, and the accuracy is tested on the test set. Pool21 pooling is used regardless of whether it is max pooling or average pooling, and the test accuracy exceeds 95%, indicating that Pool21 pooling tends to rely on vertical position features for image classification. Pool1 and Pool12 pooling have an accuracy of nearly 50%, indicating that these two pooling methods do not rely on position features for classification, because Pool1 shields position features, while Pool12 only extracts horizontal positions that are task independent. Characteristics. The testing accuracy of Pool1+Pool21 pooling is between 50% -95%.

After experiments, the results of the proposed method in terms of accuracy, Macro F1, and SPS are shown in Table 1.

Table 1. Model running results

	ACC	Macro F1	SPS
Lightweight-MobileNet	96.7%	0.9585	21

Through the above methods, the use of intelligent algorithms in video editing can improve dataset migration performance and enhance the anti-interference ability of foreground shots.

3.3 Intelligent Mirror Operation Classification Experiment

Computer vision is the application of machine learning in the field of vision, similar to the eyes of machine learning. Its purpose is to collect images or videos, analyze and process them, and obtain the information we need from them. Facial expression recognition is an important research object in intelligent human-computer interaction technology. It is also an important technical foundation for post production video editing in film and television.

Facial expression recognition is the process of feature extraction and classification of facial expression information through computers. It enables computers to learn human facial expression information, infer human psychological state, and achieve intelligent interaction between humans and machines. Video editing has very important application value and significance in contemporary life, and movies, TV dramas, online videos, etc. all require video editing processing.

As shown in Fig. 1, the core layer constructed by MobileNet is a depth separable filter, and depth separable convolution is a form of decomposition convolution. The standard convolution operation directly extracts features from the input and combines them into a series of outputs, while depthwise separable convolution divides this process into two layers: one layer is depthwise convolution, which is used to extract features separately for each input channel; The first layer is point by point convolution, which combines the output of the previous step with a 1x1 convolution. This decomposition has a significant effect on reducing computation and model size.

In order to improve the model's ability to recognize faces and expressions during video editing, while minimizing the impact on facial recognition accuracy, this paper adopts MobileNe [9] and GoogLeNet [10] to integrate and obtain a new video automatic editing network structure. The structural model is shown in Fig. 4.

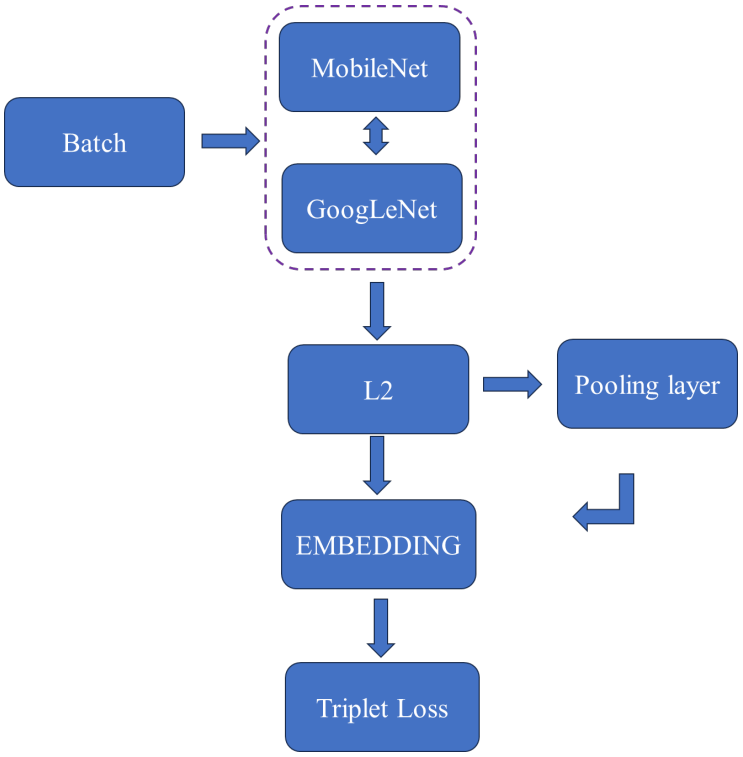

Fig. 4. Improved network structure

Batch refers to the input face image samples that have been detected and cropped to a fixed size through face detection. Then, the lightweight model MobileNet is used for feature extraction, followed by L2 feature normalization. Finally, the Triplet loss function is used for classification, so that the feature distance between the same identities should be as small as possible, while the feature distance between different identities should be as large as possible.

