The Implementation of Image Conceptualization Split-Screen Stitching and Positioning Technology in Film and Television Production

Zhouzhou Deng[1*], Rongshen Zhu[2]
School of Arts and Science, Chengdu College of University of Electronic Science and Technology of China, Chengdu, 611731, China[1]
School of Computer Science, Chengdu College of University of Electronic Science and Technology of China, Chengdu, 611731, China[2]

***Abstract*—In order to study the technology of image conception, splitting, stitching and positioning in film and television production, this paper first discusses the relevant research literature, then designs an improved biomedical image segmentation convolution network model applied in film and television production, and then verifies the effectiveness of the proposed model. Ultimately, the paper summarizes the research findings. Aiming at the problem that the traditional image mosaic positioning model has poor robustness because of its insufficient ability to extract features and inaccurate segmentation and positioning areas, this study proposes a biomedical image segmentation convolutional network model that is based on dense block and void space convolutional pooling pyramidal module. Additionally, an attention mechanism is introduced to enhance the biomedical image segmentation convolutional network model. The results show that the accuracy, recall, and F1 value of the biomedical image segmentation convolutional network model are 96.48%, 95.24%, and 95.96%, respectively, on the Colombian uncompressed image stitching detection dataset, and the accuracy, recall, and F1 value of the improved biomedical image segmentation convolutional network model are 98.19%, 96.23%, and F1 value of 97.21%. In summary, the improved convolution network model for biomedical image segmentation has excellent performance, and it has certain application value in image conception, mirror splicing and positioning in film and television production.**

***Keywords*—Convolutional neural network; attention mechanism; null space convolutional pooling pyramid; spatial rich model; dense block**

I. Introduction

The field of cinema and television production has encountered novel challenges and made significant strides as a result of the ongoing advancements in science and technology [1-2]. As an essential pre-production step in film and television, the split screen is an intermediate medium for converting text into a three-dimensional audio-visual image and presenting it in a pictorial form [3-4]. Due to the increased functionality and user-friendliness of image editing software, producers can easily select areas of interest from other images to be cut and spliced into the split-screen image, which brings great convenience for creating the split-screen image, but also easily brings trouble to the film and television copyright [5]. To enhance split-screen imagery creativity in film and television production and reduce the risk of copyright, it's crucial to prioritize research on image stitching and positioning technologies [6]. Image stitching is a semantic segmentation of features by dividing the stitching area into one category and the real area into another category. The CNBIS model is capable of extracting the stitched region from the real region. The traditional CNBIS model can achieve good results only when it is faced with simple semantic information of the same kind of content. However, in image mosaic and positioning, the mosaic area often comes from different semantic interference information, resulting in inadequate feature extraction and inaccurate segmentation and positioning area, which reduces the segmentation accuracy of the model [7]. The research aims to improve image semantic segmentation accuracy and make up for the deficiency of model extraction features caused by different semantic interference information. To overcome these issues, the study implements Dense Block (DB) and Atrous Spatial Pyramid Pooling (ASPP) modules to improve the CNBIS model and introduce the DACNBIS model. The research also incorporates the Attention Mechanism (AM) to enhance the DACNBIS model and create the AM-DACNBIS model, which increases the segmentation accuracy of the DACNBIS model. The primary contribution of the research is to broaden the application of image mosaic positioning technology in film and television production. The goal is to increase the diversity of split-mirror images, improve the reduction and richness of split-mirror in films, and make the storyline and the overall picture of film and television works more complete. The research focuses on two significant innovations. The first point is to improve the traditional CNBIS model by introducing DB module and ASPP module, and improve the DACNBIS model by combining attention mechanism. The second point is that AM-DACNBIS model is divided into three parts: parameter sharing, area monitoring and edge detection, and specific feature layers are used for each part to extract the corresponding task information. The study is primarily divided into four parts. The first part reviews relevant pertinent research findings. The second part constructs the construction of DACNBIS model and AM-DACNBIS model. The third part validates the two proposed models' validity in the study. The final part concludes the research.

II. Related Work

Many image processing jobs start with the pre-processing stage of picture semantic segmentation, and many academics have written extensively about ways to increase segmentation accuracy in this process, Gao and co. To help the network better concentrate on object borders and small objects during the feature extraction process, a sensitive feature selection module was created to reweight each pixel on several channels. The findings of the experiments demonstrate that the sensitive feature selection module can aid in the semantic segmentation algorithm's high segmentation accuracy [8]. To reduce semantic information loss and improve image information, Zhou et al. proposed a semantic segmentation model based on dense convolutional separation convolution. This model also integrates multi-scale feature information for a broader perceptual field and captures more dense pixels. Simulation experiments demonstrate the superior performance and good performance of this model for image segmentation [9]. Zhang et al. proposed a system that utilizes the original semantic segmentation network to deal with the issue of missing static object segmentation occluded by dynamic scenes and combines it with image restoration [10]. Maurya et al. created a cross-form attention pyramid to extract multi-scale information using a pre-trained model in order to address the issues of feature redundancy and low discrimination in image semantic segmentation. An attention module in a spatial manner is then introduced to further improve the segmentation effect. With the addition of the attention pyramid and attention module, simulation results demonstrate that the semantic segmentation model has greater segmentation accuracy [11]. In order to tackle the problem of loss of image information caused by feature extraction in the process of image semantic segmentation, Chen constructs a semantic segmentation model with encoder-decoder as the basic structure. The research results verify the logic and effectiveness of the model [12].

