Image Classification on Histopathological Colon and Lung Cancer data

Remco Spelthan
Department of Data Science and Knowledge Engineering
Maastricht University
Maastricht, The Netherlands

***Abstract*—In this paper, we present a study on image classification of histopathological colon and lung cancer images. Using a dataset of these images, we trained 4 state-of-the-art image classification models, 2 simple CNN architectures, 2 Transformer models and 4 Hybrid Quantum Convolutional Neural Network models. We evaluated the performance using various metrics, including cross entropy loss, accuracy, precision and F1 score. We further try to give insight to why certain models choose classify the image with or without cancer, we did this using saliency maps, integrated gradients and integrated gradients with smoothgrad squared. Our results show that ResNet18 is the best performing image classification model, that there is a potential to using Hybrid Quantum Convolutional Neural Networks, and that models that correctly classify images have a more specific focus in images than models that incorrectly classify images.**

CONTENTS

I. INTRODUCTION

Image classification is a crucial task in the field of medical image analysis. It involves the automatic categorization of images based on their content, and has numerous applications, including the diagnosis and treatment of diseases. In this paper, we will focus on the use of image classification techniques for the analysis of histopathological images of colon and lung cancer. These types of images are obtained by examining tissue samples under a microscope, and they provide detailed information about the structural and cellular changes that occur in cancerous tissue. By accurately classifying these images, we can better understand the nature and progression of these diseases, and develop more effective treatments. We will present an overview of the existing methods for image classification, and evaluate their performance on a dataset of colon and lung cancer images. In addition, we will also experiment with Quantum CNN's, evaluating their performance as well. Our goal is to identify the most effective approaches for classifying these images, and to provide insights into the challenges and opportunities for improving the accuracy and reliability of these techniques. This leads to the following research questions our paper attempts to answer:

- How effective are image classification techniques at accurately classifying colon and lung cancer in histopathological images?
- What are the characteristics of histopathological images that are indicative of colon and lung cancer?

In addition to these questions, we will also be experimenting with a hybrid Quantum Convolutional Neural Network. This leads to the following research question:

- Is there an advantage to using a Hybrid Quantum Convolutional Neural Network model, in opposition to a classical method?

II. MOTIVATION

Cancer is a leading cause of morbidity and mortality worldwide [1], with colon and lung cancer being among the most common types [2]. Early diagnosis and treatment of cancer can

significantly improve patient outcomes, and histopathological images of cancerous tissue play a crucial role in the diagnosis and characterization of cancer [3]. However, the interpretation of histopathological images is a complex and time-consuming task, and is often subject to inter- and intra-observer variability [3].

In recent years, there has been increasing interest in using image classification techniques to automatically identify and classify cancerous tissue in histopathological images. These techniques have the potential to improve the accuracy and efficiency of cancer diagnosis, and could potentially be used to assist pathologists in their work [4].

In this paper, we aim to provide a comprehensive overview of the state-of-the-art in this area and to identify potential avenues for future research.

III. State of the Art

Four of the most widely used and state-of-the-art CNN architectures for image classification are ResNet18, AlexNet, SqueezeNet, and VGG16. These architectures have achieved impressive results on a variety of image classification tasks, and have been widely adopted in the field.

a) ResNet18: ResNet18 is a deep CNN architecture that was developed by He et al. (2015) [5]. It is characterized by the use of residual connections, which learns residual functions with reference to the layer inputs, instead of learning unreferenced functions. This way the residual networks let the stacked layers fit a residual mapping. This enables the network to learn more complex features and improve performance on tasks with deeper architectures.

b) AlexNet: AlexNet is another widely used CNN architecture that was developed by Krizhevsky et al. (2014) [6]. It is characterized by the use of large convolutional filters, a large number of parameters nad max pooling and dense layers. This enables the network to learn complex features and achieve high accuracy on image classification tasks.

c) SqueezeNet: SqueezeNet is a lightweight CNN architecture that was developed by Iandola et al. (2016) [7]. It is characterized by the use of 1 x 1 convolutional filters and a reduced number of parameters, which enables it to achieve high accuracy on image classification tasks with a smaller model size and faster training times.

d) VGG16: VGG16 is a deep CNN architecture that was developed by Simonyan and Zisserman (2015) [8]. It is characterized by the use of small 3 x 3 convolutional filters and a large number of parameters, which enables it to learn complex features and achieve high accuracy on image classification tasks.

These CNN architectures have been widely used in the field of image classification, and have demonstrated strong performance on a variety of tasks. In this paper, we will evaluate the performance of these architectures on the identification and classification of colon and lung cancer in histopathological images.

IV. Methods

a) CNN1 & CNN2: These 2 simple CNNs used in this study were designed to be straightforward and easy to implement, with a minimal number of layers and parameters. These models are intended to be used as a baseline to compare against the more complex State of the art and Vision Transformer models. The exact specifications of the networks are the following: CNN1 has 2 convolutional layers, where each have a kernel size of 5. Between each convolutional layer, there is a max pooling layer, with a kernel size of 2. Following these layers, we apply linear transformations until we get our desired output dimension. We use ReLu as our activation function. The only difference between CNN1 and CNN2 is that CNN1 uses 3 convolutional layers instead of 2, with of course an additional pooling layer added.

b) ViT & SimpleViT: ViT is a type of transformer-based neural network architecture that has achieved impressive results on a variety of tasks, including image classification [9]. It is characterized by the use of self-attention mechanisms and a relatively large number of parameters, which enables it to learn complex features and achieve high accuracy on tasks.

SimpleViT is an update of ViT that was designed to be more efficient and easier to implement, with a reduced number of parameters and a simpler architecture [10]. In this simplified architecture there is no dropout, a smaller batch size, and a simpler linear transformation at the end. All these changes result in the SimpleViT method to be trained faster, and potentially better.

c) Classic CNN: This method exists mainly to use as a base and comparison to our Hybrid QCNN. We have 2 convolutional layers, where each layer has a kernel size of 5. The max pooling layers are between the convolutional layers, with a kernel size of 2. After these layers we apply linear transformations until we get our desired output dimension. We use ReLu as our activation function.

d) Hybrid QCNN: Our Hybrid QCNN is a combination of a classical Convolutional Neural Network, and a Quantum Neural Network layer. We use a relatively simple Convolutional Neural Network as our base, and then interchange a classical part with a Quantum layer. The structure of our network is the same as our Classic CNN, but after our convolutions and max poolings we linearly transform the tensor to size 3. We then apply this tensor to a Quantum Circuit, then measure this quantum circuit and linearly transform this result to the required output dimension. The Quantum layer is structured as follows: We use the ZZFeatureMap to encode our tensor of size 3 onto a Quantum Circuit with 3 qubits. This maps our tensor to a Hilbert Space where there is potential for a quantum advantage, due to the entanglements of the qubits provided by the ZZFeatureMap [11]. We then still need a method to give weights to each qubit. For this we use the RealAmplitudes circuit, which can be used to map parameters in our quantum circuit, thus giving us the ability to update weights, just like a classical Neural Network. The entire Quantum Circuit is then used in combination with the EstimatorQNN from Qiskit

[12], which allows us to update our weights accordingly to our input gradient. We use 4 different Hybrid QCNNs, denoted as QCNN1, QCNN2, QCNN3 and QCNN4. The difference between these models, is the repetition of circuit architectures, namely the ZZFeatureMap and RealAmplitudes. This has the potential of giving us a quantum advantage, since the quantum circuit is more complex to simulate. QCNN1 has 1 repetition for the ZZFeatureMap and 1 Repetition for the RealAmplitudes circuit. QCNN2 is the same, but has 2 repetitions for the RealAmplitudes circuit. QCNN3 has 2 repetitions for the ZZFeatureMap, but only 1 for the RealAmplitudes circuit. QCNN4 has 2 repetitions for both the ZZFeatureMap and the RealAmplitudes circuit.

