

# Application of Three-dimensional Coding Network in Screening and Diagnosis of Cervical Precancerous Lesions

Yiming Pan<sup>1,\*</sup>, Zepeng Shen<sup>2</sup>, Yuhang He<sup>3</sup>, Kuo Wei<sup>4</sup>, Yuchen Zhang<sup>5</sup>

<sup>1</sup> Computer science, individual contributor, Austin, TX, USA

<sup>2</sup> Network Engineering, Shaanxi University of Technology, Shaanxi 723001, China

<sup>3</sup> Computer Science and Technology, Tianjin University of Technology, Tianjin, China

<sup>4</sup> Computer Science, Individual Contributor, Shenzhen, China

<sup>5</sup> Biostatistics, Columbia University, NY, USA

\* Corresponding author: Yiming Pan

**Abstract:** Globally, the incidence of cervical cancer ranks fourth among female malignant tumors, seriously threatening the physical and mental health of women. Early detection and early treatment can greatly reduce the mortality of cervical cancer "cytology /HPV test, colposcopy and cervical biopsy" is the main method for clinical diagnosis of cervical cancer. The progress of medical technology has significantly improved the early diagnosis of cervical cancer, but due to various factors, there are still many cases of missed diagnosis and misdiagnosis. In recent years, artificial intelligence has developed rapidly in the medical field, and has also shown good applicability in the screening and diagnosis of cervical cancer. In this paper, we design an mcs-SEM structure that contains both channels and space compression excitation modules, which can retain comprehensive spatial information at a low computational cost. Then, the structure is embedded in a three-dimensional coding network to realize the combination of two-dimensional convolutional neural network and three-dimensional spatial information, so as to predict the DVH in the complex distribution scenario of multi-endangered organs with higher accuracy.

**Keywords:** Cervical Cancer; Deep Learning; Screening; Three-dimensional Coding Network.

## 1. Introduction

The application of artificial intelligence assisted diagnosis system has improved the diagnosis accuracy of cervical cancer related diseases, reduced the impact of subjective factors and the missed diagnosis rate of positive cases, helped clinicians save time and energy in practical work, and is expected to solve the problem of unequal distribution of medical resources in China. Based on domestic and foreign research results, this paper will discuss the application of artificial intelligence in cervical cancer screening and diagnosis from the four aspects of HPV detection, cytology, colposcopy, and imaging examination, and propose the challenges faced by artificial intelligence in cervical cancer screening and diagnosis in the future [1][2][3].

Intensity modulated radiotherapy (IMRT) is a very accurate radiotherapy technique, which can achieve accurate dose irradiation of target area by adjusting the intensity of the field, so it is widely used in the treatment of cervical cancer. But its formulation is a complex process that requires precisely delivering high doses of radiation in three dimensions while minimizing damage to surrounding healthy tissue. Based on this, we propose the dose volume histogram as an important evaluation tool, which can provide radiation treatment ratio information, so that physicians can quickly develop high quality radiotherapy plan[4]. However, the traditional DVH prediction model can only describe the spatial geometric relationship between a single organ at risk and the target area, and it is difficult to extract comprehensive spatial geometric information of patients, thus affecting the accuracy of prediction results. In 2019, Dan Nguyen et al. proposed that 3D-Unet convolutional neural networks could be used to

achieve voxel-level dose prediction. However, it can only predict the DVH curve of one organ at risk, which is not efficient [5].

In recent years, artificial intelligence (AI), relying on its strong learning ability, has been gradually applied to the field of clinical medicine, especially in the field of disease diagnosis, AI has shown strong potential. At present, a number of studies have applied AI to the screening and diagnosis of cervical cancer and precancerous lesions, aiming to improve the accuracy of disease diagnosis, reduce the rate of missed diagnosis, and save more time and precision for clinicians

Force. This paper will discuss the application of AI in cervical cancer screening and diagnosis from four aspects: HPV detection, cytology, colposcopy, and imaging.

