

ALZHEIMER'S DISEASE PREDICTION USING DEEP TRANSFER LEARNING ON MULTILAYER OCTA IMAGES

Alice Chen

Department of Computer Science
Columbia University
New York, NY 10025, USA
yc3877@columbia.edu

ABSTRACT

Alzheimer's disease stands as a prevalent neurodegenerative disorder typically identified through subjective and time-consuming clinical assessments. The prevailing methods aimed at automating this process predominantly center on the classification of Magnetic Resonance Imaging (MRI) images, which can be costly. This project endeavors to diagnose Alzheimer's disease by utilizing deep learning models applied to Optical Coherence Tomography Angiography (OCTA) images. Despite containing comparatively less information than MRI images, OCTA images offer enhanced accessibility and reduced cost. This project aims to build upon these findings by employing deep learning models to classify multilayer OCTA images. The models used in this project are based on popular classification models: ResNet and VGGNet. The models were trained using transfer learning on the ImageNet dataset and evaluated using five-fold cross validation. The best achieving model was transfer learning on VGG16 with an accuracy of 0.69 and AUC of 0.67. The results of this project are comparable with the results using Machine Learning techniques and demonstrate that deep learning models can be used to classify OCTA multilayer images.

1 INTRODUCTION

Alzheimer's disease, a prevalent neurodegenerative disorder, is traditionally diagnosed through subjective and time-consuming clinical assessments. Previous efforts in Alzheimer's Disease Classification primarily centered on Magnetic Resonance Imaging (MRI) images, and many have shown remarkable performance, achieving up to 0.95 accuracy (Yang et al., 2018; Singh et al., 2020). However, this study seeks to diagnose Alzheimer's disease using deep learning models applied to multilayer Optical Coherence Tomography Angiography (OCTA) images, which, despite containing less information than MRI images, offer increased accessibility and reduced cost. Several publications have demonstrated the potential of OCTA images for Alzheimer's Disease diagnosis Turkan & Tek (2022). Yet, exploration of analogous classification techniques using retinal images remains limited. Wisely et al. (2024) incorporated multimodal retinal images with patient data, achieving an AUC of 0.836 with deep learning models. However, reliance solely on OCTA data reduced the model's AUC performance to 0.58, highlighting limitations due to a small and imbalanced dataset. Recent work by Wang et al. (2022) leveraged machine learning models on OCTA images, with XG-Boost yielding the highest accuracy of 0.74 and an AUC of 0.69 following an assessment of OCT features' statistical significance.

This project builds upon these findings by employing deep learning models for classifying multilayer OCTA images. Specifically, ResNet and VGGNet-based models, trained using transfer learning on the ImageNet dataset and evaluated via five-fold cross-validation.

2 METHODOLOGY

In this section, we will provide more details on the dataset used in this study, the preprocessing steps, and the model architecture.

2.1 DATA PREPROCESSING

The dataset used was collected from Chang-Gung Medical Foundation and consists of patients with Alzheimer's disease and patients without Alzheimer's disease. The images were taken from both the left and right eye, each eye contains a total of 4 Retina images. We trained each model using each layer individually. Then, we concatenated image of Retina Layer d, r, s. Due to limited number of data, we considered data from different eyes to be a different sample. The total number of samples is 287, with 133 positive examples, and 154 negative examples. The image in each layer is 2D has resolution 400x400. Each image was applied with enhanced contrast by a factor of 2. Since the data consists of data from both the left and right eye, data from the left eye was mirrored so that all images are in the same perspective. All images were cropped at the center to 340x340 to remove metadata at the edge of the images. To reduce noise, a gaussian filter with kernel size 3 was applied to all the images. The images were next normalized to have mean 0.4 and standard deviation 0.5, then resized to 224x224. The dataset was split into 80% training and 20% testing. The training set was further split into 80% training and 20% validation. Five-fold cross validation was used to evaluate the models. Figure 1 shows sample images from each of the 4 layers after preprocessing. Note that some images, may contain dark regions or unnatural lines due to errors in the image capturing process.