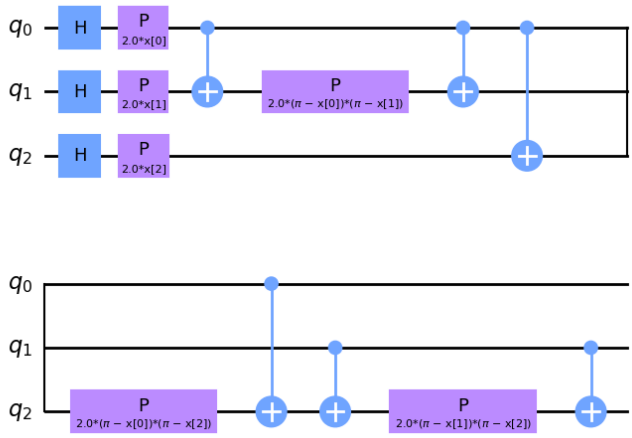

Fig. 1. ZZFeatureMap circuit with 3 qubits, which encodes the data

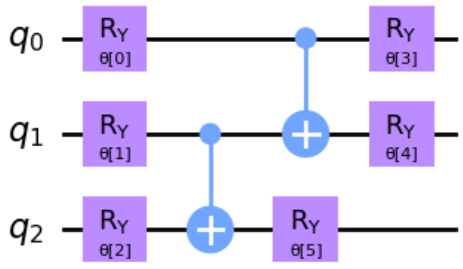

Fig. 2. RealAmplitudes circuit, which allows us to use parameters to update the circuit like a Neural Network

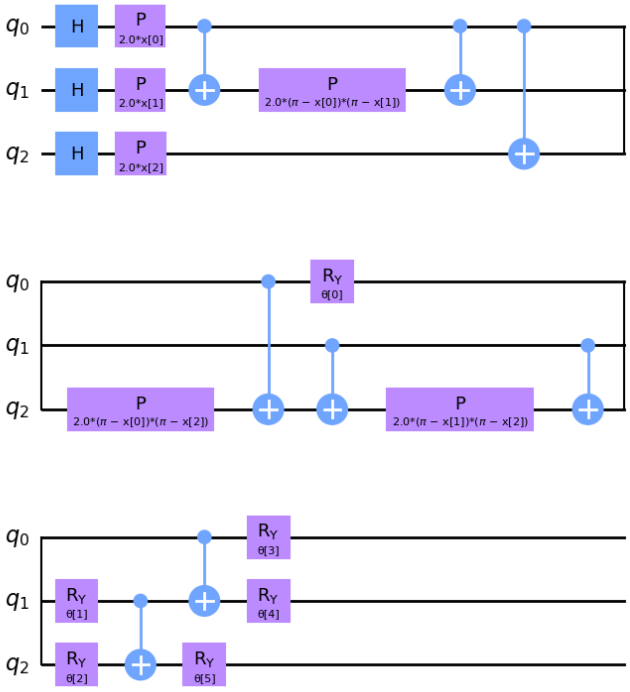

Fig. 3. A combination of the ZZFeatureMap circuit and RealAmplitudes circuit.

V. Data

A. Dataset

For our study, we use the Lung and Colon Cancer Histopathological Image Dataset (LC25000) [13]. This dataset was created specifically to address the need for Machine Learning solutions in the pathology field. It contains in total 25,000 colour images in 5 classes, where each class contains 5,000 images of the following histologic entities: colon adenocarcinoma, benign colonic tissue, lung adenocarcinoma, lung squamous cell carcinoma, and benign lung tissue. All of these images are de-identified, HIPAA compliant and validated. For our study, we discard the lung squamous cell carcinoma, so that we can treat both the lung and colon data as a binary classification problem.

The images were generated from a total of 500 images of lung tissue, and 500 images of colon tissue. These images were then augmented by using left and right rotations up to 25 degrees, and a possible horizontal and vertical flips. Through these augmentations, the dataset was expanded from 1,000 images to 20,000 images for our use.

B. Potential Problems & Solutions

For the following potential problems with histopathological data, we refer to the paper Machine Learning Methods for Histopathological Image Analysis [4], which outlines a overview of papers which describe problems and solutions for histopathological image analysis. From this, we have derived the following problems and solutions:

a) Inconsistency in size and magnification: Every image from a histopathological dataset without augmentations will differ in size and magnification, but since our data is generated from only 250 images per category, there could be a potential problem on overfitting on the detail for images. Therefore we randomly apply random zoom-ins on images that crop images, giving us a variation in detail, making the dataset less prone to overfitting.

b) Image size: Given that our images from the data set are of size 768 x 768, which is too large for our hardware too handle, and that some images are randomly cropped, we resize all images to 128 x 128.

c) Colour variation: As specified in Machine Learning Methods for Histopathological Image Analysis [4], histopathological images are prone to have varying colour samples, which tend to be inaccurate. Therefore, we randomly apply a method that randomly changes the brightness, contrast and saturation.

d) Detail: Different histopathological samples can have varying degrees in detail, due to stains on the glass or less precise instruments. To account for this, we randomly apply gaussian blur.

e) Inconsistent Colour due to different dye use: With histopathological samples, different dyes are used to colour in the samples. Since we extrapolate images from a single image, it is possible that we can overfit on a certain colour, instead of more useful features of the image. We noticed that in the lung dataset, a significant amount of the lung adenocarcinoma data had wildly different colours than the benign lung tissue data. This leads to overfitting on the colour of the sample instead of features of the cells, which is not what we want. To account for this problem, the best way seemed to be to transform the entire dataset into grayscale, keeping the colour information partially out of the data. This in combination with our method that randomly transforms the colour of pixels, the information of the colour of the image is altered enough such that we consider the problem is alleviated.

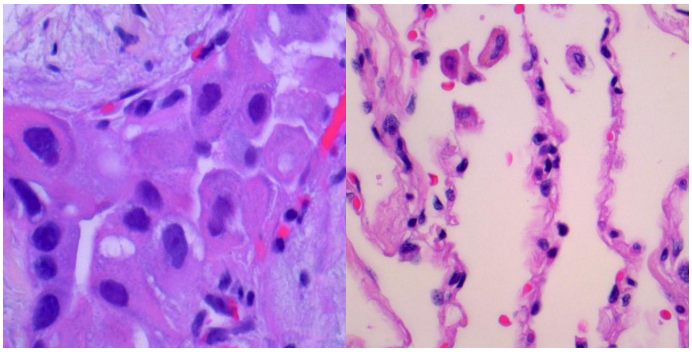

Fig. 4. Difference between dyes of images with cancer (left) and images without (right).

VI. Experiments

In this section we will be explaining which experiments are to be done, and what exact details this experiments entail. We will be explaining 2 sets of experiments, each split into 2 categories, for the State of the Art and CNN1, CNN2, ViT and SimpleViT approaches, and one for all Quantum CNN Hybrid methods, denoted as QCNN1, QCNN2, QCNN3 and QCNN4, together with the comparable Classical CNN. The first set of experiments will compare the train and validation loss for both the lung and colon cancer dataset, and will report the resulting measures on the test set. We do the same for the Quantum approaches, but we will need to use a smaller partition of the data, since Quantum Circuits are exponentially difficult to simulate on a classical computer. In the second set of experiments we will be comparing the explainability of the approaches using saliency maps, Integrated Gradients and Integrated Gradients with SmoothGrad. For all experiments we will be using a train test validation split of 30/35/35, because of the large number of images augmented from the same base image.

A. Training and Testing Measurements

1) Classical Experiments: For this experiment, we experiment with the following models: ResNet18, AlexNet, SqueezeNet, VGG16, CNN1, CNN2, ViT and SimpleViT. We run 10 epochs to train for each of these models. The State of the Art models are not pretrained. As our criterion, we use Cross Entropy Loss, and as optimizer we use the Adam optimizer. Our learning rate is set to 0.0001. We have a batch size of 20. We test with our validation set after every training set, to see whats happening with data that isn't trained on during training. After training is done, we evaluate the models on our test set, and measure the Accuracy, Precision and F1 score. We perform these experiments on both the lung and colon dataset.