In the model, the loss function used in this article is based on the Triplets maximum boundary nearest neighbor classification algorithm to train the neural network, and the network directly outputs a 128 dimensional vector space.

The loss function minimizes the feature distance between identical identities and maximizes the feature distance between different identities, so that the distance between points in Euclidean space corresponding to the features of two images directly corresponds to whether the two images are similar. The purpose of Triplet Loss is to embed facial image X into the D-dimensional Euclidean space, ensuring that the distance between a specific person's image (reference image) and its own other images (positive values) is closer than when the person's image is compared to other people's images (negative values).

$$\left\| f\left(x_i^a\right) - f\left(x_i^p\right) \right\|^2 + a < \left\| f\left(x_i^a\right) - f\left(x_i^n\right) \right\|^2 . \tag{3}$$

$f\left(x\right) \in R^d$ - Euclidean space;

x_i^a - Reference image;

x_i^p - Positive image;

x_i^n - Negative value image.

$$\mathrm{H} = \sum_i^N \left[\left\| f\left(x_i^a\right) - f\left(x_i^p\right) \right\|^2 - \left\| f\left(x_i^a\right) - f\left(x_i^n\right) \right\|^2 + a \right]. \tag{4}$$

The meaning of the above equation is to optimize triplets that do not meet the conditions; For triplets that meet the conditions, set them aside and ignore them. The selection of Triples is crucial for the convergence of the model. In practical training, it is unrealistic to calculate the maximum and minimum distance between images across all training samples, and it may also be difficult to converge due to incorrect labeled images. Therefore, this article sets every 64 samples as a Mini Batch and uses online generation to filter Triplets in each Mini Batch. In each Mini Batch, 2 face images are selected as positive samples for a single individual, and other face images are randomly selected as negative samples. To avoid premature convergence of training caused by improper selection of negative samples, the negative sample selection method is as follows:

$$\left\| f\left(x_i^a\right) - f\left(x_i^p\right) \right\|^2 < \left\| f\left(x_i^a\right) - f\left(x_i^n\right) \right\|^2 . \tag{5}$$

This article uses the AdaGrad optimizer to train the MobileNet model using the random gradient descent method, with a learning rate of 0.02. After training on the CPU cluster for 300 hours, the loss function significantly decreases, and the boundary value α Set to 0.2. To test the improved MobileNet's ability to capture characters from video clips from the movie "Tunnel War". The character image is shown in Fig. 5.

Fig. 5. Facial recognition results

On the left is the image of the input model, and on the right is the recognition result.

4 Computer Assisted Intelligent Video Rendering

In the process of video production, we often need to incorporate early video materials with unclear image quality and single color into the video editing effect. Therefore, in order to improve the overall quality of the video, low-quality video segments with monotonous colors need to be rendered. Therefore, this section discusses how to use computer intelligent algorithms to assist in automatic video rendering [11].

Deep learning mainly provides feature extraction and color prediction functions in the video rendering stage. It divides the color space into continuous color spaces, converts the color regression problem into a classification problem, and then trains on a massive collection of grayscale color images in videos. The trained deep model automatically renders new photos without the need for users to specify specific colors or reference images in specific areas [12].

This article takes early black and white video materials as an example. Due to the lack of color information, it is not feasible to use color semantics to classify images in black and white videos. In addition, the semantic information of object colors is rich, making modeling very difficult. For a black and white object, its original color has multiple possibilities. To address this issue, a measurement method for item similarity in grayscale images is adopted. A similarity model is constructed by using the brightness information of the image, and the semantic features are used to calculate the correspondence between the target frame and the reference frame. The target frame is colored based on the color of the reference frame, so that the semantic features extracted by the model are independent of the color information of the image, thereby alleviating the cumulative prediction error of object color propagation on the video time scale. The schematic diagram of the working principle of similarity measurement modeling is shown in. This article selects a segment from the movie "Tunnel War" as the rendering object and uses it as video material after rendering. The rendering process is shown in Fig. 6.