CNBIS model is a very famous segmentation network model in the field of image segmentation, which has been widely used in various fields. In order to accurately divide non-enhanced tumor, enhanced tumor, tumor core and undamaged region in brain images, Teki et al. used CNBIS model to realize semantic segmentation of brain tumor images. The research findings demonstrate that the CNBIS model performs well in segmenting simple semantic information. However, in the presence of complex semantic interference information, the model exhibits insufficient feature extraction capability and low segmentation accuracy [13]. Singh et al. designed an improved Deep-CNBIS model for semantic segmentation of images using satellite images to extract vegetation cover. The simulation experiments show that the model has superior performance in semantic segmentation accuracy of satellite images [14]. Tiwari et al. proposed an improved CNBIS model for segmenting vehicles, which segmented the input images by successive encoding and decoding steps. The simulation results show that the improved CNBIS model outperforms other segmentation models in terms of segmentation accuracy [15]. Cheng et al. designed a separated convolutional CNBIS model combining convolutional downsampling. The outcomes of the simulation experiment demonstrate that the model is capable of quickly and easily detecting fabric defects with high accuracy [16]. Mahmoud et al. developed a deep learning model based on the combination of CNBIS model. The simulation experiments demonstrate that the model performs significantly better than the traditional CNBIS model, exhibiting a higher accuracy rate [17]. Abdelrauof et al. proposed to use multi-gated expansion starting blocks. It is evident from the experimental results that the addition of multi-gate expansion start blocks improves the CNBIS model performance and demonstrates exceptional capability [18].

In summary, there are many research results on image semantic segmentation and CNBIS model applications. However, most of the image semantic segmentation studies use small data set samples, which do not sufficiently meet the deep learning requirement for extensive data training. The traditional CNBIS model has insufficient feature extraction capability, which leads to inaccurate image localization and poor robustness. The paper suggests the DACNBIS model to overcome the aforementioned issues and introduces an attention mechanism to enhance the DACNBIS model, resulting in the AM-DACNBIS model.

III. DACNBIS Model and AM-DACNBIS Model Construction in Image Stitching Localization Detection

The purpose of image stitching, which separates the stitching area from the real area so that they are presented separately, can be thought of as a unique sort of semantic segmentation.

A. Construction of DACNBIS Model Based on CNBIS Model

As the initial stage of rendering text into images in film and television production, splitting is the process of dividing a film script into a series of shots that can be filmed and presented as images. Convolutional Neural Networks (CNN) are a deep learning model with learnable weights and bias constants that are suitable for processing image data processing through supervised learning [19-22]. The convolutional layer can extract features from images. Equation (1) expresses the parameters of this layer.

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * Kernel_{ij}^l + b^l\right) \quad (1)$$

In equation (1), x_j^l is the j neuron of the l layer, $Kernel$, *, and $f(\square)$ represent the convolution kernel, convolution operation, and nonlinear excitation function, respectively, a and M_j represent the bias term and the number of inputs of the j neuron, respectively. The pooling layer can reduce the risk of overfitting by reducing the dimensionality. The excitation layer introduces nonlinear features to the neural network, enabling it to approximate any nonlinear function. The activation function image is shown in Fig. 1.

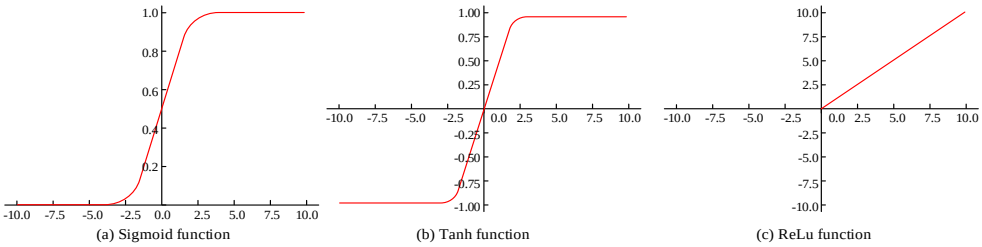


Fig. 1. Schematic diagram of three activation functions.

