LOCALLY RUN WEB-BASED APP FOR INTERPRETABLE BREAST CANCER DIAGNOSIS FROM HISTOLOGY IMAGES

Jeremiah O. Fadugba;Oluwatoyin Y. Sanni;Moshood O. Yekinni;

;

IDRC Grant/ Subvention du CRDI: 109187-002-Laying the foundations for artificial intelligence for development (AI4D) in Africa

Locally-run Interpret-able Breast Cancer diagnosis from Histology Images

Jeremiah O. Fadugba, Oluwatoyin Y. Sanni, Moshood O. Yekinni

March 31, 2021

1 Introduction

In recent years, Artificial Intelligence (AI) has made tremendous advances in identifying diseases from radiology images. Convolutional Neural Networks (CNNs), a class of deep learning algorithm trained on large volumes of labelled radiological images, have led these advances. Various results has shown that CNNs improves the speed, accuracy and consistency of diagnosis [7].

However, the adoption of deep learning diagnostic system by healthcare practitioners is prevented by two major challenges: 1) interpreting the prediction outputs from a deep learning network is not trivial, and 2) Privacy of patient data is not guaranteed when using online services that provides deep learning models. These challenges are why healthcare practitioners remain wary of using AI-driven diagnostic tools [10]

A medical practitioner cannot fully trust the CNN network except it can explain its reason for its decision, semantically or visually. Earlier methods in machine learning are transparent in how they compute the predictions but deep learning models are not so. Deep learning models automate the hand crafted feature engineering and hence no knowledge of how the predictions are computed. Diagnosing with CNN involves studying image regions that contribute most to prediction outputs at the pixel level. In interpretability, we expect the CNN to explain its decision at the object-part level. Given an interpret-able CNN, previous work reveals the distribution of object parts that are memorized by the CNN for object classification [14].

In addressing the second motivation behind this work, a CNN model is being deployed through a server client architecture which requires the data to be sent online to the model for prediction. Deep learning models are large in memory and computation. Hence, they need large computing power like GPUs that requires an existing remote servers. To get predictions, doctors have to upload the patient radiological scan through the internet exposing it to the risk of data privacy. What we have done in this work instead, is to use solutions that make such models run locally thereby solving the issue of privacy. This technique also solves the challenge faced in developing countries where access to internet could be expensive.

2 Literature Review

In Breast Histopathological Image analysis, various approach have been used for image segmentation, classification and feature extraction. Artificial Neural Network, especially deep learning have been widely used for these tasks. This review is limited to the related work that uses the BACH dataset [1]

In the experiment of [4], achieving an accuracy of 85% for the muliclass and 95% on a binary class (carcinoma or non-carcinoma) by using an Inception-V3 based deep learning network. Patches were extracted from the sample based on the density of nuclei present and reject other patches that did not meet up to the threshold of nuclei density.

[9] proposed a patch based technique that consists of a patch-wise convolutional neural network (CNN) and an image-wise CNN. The former acts as an auto-encoder that extracts the most salient features while the later acts as a classification technique by first extracting the global information in the image. A similar work to this two stage CNN is in [6] that uses AlexNet as a feature extraction technique and the second sage uses a support vector machine (SVM) as a classification achieving a 98.4% accuracy.

The use of transfer learning based approach in [13] was employed via Inception-V3 and ResNet-50 pre-trained on ImageNet database. Similarly, [3] uses Inception and ResNet-V2 without the use of patch extraction or any data augmentation to achieve a test accuracy of 90%.

Finally, [5] proposed a transfer learning approach with global pooling for the multi-classification problem. In this work, they made use of a pre-trained Xception network and a global average pooling was used on the extracted feature from a convolutional layer after the mac-pooling layers. Patch extraction is a common technique used in various literature during data processing. It s worth mentioning that this work did not make use of patch extraction as used by others but rather rely heavily on various data augmentation techniques after the images were down-sampled.

3 Experiment and Results

This section details Some of the experiments and result performed in this work.

3.1 Dataset

We made use of the high resolution H&E breast histology data from the Breast Cancer Histology Challenge (BACH) 2018 repository [2].The dataset consists of 400 images that are evenly distributed between four categories; (i) normal (ii) benign, (iii) insitu carcinoma and (iv) invasive carcinoma. Each image in the dataset is of RGB color channel with size of 2048 x 1536 pixels and a pixel scale of $0.42\mu m$ x $0.42\mu m$ all represented in .tiff format.