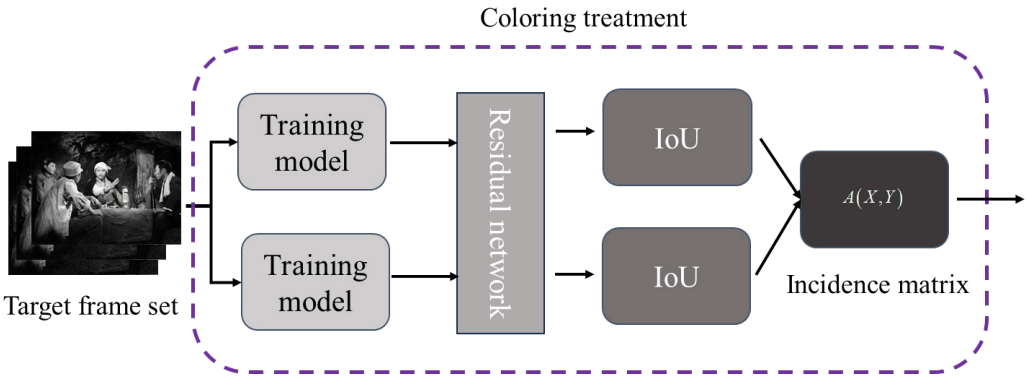

Fig. 6. Color processing model

There are three stages for the post production restoration of "Tunnel War": screen restoration, artificial intelligence rendering, and artistic coloring. In the image restoration stage, digital restoration technology is used to scan the film in 4K ultra high definition using a glue rotation scanner to obtain digitized images and sound files stored in an interconnected disk array. Then, repair software is used to eliminate scratches, mold, shaking, stretching and other issues on the image, restoring it to its optimal state.

Iterative optimization in the rendering process is the core key to the entire rendering process, mainly focusing on keyframe production, rendering result evaluation, and local optimization. The industrial process of black and white video rendering includes video restoration, video rendering, video coloring, and other stages. For rendering tasks, the upstream task is generally video restoration, while the downstream task is video coloring. To achieve video rendering, it is necessary to establish corresponding datasets, train artificial intelligence rendering models, and establish a rendering iteration optimization process [13].

1) Model training

Before rendering with a model, it is necessary to first build and train the model. Traditional rendering models are unable to complete the task of rendering black and white films well, mainly because the entire movie contains

a large number of different scenes, and a set of models is difficult to adapt to all scenes. In addition, machine learning methods also have an unavoidable issue of forgetting, which means that when the learning model learns to render a scene, it will forget the content learned in the previous scene when learning the next scene. Therefore, the entire film is manually divided into 172 segments for specialized training, and a specialized training set is established for training by categorizing them. Each type of scene corresponds to a model, such as natural scenery, architectural outdoor scenes, indoor scenes, etc. Through specialized training, enhance the pertinence and accuracy of film rendering.

2) Keyframe extraction

Use a semantic based video keyframe extraction method. Firstly, the video keyframes are preliminarily extracted using the hierarchical clustering algorithm. Then, semantic related algorithms are combined to perform feature comparison on the preliminarily extracted keyframes, remove redundant frames, and determine the keyframes of the video.

3) Manual confirmation of keyframes

The keyframe automatic rendering model is developed and implemented based on a deep neural network model. The model comprehensively considers feature loss, which is the deviation between the output image of the generator and the real image, as well as the loss of the discriminator, to generate keyframes with good quality.

4) Keyframe automatic rendering

The keyframe automatic rendering model is developed and implemented based on a deep neural network model. The model comprehensively considers feature loss, which is the deviation between the output image of the generator and the real image, as well as the loss of the discriminator, to generate keyframes with good quality.