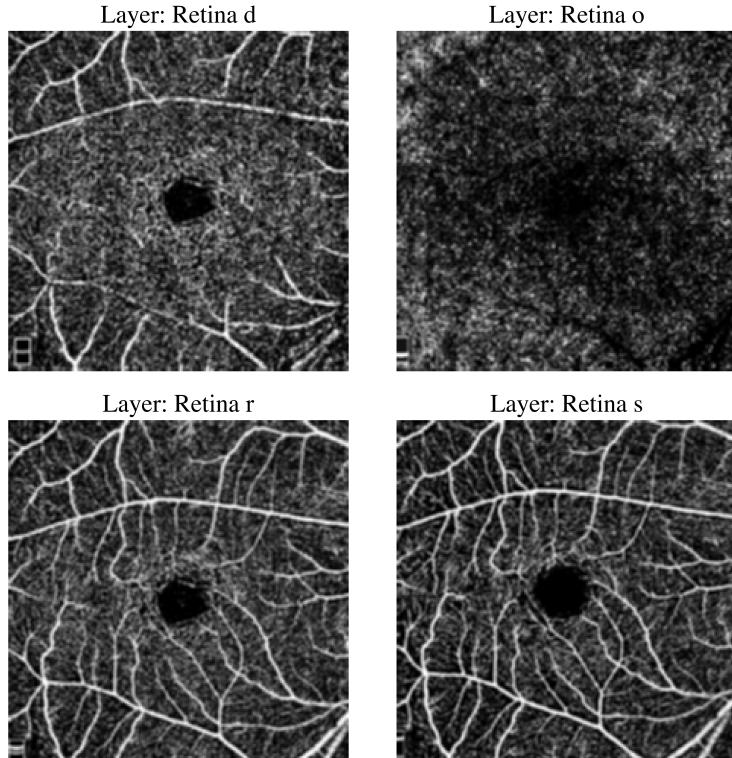


Figure 1: These 4 images illustrate the 4 layers of OCTA images in each sample after preprocessing. Retina d, o, r, s, represent deep, outer retina, choriocapillaries and superficial.

2.2 MODEL ARCHITECTURE

Due to the limited data used in this study, we applied transfer learning using models pretrained on the ImageNET dataset. Early stopping was used such that the best model saved was based on the lowest validation loss. All models were implemented using PyTorch.

2.2.1 VANILLA CONVOLUTION NEURAL NETWORK (CNN)

The Vanilla CNN model consists of four convolution blocks, each convolution block consists of a convolution layer with kernel size three, stride one, padding one, followed by a ReLU layer, a max pooling layer with kernel size two, stride two. The convolution blocks are followed by a fully connected layer with 512 hidden units, followed by a ReLU layer and a sigmoid layer. The model was trained using binary cross entropy loss and Adam optimizer with learning rate 1e-5 and batch size 8.

2.2.2 RESNET

We used a ResNet model as the base model and added a fully connected layer and a sigmoid layer at the end of the model to serve as a baseline model for comparison. All models were trained using binary cross entropy loss and Adam optimizer with learning rate 1e-5 and batch size 8. All models were trained with the original ResNet parameters frozen for 5-10 epochs, then trained with all parameters for 5-10 epochs.

2.2.3 VGGNET

We used a VGGNet model as the base model and added an additional classifier consisting of a Dropout Layer, a Linear Layer, a ReLU Layer, another set of Dropout, Linear, ReLU, followed by a Linear Layer to a sigmoid layer. The model was trained using binary cross entropy loss, Adam optimizer with learning rate 1e-5 and batch size 8. The VGG16 model was trained with the original VGG16 parameters frozen for 5 epochs, then trained with all parameters for 5 epochs. The VGG19 model was trained with the original VGG19 parameters frozen for 10 epochs, then trained with all parameters for 2 epochs.

