Detection of Specular Reflection on the Smart Colposcopy Images Using Fine-Tuned U-Net Model

Jennyfer Susan M.B.,[a] *and Subashini P.*[b]

[a]Avinashilingam Institute for Home Science and Higher Education for Women, India.
[b]Avinashilingam Institute for Home Science and Higher Education for Women, India
E-mail: [a]jennyfersusan26@avinuty.ac.in, [b]Subashini_cs@avinuty.ac.in

Abstract

Cervical cancer is the most prevalent genital malignant tumor, causing a serious health risk to women. Smart colposcopy is the screening procedure used to identify the early stages of cervical lesions. Due to physiological mucus in the human body, the bright luminous is generated and appears in the high-definition image produced by smart colposcopy. The highlighted area in the colposcopy images resembles similar to the metaplasia epithelium and causes difficulty during visual analysis. This paper presents a deep learning strategy for identification of the luminous spots in colposcopy images using the finetuned U-Net model with customized hyperparameters. The specular reflection region is labeled based on the intensity range and trained with the U-Net model to distinguish the specular region from the non-specular region on the smart colposcopy images. With a loss value of 0.8234, the automated technique correctly recognizes the specular reflection with an accuracy of 97.28%.

Keywords: Smart Colposcopy Images, Specular Reflection, U-Net Model

Introduction

Cervical cancer is a highly prevalent malignancy that affects women around the world. In the year 2020, it is estimated that 604,000 cases are newly reported, and 342,000 deaths are reported by the World health organization (WHO). Due to inadequate knowledge of cervical cancer, persons in rural areas are more likely to be affected by cervical cancer. Initially for screening the traditional pap smear method is used for screening. Due to the difficulty in providing the laboratory structure in developing countries very few screening programs are provided to the rural and tribal people. To provide simple effective screening smart colposcopy is used to identify the abnormalities in the cervical region. Smart colposcopy is the device used to capture high-quality images and videos for the visual inspection of cervical images. The physician analyzes the color density of the metaplasia epithelium to detect neoplasm and helps analyze the severity of lesion on the cervical images. During the analysis, the specular reflection appears on the cervical images affecting the diagnosis process and significantly decreasing the performance of the smart colposcopy images. Various image processing techniques are applied to detect the specular reflection on the cervical image, but few are automated. Meslouhi et al. (2011) proposed the identification of the reflection region on the colposcopy images by using the chromatic properties of the reflection area which is formulated from the properties of the dichromatic reflection model approach. After identifying the inpainting method is applied to the detected region to remove the glare region on the cervical images. Wei et al. (2019) modified the traditional U-Net segmentation model for the automatic segmentation of seabed mineral images. The model is applied to the grayscale microscopy and seabed mineral images dataset. The experimental method achieves 0.916 for the EM dataset and provides higher quality in the segmentation of the mineral images. Attard et al. (2020) proposed the semantic segmentation deep learning model for the localization of the specular highlights. The proposed method is trained and tested with the reflected affected image dataset. The model achieved a frequency weight of 0.98 and a mean iou of 0.80. Liu et al. (2021) suggested a novel deep learning framework to fix the highlighted region on the endoscopy images using a surgical fix deep neural network (SFDNN). Reflective region and the anti-reflection regions are labeled before training through the deep learning model. It identifies the defects and is further used for refilling the missing portion of the detected region. Jus et al. (2018) presented the U-Net deep learning model for the segmentation of the iris, and a modified U-Net is proposed by Ayalew, et al. (2021) for the segmentation of the ultrasonic lung's images. Based on the analysis of the related paper the U-Net is suitable for the segmentation process in many medical images. The next section discusses the segmentation of the specular reflection on the smart colposcopy images for automatic identification.

So, in this paper, proposed the detection of specular reflection on the smart colposcopy images using the finetuned U-Net model. In section 2 discuss the methodology and in section 3 Results and analysis and finally the conclusion of the paper.