## 2. Related Work

Among women in developing countries, cervical cancer is the second most common cancer (15.7 per 100,000 women) and the third most common cause of cancer death (8.3 per 100,000 women). Brachytherapy (BT) is a necessary part of local advanced cervical cancer treatment, in the cervical lesions of higher doses of radiation at the same time, will also cause some damage to the surrounding normal tissue. Compared to conventional brachytherapy, image-guided brachytherapy (IGBT) can increase the dose of the target area more effectively, while accurately assessing the dose received by the surrounding tissue [6-8]. Several observational studies suggest that IGBT is superior to conventional brachytherapy in terms of disease control rate, overall survival, and complication control. Although IGBTs based on Magnetic resonance imaging (MRI) are the gold standard for cervical

cancer brachytherapy, CT images provide information on the location of the applicator and can be used to map the Organ at risk (OAR). It can replace MRI to some extent. The deformation caused by the implantation of the vaginal applicator makes it impossible to register with pre-treatment MRI. In addition, MRI-based IGBTs have obstacles such as limited application range, high cost and long scanning time, resulting in wider popularity of CT simulators and easier implementation of CT-based IGBTs. The extent of tumor invasion can be determined by gynecologic MRI examination before or during treatment, which makes CT-based IGBT more accurate and can be used as an alternative to MRI-based IGBT[9].

In recent years, one of deep learning algorithm of convolution neural network (convolutional neural network, CNN) has been successfully applied to the automatic segmentation of medical image. In the field of radiation therapy, various CNN models have achieved automatic mapping of CTV or OARs for brain tumors 211, head and neck tumors 125.2, esophageal cancer 127, rectal cancer 28.29, etc. Although some studies have explored the performance of the CNN model for auto-mapping in locally advanced cervical cancer, most of them focused on external exposure. Three-dimensional Stereoscopic Encoding Network (TSEN) model, which combines three-dimensional spatial information with two-dimensional convolutional

neural networks, from

It can effectively predict DVH of all organs at risk in cervical cancer radiotherapy program[10-12]. In this model, in order to reduce computing costs and improve network performance, We propose a multi-scale channel and spatial Squeeze & Excitation Module (mcs-SEM) based on expansive convolution and integrate it into the network architecture. Through this network model, the original image information can be used to predict the DVH of multiple organs at risk in new patients.

### 3. Methodology

In this paper, it is proposed that the backbone network of TSEN is DenseNet, and its basic module is Dense Block, which is composed of multiple convolutional layers, and the input of each convolutional layer is the output of all the previous layers, so that the output of dense block contains all the input feature information, which is conducive to feature reuse and enhancement.

In the dense block, in order to reduce the size of the feature map, 1x1 convolution layer is generally used to reduce the dimension, and then 3x3 convolution layer is used to extract the feature, whose output can be the subsequent dense block or the whole network output.

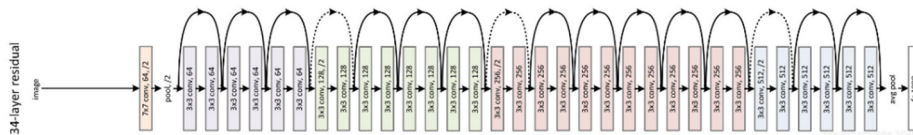


Figure 1. ResU-Net

The common algorithm in the CNN model is as follows:

volume layer (Convolution) : It is used for the extraction of line features, that is, the extraction of line features in a region on the input data. For example, when we input an image of size 3 2 x 3 2 x 3 (3 is the depth of the image, i.e. R G B color), by constructing a 5 x 5 x 3 receptive field (F i l t e r), We can get a characteristic plot of 2 8 x 2 8 x 1. Each volume layer generally has multiple different receptive fields to extract different characteristics. By means of multiple volumes, the characteristic maps of deeper layers can be obtained.

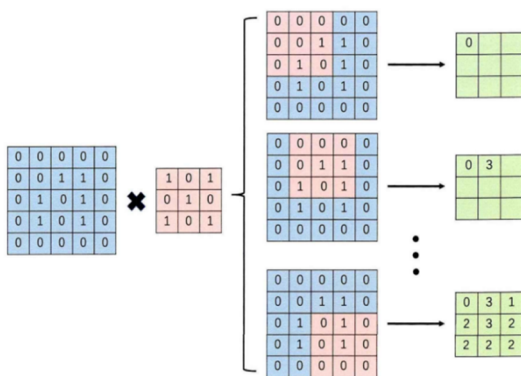


Figure 2. Convolution operation

Pooling: It is used to compress the feature map, on the one hand, to make the feature map smaller and reduce the calculation amount, on the other hand, to compress the feature. pooling operations mainly include Max pooling and Average pooling. The process of the former is shown in Figure 3. Using a 2x2 receptive field, we compress the data to a quarter

of its original size. This method of shrinking the feature map has the potential to affect the accuracy of the model, which can be compensated for by increasing the depth of the feature map. Generally, the eigenvalues of edges are large, so we use maximum pooling in ResU-Net.

#### 3.1. Model Architectures

Each Dense module contains two 3x3 convolution. For all convolution layers, we set the filling parameter to 1, which guarantees that each convolution layer in the module can eventually produce a feature map of the same size.