2) Quantum Experiments: For this experiment, we experiment with the following models: QCNN1, QCNN2, QCNN3, QCNN4 and Classical CNN. We run 30 epochs for each algorithm, this differs from the first set of experiments, since we want to see what happens when Quantum Circuits are trained longer. We also have a smaller dataset than our first set of experiments, since Quantum Circuits are harder to simulate. We randomly select 100 samples from our lung dataset to perform our experiments on. We use the same criterion and learning rate as in the first set of experiments. We have to set our batch size to 1, again thanks to the difficulty of simulating Quantum Circuits. We also use the validation set in the same way. We only test these methods on the lung dataset.

B. Saliency Maps and Explainability

1) Classical Experiments: For this experiment, we experiment with our best and worst performing model: ResNet18 and CNN1. We only perform these experiments with 2 models, due to time and hardware constraints. We only can run these tests on 4 images, again due to time and hardware constraints. For each of these images and models, we compute a class saliency map, specific to the given image, and predicted class. We do this as described in Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps

[14]. For further visualizations, we also provide Integrated Gradients, which calculates the gradient of the model's output with respect to its input features. We do this as described in Axiomatic Attribution for Deep Networks [15]. We also do the Integrated Gradients again, but with the smoothgrad square option enabled. This adds gaussian noise with a standard deviation of 0.2, to the input image 100 times, computes the attributions for 100 images and returns the mean of the squared attributions across 100 images, as described in SmoothGrad: removing noise by adding noise [16].

As previously stated, we have 4 images to run these tests on. We have 2 images for both the lung and colon dataset respectively. The first picture we provide for testing is a picture which is labeled as containing cancer, which ResNet18 classified correctly, but CNN1 didn't. The second picture we provide for testing is a picture which is labeled as not containing cancer, which ResNet18 correctly classified, but CNN1 again didn't. Do this for both the colon and lung dataset.

2) Quantum Experiments: For this experiment, we experiment with our best performing Quantum Model, QCNN4 and our Classical CNN model. We again perform these experiments with 2 models, due to time and hardware constraints. We only run the explainability methods on 1 image, which has been classified as having cancer, where QCNN4 wasn't able to classify the image correctly, but the Classical CNN model was.

C. Hardware & Software

Our experiments are performed on a Windows 11 machine using 16.0 GB of RAM, AMD Ryzen 7 5700G, clocked at 3801 Mhz. GPU: GeForce RTX 3070Ti. Experiments are using CUDA when possible. We use the Python programming language, Qiskit for our Quantum Experiments, the PyTorch framework for our Deep Learning models and Captum for our Saliency Maps and Explainability Experiments.

VII. Results

A. Training and Testing Measurements

1) Lung Dataset: In this section we provide the results from our experiments on the lung dataset.

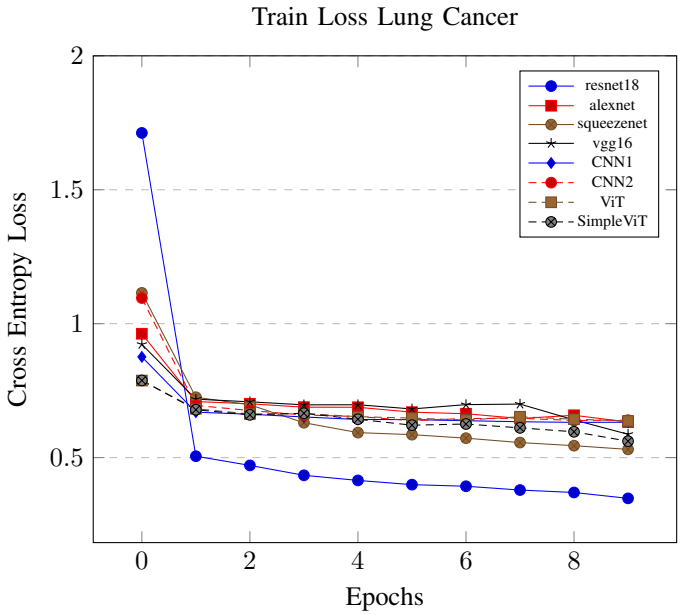

Fig. 5. Train Loss Lung Cancer

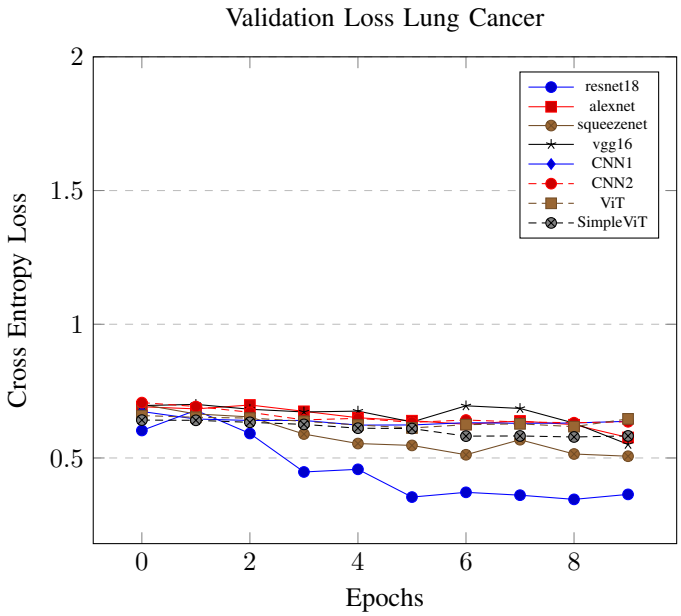

Fig. 6. Validation Loss Lung Cancer

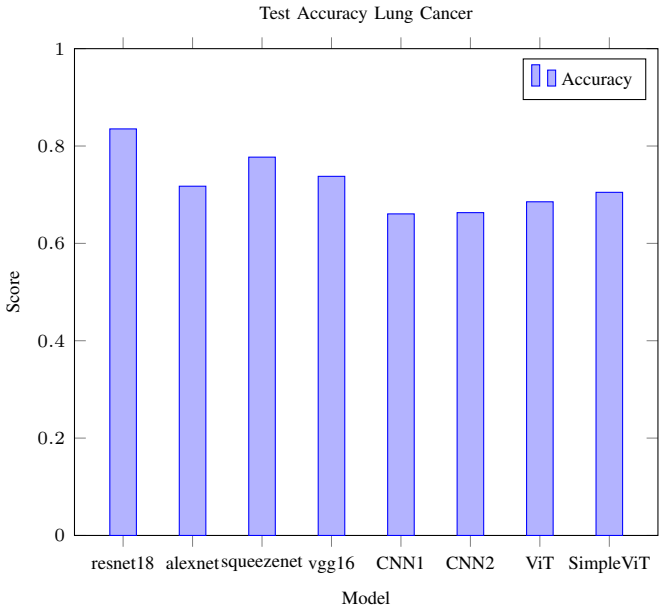

Fig. 7. Test Accuracy Lung Cancer

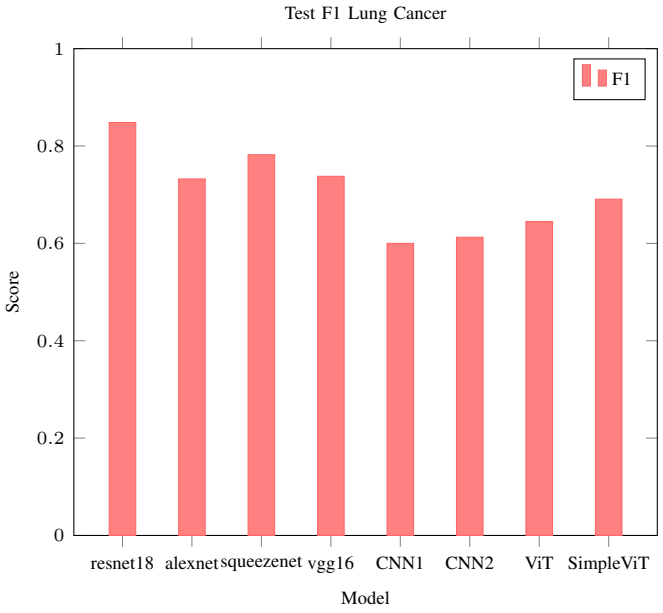

Fig. 9. Test F1 Lung Cancer

2) Colon Dataset: In this section we provide the results from our experiments on the colon dataset.

