Comp150-DNN Project Proposal

Jeremy Shih, Fabio Vera

March 15, 2019

1 Problem Description

Recent advancements in deep learning coupled with the computational power of modern computing has impacted many areas of application including computer vision for healthcare. In Haenssle et al. [2018], Google's Inception v4 convolutions neural network (CNN) outperformed a group of international dermatologists in classifying dermoscopic images for melanoma, which is a type of skin cancer. Moreover, even when dermatologists were provided with additional clinical performance, the CNN still performed up to par. Thus, Haenssle et al. [2018] showed that a tool like Google's CNN would at the very least provide significant support to dermatologists. However, di Ruffano et al. [2018] states that "diagnosing a skin cancer when it is not actually present (a false-positive result) might result in unnecessary surgery and other investigations and can cause stress and anxiety to the patient." Surgeries can cost to upwards of 2,000 USD. Thus, false positives from computer vision tools may come at great cost.

However, using computer vision tools such as Google's CNN help provide much needed efficiency, especially dermatologists under severe workloads. Therefore, we propose to utilize Bayesian techniques to help introduce uncertainty into computer vision tools. More specifically, we hope to apply Bayesian CNN's (through various inference methods) to this subject area. By introducing uncertainty into the CNN, the CNN can act as a powerful screen for skin cancer when uncertainty is below a certain threshold and also flag samples that may require additional human effort when uncertainty is high. Thus, a institution may introduce their own standards to optimize the use of such computer vision tools.

2 Dataset

Our chosen dataset (Tschandl [2018] is comprised of 10000 images of pigmented skin lesions, each with their own ground truth diagnosis. The task is to properly assign a diagnosis given an image, as well as gather a notion of uncertainty in the chosen diagnosis.

The dataset can be found on Kaggle at:
https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000

3 The Plan

The plan is to compare 3 different kinds of training methods on a particular CNN architecture. In particular, we will be using two Bayesian methods and a vanilla method. These methods are Bayes by Back-Propogation, Bayesian Dropout, and a regular CNN. We will compare the effectiveness of these different methods in a few different ways. First, we will compare the validation accuracy of the models. Next, we will compare the performance of the models when attributing costs to misclassifications, e.g different costs for false positives and false negatives. Finally, we will compare efficiency of the models by comparing training time.

We will then provide a discussion of the pros and cons of the 3 different models.

4 Review of Current Methods

We will explore two inference methods for Bayesian CNN's: Bayes by Back-Propogation by Blundell et al. [2015] and Dropout as Bayesian Approximation by Gal and Ghahramani [2016]. Both of these methods infer a posterior distribution of network weights instead of a point estimate achieved through regular optimization. Bayes by Back-Propogation shows that variational inference can be achieved through the usual back-propogation technique used to optimize neural networks. Dropout as Bayesian Approximation utilizes a popular regularization technique called Dropout to approximate the posterior. Moreover, Laumann et al. [2018] has showcased how these inference methods can be used on not only a vanilla neural network but also a convolutional neural network. These methods allow for more practical inference of the posterior and we hope to explore the pros and cons of using each inference method in a practical application.

References

Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight uncertainty in neural networks. *arXiv preprint arXiv:1505.05424*, 2015.

Lavinia Ferrante di Ruffano, Yemisi Takwoingi, Jacqueline Dinnes, Naomi Chuchu, Susan E Bayliss, Clare Davenport, Rubeta N Matin, Kathie Godfrey, Colette O'Sullivan, Abha Gulati, et al. Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. *Cochrane Database of Systematic Reviews*, 1(12), 2018.

Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059, 2016.

HA Haenssle, C Fink, R Schneiderbauer, F Toberer, T Buhl, A Blum, A Kalloo, A Ben Hadj Hassen, L Thomas, A Enk, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 29(8):1836–1842, 2018.

Felix Laumann, Kumar Shridhar, and Adrian Llopart Maurin. Bayesian convolutional neural networks. *CoRR*, abs/1806.05978, 2018. URL `http://arxiv.org/abs/1806.05978`.

Philipp Tschandl. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, 2018. URL `https://doi.org/10.7910/DVN/DBW86T`.