Thus, the  $i$ th layer output in the module can be expressed as the formula:

$$x_i = O_i([x_0, x_1, \dots, x_{i-1}]), i = 1, 2, 3, 4 \quad (1)$$

Where  $O_i(\cdot)$  refers to a nonlinear transformation that increases the complexity of the model. Specifically, after the batch normalization layer, an excitation with a parametric rectified linear unit as the activation function is connected. The live layer is finally connected to a 1x1 volume layer.  $[x_0, x_1, \dots, x_{i-1}]$  refers to the mapping feature output at layer 0, ...,  $i-1$ , and sets the DenseNet backbone network growth rate to 10. In this study, DenseNet is used as a model to set special hyperparameters. Its dense network can well solve the overfitting problem caused by training on relatively few medical data sets, which also shows that the network has excellent generalization ability. The original U-Net could only extract relatively low-level features, such as edge and intensity contrast. This problem can be solved by increasing the depth of the network, but the increase of the depth of the network brings two problems: first, the gradient disappears,

which makes it difficult to converge; The second is degradation, that is, the appropriate depth continues to increase the number of layers, and the more errors generated during training, the accuracy of the model will decline. To this end, we joined Residual Block to design a new ResU-Net, which has more branches connecting the input directly to the following layer than the ordinary directly connected convolutional neural network.

### 3.2. DataSet Introduction

#### 3.2.1. Space Squeeze&Excitation,s-SE

The space Squeeze&Excitation (s-SE) module was originally proposed in the 2017 ISVRC competition. It can effectively learn the interdependencies between similar features and use these dependencies to strengthen the representational ability of the network. Specifically, the s-SE module reduces the number of channels through a space compression, then uses a fully connected layer to learn the importance weight of each channel, and finally applies the weight to the compressed channel to obtain a weighted feature representation. In this way, the network can actively focus on the important feature channels, and ignore the unimportant feature channels to improve the accuracy and generalization of the network. The s-SE module has been widely used in various computer vision tasks, and has achieved good results.

In each module of the s-SE module structure, when we extract the spatial geometric features of a set of CT images, we can follow the following steps: First, the channel information should be compressed to reduce the number of channels and retain the most important information. Second, the incentive operation of the spatial part is necessary. In order to obtain the index of the importance of spatial position, Excitation operations can be used to activate each spatial position in the feature map. Then, it is feature extraction. The CT images after compression and excitation processing are input into five Dense modules to generate feature maps.  $M =$

$$\begin{bmatrix} m^{1,1} & m^{1,2} & \dots & m^{u,v} & m^{L,N} \end{bmatrix}. \quad (2)$$

Here,  $L \times N$  is the size of the feature map, and  $u$  and  $v$  are the spatial position coordinates of the feature map. Finally, space compression. In order to reduce the dimension of feature map,  $1 \times 1$  convolution is used for space compression  $Nsp$ , and the output feature  $p$  is obtained.

$$p = Nsp \times M \quad (3)$$

In the above formula,  $Nsp \in \mathbb{R}^{1 \times 1 \times c}$ . For  $p$ , we finally run it through an activation function to normalize its result to between 0 and 1.

$$M_{s-SE} = [\sigma(p_{1,1})m^{1,1}, \dots, \sigma(p_{L,N})m^{L,N}] \quad (4)$$

In the above formula, the magnitude of  $\sigma(p_{u,v})$  represents the importance weight of the spatial position coordinates  $u$  and  $v$ , which are important parameters for the model. When the training process continues to iterate, the module will gradually ignore the non-important positions and actively focus on the important spatial positions. This is how the s-SE module works

### 3.3. Data Analysis and Preprocessing

The TSEN structure consists of three component modules:

1. DenseNet module of intensive convolutional neural network. As the backbone network of TSEN, it can extract three-dimensional spatial features from three different types

of images and retain them, as shown in the figure below

Main spatial information.

2. Translation layer module. After the module is set as a batch normalized layer in the TSEN structure, a convolution filter with a size of  $1 \times 1$  is connected for convolution, and then an average pooling layer is connected for dimensional compression, which can ensure that the final output of each Dense layer has the same size feature map.

3.mcs-SEM module. The module integrates and calibrates the three-dimensional spatial information extracted by the convolutional filter, which can not only retain comprehensive spatial information at a low computational cost, but also effectively extract the contour features of different shapes of endangered organs.

