

Research Methodology

Medical image processing and analysis have seen a notable upsurge in attention in the past couple of years. Medical image processing is a powerful tool for identifying reflection glare regions and enhancing the quality of medical images. Much of the existing research has primarily centered on tasks like segmenting cancer regions and grading these images. However, a notable need for more focus has been addressing image quality, especially SR artifacts in colposcopy images. These SR artifacts are often present in the images, causing diagnostic challenges and negatively impacting the accuracy of the grading process. The image quality issue has received limited attention, with most research efforts being dedicated to refining models for image grading and segmentation. Therefore, this session will discuss the methodology employed to improve image quality by identifying and removing glare regions from the images.

3.1. Research Design

Specular reflection is the artifacts that the researchers ignored during the medical image analysis. In order to accomplish the study objectives, the research methodology suggests several of ways for analysing the optimal solutions. The deep learning algorithm and medical image processing techniques are used to meet the desired objectives. The methodology adopted for the proposed research work will be presented in four phases, as listed below:

Phase I: Specular Reflection Identification

Phase II: Segmentation of Specular Reflection

Phase III: Inpainting for Image Enhancement

Phase IV: Enhanced Image Grading

The solution proposed in this research work is efficient and viable for enhancing the quality of smart colposcopy images. Figure 3.1 depicts the research design.