Methodology

To detect the specular reflection on the smart colposcopy images initially the glare regions are labeled using binary masking. The labeled images are trained using a simple convolutional neural network model and Fine-tuned U-Net models for the identification of the glare region of the specular reflection on the smart colposcopy images.

Proposed Binary Masking

Smart colposcopy is a mobile-based digital device for the screening of cervical regions. During the analysis, colposcopy images are affected by the white bright region pixel causing luminosity on the surface of the images. The proposed masking method is formulated from the intensity value of the colposcopy images. The RGB images are initially converted to grayscale images. The intensity range of the pixel ranges from 0 to 255 where 0 represents the darkest pixel value of the image and 255 represents the whitest pixel value of the images. On analyzing the intensity value, the pixel value ranges from 191–255 falls in the specular reflection region and pixels from 0–190 are labeled as the non-specular region. The specular reflection pixel is rescaled and set as 0 and another region is set as 1 for creating masking images on the smart colposcopy images as in equation 1. The *f(x)* represents the rescaled pixel value of the cervical images.

$$f(x) = \begin{cases} 1 = Non\ specular\ Reflection \\ 0 = Specular\ Reflection \end{cases} \quad (1)$$

Network Architecture for the Detection of Specular Reflection:

The rescaled input images and the original images are trained using the simple convolutional neural network and finetune U-Net model for the detection of specular reflection on cervical images.

Simple Convolutional Neural Network

The labeled images and original images are trained using a convolutional neural network (CNN) to predict the specular reflection pixel on the cervical images (Susan et al., 2022). The max-pooling layer reduces the feature map dimension by half of the size of the input image. The maximum value of the convolutional process is the max-pooling of 3x3 with stride values of 1, 2, and 3. The thirteen convolutional layers with max-pooling are applied to each convolutional layer. The average pooling is concatenated to create the original images in the output images. The convolutional transpose is applied along with the convolutional and concatenate layer to segment the desired pixel on the cervical images. The batch normalization layer is used to normalize the value of the cervical images, and the dropout layer is utilized to reduce overfitting of the training model. In the hidden layers, the sigmoid activation function is employed for binary segmentation of pixel-like 1 and 0 from the smart colposcopy images. The workflow for the detection of specular reflection using simple convolutional neural network is shown in Fig. 1. The model constructed is trained and tested for the identification of specular reflection on the colposcopy images.

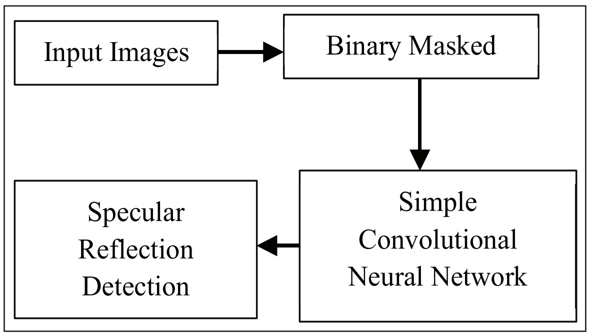

Fig. 1. Flowchart for the Detection of specular Reflection using Simple CNN mode

Fine Tuned U-Net Convolutional Neural Network

Fine Tuned U Net Model

The U-Net model is built for the segmentation of medical images (Ronneberger, et al., 2015). The main advantage of this method is attaining higher segmentation accuracy with minimum trained dataset. It consists of a contracting path also known as the encoder and an expansion path representing the decoder part of the images. The contracting and the expansion section consist of four convolutional blocks and each convolutional block has two convolutional layers. The convolutional layer of the U-Net model is represented in equation.2. Each convolutional kernel is set as (3 × 3) with the max polling layer size of (2x2).

$$F_{i,j} = f\left(\sum_{m=0}^{2}\sum_{m=0}^{2} W_{m,n} I_{x+m,y+n} + W_b\right) \quad (2)$$