We used RT-Structure data to obtain the 3D target contour and 3D scanning map of the contours of the organs at risk, and then calculated the 3D distance image, and then extracted these images together through the feature extractor, converted into feature vectors and sent to the TSEN network. The spatial features were extracted by convolutional neural network, and then the feature fusion was calibrated by mcs-SEM. Finally, the predictive DVH feature vectors of multiple organs at risk were generated by the output module of TSEN backbone network.

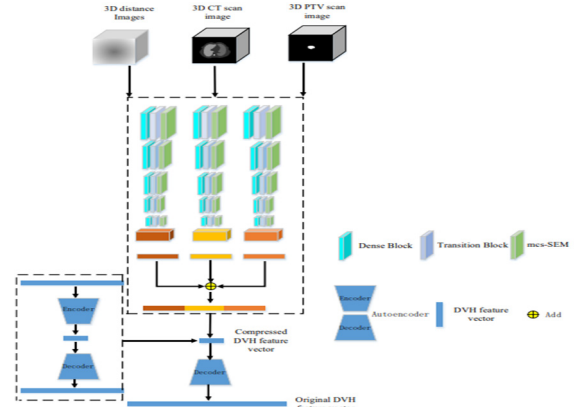


Figure 3. TSEN structure

## 4. Conclusion

The main content of this study is to develop a DVH prediction model based on TSEN structure for the dosimetric evaluation of cervical cancer radiotherapy program. The traditional DVH prediction method is based on the artificial design of geometric operators, this method has great limitations in dealing with the complex three-dimensional relationship of multiple organs at risk. Therefore, this paper designed an mcs-SEM structure, which includes both channels and space compression excitation modules, which can retain comprehensive spatial information at a low computational cost, and use distance images to characterize spatial geometric relationships of multi-endangered organs, directly using three-dimensional image data as the original input of the network. Thus, the limitation brought by manual design of geometric operators is avoided. In conclusion, this study provides a new method for dosimetric evaluation of cervical cancer radiation therapy plans, which can not only better characterize the complex three-dimensional relationship of multiple organs at risk, but also predict DVH more accurately, helping radiologists to evaluate the quality of cervical cancer radiation therapy plans.

In summary, the application of AI in cervical cancer screening and diagnosis can improve the specificity and

sensitivity of cervical cancer screening and diagnosis. In the first stage of screening, AI-assisted HPV testing and cytological examination of cervical detachment can improve diagnostic accuracy and reduce unnecessary colposcopic triage. AI-assisted cytology diagnosis is not limited by time, and can be used to screen a large number of normal cervical smears. According to statistics, cervical cytology results are negative, accounting for more than 99.5%; At the same time, avoid the influence of subjective factors, reduce missed diagnosis and misdiagnosis; It is expected to solve the problem of the lack of medical resources in remote areas and realize the network remote assisted diagnosis. AI-assisted colposcopy can also avoid the limitations of subjective factors, classify colposcopic pictures and mark cervical lesions, guide colposcopists to perform biopsies, and reduce missed diagnosis. In imaging examination, AI model not only has the function of assisting screening, but also can predict vascular metastasis, treatment response and recurrence, and assist clinicians to adjust treatment plans.

The limitations of this study are as follows: First, it is a single-center study with a small sample size. Due to differences in sketching principles, CT scanning modes, and emitter placement methods, the generated AI model may not meet the sketching standards and habits of other centers, resulting in poor generality. In the future, multi-center cooperation is needed to jointly determine the delineation criteria, and train the model with a larger multi-center data set to make it more versatile. Second, the clinical score is a subjective judgment of the evaluator, and although most of the sketched results can be used clinically, they still need to be considered in the context of the actual dose distribution and normal tissue toxicity. Finally, the amount of data used to evaluate WDSC is small, and although WDSC and AD2cc are significantly correlated, larger sample size validation is needed. Finally, this study did not use the ResU-Net model trained by Fletcher emitter IGBT to calculate WDSC, so it is not possible to verify the correlation between WDSC and D2cc differences.

## Acknowledgments

Image recognition technology holds crucial application value in the field of mechanical industry, such as in the robotic arms on automobile production lines. We appreciate the specialized knowledge provided by Bo Liu and Chang Che in mechanical and computer vision-related areas. Their research in the detection of defects in industrial castings and vehicles' components using artificial intelligence provides us with significant guidance in this field.