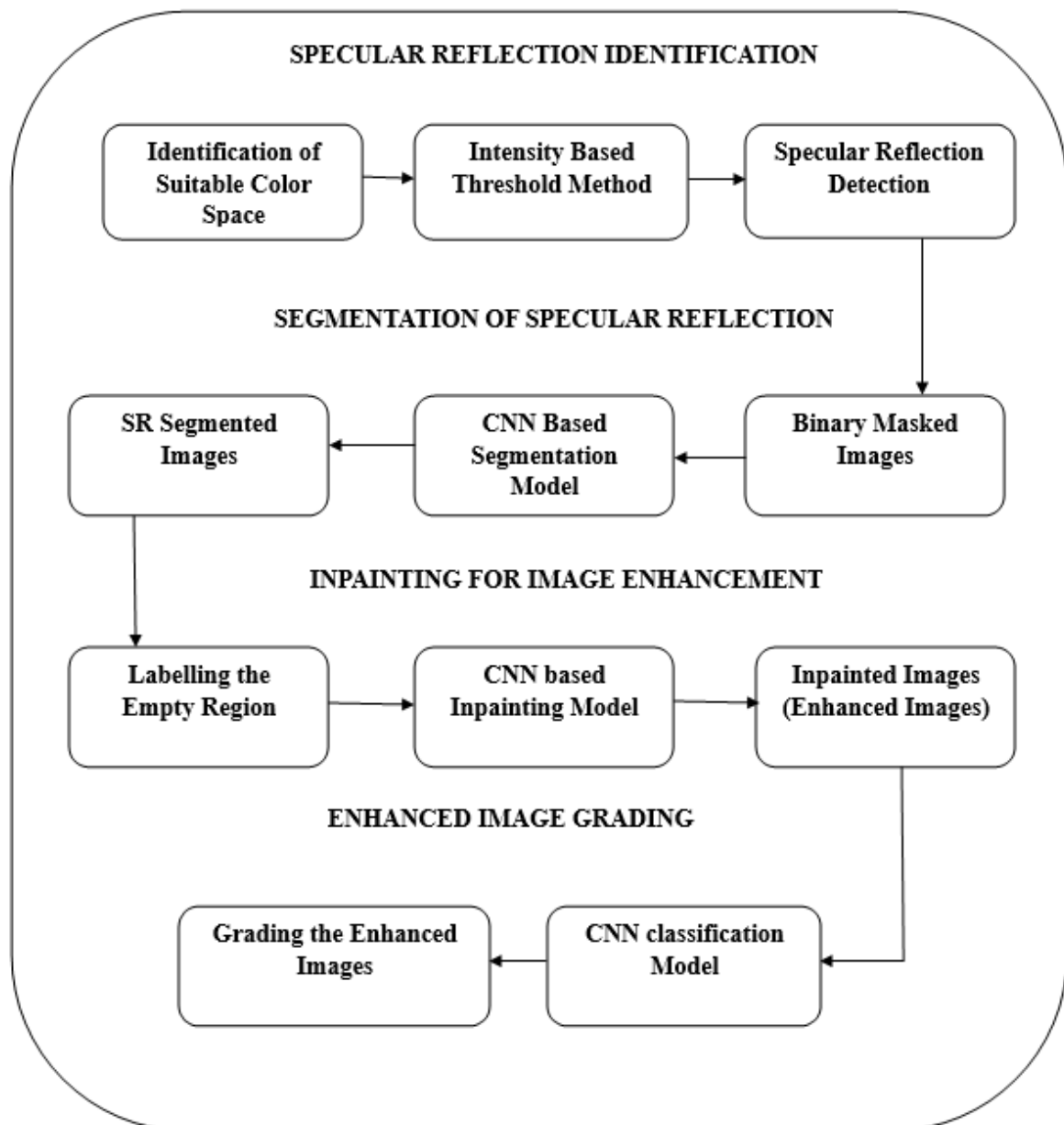


Figure 3.1 Proposed Research Methodologies

3.2. Phase I: Specular Reflection Identification

Researchers initially employed the Dichromatic Reflection Method (DRM) in the early stages of SR detection. However, it became evident that DRM had limitations, mainly when dealing with regions of similar color. This constraint made it challenging to apply the DRM directly to medical images [126]. In response to this challenge, a gradient-based approach was introduced for SR detection, primarily relying on various color space models. These color spaces, including RGB, HSV, XYZ, YUV, and Lab, offer distinct advantages in detecting SR in medical images. Each color space comprises channels with unique

properties, some of which align closely with SR characteristics. These specialized properties within specific color space channels prove valuable when identifying and isolating SR regions within medical images. This novel approach provides a more robust and versatile solution for SR detection in medical image analysis, especially in cases where similar color regions may confound traditional methods. Consequently, this methodology represents a significant advancement in medical image processing and quality improvement.

The primary goal of this research is to identify the reflection regions within smart colposcopy images. To accomplish this, a thorough investigation is conducted into various color spaces to determine which is best suited for identifying glare regions. Once the optimal color space is identified, a thresholding technique is employed to precisely delineate and locate the reflection region within images. In context of intensity-based thresholding, bicubic interpolation emerges as a valuable technique for upscaling or resizing images by considering the values of surrounding pixels. This process entails estimating the values of new pixels by examining a 2x2 grid of neighboring pixels. The algorithm utilizes these nearby pixel values to create a bicubic polynomial interpolation function, which is then employed to calculate the pixel value at the desired output position. This technique aids in achieving more precise image resizing and enhancement.

A comprehensive approach is conducted by the researcher by leveraging every color space channel to identify reflection regions within digital images [67]. It is beneficial because, in some cases, each channel may contain small reflection regions even if it does not convey distinct color information. As such, in their proposed method, each channel of the color space is harnessed to identify glare regions. The crux of this approach lies in calculating the intensity value of each channel, which effectively represents the average intensity or grayscale value within the selected color images, such as those in the RGB and XYZ color spaces. The digital images typically possess intensity levels between 0 and 255, where 0 corresponds to the darkest portions of the image, and 255 corresponds to the lightest areas. In the case of the RGB color space, the average intensity values typically fall within the range of 152 to 168. Consequently, the manual analysis of intensity values for images in the RGB color space commences from a threshold of 150. Similarly, the average intensity values typically range from 160 to 179 for the XYZ color space. Hence, the manual analysis of image intensity values in the XYZ color space begins with a threshold of 160. In summary, this research approach leverages bicubic interpolation and intensity

analysis across color channels to effectively detect glare regions in smart colposcopy images, enhancing the grade of cervical image analysis.