Traditional rendering models are mostly aimed at rendering a single black and white image, and the rendering problem of black and white films poses new requirements for the model, which is to ensure that in addition to completing the rendering of a single frame image, the continuously played image sequence is visually coherent and stable. Traditional image rendering models do not take into account the continuity of image sequences, so even if adjacent frames of images are very similar, the model may consider them to be different images, which may result in unrelated rendering results. Playing such images together will result in color jumps, flickering, and other issues. In addition, traditional rendering models are mostly unsupervised learning models. It is not possible to specify the correct color for a specific item. For example, an apple can be painted in yellow, green, blue, and so on. Regardless of the color, it is considered complete rendering, and to ensure that the target is painted in the specified color, reference frames are needed to guide it [14].

5) Artificial refinement of color keyframes

Using an automatic rendering model to render reference frame images can solve the color output problem of most pixels in the reference frame from grayscale values to color values, but some details need to be further refined through manual optimization.

The rendering model uses reference frames as a guide to convert black and white images into color images according to the color rules of the reference frames. During the rendering process, due to the presence of multiple reference frames, the rendering team also sorted them in their original order in the clip and set up a sliding window. The black and white frames to be rendered can only see the reference frames within the window, and the similarity with the reference frames within the window will be calculated to select the most similar image for learning, completing the transition from black and white to color [15]. The sliding window gradually moves along with the rendering progress, completing the rendering of all black and white images of the entire lens. In addition, due to the manual rendering of reference frames, there may inevitably be slight differences between frames, which can lead to color flickering at the connection points of segments rendered based on different reference frames. Therefore, the rendering team will also smooth out the lens connection points to eliminate or reduce jitter.

5 The Application of Computer Multimedia Technology in Post production Audio Processing of Film and Television

In practical environments, there are often a large number of different types of noise that interfere with the target signal during the acquisition and transmission of audio signals. There will be strong background noise mixed in the audio signal, causing the essential characteristics of the audio signal to be seriously masked [16].

In film and television reviews, audio processing generally involves various effects processing, such as noise reduction, equalizer, reverberation, choir, and distortion. Processing audio signals through algorithms and filters to make them more pure and alter their timbre, time-domain, and frequency-domain characteristics. Digital audio processing technology can also achieve stereo and surround sound effects, creating stereo and surround sound effects through synthesis, translation, and mixing methods, enhancing the spatial and three-dimensional sense of audio. There are many uncontrollable factors in outdoor reality show programs, outdoor short videos, and other recording scenes, which may lead to good shooting but also noise. Therefore, this article introduces the use of intelligent digital multimedia technology to deal with uncontrollable noise in the on-site environment and preserve the most desired on-site sound effects.

The conventional noise reduction methods in videos use spectral subtraction, threshold cutting, and other methods to detect and process abnormal audio files such as additive noise and echo. However, most of them rely on manual inspection. After the inspector repeatedly listens to the audio file to capture noise, the corresponding calculation and processing are carried out. This is undoubtedly a serious waste of time and labor costs, while also causing great damage to human hearing and hearing.

This article uses artificial intelligence multimedia technology to process noise in audio files. The audio file is divided into two parts, time-domain and frequency-domain. In order to reduce the loss of hidden information in both fields, this paper considers designing a noise monitoring model that combines the advantages of CNN and RNN processing in both time-domain and frequency-domain [17]. The original audio file has two axes, with the horizontal axis representing seconds (time or frame rate) and the vertical axis representing signal strength. Because information can only be extracted from the vertical axis, audio signals are processed as one-dimensional data. The advantage of CNN convolution kernels lies in matrix operations. Therefore, the most important thing is to select a suitable type of data as input and hand it over to the model for training. Most CNN convolutional models can be used for CNN layer parts in CR-DNN structures, such as AlexNet and LeNet-5, depending on the type and complexity of input and noise types. Because each convolution operation within the CNN layer can generate a feature map, thereby representing the data features to be learned. The model structure is shown in Fig. 7.