Processing: Two key processing technique was used (i) staining normalization [8] which solves the appearance variability in histopathology images and

(ii) random H&E augmentation [11] that adjust the RGB color space of the tissue into H$E color space.

3.2 Data Augmentation

Considering the low data sample we have and a multi-class problem, it get difficult for DNN to perform well. Data augmentation is a technique to artificially expand the size of our dataset by creating a modified version of the initial data. In this work, we made use of In techniques such as horizontal and vertical flips, rotation, contrast adjustments and brightness correction. These were applied to enlarge the dataset and improve the classification performance

3.3 Classification Method

The task of this work is not to achieve the state-of-the-art method in classification of breast cancer diseases however, our methodology lies in having a strong baseline classification. Various approached were made to achieve a higher accuracy and performance for the multi-class problem. Most of the approaches made were re-implementation of some papers reviewed in earlier section.

3.4 Experiment

The experiment in this work followed a standard procedure where we tested out different models on the data. With the pre-processing stage constant along all experiments, we report our results on (i), models with augmentation techniques and (ii), models without the use of augmentation. In the former, we use the augmentation as described in 3.2 which results in having 560 data samples for training and validation across in each category and a hold out test set of 20 samples. While the latter make use of the 80 samples for training and validation for each category with 20 held out test set. The models are trained using keras framework with a learning rate of 0.01 using stochastic gradient descent optimizer with momentum of 0.9. For callbacks, we set learning rate scheduler with a factor of 0.5. The experiment was performed on a NVIDIA Tesla P100 machine.

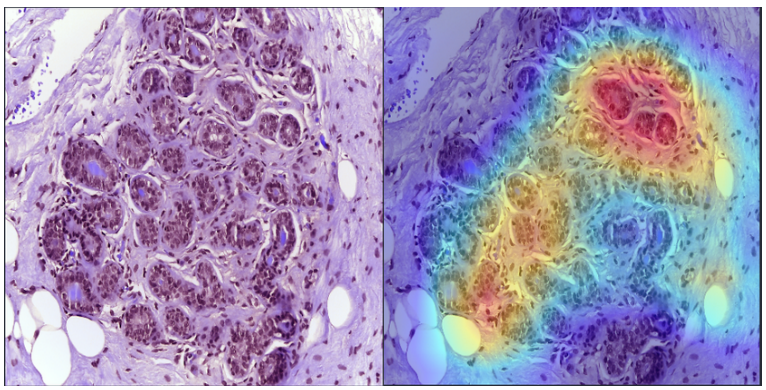
	With Augmentation			**W/O Augmentation**		
Model	**Precision**	**Recall**	**Accuracy**	**Precision**	**Recall**	**Accuracy**
XceptionNet	0.83	0.22	0.62	0.79	0.21	0.60
XceptionNet_GAP	0.65	0.19	0.55	0.64	0.16	0.53
InceptionNet	1.0	0.24	0.79	1.0	0.23	0.77
InceptionNet_GAP			0.25			0.24

Table 1: Experiment results on different model comparison with and without data augmentation.

On all experiments, we observe that all models gave a high precision but low recall despite having high accuracy. This performance can be attributed to the inadequate samples, the quality of the augmentation techniques and also, multi-class problems are difficult in a low data regime.

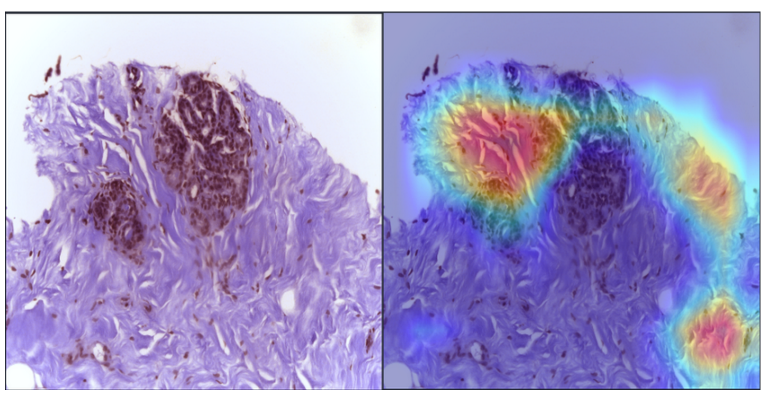
3.5 Interpretability Method

As mentioned earlier in 1, we want to take a step further in this work to explain the decisions behind our model predictions. In other words, making our algorithm transparent. One technique of doing so is through visually exploring the gradients of the computations by using Gradient Weighted Class Activation Mapping (GRAD-CAM) [12] It works by looking at the gradient flow of any target class in the final convolutional layer. It produces a coarse localization map showing the regions that is most important in the image for the predicted target this is shown in figure **??**.