The Sigmoid function, Tanh function, and ReLu function, respectively, are depicted in Fig. 1 where the expression of the Sigmoid function is shown in equation (2).

$$f(x)=\frac{1}{1+e^{-x}} \tag{2}$$

The value range of equation (2) is $(0,1)$, which corresponds to the probability value range $(0,1)$. Equation (3) contains the Tanh function's expression.

$$f(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{3}$$

The output of equation (3) is centered on 0, which can play the effect of data centering. Due to the saturation region of Sigmoid function and Tanh function, the gradient is prone to gradient disappearance when back propagation, and this phenomenon becomes more and more obvious. This issue is mitigated by the ReLu function, which is expressed mathematically in equation (4).

$$f(x)=\begin{cases}x, x>0\\0, x<0\end{cases} \tag{4}$$

The positive activation value derivative of the ReLu function in equation (4) is 1. In order to accomplish the image classification task, the fully connected layer must integrate the input image data and map the feature map from the convolution layer into a fixed-length feature vector, as shown in equation (5).

$$x_j^l=f\left(\sum_{i\in M_j}x_i^{l-1}*w_{i,j}+b_j\right) \tag{5}$$

In equation (5), $w_{i,j}$ and b_j denote the weight and offset between neurons, respectively. CNBIS, as a special semantic segmentation algorithm in the field of deep learning, is a derivative model of CNN and capable of addressing image stitching and localization issues.

The CNBIS model in Fig. 2 adopts a fully symmetric coding-decoding structure. The coding section is downsampled four times, and the decoding link is corresponding to the four stages of the coding link, each of which incorporates the feature information corresponding to that from the coding process. This results in the final feature map is restored to the original image size. Due to the small sample of dataset about image stitching at this stage, the neural network is prone to the risk of overfitting, the study introduces the DB and ASPP modules to improve the CNBIS model, resulting in the DACNBIS model. The DB module is a CNN with tightly connected nature. The feature map resolution of each layer is of the same size, so the channels of each layer can be stitched together in dimension. Suppose the output channel of the l layer inside the DB module is X_l, then the expression of X_l is shown in equation (6).

$$X_l=H_l\left[X_{l-1},\ldots,X_1,X_0\right] \tag{6}$$

In equation (6), H_l is the nonlinear transformation function of the l layer, $[\square]$ indicates that all output feature maps of the $X_0,X_1,\ldots,X_{l-1}$ layer are combined by channel, and the expression of the network input X_i of the i layer is shown in equation (7).

$$X_i=K_0+(i-1)\times K \tag{7}$$

In equation (7), K is the number of output channels of the nonlinear function H, and K_0 is the number of input channels. The dense jump connection implemented inside the DB module causes the number of channels after the l layer to become large. The study adopts a convolution of 1×1 before the convolution of 3×3 in the DB module. The ASPP module aims to minimize the loss of accuracy generated by the feature map in the process of recovering the resolution size of the original image. It employs four varied expansion rates of the null convolution to capture multi-scale information. The specific structure of DACNBIS model is shown in Fig. 3.

In Fig. 3, the study introduces the DB module to replace the original CNBIS model upsampled by the 3×3 convolutional network. This enhances feature reuse and propagation, increasing the richness of feature extraction. The study also introduces the ASPP module to replace the fourth sampling session of the original CNBIS model. This module extracts multi-scale information, expands the network perceptual field, and improves the model's segmentation effects on spliced regions of different sizes.

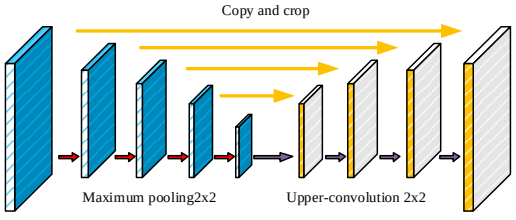


Fig. 2. Schematic diagram of CNBIS model structure.

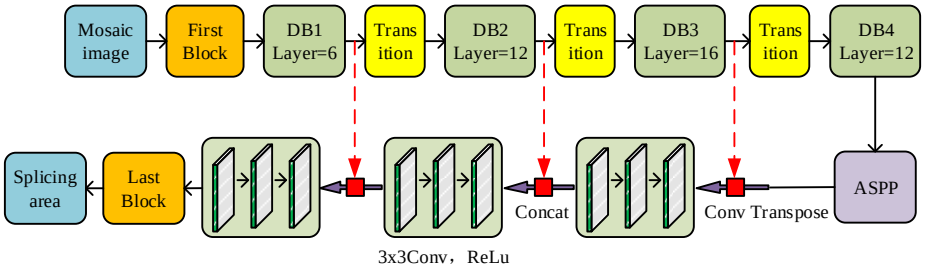


Fig. 3. Schematic diagram of DACNBIS model structure.

B. *Construction of AM-DACNBIS Model based on DACNBIS Model*

In order to avoid some redundant and useless low-level semantic information from interfering with the decoding process of the DACNBIS model, thereby affecting the influence of image sub-screening conception in the film and television production stage, the research combines AM to improve the DACNBIS model, resulting in the AM-DACNBIS model. It is based on the principle that humans selectively focus on the more interesting and informative visual areas when observing a scene, thus ignoring other irrelevant areas and thus improving the utilization of visual information. The study draws inspiration from AM and introduces a Global Attention Upsampling (GAU) module to provide global context as a low-level guide, see Fig. 4 [20].