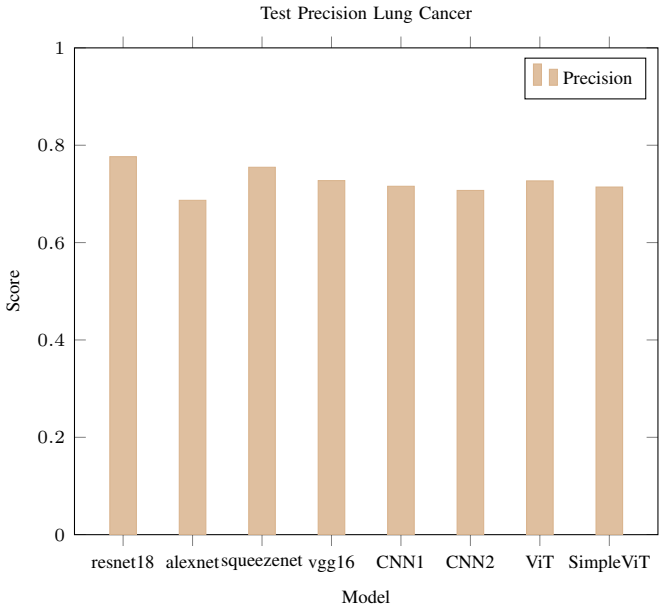

Fig. 8. Test Precision Lung Cancer

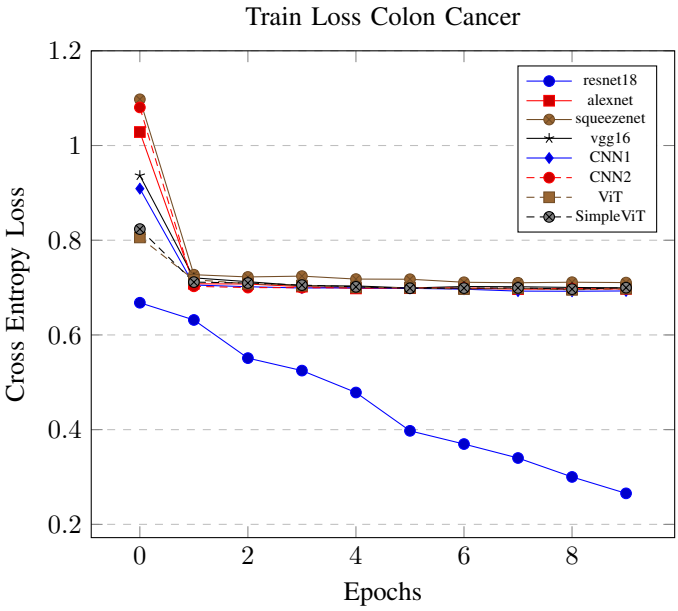

Fig. 10. Train Loss Colon Cancer

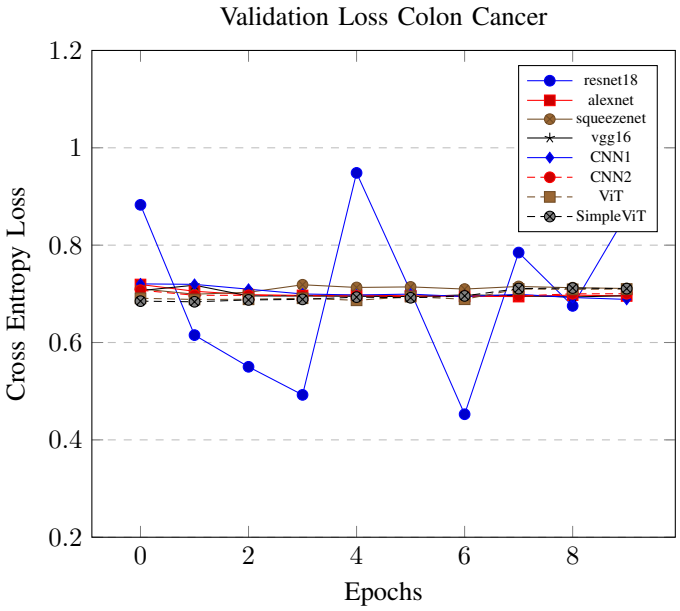

Fig. 11. Validation Loss Colon Cancer

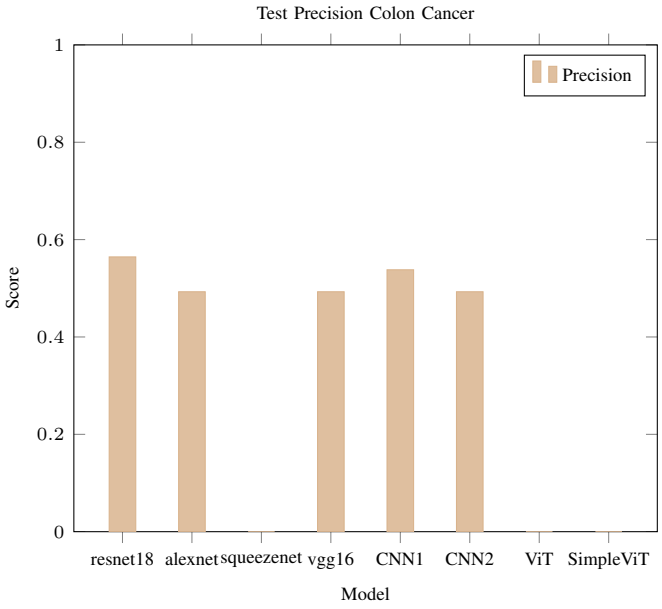

Fig. 13. Test Precision Colon Cancer

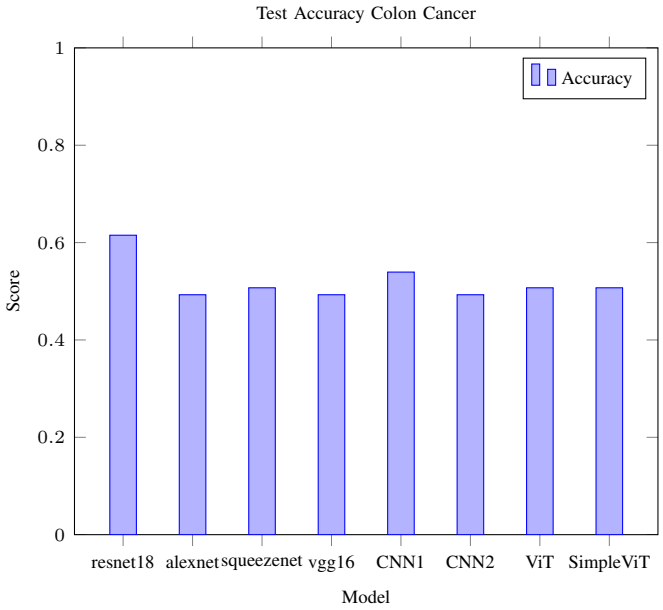

Fig. 12. Test Accuracy Colon Cancer

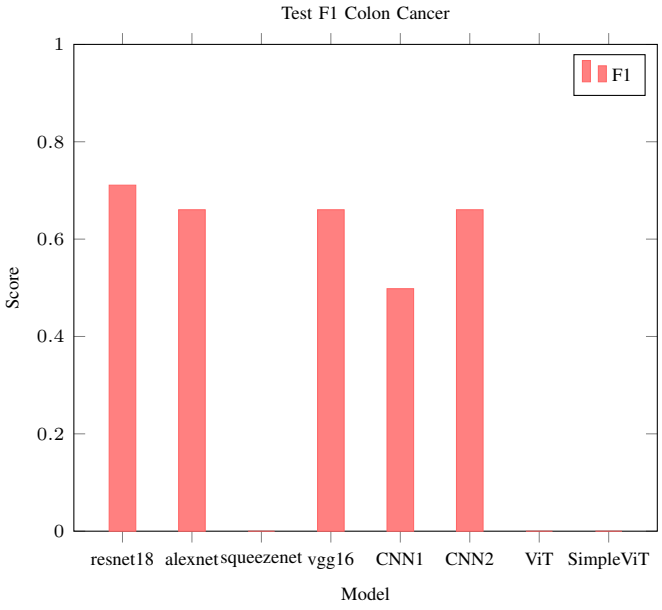

Fig. 14. Test F1 Colon Cancer

3) Quantum Experiments: In this section we provide the results from our experiments with the Quantum models.

