CHAPTER 6

ARTEFACT RESTORATION AND POLYP CLASSIFICATION

6.1 INTRODUCTION

The third leading cause of mortality in India and the fourth leading cause of cancer-related death worldwide is Colorectal Cancer (CRC). This disease is one of the deadliest malignancies in the world (Gajanand Sharma et al., 2020; Siegel et al., 2021). According to research, genetic and epidemic factors account for the growth of adenomas in 85% of colorectal cases. Endoscopic removal of such polyps can lower this risk (Strum, 2016; Davidson et al., 2021). Polyps are categorized into two main kinds they are nonneoplastic and neoplastic. Nonneoplastic types of polyps include hyperplastic, inflammatory and hamartomatous polyps. Nonneoplastic polyps typically do not become cancerous (IJspeert et al., 2016; Ijspeert et al., 2017).

Similarly, neoplastic polyp includes adenomas as well as serrated types and this type of polyp has a higher possibility of turning into cancer. A study confirms that hyperplastic polyps in the colon and rectum do not need invasive procedures. They are considered non-malignant (Rex et al., 2011). Therefore, accurate polyp classification and segmentation are vital in digital endoscopy. Such system reduces the efforts of patients and clinicians (Hassan et al., 2010; Sikka et al., 2008).

Although CRC is highly dangerous, it develops into cancer after a long period. Cancer development can be avoided if such CRCs are identified and treated in the early stages. To date the clinicians, depend on colonoscopy to detect CRC's. Hence detection, classification and segmentation all plays a vital role. Especially categorizing nonmalignant and malignant polyps is important (R. Zhang et al., 2017). Amidst of technology advancement, advanced researches are present in the field characterization of the polyp. There is a huge difference in how the clinician characterizes them (Lakhani & Sundaram, 2017; Mnih et al., 2015; Sainath et al., 2015).

Recent developments based on DL have given a ray of hope. With the assistance of DL, many tasks in the healthcare industry have become easier (Ciompi et al., 2017). The diagnosis and categorization of polyps have been extensively studied nowadays. However,

most of the reported studies discuss CAD, especially in the field of detection and categorization (Chen et al., 2022; Singh, 2021; X. Zhang & Wang, 2021). Such studies assist researchers in inspiring and bringing out accurate results. Figure 6.1 shows normal endoscopic images and the images with an instance of a polyp as abnormal.