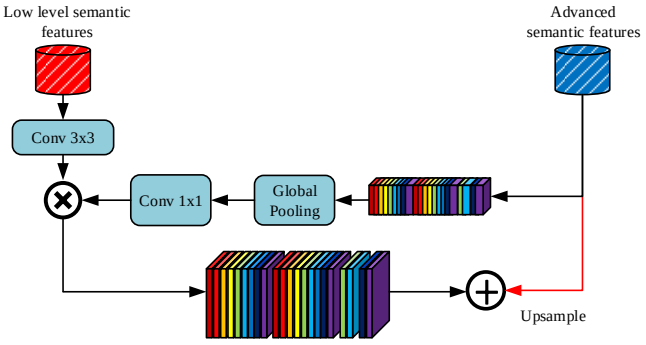


Fig. 4. GAU module schematic diagram.

In Fig. 4, GAU first convolves the low-level semantic features with 3×3. Since the AM-DACNBIS model ignores the differences between the essential attributes of images when performing image input, which in turn leads to large differences between the extracted real images and the stitched images and affects the splitting effect in the film and TV production process. In order to reduce the interference situation of different semantic information in the stitching region on the network's extracted features, the study implements the Spatial Rich Model (SRM) filter in the input stage to analyze the image features.SRM is a high-latitude steganographic model, when the image hides secret information, SRM will destroy the image attributes and extract the feature information from these images that are hidden information. Suppose the image pixel is X, the expression of image residual $R_{i,j}$ is shown in equation (8).

$$j=1,2,\ldots,n_2, i=1,2,\ldots,n_1, R_{i,j}=\hat{X}_{i,j}\left(N_{i,j}\right)-cX_{i,j} \quad (8)$$

In equation (8) n_1, n_2 denote the pixel points in horizontal and vertical directions respectively, $N_{i,j}$ is the field without the central pixel $X_{i,j}$, $\hat{X}_{i,j}\left(N_{i,j}\right)$ refers to the estimated value of the central pixel $X_{i,j}$ in the region $N_{i,j}$, c is the residual order, when the magnitude of the residual pixel is large, the pixel correlation there will be reduced, so the study also needs to quantize and truncate the residual image, as shown in equation (9).

$$R_{i,j} \leftarrow Trunc_T\left(round\left(\frac{R_{i,j}}{Q}\right)\right), j=1,2,\ldots,n_2, i=1,2,\ldots,n_1 \quad (9)$$

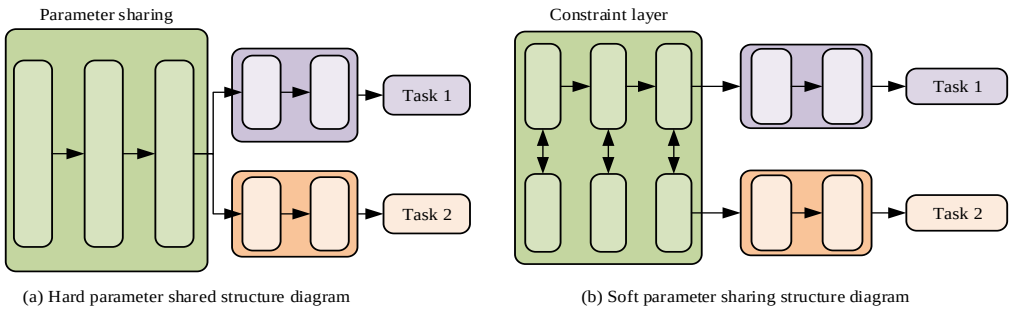
In equation (10), Q is the quantization step, *round* indicates rounding, and $Trunc_T$ means the elements are truncated one by one by the threshold T, where $Trunc_T\left(x\right)$ is expressed as shown in equation (10).

$$Trunc_T\left(x\right)=\begin{cases}x, x\in\left[-T,T\right]\\ Tsign\left(x\right), x\notin\left[-T,T\right]\end{cases} \quad (10)$$

In equation (10) $sign\left(x\right)$ is the symbolic function and the high-pass filter extracted from the SRM can be used for the extraction of RGB image noise, as shown in equation (11).

$$K=\frac{1}{2}\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\0&1&-2&1&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix} \quad (11)$$

The residual image resulting from the extraction of the SRM filter in equation (11) highlights edge features of the stitched region while suppressing other contents. In order to maintain sharing in the feature extraction process and minimize the risk of network overfitting, the study comprises of a multi-task learning output from the branching task of stitching edge localization using the DACNBIS model. According to Fig. 5, the hard parameter sharing and soft parameter sharing categories best describe the multitask learning structure.



(a) Hard parameter shared structure diagram (b) Soft parameter sharing structure diagram

Fig. 5. Schematic diagram of multi-task learning structure.