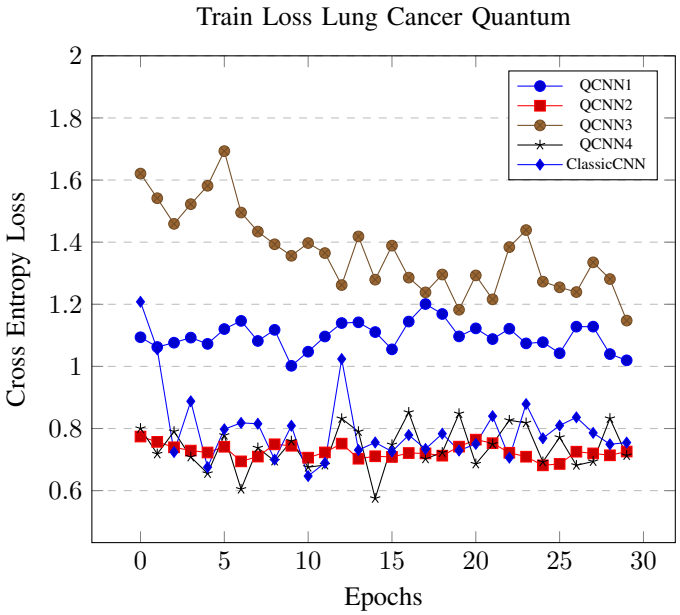

Fig. 15. Train Loss Lung Cancer Quantum

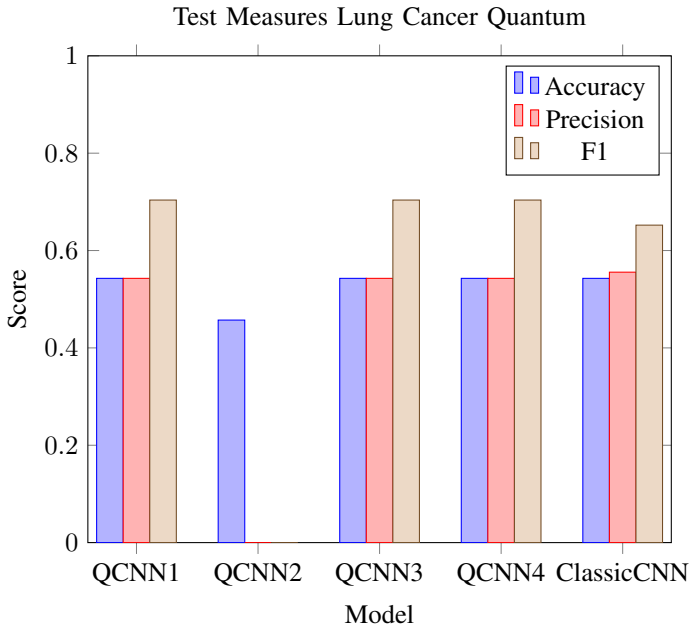

Fig. 17. Test Measures Lung Cancer Quantum

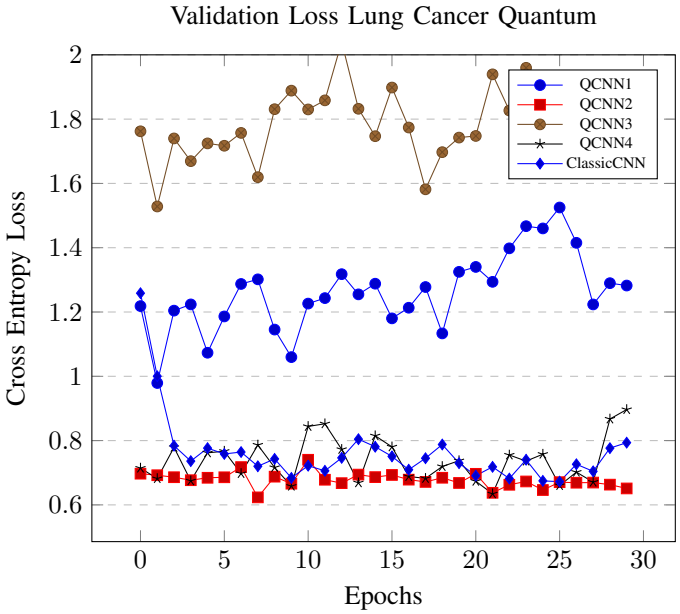

Fig. 16. Validation Loss Lung Cancer Quantum

B. Saliency Maps and Explainability

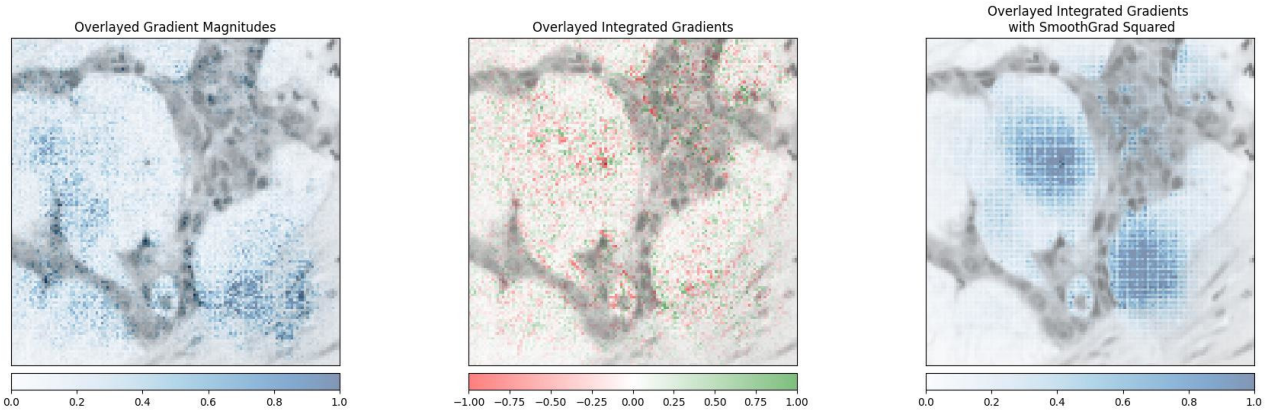

Fig. 18. ResNet18, Colon without cancer, classified correctly

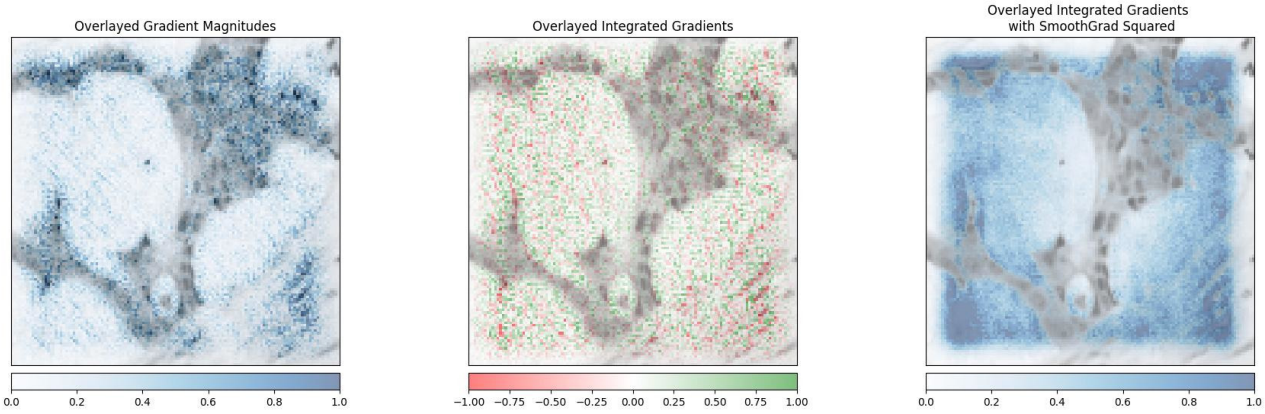

Fig. 19. CNN1, Colon without cancer, classified incorrectly

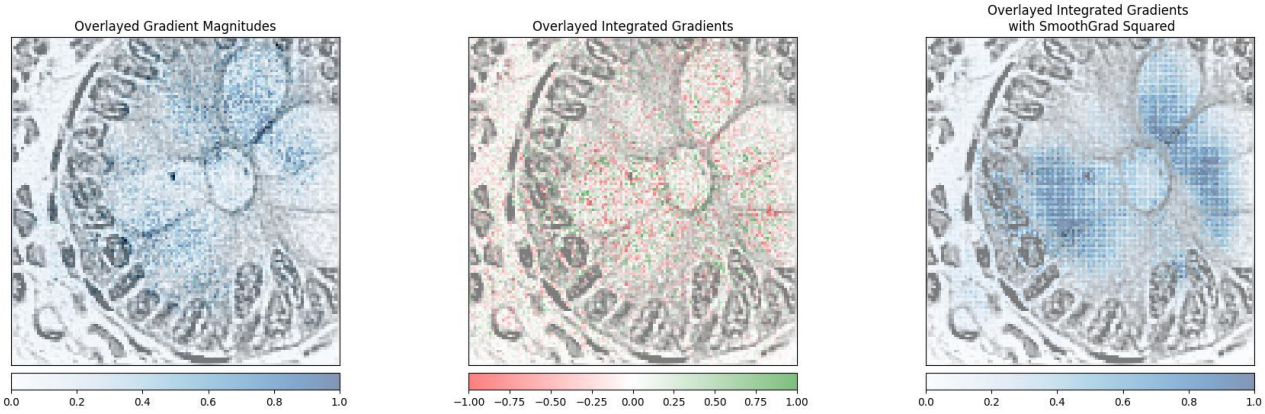

Fig. 20. ResNet18, Colon with cancer, classified correctly

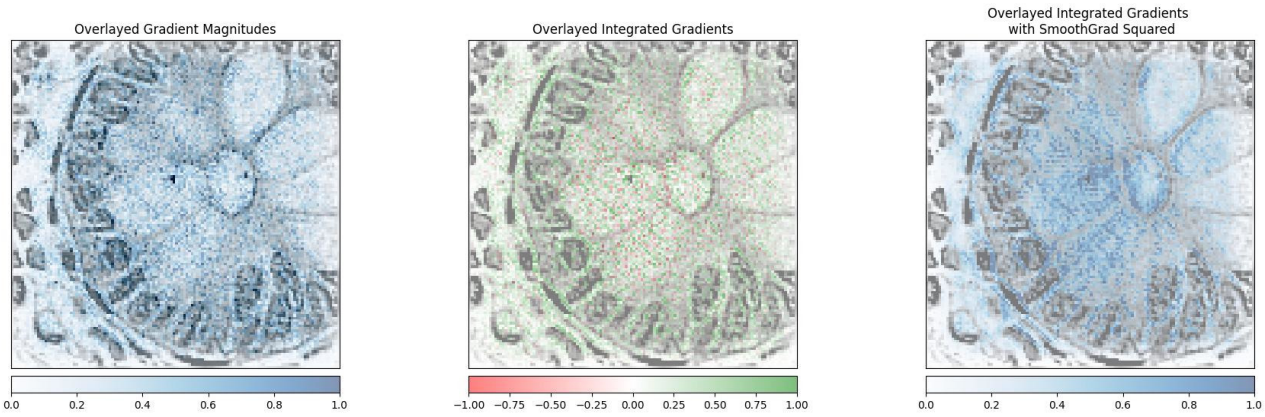

Fig. 21. CNN1, Colon with cancer, classified incorrectly

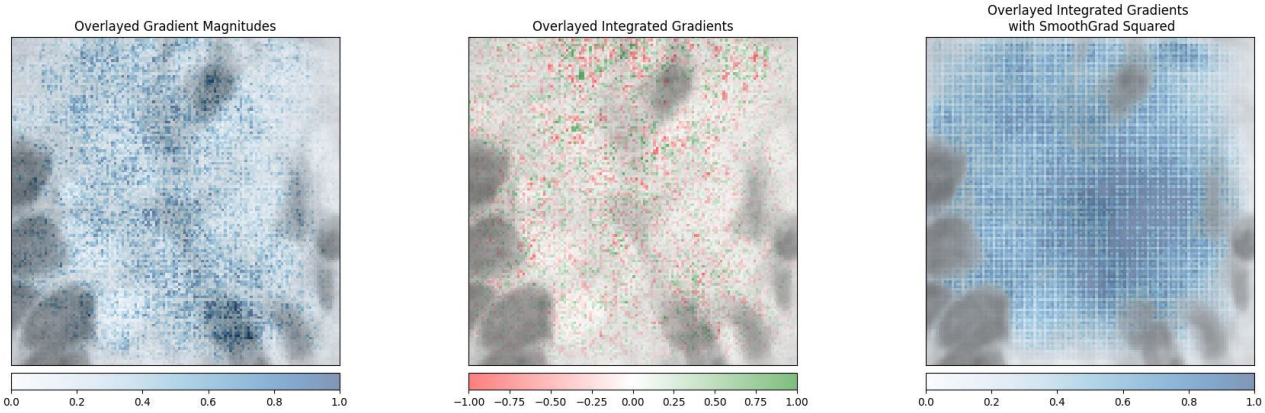

Fig. 22. ResNet18, Lung without cancer, classified correctly

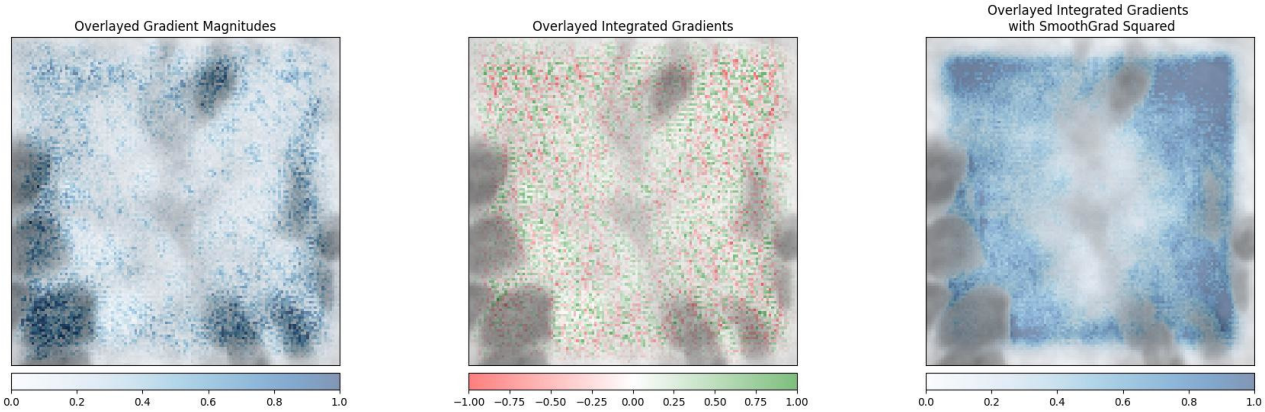

Fig. 23. CNN1, Lung without cancer, classified incorrectly

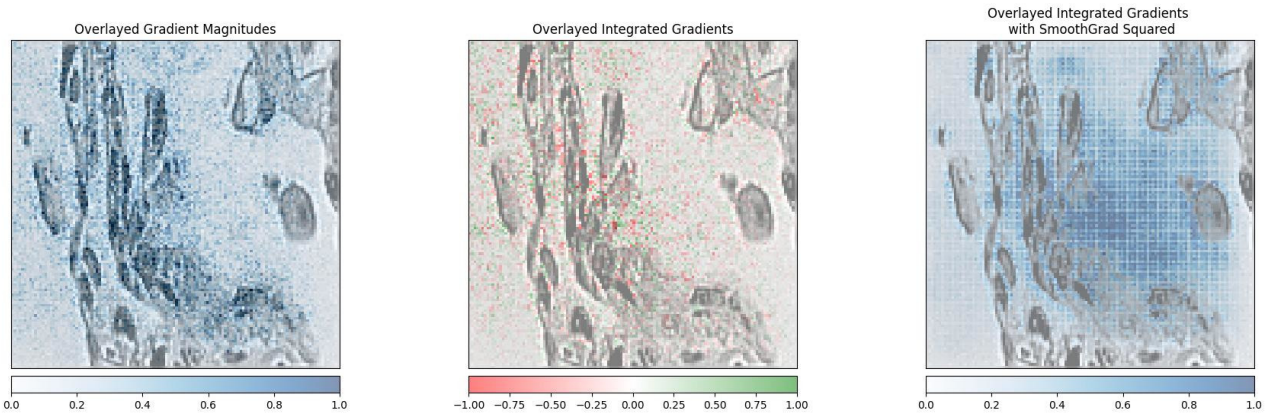

Fig. 24. ResNet18, Lung with cancer, classified correctly

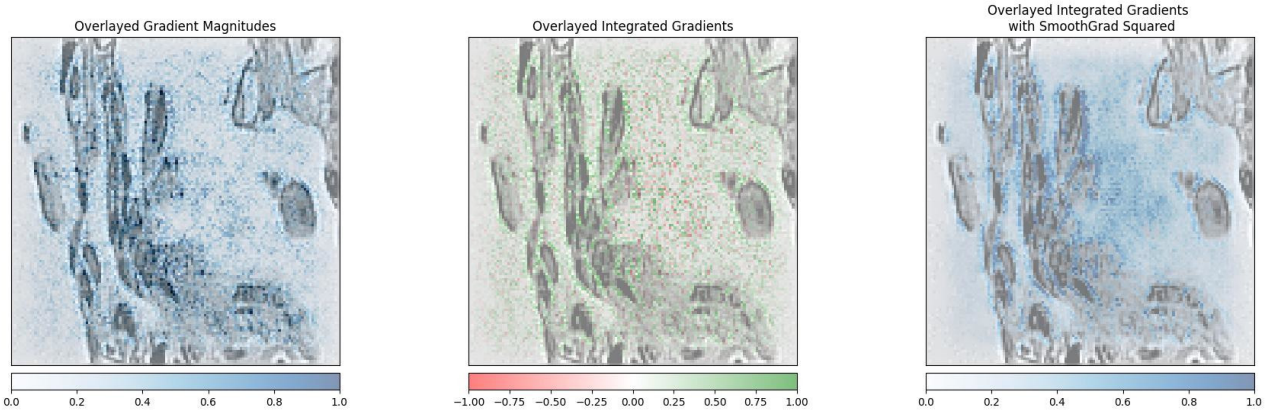

Fig. 25. CNN1, Lung with cancer, classified incorrectly

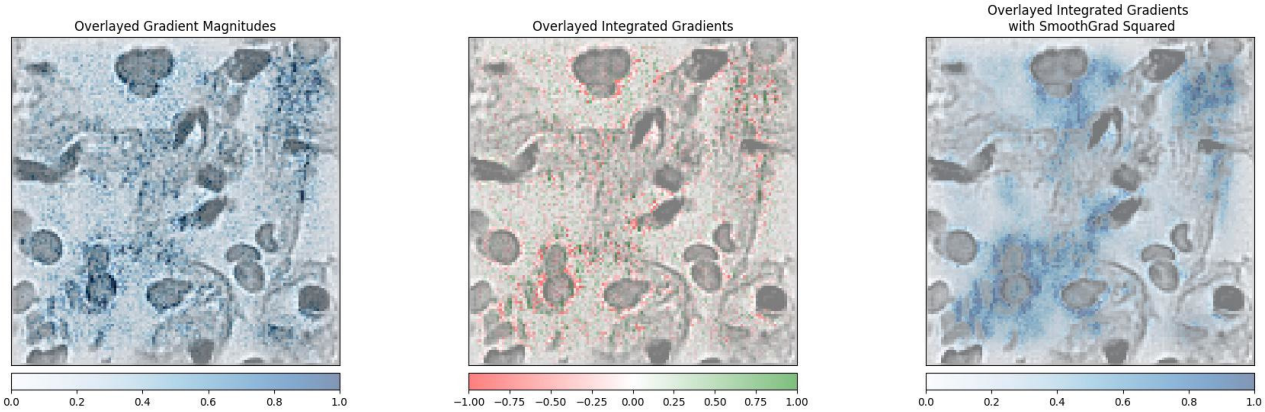

Fig. 26. Classic CNN, Lung with cancer, classified correctly