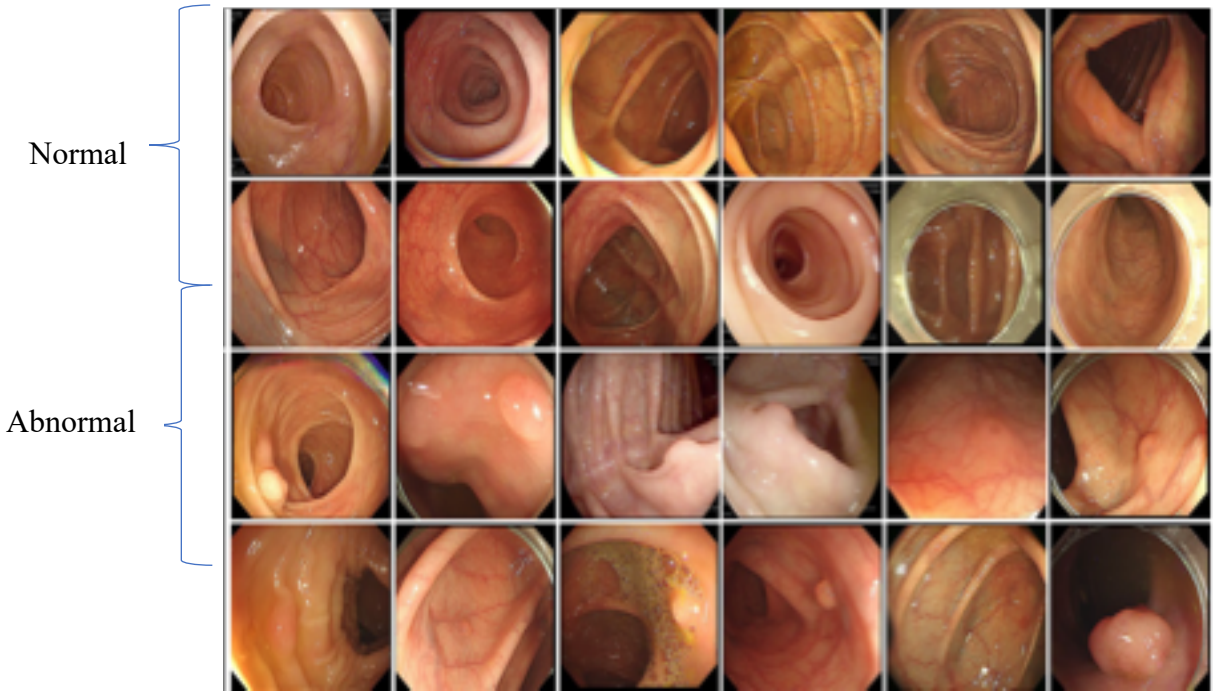


Figure 6.1 Endoscopic Images with and without an Instance of Polyp (Kim et al., 2021)

This study focus on the technologies that if it is assisted by image restoration then performance of the disease classifier would improve. Such technology may assist clinician to take lesser time for diagnosis.

In order to perform the simulation the whole process is divided into four stages. First stage is to prepare a dataset. Second is to train a DL based segmentation algorithm to segment the artefacts present in the input image. Once a binary mask is obtained from the segmentation algorithm, the third step is to restore the artefacts. The last stage is to pass the restored image to a trained classifier. The classifier classifies whether the input image has an instance of polyp or not. Figure 6.2 depicts the overall flow of the simulated study.

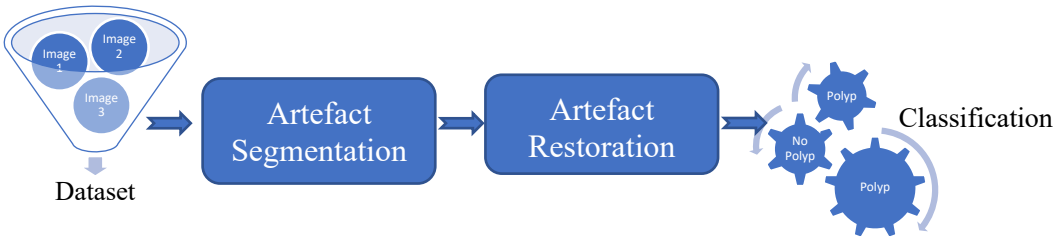


Figure 6.2 Overall Architecture of the Proposed Simulation Study

6.2 DATASET

The dataset for polyp classification is extracted from two the different image sources. Kvasir SEG dataset for endoscopic images containing an instance of a polyp. The second source is from the EAD dataset. Selective images are chosen from the EAD dataset for non-polyp images. Hence the combined dataset contains 1000 images with an instance of the polyp and 1000 images without the polyp. All the images are resized to 224 x 224 before feeding into the network.