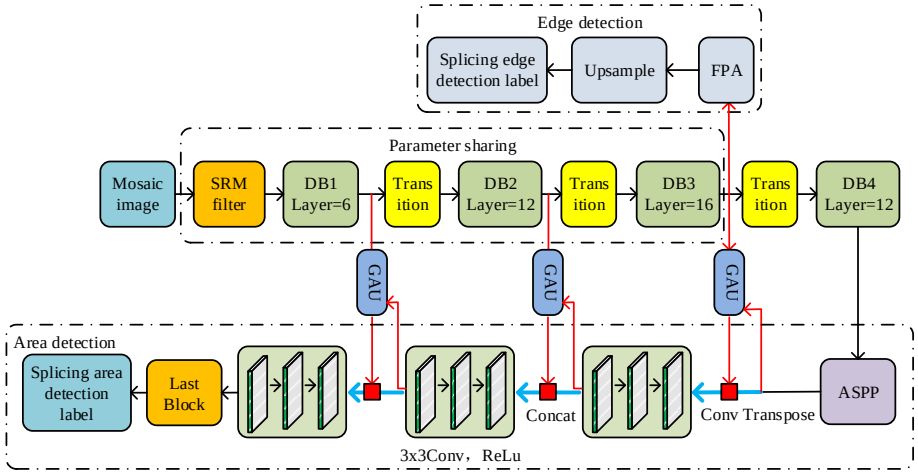


Fig. 6. Structural diagram of AM-DACNBIS model.

Fig. 5(a) shows the hard parameter sharing structure, in which the bottom layers of the network's input share parameters. This is known as bottom parameter sharing. Different learning tasks present different branches after sharing, and these tasks are trained in parallel with each other, and the feedback action is performed through the loss function each learning task. Fig. 5(b) shows the soft parameter sharing structure, in which every task has independent models and parameters. Every model can access the internal information of other models and regularize the distance between model parameters to ensure the similarity between parameters. Since the soft parameter sharing mechanism has separate models among multiple tasks. A schematic representation of the final AM-DACNBIS model structure is shown in Fig. 6. The study also embeds the data information of multiple tasks into a single semantic space and extracts the relevant task information for each task through a specific feature layer.

The model structure in Fig. 6 is mainly divided into three parts: parameter sharing, region monitoring, and edge detection. For the parameter sharing part, the AM-DACNBIS model adds SRM filters and completes the shared training for both region detection and edge detection tasks through the DB module in the pre-input phase of the network. It then embeds the information of both tasks into the same semantic space. For the region detection part, the study introduces the GAU module in the decoding process of the AM-DACNBIS model to provide global context as the underlying guidance to improve the sensitivity of important feature information. For the edge detection part, the study invokes the feature pyramid attention module in the edge branch and learns better feature representation at the parameter sharing layer through this module. To evaluate the AM-DACNBIS performance, the study presents the performance metrics of accuracy, recall, and F1 value for testing, as shown in equation (12).

$$\begin{cases} R = \dfrac{TP}{FN + TP} \\ F1 = \dfrac{2PR}{R + P} \\ P = \dfrac{TP}{FP + TP} \end{cases} \tag{12}$$

In equation (12), P and R represent the precision and recall, respectively. TP is the actual stitched area pixels and the predicted stitched area pixels; FP is the actual real area pixels and the predicted stitched area pixels; FN is the actual stitched area pixels and the predicted real area pixels.

IV. Analysis of the Results of DACNBIS Model and AM-DACNBIS Model in Image Stitching Localization Detection

The section focuses on the examination of the experimental results of the DACNBIS model and the AM-DACNBIS model. To confirm the validity of the DACNBIS model and the AM-DACNBIS model, comparative tests were conducted on various data sets utilizing different model sets.

A. Experimental Data Preparation

The algorithmic model was developed with the PyTorch deep learning framework, and the hardware environment for the experiments was a workstation running Windows OS. This was done to compare the performance of the proposed models. With a stochastic gradient descent network training optimizer, a binary cross-entropy loss function, an initial learning rate of 0.01, momentum of 0.9, weight decay of 0.0005, and performance metrics of accuracy, recall, and F1 value. The Chinese Academy of Sciences Institute of Automation 1 (CASIA1) dataset, CASIA2 dataset, Columbia Uncompressed Image Splicing Detection (CUIS) dataset, and the Chinese Academy of Sciences Institute of Automation 1 (CASIA2) dataset are used for the study. For the experiments, the Image Splicing Detection (CUISD) datasets are utilized, and the specific experimental data are divided as shown in Table I.

TABLE I. PARTITION RESULT OF DATA SET

Data Set	CASIA1	CASIA2	CUISD
Training set	1050	7000	175
Verification set	300	2000	50
Test set	150	1000	25

1500 samples total for the CASIA1 dataset, 10,000 samples total for the CASIA2 dataset, and 250 samples total for the CUISD dataset are shown in Table I, where the training, validation, and test sets account for 70%, 20%, and 10% of the corresponding datasets. To enhance the understanding of the datasets, the schematic diagram of some data sets is selected, as shown in Fig. 7.

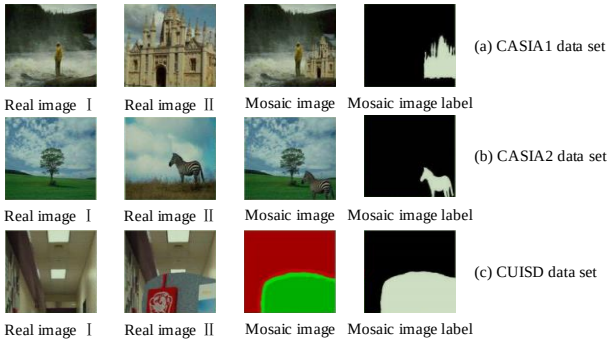


Fig. 7. Sample diagrams of three data sets.