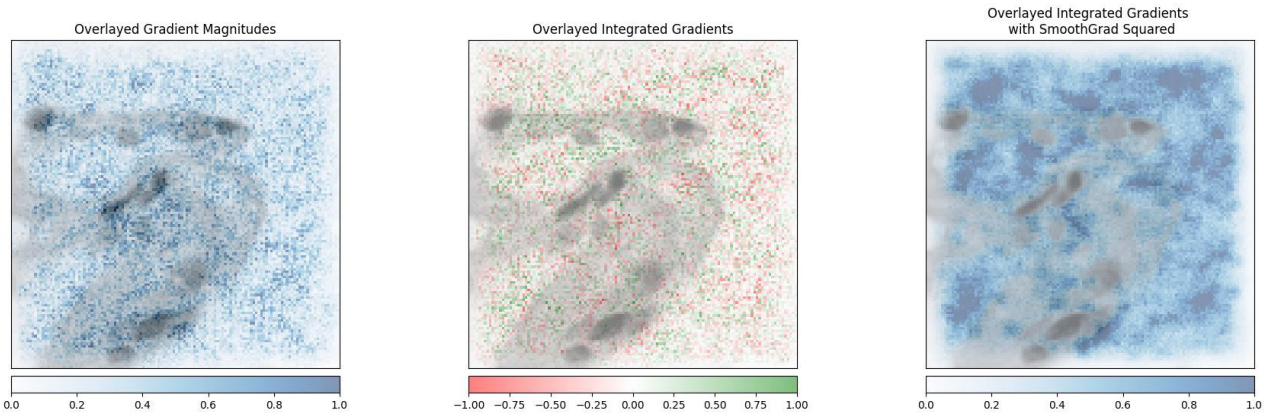

Fig. 27. QCNN4, Lung with cancer, classified incorrectly

VIII. Discussion

In this paper, we explored the use of image classification techniques for the identification of histopathological colon and lung cancer. Our results showed that the use of different convolutional neural network architectures were effective for classifying colon and lung cancer images.

One goal for this paper was to identify the most effective approaches for classifying histopathological images for colon and lung cancer, and provide insights into challenges with these models.

The results indicate that ResNet18 is the best performing model for both the lung and colon dataset. In the Lung dataset, ResNet18 performs all other models for both the training loss, and validation loss, giving it a clear advantage. We see that for both the test accuracy, precision and F1 score, ResNet18 outperforms all models. The second best model is the SqueezeNet model. We see that it ranks second in the both the training and validation loss. We also notice that it ranks second in the test accuracy, precision, and F1 scores. Our worst performing model is not so clear, however. We notice that CNN1 seems to be performing the worst in the accuracy and F1 metric, but that alexnet performs worse in the precision metric.

For our Colon dataset, we notice again that ResNet18 is the best performing model in our training loss, but we see that during the validation , it wildly varies per step. We notice here that ResNet18 probably overfits on a partition of the data, but that all other models do not. Even though ResNet18 seems to overfit on the training data, we still see that it performs the best out of all methods in the test set, having the highest accuracy, precision and F1 score. Our second best method can not be decided from the train and accuracy loss, since these are all almost equal. We see that in the test dataset, the second best method is still very unclear, since most of the graphs that have numbers are all fairly equal. Therefore we say that a second best method remains undecides. Our worst performing methods is a mix between SqueezeNet, ViT and SimpleViT. We choose these since they all were unable to correctly identify images were there was no cancer. This is why the accuracy is around 50%, but both the precision and F1 metric are 0, since it wasn't able to identify any image without colon cancer.