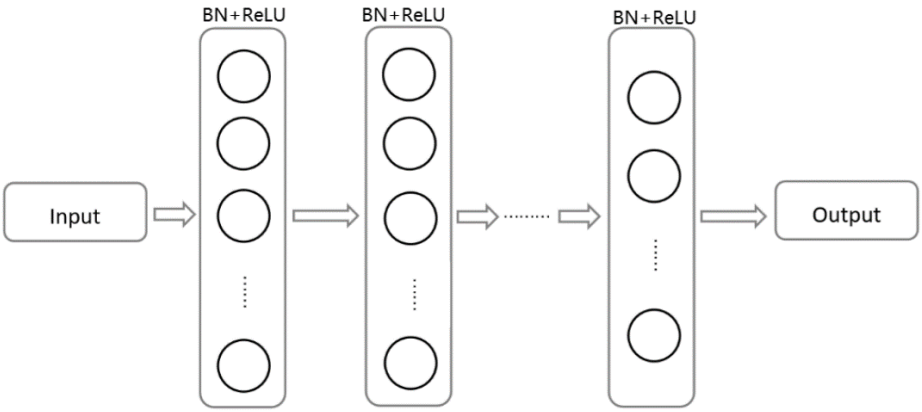

Fig. 7. Noise processing model

Using CNN structure as an encoder to automatically extract features from audio files. Then, it will share parameters with the Conv encoder, which is a pair of twin networks with the same structure but different parameters. The Conv decoder is a mirror image of the Conv encoder, and here we refer to the structure of the SEGAN model using a step connection method. This means that each layer in the Conv encoder can skip other layers and directly connect to the matching layer in the Conv decoder, which is more suitable for producing more realistic results. If only the original file is used as input to the streaming network, it may cause compression bottlenecks and result in the loss of many feature information [18]. Step connections can more effectively ensure that the information obtained from analyzing audio can be transmitted to the decoding layer stage.

Set the loss function as follows during the process:

$$\min_G V(G) = \frac{1}{2} E_x \left[\left(D\left(G(z,x),x\right) - 1 \right)^2 \right] + \lambda \left\| G(z,x) - x \right\|. \quad (6)$$

In summary, the algorithm flow for noise processing is as follows:

Algorithm 1. Perform downsampling preprocessing from 48kHz to 16kHz on all audio files

INPUT: Video file;
OUTPUT: ACC(X)
STEP1: Downsampling the audio in the original video to 16kHz;
STEP2: Freeze generator, single execution discriminator to train classification network;
STEP3: Freeze the discriminator, execute the generator training process, update Gaussian noise through a mirror network to generate more realistic simulation data, and fit the data distribution of clean audio;
STEP4: Repeat steps 2 and 3 to guide the loss towards Nash equilibrium;
STEP5: Save the parameter model file;

end.

6 The Application Strategy of Computer Multimedia Technology in the Promotion of Film and Television Post Production Courses

The course "Multimedia Video Post Production" is a professional course related to film and television media in vocational colleges. It integrates various technologies such as film and television editing, film and television special effects, and synthesis under the guidance of theories such as film and television basics and non-linear editing. It is a highly integrated course with artistic and technical aspects. As it is a necessary process in film and television post production, video post production directly determines the training quality of film and television post production personnel. However, most students who have completed this course only know how to operate and edit software, which cannot meet the overall requirements of the project, let alone the needs of social enterprises. The reason for this is that the knowledge learned by students in the classroom is relatively superficial and disconnected from social enterprises. In response to this situation, most vocational colleges have carried out corresponding teaching reforms on the course of "Multimedia Video Post production" to meet the needs of the meeting [19].

The course objective is to cultivate high-quality, highly skilled and practical talents for the modernization and economic development of society. Skilled talents such as computer application technology, graphic and image processing technology, audio and video production technology, and video editing are required. The multimedia video post production course mainly cultivates talents engaged in film and television editing and production. The social work targets include television stations at all levels, film and television communication companies, advertising and cultural companies, wedding companies, etc., with a wide range of employment opportunities. The main job responsibilities of the trained talents include video editing and production, film and television special effects and production, film and television opening and production, film and television packaging, etc. In the process of cultivation, in addition to cultivating students' skills, attention should also be paid to cultivating their innovative consciousness, as people's demands for the visual beauty of film and television are increasingly high. The current teaching model is not yet sufficient in this area, especially in terms of cultivating innovative abilities for students. Project based teaching has certain advantages in teaching and can meet the training objectives of the multimedia video post production course in vocational colleges.