3 RESULTS

The model was first trained and evaluated on different layers of the OCTA images to understand which layers may be most informative. The data was trained using a Vanilla CNN and a pretrained ResNet18 and ResNet50. The test results are shown in Table 1 and the training results are shown in Figure 3. Both ResNET models are able to outperform the Vanilla CNN, which indicates that transfer learning may be beneficial in this scenario. The performance of ResNet18 is better than ResNet50, which is likely due to overfitting. The performance of ResNet18 and ResNet50 both exceed the results from Wisely et al. (2024), likely because the dataset used in this study is larger and better balanced.

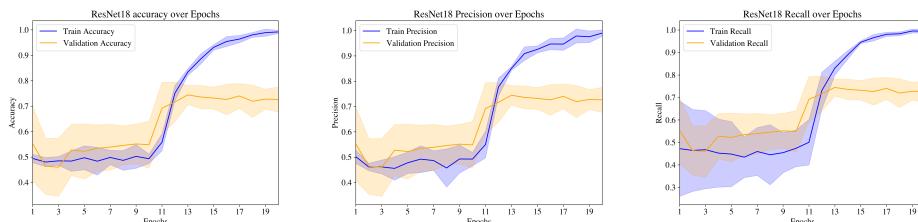


Figure 2: ResNet18 train (blue) and validation (orange) accuracy curve (left) precision curve (center) and recall curve (right) with mean and standard deviation. This model was trained only on the Retina Layer R.

Table 1: Comparison of VanillaCNN, ResNet18 and ResNet50 when trained on different layers of the OCTA image. All values are evaluated on the test set using five-fold Cross Validation. ACC is accuracy, Precision is precision score, Recall is recall score, F1 is F1 score, AUC is area under the receiver operating characteristic (ROC) curve.

| Model | ACC | Precision | Recall | F1 | AUC |
|----------------|-----------------|------------------|-----------------|-----------------|-----------------------------------|
| Layer R | | | | | |
| VanillaCNN | 0.59 ± 0.01 | 0.61 ± 0.03 | 0.31 ± 0.05 | 0.41 ± 0.04 | 0.56 ± 0.01 |
| ResNet18 | 0.67 ± 0.03 | 0.67 ± 0.03 | 0.59 ± 0.06 | 0.62 ± 0.05 | 0.67 ± 0.03 |
| ResNet50 | 0.63 ± 0.04 | 0.63 ± 0.06 | 0.54 ± 0.13 | 0.56 ± 0.08 | 0.62 ± 0.04 |
| Layer D | | | | | |
| VanillaCNN | 0.57 ± 0.06 | 0.56 ± 0.12 | 0.36 ± 0.14 | 0.42 ± 0.12 | 0.56 ± 0.06 |
| ResNet18 | 0.65 ± 0.06 | 0.65 ± 0.08 | 0.51 ± 0.14 | 0.56 ± 0.11 | 0.64 ± 0.06 |
| ResNet50 | 0.67 ± 0.06 | 0.68 ± 0.10 | 0.57 ± 0.12 | 0.61 ± 0.08 | 0.66 ± 0.06 |
| Layer O | | | | | |
| VanillaCNN | 0.59 ± 0.08 | 0.60 ± 0.07 | 0.91 ± 0.09 | 0.72 ± 0.05 | 0.53 ± 0.06 |
| ResNet18 | 0.65 ± 0.07 | 0.71 ± 0.10 | 0.67 ± 0.12 | 0.68 ± 0.09 | 0.65 ± 0.06 |
| ResNet50 | 0.61 ± 0.09 | 0.65 ± 0.08 | 0.65 ± 0.12 | 0.64 ± 0.10 | 0.59 ± 0.09 |
| Layer S | | | | | |
| VanillaCNN | 0.62 ± 0.01 | 0.68 ± 0.05 | 0.36 ± 0.10 | 0.46 ± 0.09 | 0.61 ± 0.02 |
| ResNet18 | 0.66 ± 0.03 | 0.68 ± 0.07 | 0.50 ± 0.10 | 0.56 ± 0.07 | 0.65 ± 0.03 |
| ResNet50 | 0.65 ± 0.03 | 0.65 ± 0.06 | 0.56 ± 0.11 | 0.58 ± 0.07 | 0.65 ± 0.02 |