Fig. 7(a) is a schematic diagram of the CASIA1 dataset with stitching areas of different sizes and arbitrary boundaries such as circles, triangles and rectangles. Additionally, Fig. 7(b) illustrates the CASIA2 dataset, which is an upgrade of the CASIA1 dataset with more data and better production. Fig. 7(c) displays the CUISD dataset which provides labels with red and green colored edge templates, and the edge labeling error is larger.

B. Performance Analysis of DACNBIS Model

To verify the performance of DACNBIS model, the study conducted comparison experiments using Fully Convolutional Networks (FCN), CNBIS, and Pyramid Scene Analysis Network (PSAN) models. The number of model iterations was increased from 10 to 100 in epoch.

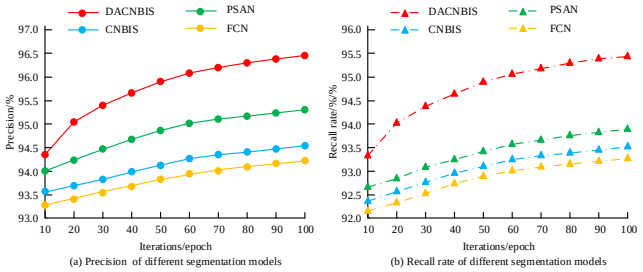


Fig. 8. Accuracy and recall results of four models.

The change curves indicated in Fig. 8 illustrate the precision rate and recall rate of the four models. As the number of iterations increases, the precision rate and recall rate of the four models improve. In Fig. 8(a), the accuracy rate for the four models is displayed, showing that the DACNBIS model has a precision rate of 96.48%. The precision rates for the PSAN, CNBIS, and FCN models are 94.87%, 94.23%, and 93.75%, respectively, at 100 epochs. Fig. 8(b) shows the recall variation curves of the four models, and the recall rates of DACNBIS, PSAN, CNBIS, and FCN models are 95.24%, 93.52%, 93.17%, and 92.91%, respectively, when the number of iterations is 100 epoch.

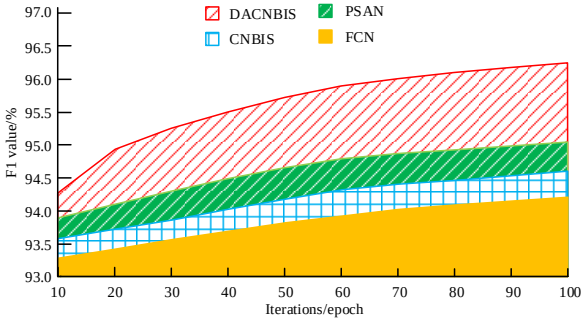


Fig. 9. F1 value results of four models.

Fig. 9 shows the results of the F1 values of the four models, which all increase with the number of iterations. When the number of iterations is 10 epoch, the F1 values of the DACNBIS, PSAN, CNBIS, and FCN models are 94.30%, 93.81%, 93.54%, and 93.26%, respectively. When the number of iterations is 100epoch, the F1 value of DACNBIS model is 95.96%, surpassing the 94.58% of the PSAN model, the 93.99% of the CNBIS model, and the 93.71% of the FCN model. In summary, the DACNBIS model proposed in the study has higher segmentation accuracy compared to other models and performs well in the field of split-screen image stitching localization in film and TV production.

C. Performance Analysis of AM-DACNBIS Model

In order to verify the validity of AM-DACNBIS model, literature [14], DACNBIS model and AM-DACNBIS model were set up for comparative experiments.

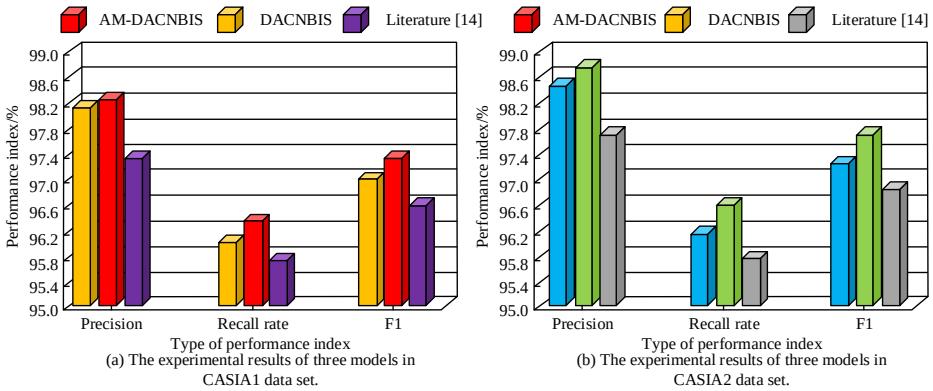


Fig. 10. The experimental results of three models in CASIA1 and CASIA2 data sets.