6.1 Analysis of Learning Situation

The main target audience for training is vocational college students, who have the following characteristics:

1) The significant differences in student personality are accompanied by the continuous expansion of enrollment in vocational colleges. The quality of students in vocational colleges is uneven, and their grades are gener-

ally low, with a large gap between grades. Most students come with the necessary attitude to obtain a graduation certificate, with low learning interest, insufficient motivation, lack of the necessary learning confidence, and no clear learning goals.

2) Poor self-control and weak motivation. Most students from vocational colleges come to school with the attitude of finding a school and eventually obtaining a graduation certificate, without clear goals and no interest in learning. I have found that most students have poor self-discipline during class, and for strict teachers, they have to sit in the classroom to avoid failing. For teachers who are not strict, the scene is even more spectacular, chatting and sleeping. Especially in this era of advanced internet, students have a mobile phone and sit for a long time without anything to do. They just want to play with their phones. Some students play with their phones for a whole class, and once the teacher becomes strict, students will have a rebellious attitude. For teachers who don't have strict requirements, students love direct wing classes the most. The phenomenon of sending seeds is common in vocational colleges. In order for students to truly integrate into the classroom and learn better, there must be an effective teaching method.

6.2 Course Content Setting

The multimedia video post production course is a highly integrated course that combines theory and practice, art and skills. The course content is shown in Table 2.

Table 2. Content setting of the course

Chapter	Primary coverage
Chapter 1 Theoretical Foundations of Film and Television Post Production	Preliminary understanding of film and television fundamentals, mastery of professional terminology and development status in film and television post production.
Chapter 2 Introduction to Post Production Software	Understand the functions and usage of video editing software
Chapter 3 Video Editing	Master video editing skills; Editing ideas; Transition settings; Principles and techniques for lens assembly.
Chapter 4 Video Special Effects	Master the production of some simple video special effects, including the production of opening and ending credits, as well as video special effects and video special effects.
Chapter 5 Subtitle Design and Production	Master the production methods of static and dynamic subtitles; Master subtitle design and video subtitle production for various themes; I will add and design subtitles for feature films and various accessories.
Chapter 6 Audio Processing	Master audio processing skills; Noise reduction techniques; Balance technique; Audio video synthesis technology.
Chapter 7 Video Output	Master various video formats and video output methods. Understand the characteristics and commonly used video formats of each type of video.

6.3 Analyzing Film and Television Production Project Arrangements Through Real Case Studies

In project teaching, choose "Ceramic Art Design" video editing and production as an example for teaching. The teaching of this project is divided into three stages: the first stage is preparation before the project. The second stage is the implementation of the project. The third stage is the evaluation of the effectiveness of project-based teaching. Divide each stage into several main stages. Teachers organize teaching according to specific stages.

A company in Beijing filmed a video of the ceramic hand drawing process for a teacher majoring in art and design at our institute. The materials have been filmed and there is a detailed shot book available. Now we have progressed to the post video editing stage, where the video editing task is completed by the class students. In order to cultivate students' ability in international editing, it is necessary to edit this video into a complete ceramic hand-painted demonstration video based on the captured video and the storyboard script.

Clarify production requirements. There is a complete opening, necessary subtitles, and detailed process commentary. The company will select a good set of videos produced by classmates as the final product for display, and there will be a certain amount of financial rewards. By citing the project, students' interest and hobbies in

learning are stimulated, and a corporate environment is created for them, giving them a sense of achievement and competition after completing the video. This can stimulate students' enthusiasm and creativity, which is more conducive to their learning.

Display company project demonstration videos. By showcasing demonstration films, provide students with clear production standards and references. The standards should not be lower than those of demonstration films, the clarity of the video should not be too low, and the recording of commentary should be clear. Simultaneously describe the key and difficult points in the project production process.

Student grouping. In the student grouping stage, according to the principle of voluntary participation of students and the teacher's grasp of the actual situation of students, the teacher guides the organization of student grouping. Divide the students in this class into five groups, with each group consisting of 4-5 members, and generally no more than 5 people.

In the second stage, the following tasks should be completed:

The team leader organizes members to discuss the specific sub tasks of the project, divides the project into several tasks, and each member is responsible for the corresponding sub tasks, and develops a project plan. In the process of division of labor, the team leader should solicit opinions from team members, leverage their strengths, and discuss the division of labor before filling out the project division of labor form and project plan.