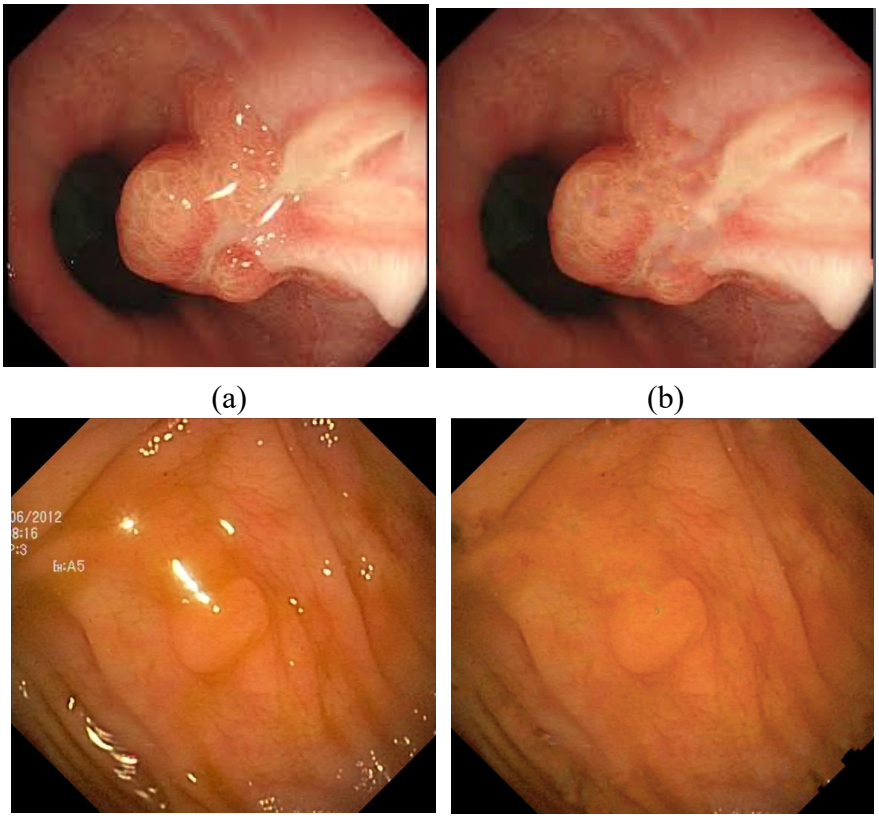
6.3 ARTEFACT SEGMENTATION

This research proposes a method to segment specific endoscopic artefacts. The architecture used for artefact segmentation is detailed in the chapter 5. The binary mask output from the segmentation algorithm contains binary mask of five different artefacts. The masks are selectively chosen for artefacts such as, specular reflections, saturation and selective miscellaneous artefacts. The selected masks are used for further process like restoration.

6.4 IMAGE RESTORATION

Image restoration is the art of extracting the neighbouring pixel information and filling it in a specific part of an image. The application of such a technique include the restoration of an image which is damaged. This technique can also be extended to remove a selected portion from an image and replace it with surrounding pixel information. Such techniques can be extended to various images, from natural to medical images. In the field of

healthcare, there are various modalities of imaging. In every modality, the clinician depends on specific imaging hardware. Advancement in hardware usage for every specific function have advantages and disadvantages. One such disadvantage is the occurrence of artefacts. These artefacts occlude the imaged organ leading to a false diagnosis and many other associated degradations. Especially in the field of endoscopy, artefacts play a dominant role. Restoring one or many of those artefacts is a challenging role for many researchers to date. Figure 6.3 shows an endoscopic image before and after restoration. It is apparent that the restored image shown in Figure 6.3 (b) and (d) is considered to be much more useful for post-processing like report generation and information extraction than the original image in the Figure 6.3 (a) and (c).



(a) (b)

(c) (d)

Figure 6.3 (a)-(d) Endoscopic Images Before and After the Restoration of Artefacts (Specular Reflection)

To attain a restored endoscopic image, there are a various limitations. The first limitation is that more than one artefact is present, and one algorithm cannot effectively work

on all different artefacts. Secondly, neither traditional algorithms nor DL-based restoration algorithms can work well on all artefacts, so frames must be segregated based on the artefact present in the image and then restored one by one. A separate pipeline must be designed to achieve the same. A few endoscopic artefacts with similar characteristics are identified as an initiation, and traditional algorithms are used to restore the same. The artefacts identified include specular reflections, saturation and a few miscellaneous artefacts. A simple thresholding technique would work fine, but the traditional algorithm fails to categorise the underlying tissue or other artefacts in a few cases where the specular reflections fall on an instrument. Hence relying on a DL-based algorithm gives a ray of hope.

Regarding real-time vision systems, specular reflections and saturation distract the field of view. Hence removing such specular reflections holds an important role in the process. Image restoration is one of the best methods to restore the specular reflections. Detecting specular reflection present in an endoscopic image can be easily performed even with the image segmentation algorithms, as the pixel values of the affected region and the background have a vast difference. The specular reflections are bright spots of light. Hence a simple algorithm is enough to detect and restore. Many researchers proposed research results in this area. Such results detect and segment the boundaries of specular reflections. Later the same can be restored with image restoration algorithms.

Similarly, saturation is an artefact that holds bright pixel areas but covers more area than specular reflections. Most miscellaneous artefacts also have similar properties as specular reflections and saturation. All the three artefacts are restored using a simple traditional restoration algorithm.

