CSCI 1430 Final Project Report:
Diabetic Retinopathy: Transfer Learning using Inception V3

tf Diabetic Retinopathy: Fengyi Jiang, Venkata Shubhang Kandiraju, Haotian Fu, Joey Bai.
Brown University

1. Introduction

Diabetic retinopathy (DR) is an eye disease that arises due to diabetes and is a leading cause of blindness around the world. It's a hard disease to catch early on, when it's most treatable, due to lack of significant symptoms.

DR contributes to significant global burden to due to the widespread prevalence of diabetes - 422 million patients around the world. Out of these patients with diabetes, 146 million people have some form of DR.

This makes DR a major problem around the world that needs to be tackled soon. One of the ways to reduce blindness caused by DR is to catch the disease early and that can be done by leveraging technology at the right places in the healthcare system.

Here in this paper we aim to showcase a deep learning based solution to automatically detect presence of DR in the retina. We feel that a solution like this can be very beneficial to catching DR early on in places where the patients to ophthalmologists ratio is very low. Eventually, as rates of early diagnosis of DR go up so will the cases of blindness as early stage disease can be treated effectively.

2. Related Work

Feng Li et al.[1] have previously implemented a model similar to ours. They implement a deep learning model based on the Inception-v3 network to detect diabetic retinopathy in retinal images. They reach a classification accuracy of 93.49 percent.

3. Method

3.1. Data Preprocessing

We preprocess the retina image using normalization, downsampling, and auto-cropping. We show the original input images in Figure 12. For this dataset, we focus on five categories of labels, which rate each image for the severity of diabetic retinopathy from level 0 to 4.An important issue we can see from the visualization is that some images (i.e. the 2nd and 4th images) have large dark and uninformative areas around the eyes. This may confuse the classifier and

Figure 1. Original retina image

Figure 2. Retina image after auto-cropping

have a bad effect on our training process. Thus we write an auto-cropping function to crop out those dark areas. The result images for comparison are shown in Figure 2. Then we use the linear scaling to normalize all the pixels of the images to values between 0 and 1. After normalization, we resize all the images to be consistent to the required input size of Inception-V3/VGG model we used later. We also try subtracting the mean of all the images from the training set to center the data around zero mean.

Recognizing the diabetic retina require professional knowledge on retina pathology and As shown in the Figure.1 above, differences between retina with different severity level can be barely recognized. As we have seen in the lectures, hand-designing features for this recognition task might be too complicated to finish. Thus we rely on the the power of deep learning and convolutional neural network.

3.2. Transfer Learning

We are using Inception V3, a publicly available keras model trained on ImageNet designed for image classification task. The idea we are implementing is transfer learning. We are taking the knowledge Inception V3 has over the ImageNet dataset and apply it to our much smaller retina pathology data set.

We hope the Inception V3 can help us extract the low-level, as well as high-level features inside the retina images. Since there might be some commonalities in the low level features between ImageNet and Retina dataset, we hope to bypass the process of hand-designing features for extraction.

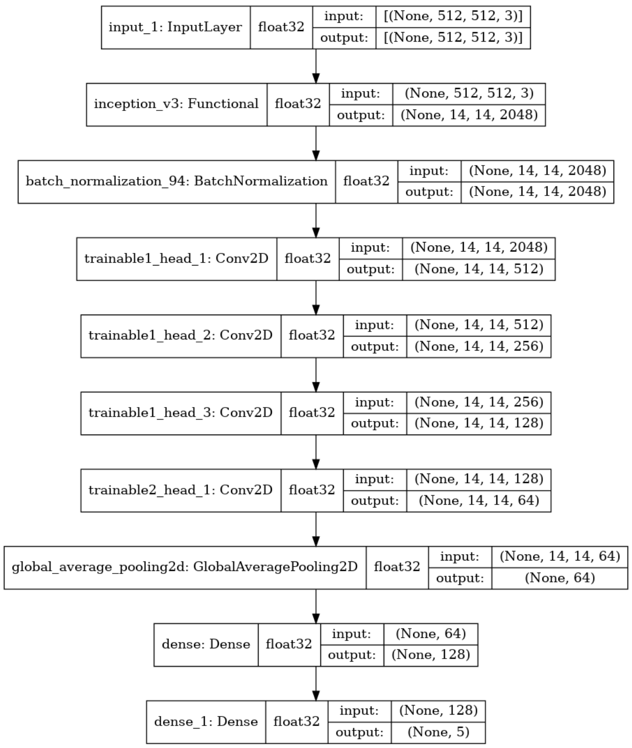


Figure 3. Combined Neural Network Architecture with Inception V3 at the second layer and custom trainable head attached.