Fig. 10 shows the experimental results of three models in CASIA1 and CASIA2 data sets. Fig. 10(a) shows the results of accuracy, recall and F1 value of the three models on the CASIA1 data set. The accuracy of AM-DACNBIS model is 98.19%, and that of study [14] and DACNBIS model is 97.21% and 97.97% respectively. The recall rates of AM-DACNBIS, study [14] and DACNBIS models are 96.23%, 95.68% and 95.87% respectively, and the F1 value of the AM-DACNBIS model is 97.21%, which is higher than that of study [14] and DACNBIS model, which is 96.57% and 96.92%. Fig. 10(b) shows the results of accuracy, recall and F1 value of the three models in CASIA2 data set. The accuracy of AM-DACNBIS model is 98.76%, and that of study [14] and the DACNBIS model is 97.59% and 98.45% respectively. The recall rates for the AM-DACNBIS, study [14] and DACNBIS models are 96.54%, 95.51% and 96.01% respectively. The F1 value for the AM-DACNBIS model is 97.65%, which is higher than that of study [14] and the DACNBIS model, which are 96.73% and 97.23%.

Fig. 11 shows the detection results of the DACNBIS model and AM-DACNBIS model on two data sets; Fig. 11(a) displays the outcomes for both models in the CASIA1 dataset. The detection results for both models in the CASIA2 dataset are shown in Fig. 11(b). The figure shows that the AM-DACNBIS model has better performance in detection and can significantly amplify the sensitivity of key features, and can filter out key feature information. To further verify the robustness of the AM-DACNBIS model, the study performs two operations of compression and Gaussian blurring on the CASIA1 test set images, where the image compression factors are 95, 90, 80, and 70, and the standard deviations of Gaussian blurring are set to 0.5, 1.0, 1.5, and 2.0. The results of the model tests are shown in Table II.

Table II shows that the accuracy, recall, and F1 values of both models decrease as the picture compression factor drops and the Gaussian fuzzy standard deviation rises. The AM-DACNBIS model's accuracy, recall, and F1 values are 76.58%, 59.63%, and 62.99%, respectively, when the image compression factor is 70, while those of the DACNBIS model are, respectively, 70.28%, 40.21%, and 50.31%. When the Gaussian fuzzy standard deviation is 2.0, the accuracy rate of AM-DACNBIS and DACNBIS models are 85.07% and 78.74%, the recall rates of AM-DACNBIS and DACNBIS models are 85.41% and 74.31%, and the F1 values of AM-DACNBIS and DACNBIS models are 85.21% and 76.64%. The combined results show that AM-DACNBIS can effectively reduce the risk of overfitting and is more suitable for split-screen image conception in the film and television production than the DACNBIS model.

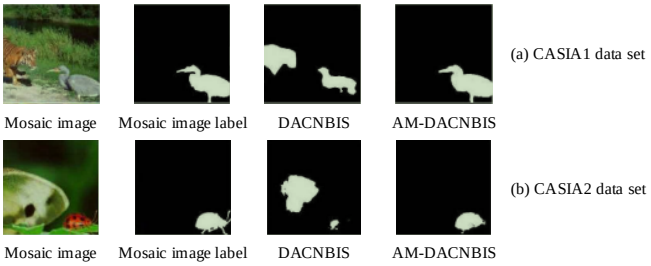


Fig. 11. Test results of two models in CASIA1 and CASIA2 data sets.

TABLE II. DETECTION OF TWO MODELS UNDER DIFFERENT COMPRESSION FACTORS AND GAUSSIAN FUZZY STANDARD DEVIATION ATTACKS

Model Type		**AM-DACNBIS**			**DACNBIS**		
Performance index		***Precision***	***Recall***	***F1***	***Precision***	***Recall***	***F1***
Compressibility factor	95	87.95%	78.27%	82.81%	87.03%	70.54%	77.74%
	90	86.94%	67.61%	75.94%	80.51%	62.41%	70.24%
	80	81.47%	60.41%	64.12%	79.74%	44.09%	55.89%
	70	76.58%	59.63%	62.99%	70.28%	40.21%	50.31%
Gaussian fuzzy standard deviation	0.5	87.41%	85.54%	86.48%	86.31%	78.79%	82.45%
	1.0	87.02%	86.21%	85.84%	84.29%	77.59%	80.81%
	1.5	86.83%	85.77%	85.64%	80.87%	75.47%	78.54%
	2.0	85.07%	85.41%	85.21%	78.74%	74.31%	76.64%

V. CONCLUSION

As image editing software becomes more functional and easier to use, the stitching and positioning techniques for splitting images, a crucial step in pre-production for film and TV, is imperative. The traditional stitching and positioning techniques have limited segmentation accuracy and poor robustness due to inadequate feature extraction capabilities in the sampling process and inaccurate positioning segmentation for different shapes. To address the above problems, a DACNBIS model based on DB module and ASPP module is proposed. The results show that when the number of iterations is 100 epoch, the accuracy, recall and F1 values of DACNBIS model under CUISD dataset are 96.48%, 95.24% and 95.96%, respectively, which are higher than 94.87%, 93.52% and 94.58% of PSAN model. Under the CASIA2 dataset, the AM-DACNBIS model exhibited higher accuracy, recall, and F1 values than the DACNBIS model with scores of 98.76%, 96.54%, and 97.65%, respectively, which were improved by 0.31%, 0.53%, and 0.42%. Under the CASIA2 dataset, the F1 values of the AM-DACNBIS model are 62.99% and 85.21% when the image compression factor and Gaussian fuzzy standard deviation are 70 and 2.0, respectively, which are higher than those of the DACNBIS model by 50.31% and 76.64%. In summary, the DACNBIS model proposed in the study performs well with the AM-DACNBIS model, but in the CASIA1 and CASIA2 datasets, the AM-DACNBIS model performs significantly better than the DACNBIS model and is more suitable for the image splitting conceptualization applications in film and television production. However, there are still shortcomings in the study, and the image quality of the CUISD, CASIA1, and CASIA2 datasets is somewhat different from the demand of multi-scope image stitching, and the subsequent study will further construct a high-quality and complex professional stitching dataset.

