EOCNet:Improving Edge Omni-Scale Convolutional Networks for skin lesion segmentation

PROF. BINOD KUMAR
DESIGN CREDITS

POLKAM SHRUTHIKA
B20MT031

INTRODUCTION:

Skin lesion segmentation refers to the process of recognising and separating skin lesions or abnormalities from the surrounding healthy skin in medical images. Segmenting skin lesions is essential for detecting skin cancer. This method is used by dermatologists and other medical experts to precisely analyze and evaluate skin lesions. It is simpler to quantify the size, form, and texture of the skin lesions as well as to spot any anomalies or asymmetries by segmenting them. In order to diagnose and treat skin cancer early, this knowledge is crucial. The segmentation technique aids in pointing to the precise region of the skin lesions where cancerous cells are present by precisely defining the borders of the lesions.
Additionally, segmenting skin lesions makes it easier to follow the development of skin lesions over time. The size, form, or appearance of the lesion might change over time, and medical professionals can see this by comparing the segmented photos taken at various time points.
In general, segmenting skin lesions is an important step in the computer-aided detection of skin cancer. It makes early detection easier, enables efficient and reliable examination of skin lesions, and aids medical professionals in making judgements about additional diagnostic procedures and treatment alternatives.

PROBLEM STATEMENT:

The problem statement as explained in the paper is to improve the CNN Architecture so that it can be useful to detect the disease easier and faster. The problem statement is to develop a non- invasive method which diagnoses skin cancer. There are many methods proposed which use CNNs to segment and diagnose the cancer, but their performance needs to be improved. The challenge in performing the segmentation is varying size, shape, color like features of the lesion; along with skin pigmentations and human hair. The architectural problem with previous attempts is that identifying the location of the lesion requires a shallow network, whereas its segmentation requires a deeper network, and hence a single network is unable to perform the task accurately. Hence, the current problem reduces to developing a non-invasive method, diagnosing skin cancer better than those single network CNNs.

EOC-NET ARCHITECTURE:

The EOCNet architecture is made up of three modules:

- The encoder module,
- Boundary extraction, and
- The omni-scale module.

Here, the Omni-Scale module serves to increase the spatial resolution of the images from which features may be extracted, and as a result, the features extracted by the encoded module are enhanced.

For the purpose of segmenting skin lesions, the EOCNet architecture described in the publication "EOCNet-Improving Edge Omni-Scale Convolutional Networks for Skin Lesion Segmentation" is a deep learning model. The Edge Detection Module, Feature Pyramid Pooling Module, and Segmentation Module are just a few of the modules that make up the design.
The task of detecting the borders of skin lesions falls in the Edge Detection Module. It is composed of a number of dilated convolutional layers that discover edge features at various scales. After being combined, the outputs from these layers create an edge map that is utilized to direct the segmentation procedure.

The Feature Pyramid Pooling Module is made to pull out features from various scales of the input image. In order to collect features at multiple scales, it employs dilated convolutional layers with varying dilation rates. A pyramid pooling layer is another component of the module that combines features from several scales to create a multi-scale feature map.
The Segmentation Module generates a pixel-level segmentation map using the multi-scale feature map created by the Feature Pyramid Pooling Module. The feature map's resolution is steadily increased throughout the module's convolutional and upsampling layers until it is the same size as the input image. The output is a binary mask that shows the border of the skin lesion as the final result.

So, in order to increase the precision of skin lesion segmentation, the EOCNet architecture combines the advantages of edge detection and multi-scale feature extraction. The network can accurately segment lesions of various sizes and shapes thanks to the multi-scale feature extraction provided by the Feature Pyramid Pooling Module, which enables the network to capture features at different scales. The edge information provided by the Edge Detection Module aids in guiding the segmentation process.

EOC NET IMPLEMENTATION

1. *Image preprocessing:*

   The data is first loaded and the preprocessing techniques mentioned in the paper were implemented. They were mainly:

   1. **Flipping:** Flipping refers to reversing the image horizontally or vertically. It can be useful for creating mirrored versions of images to adjust the image orientation as required for the dataset in the paper.

   2. **Resizing:** Resizing involves changing the dimensions of an image. It is being done to reduce the image size for efficient processing. Resizing is done to normalize the images so that all the images are of the same size before feeding them into the model.

   3. **Rotating:** Rotating an image involves rotating it by a certain angle in clockwise or counterclockwise. Rotating an image can be helpful for correcting orientation or aligning images.

4. **Cropping:** Cropping involves removing a portion of the image which is irregular. Cropping is used to focus on the skin lesion areas and remove the unwanted parts.

2. *Modules implemented in the model:*

The models we have implemented from our paper are:

1. **Feature extraction module:**
   High-level feature representations from the input image are extracted by this module. The ResNet50 architecture served as the framework for the feature extractor module developed by the authors. It was primarily the encoder of the model. This helps in extracting the features corresponding to segmentation; i.e., the features which are responsible for deciding the segmentation value of a pixel.

2. **Boundary extraction module:**
   This module used DexiNed to generate the probabilistic map of the boundary of the lesion which aids in separating the boundaries of the objects from the background. DexiNed is basically an encoder-decoder architecture to detect the edges in the image. This helps in identifying and preserving the location information of the lesion in the image.

   After these steps the results from the above two models i.e the feature maps and the edge information are combined together. This will supplement the deeper boundary information of the network by including both the local and the global information.