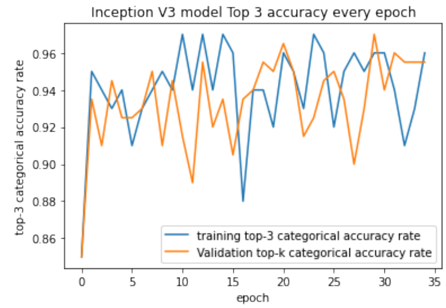


Figure 4. Top 3 Categorical Accuracy of Inception V3 Model with Cov2D layers inside the head.

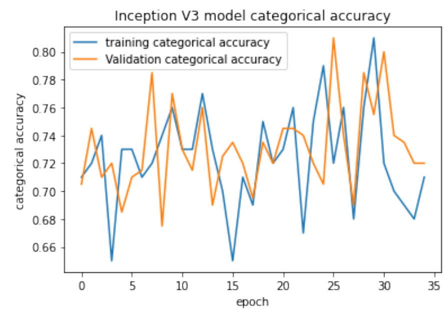


Figure 5. Overall Categorical Accuracy of Inception V3 Model with Cov2D layers inside the head.

3.3. Model Exploration: VGG16 and Inception V2

Keras core provides us with convenient access to various powerful models. We were curious about the performance of other Keras Cores, such as VGG16 and Inception V2, and included them during our training iterations for comparasion. We were specifically curious about the tradeoff between one model and another, as well as the outcome of the performance metrics.

We will present a model performance comparison between those models and elaborate the tradeoffs in the Result section below.

3.4. Neural Network Architecture

The architecture of our Neural Network, combined with base from Inception V3 and custom trainable head, is shown in the Figure 3.

Our model was using a combination of Conv2D, average pooling and batch normalization layers after the Inception V3, all laid out as sequential layer.

3.5. Google Cloud Platform and AutoML API

We used Google Cloud Platform for the Virtual Machine instances and model training. We used Google AI Platform Notebook, a service inside GCP that runs JupyterLab on top of our Virtual Machine instances for responsive coding and code visualization.

The instance we use is a VM with 4 vCPUs, 15 GB RAM, NVIDIA Tesla P4 and tensorflow 2-3/CUDA 11.0.

During the training process, we found the AutoML API provided by GCP: A implemented machine learning pipeline that trains Image Classification models automatically with provided dataset. Out of curiosity of the performance of such pipeline, we deployed the Retina dataset over Google Cloud Storage Bucket and trained the model for 16 node hours. A comparsion between our model and such AutoML model is in the Result section.

4. Results

4.1. Inception V3 Model with Conv2D Layers

Our model's performance is summarized in the plot:

Our model achieved a lowest loss 0.91 during the training process.

In comparison, we also trained the model with the Cnonvolutional 2D layer removed from the custom trainable head just for comparison.

4.2. Inception V3 Model without Conv2D Layers

Inception V3 Model without Conv2D Layers performance is summarized in the plot:

The Inception V3 without the Convolution2D layer achieved a lowest loss 1.18 during the training process. With the same hyper parameter, the added Convolution 2D layer

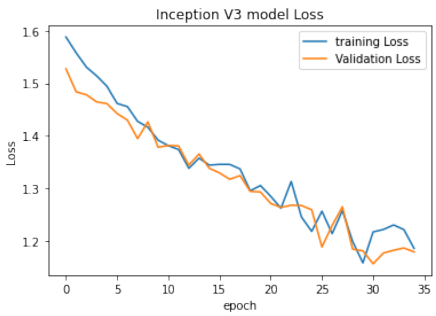


Figure 6. Loss of Inception V3 Model with Cov2D layers inside the head.