Fast Marching method (Telea, A., 2004) is used to restore the artefacts. Let us assume, p is the point to be restored. Assume a small neighbourhood $B\varepsilon(p)$, where ε is the size of the neighbourhood around the point p. In order to restore a point p, the value of its neighbour $B\varepsilon(p)$, must be known. If value of ε is small, then first order approximation of the point p of the given input image I(q) must be considered. The expression for first order approximation Iq(p), is given in the Equation 6.1, Where $\nabla I(q)$, is a gradient value at point p.

$$I_q(p) = I(q) + \nabla I(q)(p - q) \tag{6.1}$$

Next step is to restore the artefact, at the point p using the expression given in Equation 6.2. It must be iteratively applied for restoration. I(P) is the restored image. In the Equation 6.2, w(p,q) is a weighted function.

$$I(P) = \frac{\sum_{q \in B\epsilon(p)} w(p,q)[I(q) + \nabla I(q)(p-q)]}{\sum_{q \in B\epsilon(p)} w(p,q)} \tag{6.2}$$

The expression to calculate w(p,q)is given by the Equation 6.3.

$$w(p,q) = dir(p,q).dst(p,q).lev(p,q) \tag{6.3}$$

Where $dir(p,q)$ is the directional component. The expression is given in Equation 6.4. $dst(p,q)$ is the geometric distance component given by Equation 6.5. $lev(p,q)$ is the level set distance. The expression to calculate the level set distance is given in Equation 6.6.

$$dir(p,q) = \frac{p-q}{\|p-q\|}.N(p) \tag{6.4}$$

$$dst(p,q) = \frac{d_0^2}{\|p-q\|^2} \tag{6.5}$$

$$lev(p,q) = \frac{T_0}{1 + |T(p) - T(q)|} \tag{6.6}$$

The formula must be applied to all discrete point p. The movement must be from outside to inside. The algorithm must iteratively run until the whole region is restored. The code to restore the artefacts is built using OpenCV, a computer vision library.

6.5 POLYP CLASSIFICATION

6.5.1 Image Pre-processing

The images obtained are grouped into two sets of images. One contains an instance of a polyp, and the other does not contain a polyp. Enough care is taken such that the images collected are all affected by artefacts. Image restoration is performed to restore the artefacts

found in the endoscopic images, such as, specular reflections, saturation and a few miscellaneous artefacts. Hence, there are two sets of datasets one with all endoscopic images with and without polyp affected by artefacts and the other set with artefacts restored for the same set of images.

The dataset available for training is insufficient for accurate classification. Hence simple data augmentation techniques such as, horizontal flipping, zooming, rotation and width and height shift is employed. After applying data augmentation technique the data set is split up with 80% images for training and 20% for testing. The train and test set images are resized to 224 x 224 to maintain uniformity across the dataset. The CNN architecture is also designed to accept images of size 224x224. The sample augmented images are shown in the Figure 6.4 (a) – (f).

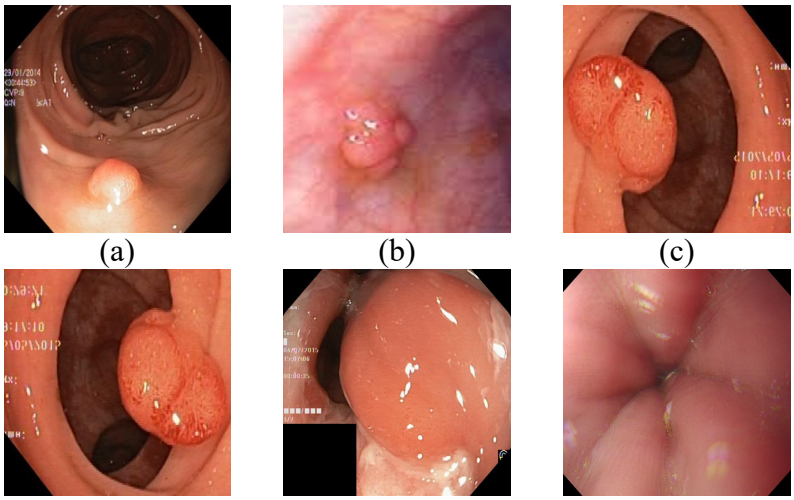


Figure 6.4 (a) – (f) Augmentation of Dataset

6.5.2 CNN Architecture

A simple CNN based classifier is designed, trained and tested on the augmented dataset. This section briefs the CNN structure adopted for polyp classification. The basic CNN consists of convolution layer, pooling layer, FC layer, drop out and activation functions. The convolution layer is considered as the basic building block of any classifier. This maximum computational load of a classifier depends on this layer. The pooling layer is

used to reduce the size of the feature map. This reduces the computation cost. There are two key types of pooling max pool and average pool. In the proposed CNN, max pooling is used.