Where $I_{x,\ y}$ represent the row and column of the images, $W_{m,\ n}$ represent the weight of the (*m*, *n*), and *F(i,j)* represents the feature map extracted from the smart colposcopy image. The rectified linear unit (ReLu) is applied to each layer of the convolutional layer in the contracting path and the batch normalization is enabled as true for each layer. It consists of the transposed convolutional layer in the expansion path. It also consists of the four convolutional layers where the padding is set as the same to get the actual size of the original images. The model is tuned for the detection of smart colposcopy images:

- The stride value of the fine-tuned U-Net model is set as (1 × 1) to consider each pixel of the image.
- The Dropout layer is attached to each block layer of the contraction path to reduce overfitting during the feature extraction process.
- The batch normalization is added in the contraction path to normalize the pixel value and to improvising the computation accuracy of the model.
- The neuron size or filter dimension is set as 8, 16, 32, 64, and 128 for the contraction path and 128, 64, 32, 16, and 8 for the expansion path of the finetuned U-Net model. The finetuned uNet model is shown in Fig. 2.

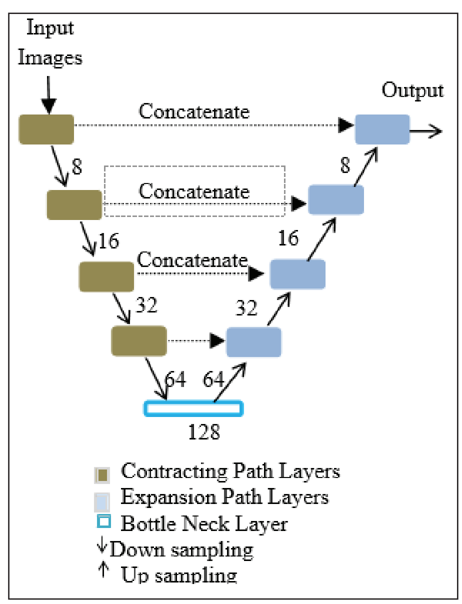

Fig. 2. The Finetune U-Net Architecture

Training and Testing for Simple CNN and Finetuned uNet Model

The input size of the images is resized as 256 × 256. The model is trained with the Adam optimizer for faster computation and the binary cross-entropy loss calculation is used for the binary segmentation of the cervical images. The model is trained with the epoch value of 10 and with a learning rate of le^{-4}. The batch size is set as 100 and trained the model with 1063 images. For the validation 200 images are selected that are not used in the training which is collected from the Kaggle dataset (Kaggle, 2017). The workflow of the fine-tuned U Net model is shown in Fig.3.

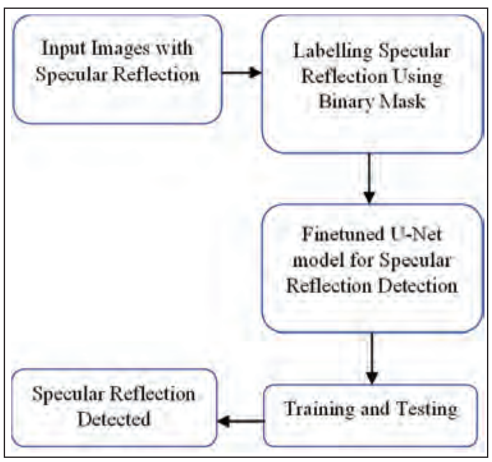

Fig. 3.

Experiment Result and Analysis

Smart colposcopy images are taken from the Kaggle dataset to train the detection of specular reflection on the cervical images. The dataset contains three types of the cervix images such as CIN1, CIN2, and CIN3 based on the severity of cancer affected on the tissue region. CIN1 is the initial stage of cancer in which one-third of the tissue region is affected by the cancer cell and these cancer cells will be higher in CIN 3 which means that the acetowhite region is higher for the CIN3 cervical cancer compared to the CIN1 and CIN2. It will be a challenging task in the CIN3 to exactly extract the specular reflection from the acetowhite region. The binary accuracy, Intersection over Union (IoU), and the dice coefficient are calculated to evaluate the performance of the model on the cervical images.

Qualitative Analysis

The qualitative analysis has been carried out on the cervical images which are collected from the Kaggle dataset.