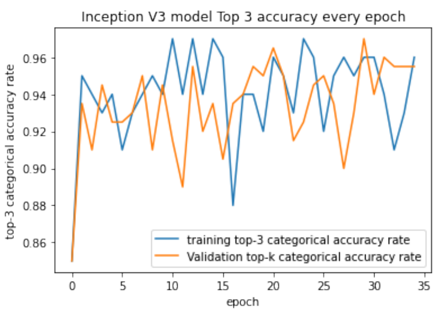


Figure 7. Top 3 Categorical Accuracy of Inception V3 Model without Cov2D layers inside the head.

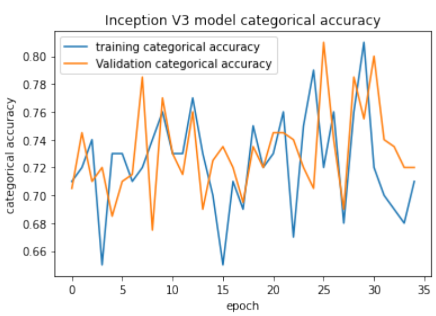


Figure 8. Overall Categorical Accuracy of Inception V3 Model without Cov2D layers inside the head.

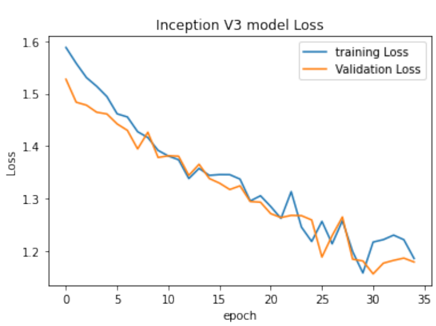
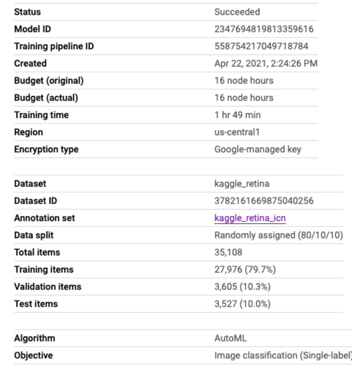


Figure 9. Loss of Inception V3 Model without Cov2D layers inside the head.

reduced the overall loss from 1.18 to 0.91, a reduction of



Status	Succeeded
Model ID	2347694819813359616
Training pipeline ID	5587542170497187840
Created	Apr 22, 2021, 2:24:26 PM
Budget (original)	16 node hours
Budget (actual)	16 node hours
Training time	1 hr 49 min
Region	us-central1
Encryption type	Google-managed key
Dataset	kaggle_retina
Dataset ID	3782161669875040256
Annotation set	kaggle_retina_icn
Data split	Randomly assigned (80/10/10)
Total items	35,108
Training items	27,976 (79.7%)
Validation items	3,605 (10.3%)
Test items	3,527 (10.0%)
Algorithm	AutoML
Objective	Image classification (Single-label)

Figure 10. AutoML Model summaries.

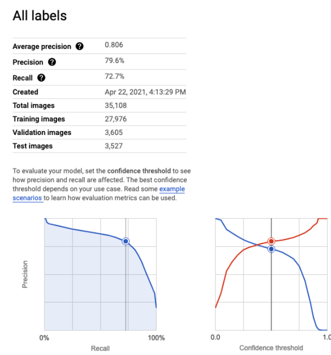


Figure 11. AutoML Model summaries.

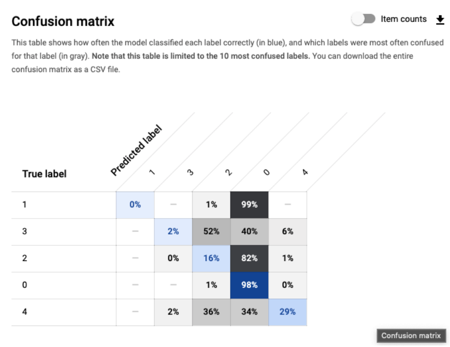


Figure 12. Confusion matrix generated by the AutoML model

22.88%.

4.3. GCP AutoML Trained Model

As stated before, we also used AutoML API from GCP for model training on the Retina data set. The results is shown in:

Although using different metrics, one interesting observation we had is that the AutoML implemented model reaches a overall accuracy of around 80.6%. Which is similar to that of our model. The confusion matrix of the model is shown in Figure 2.

However, the training time is extremely reduced. As shown in the Figure 11, the overall training time is only 1Hours and 49 minutes.