Figure 6.5 depicts the layer details of CNN architecture with the trainable parameters.

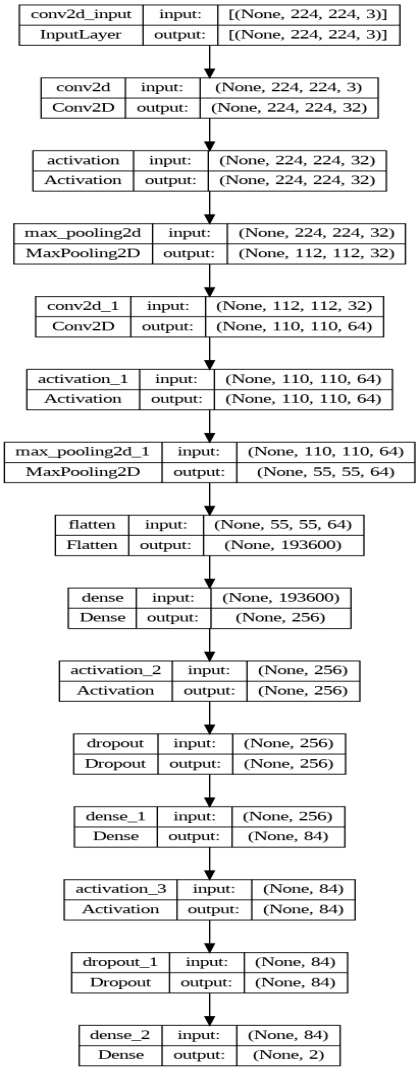


Figure 6.5 CNN Model

Dropout layer is used to drop a set of neurons during training. Dropout helps to avoid overfitting. Overfitting in a CNN means the networks learns well with training data but shows poor performance in the test data. This reduces the performance of the network. Dropout assists improving the performance by dropping a set of neurons in the forward pass. Corresponding weight updates are not included in the backward pass. Hence the co-

dependencies of the neuron units can be avoided. Drop out value of 0.25 is adopted to improve the network performance. Softmax is used as an activation function in the final layer, as there are only two classes (polyp and non-polyp). The FC layer holds the weights, biases and the neurons. This layer is placed at the output layer. The network resulted with 49,603,006 trainable parameters. The CNN architecture is modelled using Keras API. The overall model is coded using python programming language. Training of the CNN model for classification is achieved using Google co-lab.

6.5.3 Training

The classifier is trained for 20 epochs. To add the benefit of transfer learning, weights pretrained ImageNet is used. Early stopping is used during training. The training stops when the model learning does not improve for three consecutive epochs. This technique also prevents overfitting of the network. Data shuffling and batching is deployed during training. Data shuffling allows the program to shuffle the images that are passed into the classifier during training for each forward pass. Batching is done to determine number of images to be sent for training during each pass. Shuffling and batching are two different techniques implemented during training to improve the performance of the classifier. This makes the model learn more generalized features. The model file extracted after training is used to test the model. Figure 6.6 portrays the typical polyp classification structure.