Organize the materials and collect information. Under the organization of the team leader, students organize the shooting materials according to the storyboard. The organized materials should be placed in the corresponding folder according to the storyboard. At the same time, download relevant video and image resources from the internet through the material website and learning resources provided by the teacher.

Explore, collect, organize, and practice. In this section, students first need to discuss the precautions during the project production process, collect and organize materials, and complete the project production under the guidance of the teacher. During the project production process, due to the periodicity of the project, there may be many problems. When students encounter problems, they should be encouraged to solve them independently.

The third stage mainly completes the following tasks:

Work evaluation is a crucial aspect of project-based teaching. Establish a project evaluation team. The teacher serves as the group leader among the members of the evaluation group. During the scoring process, distribute a scoring sheet to the corresponding scoring personnel. The scores obtained from teacher evaluations each account for 60%, while the scores of group members account for 20%. At the same time, mutual evaluation within the group is required, and the scores obtained from mutual evaluation within the group account for 20% of the total score. The evaluation criteria are shown in Table 3.

Table 3. Evaluation criterion

Number	Evaluation project	Production requirements	Score
1	Video editing and special session settings	Reasonable video editing; Reasonable transition; Reasonable lens combination; Good visual effect;	48
2	Title designer	Reasonable font style; Appropriate font size; Correct subtitles for the commentary;	16
3	Commentary effect	The commentary closely follows the theme; Is the sound clear Reasonable volume size;	24
4	Opening and ending design	Is it due to attractiveness;	12

After the above course design, a vivid film and television post production course can be successfully completed. In order to promote film and television post production technology and methods to adapt to the current entrepreneurial wave of short video production, the following two points should be achieved in course construction and promotion:

1) Enhance digital production capabilities

There are still many shortcomings in the application of computer multimedia technology in the post production of film and television works in our country. Therefore, future development strategies need to be considered from multiple aspects, strengthening the reform and innovation of computer multimedia technology, and promoting the greater application value of advanced computer technology in the field of post production of film and television works. However, in practical applications, the vast majority of high-quality video editing software and equipment come from abroad, and the core technology is limited by foreign countries. Therefore, it is necessary for us to provide corresponding policy support in the development of related software, and at the same time, establish high-quality multimedia processing majors or order classes in higher education and vocational colleges. Provide reliable technical support and talent reserves for film and television post-processing.

2) Clarify the development direction of domestic film and television post production

Given the relatively backward hardware and software technologies in response, we should develop film and television works related to "national culture". Chinese culture is vast and profound, and there is a scarcity of text and image datasets related to classical Chinese culture. Therefore, it is necessary to root in Chinese culture, clarify the development direction of "national is the world" works, and vigorously develop Chinese characteristic film and television as well as film and television processing technology.

7 Conclusion

At present, the level of film and television post production in China is gradually improving. Innovation and breakthroughs need to be emphasized in production thinking and technological application. Computer multimedia technology should be in line with internationalization in the field of film and television post production, combined with artificial intelligence and light and shadow concepts. The application thinking of innovative technology is crucial. Therefore, enhancing the professional competence and creative thinking of technology practitioners is an important prerequisite for solving problems. At the same time, it is necessary to strengthen the research and development of independent technologies and expand resources, break free from long-term foreign film and television production technology monopolies or influences, and create a pioneer in domestic film and television post production technology, including digital film production technology, artificial intelligence embedding technology, film and television animation simulation technology, etc. Finally, strengthen the professional training of relevant technical talents, cultivate them from the perspective of ideology and skills, and improve the level of post production of film and television works in China.

This article discusses the requirements of computer multimedia technology, and then uses artificial intelligence methods in video editing and camera operation in post production, improving the level of film and television post production and expanding the ideas of film and television post production.

