Skin Lesion Segmentation and Multi-class Skin Cancer Classification

The American University in Cairo

Mira Shanouda, Ahmed Fathy

CSCE 4604 - Practical Machine Deep Learning Course

Computer Science and Engineering Department

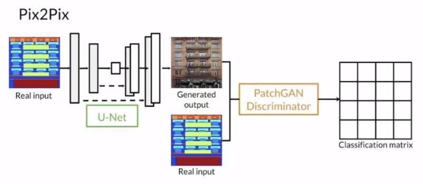
Problem Statement

With the boom of Deep Learning, the domain of applications has been expanding, and the medical field is one of the most famous domains. Medical imaging and medical diagnosis related applications are a famous example. For instance, **Skin Cancer Classification** and **Skin Lesion Segmentation** are gaining a boost in the research community.
In our project, we investigate both applications; we worked on a **CycleGAN** model for Skin Lesion Segmentation for an image-to-image translation that creates a binary mask with the segmented Lesion.
For Skin Cancer Classification, we worked on a **multi-class classifier** with the skin lesion mask as the input and a numerical class representation as the output; we tackled both aspects of the problem for the purpose of creating a pipeline for imaging and diagnosis.

Models

Segmentation (Pix2Pix Architecture):

The fully implemented model is a conditional generative adversarial network (cGAN) called Pix2Pix that learns a mapping from input images to output images. In this problem, the output image is the segment lesion of the skin cancer. The architecture of the Pix2Pix is composed of a generator and discriminator.
The generator is a U-Net based architecture that consists of an encoder (downsampler) and a decoder (upsampler). Each block in the encoder contains t: a convolution, a batch normalization, and a leaky layer. The decoder contains a transposed convolution layer, a Batch normalization, and a ReLU activation function. Between the encoder and the decoder, there are skip connections. For the discriminator, each block of it is the same as the encoder blocks.

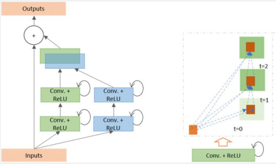


Multi-Class Classification (IRCNN):

The IRCNN architecture consists of general convolution layers, IRCNN blocks, transaction blocks, and a softmax logistic regression layer.
The IRCNN block recurrent convolution operations on different sized kernels.

Models

The transaction block applies a convolution layer, max and global average pooling and dropout layers.

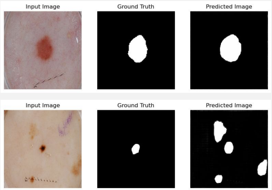


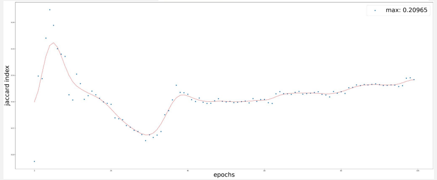
Results

Pix2Pix:

Because Pix2Pix is a model with large capacity, we were limited by the input size. On an **Nvidia Quadro 16 GB GPU**, we could only train on an input image of size **(256,256,3)**.
The original dimensions were mostly **(4000,6000,3)** leading to a significant loss of information during the resizing. Therefore, the model showed satisfying results considering the limited amount of input information.
We measured the performance mostly in a qualitative manner by comparing generated images to ground truths. And below are some good and poorly performing examples.

For the quantitative metrics, we used Jaccard Score, which we aimed to minimize as much as possible with the correlation of better performance with a lower score. A plot of the score is shown below.
The max maximum score obtained by Jaccard is **0.209**

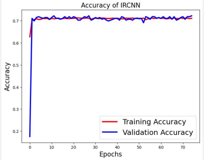


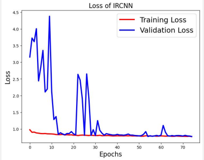


Results

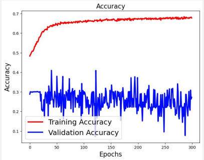
IRCNN:

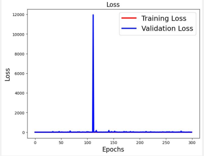
For the classification problem, we worked on the IRCNN classifier model two perspectives. Training using the binary masks from the ISIC 2017 dataset, and training using the original images from the ISIC 2019 dataset that had more training examples number of classes.
For the ISIC 2017 that contains 3 classes, the model performed considerably well considering that multi-class skin cancer classification is one of the harder problems to tackle in the medical field. For the similarities across different classes is large. We trained the model on 500 epochs and **the accuency here reached around 71% and the loss was around 0.8**. (The potting illustrates only 75 epochs).





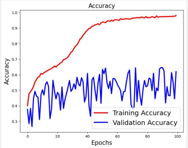
We wanted to compare each approach to determine further actions. Therefore we also trained on the ISIC 2019 dataset. Which produced lower performance because the images were not segmented. We have trained it for 500 epochs however, the model was overfitting regardless hyperparameter tuning as shown in the plotting.





Xception:

To ensure the model was not the issue, we tried training the Xception model on the ISIC 2019 images, with heavy regu-larization; unfortunately it also showed over-fitting problems after training for 100 epochs as shown in the plots.



ISIC Datasets

The International Skin Imaging Collaboration has been hosting challenges from 2016 to 2020. Since our approach considers both image segmentation and classification, we decided to consider the datasets from two different years, 2017 and 2019.

ISIC 2017:

The dataset provided 2000 images for the training data with ground truth binary masks for the lesions segmentation. In addition, all images are classified into 3 classes: Melanoma, Nevus and Seborrheic keratosis.

ISIC 2019:

This year's challenge is to classify dermoscopy images of skin legions into one of nine categories: Melanoma, Melanocytic nevu s, Basal cell carcinoma, Actinic keratosis, Benign keratosis, Dermatofibro ma, Vascular lesion, Squamous cell carcinoma, and None of the others.

Future Expansion

Considering the limitations we faced due to dataset-related issues and hardware issues, we suggest the following improvements in the future.

Pix2Pix – Increase Input Size:

Given more powerful hardware, the model can take in more information without overwhelming the GPU, which can lead to better results in creating binary skin segmentation masks.

IRCNN – Increase Input Size:

The IRCNN is a significantly large model, which can show much better results with more powerful hardware that can take in more input information.

IRCNN – Implement Residual Skip Connections:

The original model studied included residual blocks as an IRRCNN model. Implementing the model was challenging due to limited information, but for future expansions, it can prove worthwhile.

Reference

Alom, Md. Zahangir & Aspiras, Theus & Taha, Tarek & Asari, Vijayan. (2019). Skin Cancer Segmentation and Classification with NABLA-N and Inception Recurrent Residual Convolutional Networks. DOI: https://doi.org/10.48550/arXiv.1904.11126

Alom, Md. Zahangir & Hasan, Mahmudul & Yakopcic, Chris & Taha, Tarek & Asari, Vijayan. (2018). Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. DOI: https://doi.org/10.48550/arXiv.1802.06955