3.2.1 Result of Specular Reflection Identification

The experimental result is compared with the existing method, which works with the reflection detection process using different color spaces. The XYZ is highly suitable for reflection prediction, and intensity-based threshold method predicts the reflection with higher accuracy.

3.3. Phase II: Segmentation of Specular Reflection

In computer vision, utilizing CNN models holds significant importance. These models serve the crucial purpose of isolating specific regions within images. The CNN models help extract and acquire intricate features, leading to a more precise and resilient segmentation procedure. However, one of the significant challenges encountered in deep learning models is the availability of annotated data and the necessity for extensive datasets to train these models effectively. To automate the segmentation process in identification of reflection regions within smart colposcopy images, a CNN based segmentation model is employed. This model excels at performing pixel-level segmentation by utilizing masked images. In this context, where annotated images are unavailable, the intensity-based threshold method is implemented to annotate the reflection regions within the images. The outcome of this thresholding process is then transformed into binary masked images. In these binary masks, the reflection region is denoted as a white region, represented by the value 1, while the remaining areas, devoid of SR, are depicted as black regions, represented by the value 0. This binary masking technique labels the reflection and non-reflection regions within the colposcopy images. The training process for deep learning segmentation models involves using both original images and corresponding masked images as inputs. This training setup enables the model to learn and recognize the reflection regions, leveraging the binary masks as ground truth labels for guidance.

Initially, CNN segmentation models like FCN, SegNet, and U-Net models were used to segment the SR on smart colposcopy images [127][128][129]. These segmentation models are widely utilized in the medical image segmentation field, and among them, the U-Net model demonstrated superior performance due to its high segmentation accuracy even when provided with a limited dataset. The U-Net architecture is particularly well-suited for

scenarios where dataset sizes are constrained. Since there are different versions of the U-Net model, such as U-Net++ and Residual U-Net, which is systematically compared to determine the most suitable U-Net variant for effectively segmenting glare regions within smart colposcopy images [130]. The U-Net++ model exhibited the highest accuracy in predicting glare regions on comparison analysis. To further enhance the model's performance in segmenting SR in smart colposcopy images, the UNet model is fine-tuned. This fine-tuning process tailored the model to glare region segmentation in smart colposcopy images. The overall process for segmenting smart colposcopy images using these deep learning models, providing an overview of the methodology employed to achieve accurate and reliable SR segmentation.

3.3.1 Result of Segmentation of Specular Reflection

The experimental result indicates that the UNet model segments the reflection with a higher accuracy. Consequently, various variations of the UNet model are employed to assess their performance in segmenting images. This UNet ++ predict higher, so the UNet++ model is fine-tuned to reduce the over-segmentation process. The fine-tuned UNet++ model segments the reflection region with a higher accuracy.

3.4. Phase III: Inpainting for Image Enhancement

Image inpainting is pivotal in computer vision and finding applications across various research domains. This technique proves invaluable in restoring damaged or aged images, removing unnecessary objects, and filling in missing regions within an image by extrapolating information from the surrounding pixels. It uses contextual information in the vicinity and harnesses the inherent structures, patterns, and textures present in the image to generate plausible and visually coherent completions for the missing or damaged portions. Initially, the inpainting method employs various filtering methods with multiple iterations to identify and delineate the missing regions within the images. However, as the methodology evolved, the ML algorithms and advanced techniques enhanced the quality of the reconstructed images, aiming for results that are not only realistic but also visually appealing. Through comprehensive analysis, it is determined that both bilateral filters and convolutional filters play pivotal roles in the inpainting process [131]. However, in recent times, deep learning inpainting models have gained prominence, particularly those utilizing CNN models. These CNN-based models incorporate a sequence of convolutional filters to predict and reconstruct missing image regions. In conducting a comparative analysis, the

bilateral filters generate missing information with higher quality when compared to convolutional filters. This superior performance can be attributed to the unique characteristics of bilateral filters. Unlike traditional convolutional filters, bilateral filters consider spatial and intensity weights assigned to each neighboring pixel and its corresponding intensity value [132].