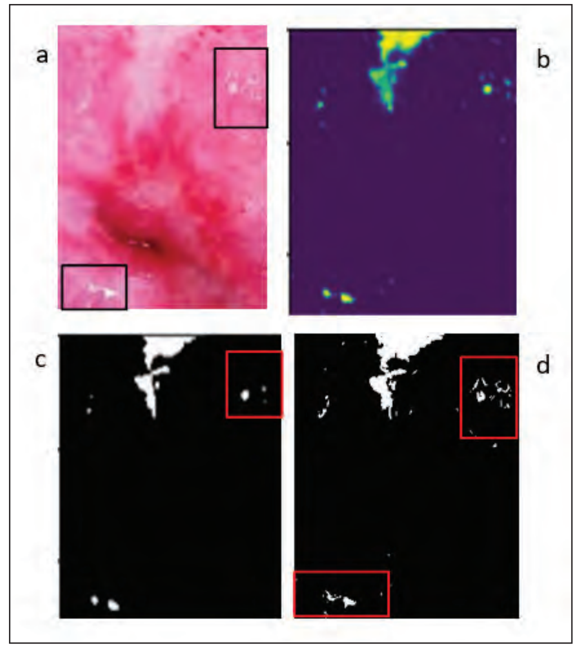

Fig. 4. (a) Original Images (b) Masked Images (c) Predicted Images using Simple CNN method (d) Predicted images using finetuned UNet model

Fig. 4a represents the original image with the specular reflection. The proposed binary masking is applied on the original to label the specular reflection on the cervical as shown in Fig. 4b. The yellow portion on the masked is labeled as the specular region and the other portion is the non-specular reflection region. The masked region was trained using the simple CNN model and fine-tuned U-Net model. Figure 4c detects the specular reflection on the cervical images but some of the small, dotted reflections are not identified as the reflection. Based on the comparison of the fine-tuned U-Net model identifies the small dots of the reflection on the cervical images as shown in Fig. 3d. Based on the qualitative analysis the fine-tuned U-Net model gives the correct prediction of the specular reflection including the small dots of the reflection.

Quantitative Analysis

For the quantitative analysis, the binary accuracy, intersection of union (IoU), Dice-Coefficient, and binary cross entropy loss value are calculated. The binary accuracy predicts the total number of correctly predicted reflection regions divided by the total number of predictions on the smart colposcopy images. The intersection of union (IoU) specifies the amount of overlap that happened between the predicted mask and the original masked images of the smart colposcopy images. The dice coefficient calculates twice the multiple of the area of overlap in the predicted and the original images divided by the total number of pixels in both images.

Table 1. Quantitative Metrics for the detection of specular reflection on the smart colposcopy images

	Binary Accuracy (%)	IoU	Dice-Coefficient	Loss value
Simple CNN[]	0.924	0.915	0.921	1.269
U-Net[]	0.972	0.967	0.977	–0.984
Fine tuned U-Net	**0.989**	**0.965**	**0.981**	**–0.980**

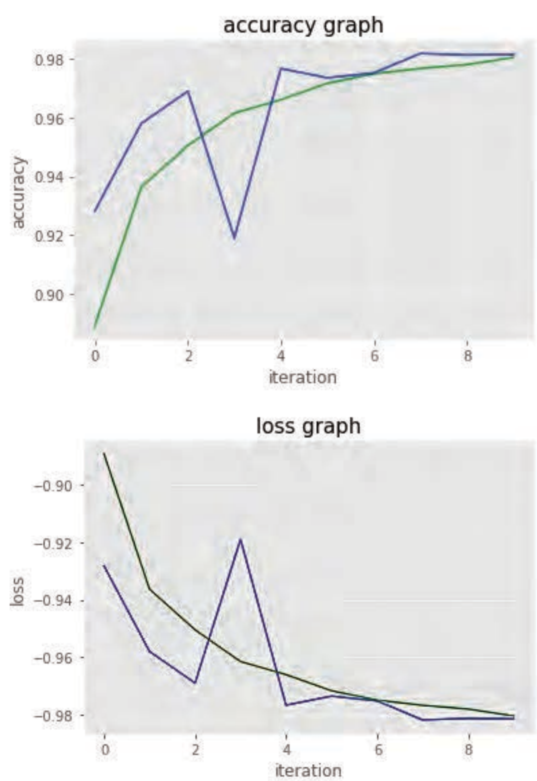

Fig. 5. The Accuracy Graph for the fine Tune Model (a) The Loss Graph of the fine Tune Model(b)