Table 2: Comparison of VanillaCNN, ResNet18 and ResNet50 when trained on concatenated R, D, S layers of the OCTA image. All values are evaluated on the test set using five-fold Cross Validation. ACC is accuracy, Precision is precision score, Recall is recall score, F1 is F1 score, AUC is area under the ROC curve.

| Model | ACC | Precision | Recall | F1 | AUC |
|--------------|-----------------|------------------|-----------------|-----------------|-----------------|
| VanillaCNN | 0.63 ± 0.06 | 0.71 ± 0.11 | 0.37 ± 0.15 | 0.45 ± 0.14 | 0.61 ± 0.07 |
| ResNet18 | 0.68 ± 0.07 | 0.68 ± 0.09 | 0.58 ± 0.15 | 0.62 ± 0.11 | 0.67 ± 0.08 |
| ResNet50 | 0.67 ± 0.06 | 0.64 ± 0.08 | 0.60 ± 0.14 | 0.61 ± 0.10 | 0.67 ± 0.05 |
| VGG16 | 0.69 ± 0.04 | 0.73 ± 0.06 | 0.53 ± 0.10 | 0.60 ± 0.08 | 0.67 ± 0.05 |
| VGG19 | 0.65 ± 0.05 | 0.66 ± 0.05 | 0.49 ± 0.17 | 0.55 ± 0.11 | 0.64 ± 0.05 |

From Table 1, we noticed that the performance was better on layers R, D and S. Thus, we further trained the model by concatenating all 3 layers of OCTA images. Table 2 shows the results from each of the models when trained on data from all 3 layers of OCTA images. ResNet18 shows the best performance among all the models. The performance of ResNet50 is worse than ResNet18, which is likely due to overfitting. The performance of VGG16 and VGG19 are similar, with VGG16 performing slightly better. The performance did not significantly improve compared to using only one layer. This may be due to the fact that each layer's feature is different, and concatenating the layers may not be the best way to combine the features. Future work includes using a weighted average from the results, or first doing feature extraction on each layer, then concatenating the features.

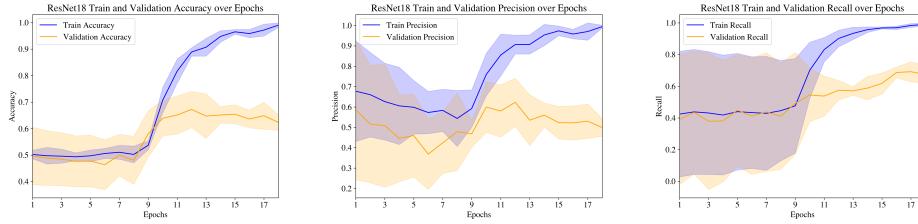


Figure 3: ResNet18 train (blue) and validation (orange) accuracy curve (left) precision curve (center) validation curve (right) with mean and standard deviation. This model was trained on the concatenated Retina Layer d, r, s.

4 CONCLUSION AND FUTURE WORK

Alzheimer's Disease is the most common type of dementia. Traditional methods are often costly and subjective. Classifying Alzheimer's Disease using OCTA images offers a solution that is time and cost efficient. This project demonstrates that deep learning models can be used on OCTA images to successfully classify Alzheimer's Disease with an accuracy of .69. The information from multiple layers also show potential on improving prediction accuracy. The results of this project are comparable with the results using Machine Learning techniques and demonstrate that deep learning models can be used to classify OCTA multilayer images. Future work includes using a weighted average from the results, or first doing feature extraction on each layer, then concatenating the features. In addition, other deep learning models can be explored such as DenseNet and EfficientNET. Lastly, the dataset used in this study is relatively small, future work includes collecting more data to improve the performance of the models.