A second goal for this paper was to experiment with a hybrid of Quantum and Convolutional Neural Networks, and see if they had any advantage or not.

In our Quantum experiments, we noticed that according to the train and test loss, it seems that the best performing Quantum methods are QCNN2 and QCNN4, where QCNN2 had 1 repetition for the ZZFeatureMap circuit, and 2 repetitions for the RealAmplitudes circuit. QCNN4 had 2 repetitions for both the RealAmplitudes circuit and the ZZFeatureMap circuit, making it the hardest to classically simulate. When we look at the metrics on the test dataset, we notice a few slight differences. We see that QCNN2 has no precision and F1 metric, this is again due to the fact that QCNN2 wasn't able to correctly identify any images without lung cancer. We do notice one interesting thing however, the F1 measure for all Quantum models (with exception for QCNN2 ofcourse), is higher than the F1 measure for the classical CNN. This could indicate that there is potential to using Hybrid Quantum Convolution Neural Network models for classifying images.

Our third goal for this study was to see if we could explain why a model chooses to classify a images as having cancer or not. This could give insight into the characteristics of histopathological images, that are indicative of colon and lung cancer.

We see that for all the image, the correct method usually seems to focus more on specific areas within the cells, as seen in Figure 20. We notice that for the the images that were classified incorrectly, the models usually seem to focus more generally, so that they don't classify on specific features of the image. This difference can be clearly seen when comparing Figure 26 and Figure 27. We see that the classical model, which classified the image correctly, focusses more on specific areas within the image, while the quantum model (seen most clearly in the Overlayed Integrated Gradients with

SmoothGrad Squared image). Seems to have a very general focus.

One potential limitation of our paper is that we didn't compare every single classification model that is there, and that we only trained our classical models for 10 epochs. It is therefore possible that other models or feature extraction methods could have achieved better results. Future research should compare more state of the art models to improve the accuracy of image classification for histopathological colon and lung cancer.

Another potential limitation is that our paper only focused on image classification and did not consider other tasks such as segmentation or object detection. Further research could explore the use of these techniques in combination with image classification to improve the accuracy and efficiency of cancer diagnosis.

Another potential limitation of our paper is that the Quantum Models used were very limited in their design, and were only tested on 100 datapoints. Future research should use more data to see the potential benefit of these models.

Another potential limitation is that the data used for this research was flawed, and multiple methods were applied to reduce the bias inherent with the data. Since there wasn't an expert available during the writing of this paper, it could be that certain important aspects of these flaws were overlooked. Future research should communicate with experts on this subject to more carefully avoid bias in the data.

Another potential limitation is that there wasn't an expert available to evaluate the correctness of the explainability part of this paper. Even though we see a pattern in what correct models classify and what they use to classify images, it could be that the models are overfitting on some other section of these images, giving us a bias. Future research should communicate with experts to explain and limit these potential flaws.

Overall, our study demonstrates the potential of image classification techniques for the identification of histopathological colon and lung cancer. Further research is needed to further address the explainability of these models, and evaluate the accuracy of the explanations. We also see that there is a potential benefit in using Hybrid Quantum Convolutional Neural Network models, and that future research should further investigate this option.

IX. Conclusion

In conclusion, this paper has demonstrated the feasibility of using image classification techniques for the identification of histopathological images of colon and lung cancer. Our results showed that the use of different Convolutional Neural Network architectures were effective at classifying histopathological colon and lung cancer images.

While there are limitations to our paper, our results suggest that image classification has the potential to be a valuable tool for the diagnosis of cancer using histopathological images.

Our results further suggest that there is potential to using Hybrid Quantum Convolutional Neural Networks for image classification, but that this method has to be explored further.

Our results also indicate that there is a clear difference in focus between a model that correctly classifies an image, and one that does not. Further communication has to be done with experts to validate this conclusion.

X. Additional Content

a) What I did for the project: I imported and trained 4 state of the art models on the histopathological dataset. I made 2 very simple CNN architectures to contrast these models, and I implemented 2 vision transformer models. I also made 4 hybrid Quantum Convolutional Neural Network models, with each having slightly different parameters, and a simple comparison classical CNN. I also cleaned the data, which I didn't see anyone doing before. I noticed that there were a few biases in the data, and tried to alleviate them using previously researched methods for this exact topic and my own ideas. I wanted to see why a model behaved the way it did, so I used saliency maps and integrated gradients to further explore this avenue.

b) What I Learned: I learned to use multiple different models in the pytorch framework, I learned to design simple CNN architectures, I learned to combine a Quantum Circuit with a classical CNN model. I learned to plot graphics into a latex format. I learned to use certain methods to see what CNN models focus on during classification.

c) Resources: I got inspired by multiple resources at the same time, some having more to do with debugging than the actual methodology itself. My contributions for all: using it on Histopathological dataset. Applied Simple CNN architectures and State of the art models to histopathological data. Neural Network Tutorial Used transformer as comparison to other state of the art methods. Transformer Package used Implemented a quantum circuit that generates entanglement, which is used at the end of a CNN. Qiskit Tutorial QCNN Applied to explain histopathological data. Captum Tutorials

d) Starting point: My experience is limited to the projects I have done in my bachelor, at DCAS. I have also participated in the IVP course, but I didn't have that much experience with histopathological data, data cleaning and preperation, saliency maps, transformers, and Hybrid Quantum Convolutional Neural Networks.

e) Time it took to work on the project: All the experiments shown above are carried out on the hardware I described in the hardware section. In total, the experiments that were carried took around 6-7 hours. This doesn't include bugs, problems I encountered and general programming time. I had to rerun these experiments around 9-10 times, making the total amount of hours I spend on these experiments alone around 60-70. The programming itself took around 30 hours. And then, given the time it took me to write this report, the presentations, and all of it added up. I would say it took around 100-120 hours to finish this project.

f) Issues and solutions: Biggest issue I faced was probably the bias in the data. I did some research into bias reduction methods specifically for histopathological datasets, and applied them to my dataset. I then also grayscale the dataset, since I noticed that different dies were used for different samples, and this gave my models an unfair advantage. I also faced that Quantum Models weren't able to run with the entire dataset, so I used a smaller, but still balanced part of the data so that I could run my Quantum Models, in comparison with another classical model on this smaller partition. I also when I started the project wanted to implement even more recent State of the art methods like CoCa. The problem I found was that were was little to no documentation of how this method worked, and the documentation that there was was hyper technical and mathematical, of which I couldn't quickly enough get a grasp on it. I sadly had to discard this idea. Another problem with CoCa was that it was only trained on very expensive GPU machines, which I don't have access to.

REFERENCES

[1] World health organization, cancer, Feb 2022.

[2] Rebecca L. Siegel, Kimberly D. Miller, Hannah E. Fuchs, and Ahmedin Jemal. Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, 72(1):7–33, 2022.

[3] S C Hiom. Diagnosing cancer earlier: Reviewing the evidence for improving cancer survival. *British Journal of Cancer*, 112(S1), 2015.

[4] Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and Structural Biotechnology Journal*, 16:34–42, 2018.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[6] Alex Krizhevsky. One weird trick for parallelizing convolutional neural networks. *CoRR*, abs/1404.5997, 2014.

[7] Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf, Song Han, William J. Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <1mb model size. *CoRR*, abs/1602.07360, 2016.

[8] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.

[9] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[10] Lucas Beyer, Xiaohua Zhai, and Alexander Kolesnikov. Better plain vit baselines for imagenet-1k. *arXiv preprint arXiv:2205.01580*, 2022.

[11] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

[12] A-tA-v et al IBM. Qiskit: An open-source framework for quantum computing, 2021.

[13] L. Brannon Thomas Catherine P. Wilson Lauren A. DeLand Stephen M. Mastorides Andrew A. Borkowski, Marilyn M. Bui. Lc25000 lung and colon histopathological image dataset.

[14] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034*, 2013.

[15] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In *International conference on machine learning*, pages 3319–3328. PMLR, 2017.

[16] Daniel Smilkov, Nikhil Thorat, Been Kim, Fernanda Viégas, and Martin Wattenberg. Smoothgrad: removing noise by adding noise. *arXiv preprint arXiv:1706.03825*, 2017.