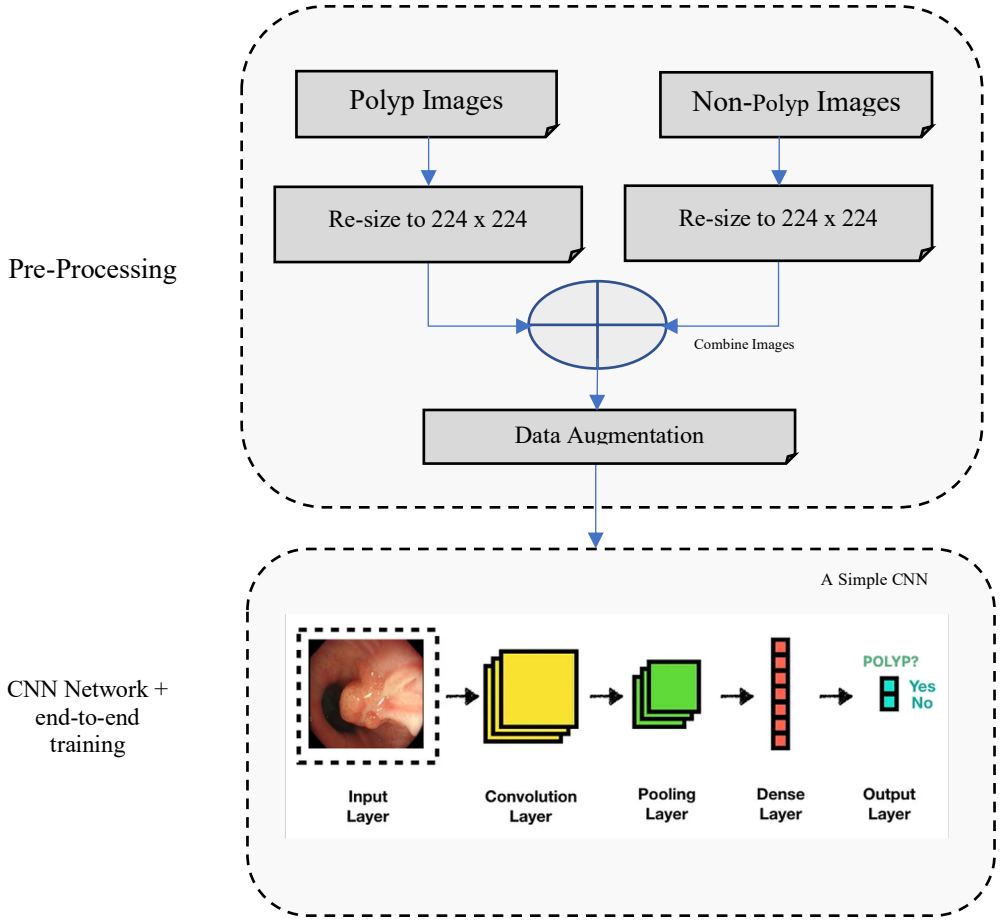


Figure 6.6 Overall Polyp Classifier

Table 6.1 briefs the parameters set for training the polyp classifier.

Table 6.1 Parameters Set for Polyp Classifier

S.No.	Parameter	Value
1	Image Size	224 x 224
2	Batch Size	32
3	Shuffle	True
4	Optimizer	adam
5	Learning rate	0.001
6	Loss	Categorical cross-entropy
7	Epochs	20

The number of epochs is set to 20. When the number of epochs are increased the network overlearn the features. Early stopping is used to predict when the loss did not improve for consecutive 3 iterations. Through the early stopping feature it is found that the algorithm could only be trained for 20 epochs for the given set of images. Hence it is decided to stop iterations at 20 epochs.

6.6 PERFORMANCE METRICS AND SIMULATION RESULTS TO EVALUATE THE POLYP CLASSIFIER

6.6.1 Test set for polyp classification

The train and test split feature from the Keras library is used to split the dataset during run time for polyp classification. 1000 images devour an instance of polyp and 1000 images for the non-polyp. Out of which 20% of images are split for testing. The percentage of the split is fed manually, and it can be varied.

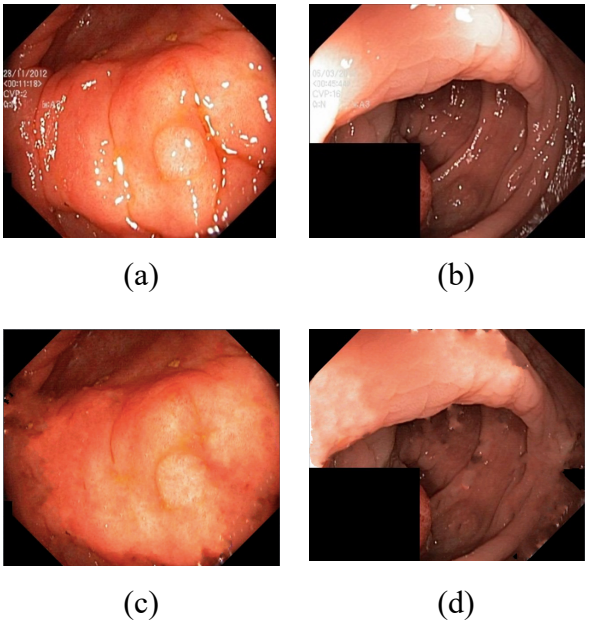
6.6.2 Performance Metrics

The performance metrics used to assess the model performance include Precision, Recall, F1- Score and accuracy. The formula to calculate F1 score, precision and recall are discussed in the chapter 5. Accuracy is calculated using the formula given in Equation 6.7.