AUTHOR CONTRIBUTIONS STATEMENT

A.C. conceived the original idea, developed the methodology, performed the experiments, analyzed the results and wrote the manuscript.

REFERENCES

- Satya P. Singh, Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan, and Balázs Gulyás. 3d deep learning on medical images: A review, 2020.
- Yasemin Turkan and F. Boray Tek. A survey on automated diagnosis of alzheimer's disease using optical coherence tomography and angiography, 2022.
- Xin Wang, Bin Jiao, Hui Liu, Yaqin Wang, Xiaoli Hao, Yuan Zhu, Bei Xu, Huizhuo Xu, Sizhe Zhang, Xiaoliang Jia, Qian Xu, Xinxin Liao, Yafang Zhou, Hong Jiang, Junling Wang, Jifeng Guo, Xinxiang Yan, Beisha Tang, Rongchang Zhao, and Lu Shen. Machine learning based on optical coherence tomography images as a diagnostic tool for alzheimer's disease. *CNS Neuroscience & Therapeutics*, 28(12):2206–2217, 2022. doi: <https://doi.org/10.1111/cns.13963>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/cns.13963>.
- C. Ellis Wisely, Alexander Richardson, Ricardo Henao, Cason B. Robbins, Justin P. Ma, Dong Wang, Kim G. Johnson, Andy J. Liu, Dilraj S. Grewal, and Sharon Fekrat. A convolutional neural network using multimodal retinal imaging for differentiation of mild cognitive impairment from normal cognition. *Ophthalmology Science*, 4(1):100355, 2024. ISSN 2666-9145. doi: <https://doi.org/10.1016/j.xops.2023.100355>. URL <https://www.sciencedirect.com/science/article/pii/S2666914523000878>.
- Chengliang Yang, Anand Rangarajan, and Sanjay Ranka. Visual explanations from deep 3d convolutional neural networks for alzheimer's disease classification, 2018.

A APPENDIX

All code can be found at <https://github.com/yuehtingchen/cv-final-project>.
Additional training results can be found below.

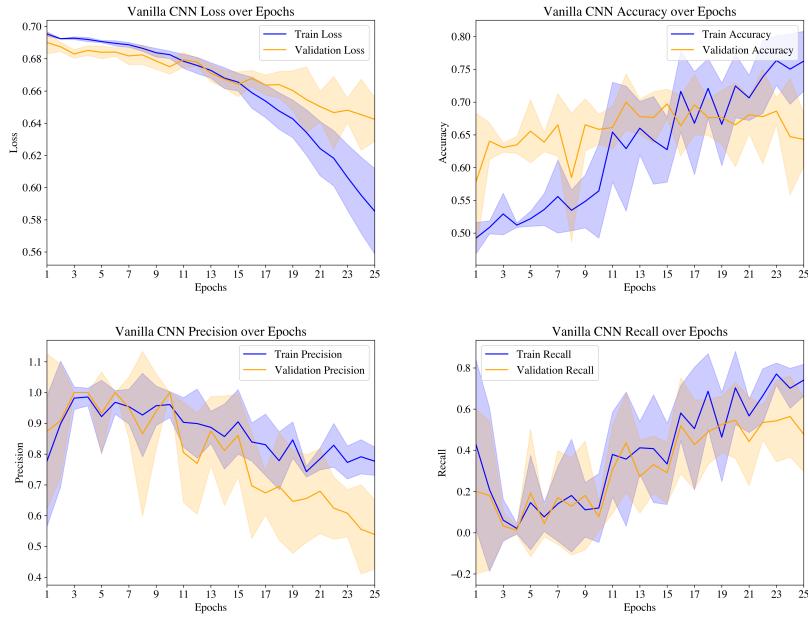


Figure 4: VanillaCNN train (blue) and validation (orange) loss curve (top left), accuracy curve (top right) validation curve (bottom left) validation curve (bottom right) with mean and standard deviation. This model was trained on the Retina Layer R.