3. **Omni scale module:**
   It basically takes multiple input data streams, and it has multiple convolutional blocks with different filter sizes, thereby creating the effect of different receptive fields of different sizes. It works as a decoder, extracting features at multiple scales, so that the problem of varying size of the lesion gets resolved.

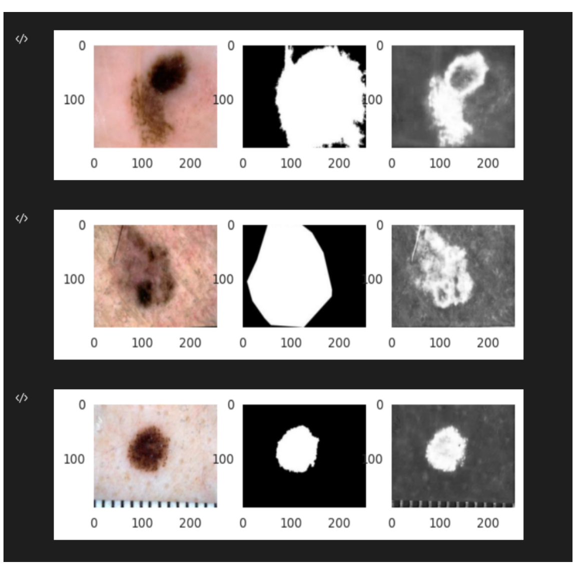
The other model implemented is the basic U-Net model so that we can compare the results of both the architectures. It is also the basic segmentation model which is being used in the research paper, which was the reason I chose to implement it.

3. Integration of the above modules:

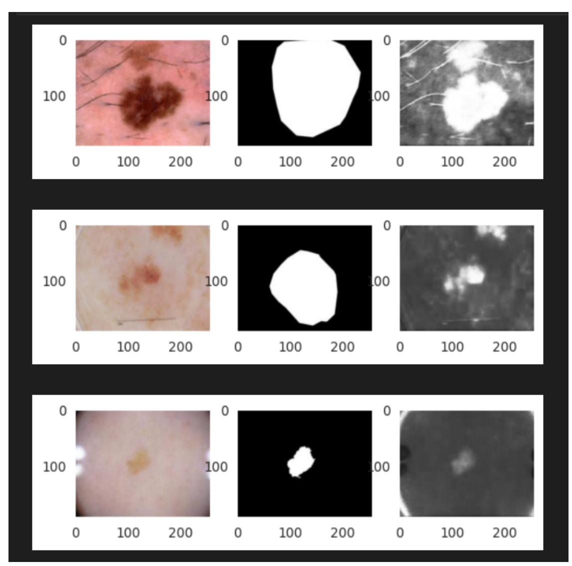
The dataset is trained using two models one is our paper implemented model and the other is a general segmentation model UNet. The paper implemented model combines the ResNet50 and DexiNed networks for skin lesion segmentation. The input image is passed through both networks simultaneously. The last three layers of ResNet50's output are extracted. The output of DexiNed is convolved and modified to match the dimensions of ResNet50's output. This is done to add the location information captured by DexiNed to the segmentation obtained by ResNet50. The combined result is reshaped and fed into the Omni-Scale Module, which consists of multiple data streams with different convolutional blocks. These blocks capture information at varied scales and help extract features from the feature-space. The output from the Omni-Scale Module is then stacked with the modified output of DexiNed. This step preserves crucial information about the location of the skin lesion. The stacked layers are convolved and upsampled to achieve the desired output image size. The model applies a Softmax layer and a Dropout layer with dropout probability of 0.5, to prevent overfitting. The training parameters include a learning rate of 0.001 and 25 epochs. The proposed model aims to provide a predicted mask for input images with skin lesions. The loss is determined after each epoch, and as the data was trained, the loss gradually decreased and eventually gave better results.

4. Results obtained:

EOC Net



UNet



The research paper proposed an improved skin lesion segmentation method by introducing the Omni-Scale module. This module outperformed the UNet architecture in terms of segmentation accuracy. The Omni-Scale module incorporated multi-scale feature extraction using convolutional blocks with different filter sizes. It effectively captured fine-grained details and contextual information. The module also integrated complementary information from other modules, preserving crucial lesion location details. By leveraging depth-wise and point-wise convolutions, the module captured complex patterns and improved discriminative power.
The Omni-Scale module provided more accurate segmentation results for skin lesions, hence enhancing its accuracy.

We have used the following evaluation metrics given in the research paper to further analyze the experimental results quantitatively:

1. Pixel-level accuracy: (AC)
2. Pixel-level sensitivity: (SE)
3. Pixel-level specificity: (SP)
4. Dice coefficient: (DC)
5. Jaccard similarity: (JS)

The outcomes of these values are:

EOC Net:

```
Accuracy:  0.7670358419418335
Specificity:  0.8495683670043945
Dice Coefficient:  0.5295370221138
Jaccard Similarity:  0.3601158559322357
Sensitivity:  0.5213677883148193
```

UNet:

```
Accuracy :  0.690789520740509
Specificity :  0.7217754530906677
Dice Coefficient :  0.5743997693061829
Jaccard Similarity :  0.4012585997581482
Sensitivity :  0.6456007790565491
```

So, we have found that the results of the EOC Net model are better as compared to UNet.