(a) Original Image

(b) Heatmap



(c) Original

(d) Heatmap

Figure 1: GRAD-CAM visualization technique

With the visualization, our aim is to guide pathologist to inspect where the algorithm focuses on in making its decision. These can be beneficial in a deployed case where the visualization can be approved as the algorithm being right or wrong.

4 Conclusion

Deep learning algorithms especially with CNNs, have improved various state of the art techniques in diagnosing diseases from pathology images. While these methods achieve comparable performance with human pathologists when trained on large volume of data, they are not sufficient to inform a pathologist for a decision. Hence, this work after having a baseline model for diagnosing disease, adds another component of visual explanation of the algorithms decision to guide a pathologist in decision making. We observe however, to improve the algorithms predicting power, careful considerations need to be paid to how histology images are pre-processed. While we made use of the latest technique in processing, much research and work need to be done in extracting patches of the slides and at the same time better ways of normalizing the data.

5 Acknowledgment

We are thankful for the grant (id: 109187-002) support from International Development Research center (IDRC) in partnership with Knowledgeforall foundation, Deep Indaba and Artificial Intelligence for development (AI4D). Thanks to Google cloud for the cloud credit support. Thanks to Joseph T. Aaron, a medical consultant for his continuous support in providing domain expertise. Special thanks to our mentor, Fernandez-Reyes, Delmiro, for his input during this project.

References

[1] Teresa Araújo, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. Classification of breast cancer histology images using convolutional neural networks.

[2] Guilherme Aresta, Teresa Araújo, Scotty Kwok, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex, Bahram Marami, Marcel Prastawa, Monica Chan, Michael Donovan, Gerardo Fernandez, Jack Zeineh, Matthias Kohl, Christoph Walz, Florian Ludwig, Stefan Braunewell, Maximilian Baust, Quoc Dang Vu, Minh Nguyen Nhat To, Eal Kim, Jin Tae Kwak, Sameh Galal, Veronica Sanchez-Freire, Nadia Brancati, Maria Frucci, Daniel Riccio, Yaqi Wang, Lingling Sun, Kaiqiang Ma, Jiannan Fang, Ismael Kone, Lahsen Boulmane, Aurélio Campilho, Catarina Eloy, António Polónia, and Paulo Aguiar. Bach: Grand challenge on breast cancer histology images. *Medical Image Analysis*, 56:122 – 139, 2019.

[3] Carlos A. Ferreira, Tânia Melo, Patrick Sousa, Maria Inês Meyer, Elham Shakibapour, Pedro Costa, and Aurélio Campilho. Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet v2. In Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny, editors, *Image Analysis and Recognition*, pages 763–770, Cham, 2018. Springer International Publishing.

[4] Aditya Golatkar, Deepak Anand, and Amit Sethi. Classification of breast cancer histology using deep learning, 2018.

[5] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters. Breast cancer diagnosis with transfer learning and global pooling. In *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 519–524, 2019.

[6] Kevin Kiambe. Breast histopathological image feature extraction with convolutional neural networks for classification. 4:4–12, 06 2018.

[7] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q. Nelson, Greg S. Corrado, Jason D. Hipp, Lily Peng, and Martin C. Stumpe. Detecting cancer metastases on gigapixel pathology images, 2017.

[8] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, Xiaojun Guan, C. Schmitt, and N. E. Thomas. A method for normalizing histology slides for quantitative analysis. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1107–1110, 2009.

[9] Kamyar Nazeri, Azad Aminpour, and Mehran Ebrahimi. Two-stage convolutional neural network for breast cancer histology image classification. *Image Analysis and Recognition*, page 717–726, 2018.

[10] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier, 2016.

[11] Arnout Ruifrok and Dennis Johnston. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol*, 23, 01 2001.

[12] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 128(2):336–359, Oct 2019.

[13] Sulaiman Vesal, Nishant Ravikumar, AmirAbbas Davari, Stephan Ellmann, and Andreas Maier. Classification of breast cancer histology images using transfer learning, 2018.

[14] Zijie J. Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, and Duen Horng Chau. Cnn explainer: Learning convolutional neural networks with interactive visualization, 2020.