A binary cross entropy loss value is calculated to determine the loss of the pixel in the predicted output. These metrics are calculated for the simple convolutional neural network for the segmentation of the specular reflection. Based on the analysis the Finetune U Net model attains the binary accuracy of 98.96% for 10 epochs as in Table 1. The loss value for each epoch is visualized in Fig. 5b and accuracy value for each epoch is visualized in Fig. 5a. The time and consumption are also important to challenge in medical images. The U-Net model took five hours and 26 minutes with a memory usage of 874 MB to train 1063 cervical images with 10 epochs. The time and memory usage are shown in Table 2. The process is computed using the GPU processor and the time

Table 2. Computational time and memory usage for Simple CNN and U-Net model

Methods	Time	Memory Usage
Simple CNN	2 hrs. 03 mins	741 MB
U-Net	5 hrs. 30 mins	874 MB
Finetuned U-Net	5 hrs. 26 mins	874 MB

Conclusion

The reflection identification on the specular reflection is a challenging task and no automated method is applied to the smart colposcopy images. Initially, the simple CNN model for the segmentation is applied but for this model, a greater number of cervical images are required to obtain higher accuracy. The U-Net model advantage is obtaining higher segmentation accuracy with minimum trained images. So, the U-Net model is fine-tuned based on the smart colposcopy images and helps predicting the specular reflection with promising accuracy with few trained images.

Acknowledgments

The author expresses deep gratitude to the "Centre for Machine Learning and Intelligence" for supporting and sharing the resource for this research work. It is supported by the Department of Science and Technology under the scheme of DST CURIE (AI) sanctioned for the "Core Research grant for Artificial Intelligence (AI)," in the period 2021–2023.

References

Attard, L., Debono, C. J., Valentino, G., and Castro, M. D. (2020). Specular Highlights Detection Using a U-Net Based Deep Learning Architecture, *Fourth International Conference on Multimedia Computing, Networking and Applications,* Valencia, Spain.

Ayalew, Y. A., Fante, K. A., and Mohammed, M. (2021). Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method, BMC Biomedical Engineering, 3(4).

Jus, L., Meden, B., Struc, V., and Peer, P. (2018). End-to-End Iris Segmentation Using U-Net, *IEEE International Work Conference on Bioinspired Intelligence*, San Carlos, Costa Rica

Kaggle (2017). Intel & MobileODT Cervical Cancer Screening. Retrived from https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening

Liu, T., Wang, C., Chang, J., and Yang, L. (2021). Specular Reflections Detection and Removal Based on Deep Neural Network for Endoscope Images. Seventh International Conference on Image Processing Theory, Tools and Applications, Montreal, Canada.

Meslouhi, O. E., Kardouchi, M., Allali, H., and Gadi, T. (2011). Automatic detection and inpainting of specular reflections for colposcopic images, Central European Journal of Computer Science. 1(3), 341–345.

Ronneberger, O., Fischer, P., Brox, T. (2005). U-Net: Convolutional Networks for Biomedical Image Segmentation, Paper presented at International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany.

Susan, M. B. J., Subashin, P., and Krishnaveni, M. (2022). Comparison of various deep learning inpainting methods in smart colposcopy images. International Journal of Computational Intelligence Studies, 11(1), 53–72.

Wei, S., Zang, S., Liu, X., and Qiu, L. (2019) An Improved U-Net Convolutional Networks for Seabed Mineral Image Segmentation, IEEE Access, 7(1), 82744–82752.