Film and television works usually use a combination of image and sound in narrative, presenting the complete story well and accurately expressing the emotions of the creators, so as to meet the ultimate requirements of the creators. In this process, the application of computer digital technology can enhance the expressive power of film and television works, but its effect on the expression and transmission of ideas and emotions is limited. Therefore, in the actual process of creation and application, computer digital technology should be reasonably used in combination with practical work needs, fully highlighting its characteristics, and improving the visual effect of computer technology in film and television works. By using digital technology to organically integrate the storyline, the theme of film and television works can be more realistic.

In the final stage, regarding the current field of film and television post production, vocational colleges elaborated on the arrangement and design of film and television post production courses, with the aim of starting from the entire industry and clarifying how to demonstrate and promote artificial intelligence as the main post-processing method in the course.

References

[1] P.-P. Li, Application Skills of Post-Production in Character Television Feature Films, Video Engineering 47(11)(2023) 90-92.

[2] Y.-T. Su, Practical Analysis of Computer Multimedia Technology in the Post-Production of Film and Television, Audio Engineering 45(5)(2021) 41-43.

[3] X.-F. Wang, Application of Computer Multimedia Technology in Post Production of Film and Television, Digital Communication World (11)(2022) 134-136.
[4] J.-L. Li, Application of artificial intelligence technology in the field of film and television production, Computer Communications 30(11-12)(2007) 2385-2400.
[5] Q. Wang, Application and Discussion of Artificial Intelligence Matting Technology in Film and Television Post-production, Information & Computer 35(15)(2023) 91-93.
[6] J.-T. Shi, H.-Y. Zhan, Improved lightweight single-person pose estimation algorithm based on MoveNet, Journal of Heilongjiang University of Science and Technology 33(4)(2023) 589-595+601.
[7] Z.-J. Wang, W.-F. Lan, X. Zhang, Z.-T. Hou, N. Jin, Cuqiu motion recognition and skill assessment based on CBAM-DCNN-BiLSTM, Journal of South-central Minzu University (Natural Science Edition) 41(5)(2022) 599-605.
[8] D.-F. Wang, L. Liu, J. Cao, G. Zhao, W.-S. Zhao, W. Tang, Semantic Segmentation Method of Autonomous Driving Images Based on Atrous Spatial Pyramid Pooling, Automotive Engineering 44(12)(2022) 1818-1824.
[9] S.-Y. Shi, J.-Z. Wei, Design of hardware accelerator based on MobileNet-SSD for object detection algorithm, Microelectronics & Computer 39(6)(2022) 99-107.
[10] Y.-B. Meng, S.-R. Mu, G.-H. Liu, S.-J. Xu, J.-Q. Han, Person Re-identification Method Based on GoogLeNet-GMP Based on Vector Attention Mechanism, Computer Science 142-147.
[11] X. Zhao, Application of Digital Color Matching Technology in Post Production of Film and TV, Electronic Technology 52(8)(2023) 340-341.
[12] L.-D. Bai, Research on the Application of Computer Multimedia Technology in Film and Television Post-Production, Audio Engineering 47(11)(2023) 75-77.
[13] H. Liu, Y.-Y. Pu, D.-H. Lv, Z.-P. Zhao, D. Xu, W.-H. Qian, Polarized Self-attention Constrains Color Overflow in Automatic Coloring of Image, Computer Science 50(3)(2023) 208-215.
[14] W.-Q. Gu, The Technology and Application of Digital Film and Television Post-editing and Special Effect Synthesis, Toys World (4)(2023) 73-75.
[15] Y. Zhou, Research on furniture product digital display platform based on virtual reality technology, Modern Electronics Technique 44(3)(2021) 172-176.
[16] T.-T. Gu, Research on the Application Route of Digital Audio Technology in Post Production of Film and Television Works, Audio Engineering 47(9)(2023) 52-54.
[17] M.-M. Zhao, Research on Computer Digital Video and Audio Processing Technology, Modern Information Technology 5(20)(2021) 87-90+94.
[18] J.-X. Xie, G.-T. Liang, D.-X. Yin, Real-time audio narrowband noise filtering method based on FFT and ADALINE network, Modern Electronics Technique 45(9)(2022) 41-45.
[19] N. Lin, Auxiliary and Promoting Effect of Short Video Operation Setting on Teaching-Taking Film and Television Post-production as an Example, Journal of Shenzhen Polytechnic 21(2)(2022) 69-72.