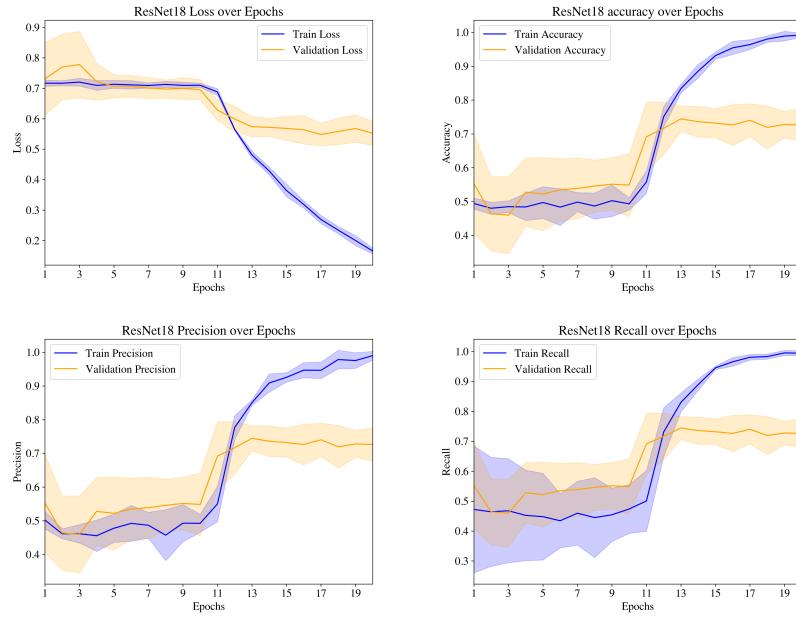


Figure 5: ResNet18 train (blue) and validation (orange) loss curve (top left), accuracy curve (top right) precision curve (bottom left) validation curve (bottom right) with mean and standard deviation. This model was trained on the Retina Layer R.

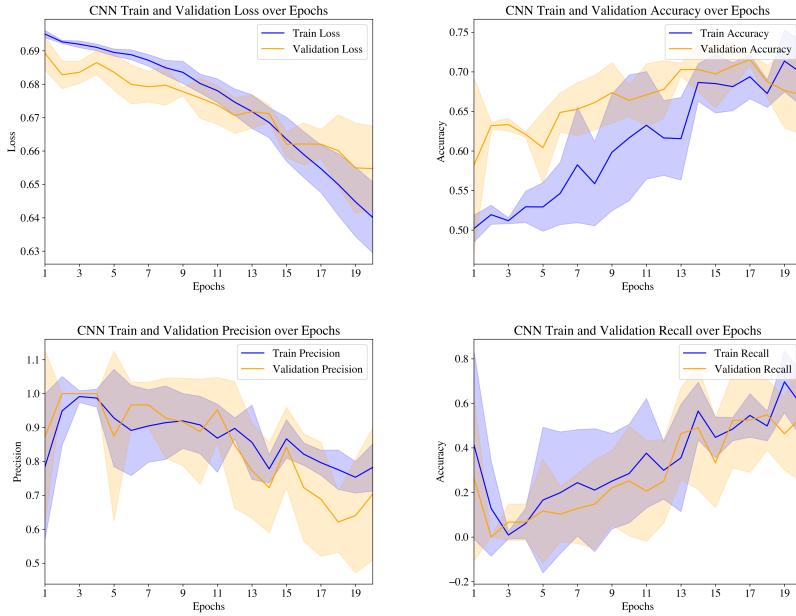


Figure 6: Vanilla train (blue) and validation (orange) loss curve (top left), accuracy curve (top right) precision curve (bottom left) validation curve (bottom right) with mean and standard deviation. This model was trained on the concatenated Retina Layer d, r, s.