5. Technical Discussion

Our method raises the question of the black-box problem. We have an algorithm which can predict disease in the retina but does not have any mechanism implemented to interpret the results that can be understood by humans easily. In healthcare, when lives or the well being of humans is involved it is essential to think about the interpretability of these algorithms. However, an algorithm like this can be useful when it is used to refer the patient to a specialist. In this case, the cost of false positive might not be too high as a specialist is in the loop to verify the diagnosis.

6. Societal Discussion

1. The socio-historical context of this project is that access to quality healthcare across the world has always been skewed towards the rich and highly educated. Due to this, poor people (low income), people belonging to certain ethnic groups, and people with low levels of education have been particularly susceptible to health ailments like diabetic retinopathy because they don't have the right access to healthcare and are unlikely to go to an ophthalmologist to get an annual retina checkup done. This could mean that a certain subset of the population might be underrepresented in the datasets and can lead to possible bias creeping into the model. Outside source.

2. Stakeholders are those who may be affected by or have an effect on your project topic. Some examples of stakeholders are a particular demographic group, residents of a particular geographic area, and people experiencing or at risk for a particular problem.
   Consider the following questions to help identify stakeholders:

   - This project currently affects people who have diabetic retinopathy, a form of eye disease caused by diabetes. It could also affect doctors as they can use a model/algorithm like this to diagnose/-triage patients to provide the best healthcare to the people who need it the most.
   - No one is going to be harmed by our research findings. This is a project where people can benefit greatly from a working solution.
   - The people who might benefit the most are the people who are at risk for diabetic retinopathy but don't have access or the infrastructure around them to seek quality healthcare.

3. The implication of the research surrounding the project topic has been to frame the goal in a way that helps the patients directly. The goal of this kind of research should be to improve the outcomes for patients rather than be a testament to the technological prowess of the researchers. A lot of the research has been around what kind of models can be used to predict the disease or how to increase the accuracy of a model on a given dataset etc. While all these are important to build a robust practical solution, the overarching goal should always be to translate these state of the art findings into practical solutions to solve the problem in real time. This entails surrounding research findings with practical context to make sure the solution that is being proposed can be used directly to improve the outcomes for the patients and society in general.

4. Privacy is definitely something that could be affected with a healthcare based AI application. If the data that is collected isn't adhering to high standards of privacy or HIPAA protocols or good cybersecurity protocols, then we infringe on the basic right to privacy of the patients whose data we have collected. Hence, with healthcare data utmost care needs to be taken to ensure that the privacy of the patients isn't compromised.

5. The data might contain biases such as not having enough data points on underrepresented groups. People in lower income communities might not have access to good healthcare or healthcare infrastructure in general to be adequately represented in the dataset. Societal biases might play an indirect role in biasing the datasets. Societal biases towards certain groups might have caused them to be more likely to belong to lower income communities which might affect their access to healthcare which in turn might affect their representation in datasets that are collected to train these algorithms. This bias can be mitigated by having inclusive data collection protocols and implementing data balancing methods.

7. Conclusion

We implemented multiple CNN models based on the feature layers of different popular models pretrained on ImageNet. Our models can detect diabetic retinopathy based on retina scans and can potentially automatize the diagnosing procedure. This can save a significant amount of medical resources and allows more patients to receive medical helps in time. With minor adjustment, our models potentially be applied to other imaging based diagnosis such as lung cancer and osteoporosis.

References

[1] Feng Li et al. "Automatic detection of diabetic retinopathy in retinal fundus photographs based on deep learning algorithm". In: *Translational vision science & technology* 8.6 (2019), pp. 4–4.

Appendix

Team contributions

Please describe in one paragraph per team member what each of you contributed to the project.

Fengyi Jiang Data preprocessing, model implementation on Inception V3/VGG16 /Inception V2 and conducting experiments to increase accuracy. Also helped in making the presentation, writing the report and compiling the code for submission.

Venkata Shubhang Kandiraju Helped in selecting the project and obtaining the dataset needed to implement the project. Worked on implementing the models specified above and code debugging. Also helped in making the presentation and writing the report.

Haotian Fu Helped in data preprocessing and trying to implement another model to compare against our original model. Helped in baseling model implementation. Also helped with the presentation and the report.

Joey Bai Implemented and trained a model based on the convolutional layers of Resnet152V2 trained on ImageNet. Also helped with the presentation and the report.