It achieves this by performing a weighted average of neighboring pixel values to compute the output pixel value. The bilateral filter takes into account both the difference in intensity values between pixels and their spatial separation within the image. This dual consideration enables the bilateral filter to excel at preserving image edges and fine details while effectively reducing noise in the image. To further enhance this approach, it has integrated the intensity value component of the convolutional filter and supplemented it with the spatial weight from the bilateral filter. This proposed hybrid filter replaces the traditional convolutional filter in the inpainting process. The novel filter is incorporated into the UNet model for training purposes. This integration has led to development of the Bilateral-based Convolutional Inpainting model, specifically designed to predict and reconstruct missing regions within images. This model leverages the strengths of both the convolutional and bilateral filters to achieve superior inpainting results. The method proposed is more comprehensive than smart colposcopy. It is implemented to different medical and digital images to check the model's performance. It is also tested with a higher masking ratio. The model fails when the masking ratio of the images is higher than 30%. An ablation study on the proposed model used different loss functions to solve the issues. The loss function is utilized in this study is pixel loss, perceptual loss, Total variation loss and style loss. Based on applying the loss function on the masked images, the proposed method and pixel loss function predict the large mask, i.e., 34-50 % with higher quality.

3.4.1 Results of Inpainting for Image Enhancement

The outcome of experiment result shows that the bilateral-based convolutional inpainting model predicts the missing region more accurately.

3.5. Phase IV: Enhanced Image Grading

The grading of smart colposcopy images refers to evaluating and categorizing the visual appearance of cervical abnormalities observed in the smart colposcopy images. One commonly used grading system is the cervical intraepithelial neoplasia (CIN) classification,

classifying abnormal cell changes into different grades based on severity. The CIN system typically includes grades CIN1, CIN2, and CIN3, with CIN3 being the most severe and potentially indicating the presence of precancerous or cancerous cells. In grading smart colposcopy images, SR poses a significant challenge, potentially leading to inaccurate grading results. To mitigate this issue, enhanced images are utilized, which involve the critical step of detecting and removing glare regions. Deep learning classification models, like DenseNet121 [133], VGG19, and EfficientNet, are highly effective for grading smart colposcopy images. It is an aspect of the DenseNet model family, which is widely used for segmentation, object identification, and picture classification among other computer vision applications.

VGG19t is part of the VGG family of models and has been extensively employed in image classification tasks [134]. It has 19 layers, 16 convolutional layers and three FC layers. EfficientNet is the balanced model with computational efficiency, and accuracy by incorporating several innovative techniques [135]. These deep learning models are robust tools for effectively grading smart colposcopy images, mainly when the enhanced images have addressed and removed glare regions. A clear justification is needed to underscore the importance of SR removal in improving the grading process. The initial grading of non-enhanced images is conducted using the classification model. Subsequently, the same classification model is employed to grade the enhanced images. On analysis, it becomes evident that the prediction accuracy significantly improves when working with the enhanced images. The deep learning classification model achieves a higher accuracy rate of 97.32% when classifying smart colposcopy images that have undergone SR removal.

3.5.1 Result of Enhanced Image Grading

The experimental result validates that the utilizes the classification models to improvise the grading of smart colposcopy images has improved the image quality through reflection detection and removal.

3.6 Summary

The key focus of the research endeavor was to develop an optimized methodology for precisely identifying and segmenting reflection areas within digital photos. The ultimate goal was to replace these segmented regions with higher-quality pixels to enhance the overall quality of the images. To accomplish this objective, the research was structured into

four distinct phases: phase 1, phase 2, phase 3, and phase 4. Each phase helped achieve the desired outcome of improved quality through reflection region detection and replacements. This part offered a thorough synopsis of the research proposal along with the methodology used in each stage.