REFERENCES

[1] Yang Y, Song X. Research on face intelligent perception technology integrating deep learning under different illumination intensities. Journal of Computational and Cognitive Engineering, 2022, 1(1):32-36.

[2] Kiani F, Nematzadeh S, Anka F A, Findikli M A. Chaotic Sand Cat Swarm Optimization. Mathematics, 2023, 11(10): 2340.

[3] Li J, Wang J. Digital animation multimedia information synthesis based on mixed reality framework with specialized analysis on speech data. International Journal of Speech Technology, 2023, 26(1):63-76.

[4] Tyagi S, Yadav D. A detailed analysis of image and video forgery detection techniques. The Visual Computer, 2023, 39(3):813-833.

[5] Olson E O. SCANNER IMAGING OF COMMON SHARP-TAILED SNAKES (CONTIA TENUIS) FOR INDIVIDUAL IDENTIFICATION. Northwestern Naturalist, 2023, 104(1):26-36.

[6] Tang Z, Ma G, Lu J, Wang Z, Fu B, Wang Y. Sonar image mosaic based on a new feature matching method. IET Image Processing, 2020, 14(10):2149-2155.

[7] Su Z, Li W, Ma Z, Gao R. An improved U-Net method for the semantic segmentation of remote sensing images. Applied Intelligence, 2022, 52(3): 3276-3288.

[8] Gao Y, Che X, Liu Q, Bie M, Xu H. SFSM: sensitive feature selection module for image semantic segmentation. Multimedia Tools and Applications, 2023, 82(9):13905-13927.

[9] Zhou E, Xu X, Xu B, Wu H. An enhancement model based on dense atrous and inception convolution for image semantic segmentation. Applied Intelligence, 2023, 53(5):5519-5531.

[10] Zhang J, Liu Y, Guo C, Zhan J. Optimized segmentation with image inpainting for semantic mapping in dynamic scenes. Applied Intelligence, 2023, 53(2):2173-2188.

[11] Maurya A, Chand S. Cross-form efficient attention pyramidal network for semantic image segmentation. AI Communications, 2022, 35(3):225-242.

[12] Chen Y. Semantic Image Segmentation with Feature Fusion Based on Laplacian Pyramid. Neural Processing Letters, 2022, 54(5):4153-4170.

[13] Teki S M, Varma M K, Yadav A K. Brain Tumour Segmentation Using U-net Based Adversarial Networks. Traitement du Signal, 2019, 36(4):353-359.

[14] Singh N J, Nongmeikapam K. Semantic segmentation of satellite images using deep-unet. Arabian Journal for Science and Engineering, 2023, 48(2):1193-1205.

[15] Tiwari T, Saraswat M. A new modified-unet deep learning model for semantic segmentation. Multimedia Tools and Applications, 2023, 82(3):3605-3625.

[16] Cheng L, Yi J, Chen A, Zhang Y. Fabric defect detection based on separate convolutional UNet. Multimedia Tools and Applications, 2023, 82(2):3101-3122.

[17] Mahmoud A S, Mohamed S A, El-Khoriby R A, AbdelSalam H M, El-Khodary I A. Oil Spill Identification based on Dual Attention UNet Model Using Synthetic Aperture Radar Images. Journal of the Indian Society of Remote Sensing, 2023, 51(1):121-133.

[18] Abdelrauof D, Essam M, Elattar M. Light-Weight Localization and Scale-Independent Multi-gate UNET Segmentation of Left and Right Ventricles in MRI Images. Cardiovascular Engineering and Technology, 2022, 13(3):393-406.

[19] Zhang Y, Lambert M, Fraysse A, Lesselier D. Unrolled convolutional neural network for full-wave inverse scattering. IEEE Transactions on Antennas and Propagation, 2022, 71(1):947-956.

[20] Cong S, Zhou Y. A review of convolutional neural network architectures and their optimizations. Artificial Intelligence Review, 2023, 56(3):1905-1969.

[21] Nematzadeh S, Kiani F, Torkamanian-Afshar M, Aydin N. Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases. Computational biology and chemistry, 2022, 97: 107619.

[22] Peng D, Yu X, Peng W, Lu J. DGFAU-Net: Global feature attention upsampling network for medical image segmentation. Neural Computing and Applications, 2021, 33(18):12023-12037.