## References

- [1] Yanagawa MEA. Diagnostic performance for pulmonary adenocarcinoma on CT: comparison of radiologists with and without three dimensional convolutional neural network. *Eur Radiol*, 2020: 1-9.
- [2] Jiang M, Zhang D, Tang S C, et al. Deep learning with convolutional neural network in the assessment of breast cancer molecular subtypes based on US images: a multicenter retrospective study. *Eur Radiol*, 2021, 31(6): 3673-3682.
- [3] Takahashi Y, Sone K, Noda K, et al. Automated system for diagnosing endometrial cancer by adopting deep learning technology in hysteroscopy. *PLoS One*, 2021, 16(3): 6248526.
- [4] Curry SJ, Krist AH, Owens DK, et al. Screening for cervical cancer: US preventive services task force recommendation statement. *JAMA* 2018, 320(7): 674-686.
- [5] Pathania D, Landeros C, Rohrer L, et al. Point of care cervical cancer screening using deep learning-based microholography. *Theranostics*, 2019, 9(26): 8438-8447.
- [6] Wong OCW, Ng IFY, Tsun OKL, et al. Machine learning interpretation of extended human papillomavirus genotyping by on clarity in an Asian cervical cancer screening population. *J Clin Microbiol*, 2019, 57(12): e00997-19.
- [7] Ai W, Wu C, Jia L, et al. Deep sequencing of HPV16 E6 region reveals unique mutation pattern of HPV16 and predicts cervical cancer. *Microbiol Spectr*, 2022, 10(4): e0140122.
- [8] Zhang H, Zhang T, You Z, et al. Positive surgical margin, HPV persistence, and expression of both TPX2 and PDL1 are associated with persistence/recurrence of cervical intraepithelial neoplasia after cervical conization. *PLoS One*, 2015, 10(12): e142868.
- [9] Chang Che, Bo Liu, Shulin Li, Jiabin Huang, and Hao Hu. Deep learning for precise robot position prediction in logistics. *Journal of Theory and Practice of Engineering Science*, 3(10):36-41, 2023. DOI: 10.1021/acs.jctc.3c00031.
- [10] Bogani C, Ditto A, Martinelli F, et al. Artificial intelligence estimates the impact of human papillomavirus types in influencing the risk of cervical dysplasia recurrence: progress toward a more personalized approach. *Eur J Cancer Prev*, 2019, 28(2): 81-86.
- [11] Hao Hu, Shulin Li, Jiabin Huang, Bo Liu, and Chang Che. Casting product image data for quality inspection with exception and data augmentation. *Journal of Theory and Practice of Engineering Science*, 3(10):42-46, 2023. [https://doi.org/10.53469/jtapes.2023.03\(10\).06](https://doi.org/10.53469/jtapes.2023.03(10).06).
- [12] Tian R, Cui Z, He D, et al. Risk stratification of cervical lesions using capture sequencing and machine learning method based on HPV and human integrated genomic. *PLoS One*, 2017, 12(12): e1778.
- [13] Z.W. Zhang, J.N. Wang: Crane Design Manual (China Railway Press, China 1998), p.683-685. (In Chinese).
- [14] Tianbo, Song, Hu Weijun, Cai Jiangfeng, Liu Weijia, Yuan Quan, and He Kun. "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition." In 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 834-837. IEEE, 2023. DOI: 10.1109/mce.2022.3206678.
- [15] C. Li, W.Q. Yin, X.B. Feng, et al. Brushless DC motor stepless speed regulation system based on fuzzy adaptive PI controller. *Journal of Mechanical & Electrical Engineering*, vol. 29 (2012), 49-52.
- [16] China National Standardization Management Committee. Specifications of Crane Design (China Standardization Press, China 2008), p. 16-19.
- [17] Lin, Q., Che, C., Hu, H., Zhao, X., & Li, S. (2023). A Comprehensive Study on Early Alzheimer's Disease Detection through Advanced Machine Learning Techniques on MRI Data. *Academic Journal of Science and Technology*, 8(1), 281-285. DOI: 10.1111/jgs.18617.
- [18] Che, C., Hu, H., Zhao, X., Li, S., & Lin, Q. (2023). Advancing Cancer Document Classification with Random Forest. *Academic Journal of Science and Technology*, 8(1), 278-280. <https://doi.org/10.54097/ajst.v8i1.14333>.
- [19] Chandran V, Sumithra MC, Karthick A, et al. Diagnosis of cervical cancer based on ensemble deep learning network images. *Biomed Res Int*, 2021, using colposcopy 2021: 5584004.
- [20] Chang Che, Bo Liu, Shulin Li, Jiabin Huang, and Hao Hu. Deep learning for precise robot position prediction in logistics. *Journal of Theory and Practice of Engineering Science*, 3(10): 36-41, 2023. DOI: 10.1021/acs.jctc.3c00031.