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \tag{6.7}$$

6.6.3 Simulation Results

Every image in the test set may or may not contain an instance of polyp. The endoscopic artefacts are restored. The same set of images are retained with restoration thus resulting two sets of test set. Figure 6.7 (a) & (b) shows the images affected by artefacts and Figure 6.7 (c) & (d) shows restored images. Both the set of images are passed into a polyp classifier. The performance of the classifier in classifying the images with and without polyp before and after artefact restorations is evaluated.



(a) (b)

(c) (d)

Figure 6.7 (a) – (d) Original and Restored Images from Test Set

During testing, it is observed that when there are more specular highlights and saturation the model missed to classify few images of polyps as polyps. The presence of multiple artefacts reduces the accuracy of the classifier. The accuracy is found to be 97%. The classification accuracy improved from 97% to 100% when restored images are passed into the classification model. Table 6.2 shows the performance of the same trained classifier on classifying the polyp and non-polyp conditions from endoscopic images before and after restoration. Figure 6.8 shows the confusion matrix of the polyp classifier.

Table 6.2 Classification Report Before and After Restoration

	Before Restoration		**After Restoration**	
Parameters	**No-polyp**	**polyp**	**No-polyp**	**polyp**
Precision	1.00	0.94	**1.00**	0.99
Recall	0.93	1.00	**0.99**	1.00
F1-Score	0.97	0.97	**1.00**	1.00
Accuracy	-	97%	**-**	100%

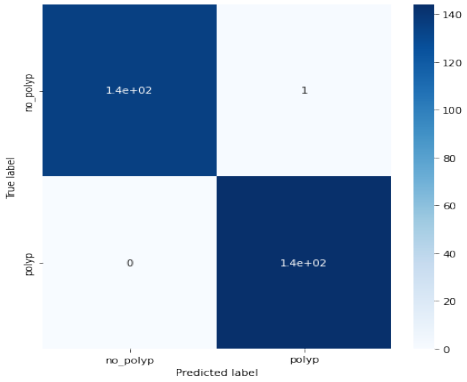


Figure 6.8 Confusion Matrix of Polyp Classifier – Testing on Restored Images

It is clear from the metrics evaluated, that the classifier performs better classification of the polyp by 3.09%, if the images are restored with **precision and F1 score of 1.00** and **recall score of 0.99**. Hence the performance of the classifier is better when artefact restored images are used than classifying using images that are not artefact restored.

It is also observed that polyp images selected for compilation of dataset includes images from Kvasir SEG dataset. Most of the polyps are protruding and the feature identification is found to be an easy task. In very few samples, the polyp is not much visible mainly due to artefact and occlusion. In such cases the classification is found to be quite challenging.

6.7 CHAPTER SUMMARY

First, a dataset for polyp classification and image restoration is recreated. The endoscopic images with an instance of polyp is obtained from Kvasir dataset. Endoscopic images without an instance of polyp is pooled from EAD dataset. The DL based segmentation algorithm proposed will segment the boundaries of artefacts and outputs a binary mask. The individual binary mask obtained corresponds to every artefact. Selectively three artefacts are chosen. The artefacts include specular reflections, saturation and miscellaneous artefacts. Using the binary masks the segmented areas are restored using fast marching algorithm.

Now a simple CNN is designed using the Keras library to classify polyps. Techniques such as, data shuffling, augmentation, batching and dropout regularization are deployed to

improve the neural network's performance. The trained network is tested with images from two different test set. The first test set contains images with an instance of polyp in approximately 50% of the images, but the artefacts are not restored.

Similarly, the same set of images are used but the artefacts are restored. That becomes test set 2. Images from the test set one and two are randomly passed into the classifier. The classifier's performance in predicting an instance of polyp before and after the restoration of the artefact is evaluated. From the evaluation it is evident that the classifier’s performance in terms of accuracy is better by 3.09% when the artefacts are restored than the images that are not restored.