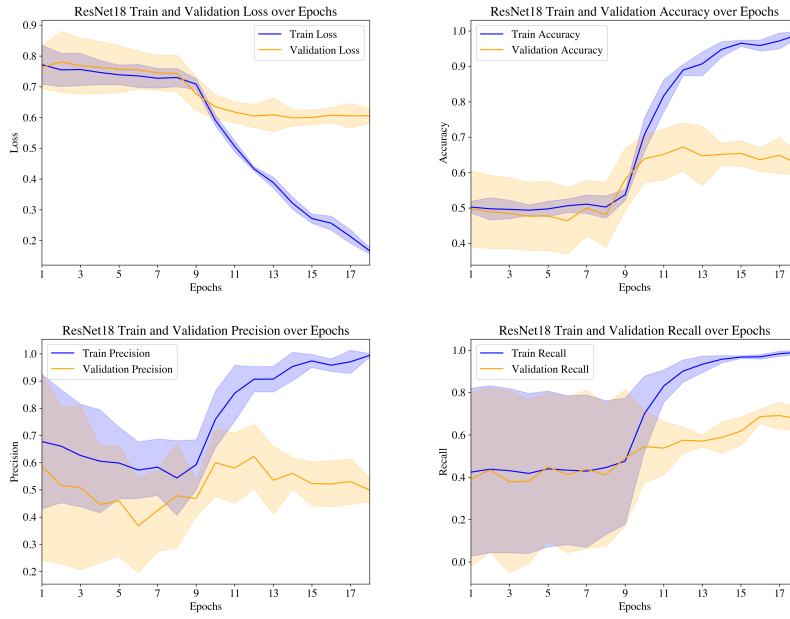


Figure 7: ResNet18 train (blue) and validation (orange) loss curve (top left), accuracy curve (top right) precision curve (bottom left) validation curve (bottom right) with mean and standard deviation. This model was trained on the concatenated Retina Layer d, r, s.

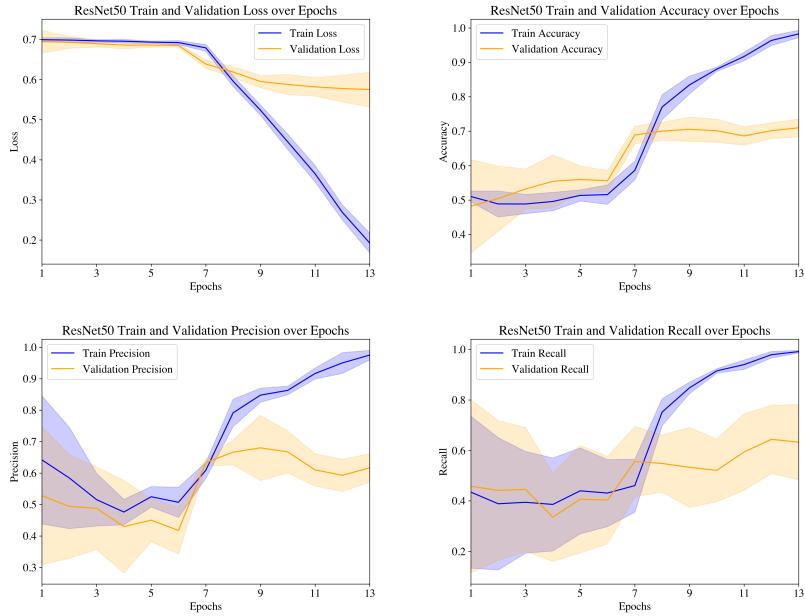


Figure 8: ResNet50 train (blue) and validation (orange) loss curve (top left), accuracy curve (top right) precision curve (bottom left) validation curve (bottom right) with mean and standard deviation. This model was trained on the concatenated Retina Layer d, r, s.