# Determining Whether Single or an Ensemble of Feature Extracting Networks are More Effective in Medically-Driven Few-Shot Image Classification Tasks

Cole M. Foster[1], Rachel M Tomasetti[1], Shengxin Luo[1], Sophia Zorek[1], Vanessa Aguiar-Pulido[1]

[1]Department of Computer Science, University of Miami, Coral Gables, FL, USA

## Introduction

One of the most pressing needs in computation is amount of data. A model is typically trained on many thousands of images in order to provide an accurate result. One problem with this approach is that it is not very useful in situations where data is harder to come by. This issue is known as the few-shot problem, which can be phrased as, "how do we maximize the accuracy of a model when data is at a minimum?" The potential answer that we are exploring is, instead of using one model to extract image features as is common in deep learning image analysis, what if we used an ensemble of multiple models and concatenated their features together? To test this approach, we will be using the EyePACS diabetic retinopathy (DR) dataset split up into 5 classes: no DR, mild DR, moderate DR, severe DR, and proliferative DR. Since this problem is very computationally difficult, it gives us a great setting to test our approach.

Objective: Determine whether single networks or an ensemble of networks for feature extraction are better suited for few-shot situations pertaining to diagnosing diabetic retinopathy.

## Materials & Methods

EyePACS Diabetic Retinopathy Dataset

35,126 eye images collected from 17,563 patients subsequently divided into 5 categories from 0 to 4 based on progression of diabetic retinopathy.

Deep and Shallow Neural Network (DaSNN)

*Feature Extracting Networks:*

A type of pretrained neural network used to analyze and process image data (Figure 1). Images are passed through different convolutional layers that filter the image into useable feature data. This feature data can be passed through a classifier head or multilayer perceptron to make a final prediction. Feature extractors will be produced using different pre-trained networks and removing their classification heads.

- Models used as feature extractors:
  - ResNet18, Resnet34, Resnet50, Resnet104, DenseNet121, DenseNet161, DenseNet169, and DenseNet201

*Multilayer Perceptron*

A multilayer perceptron (MLP) is an artificial neural network that in our case consists of three layers, an input layer that takes feature data, an output layer that contains the MLP's per-class predictions, and a hidden layer in-between that performs nonlinear transformations of the feature data to turn it into those predictions. Our hidden layer will be of size 1024 and use hyperbolic tangent as an activation function.

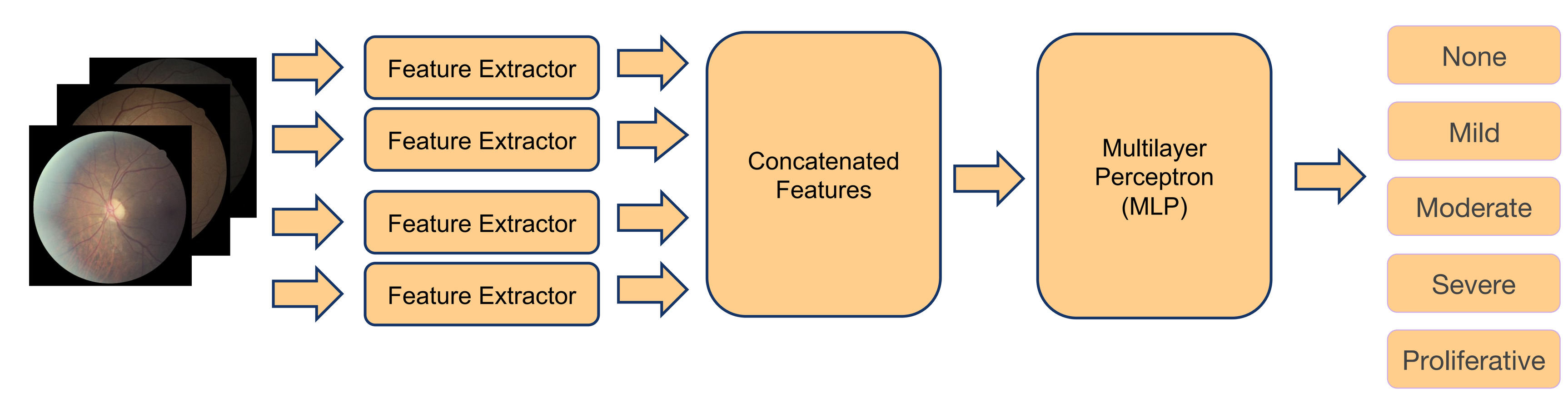


Figure 1 | **DaSNN Approach for Image Analysis**

Experimental design

We are testing 4 different data splits, with 500, 100, 50, and 10 images per class making up the training dataset, and 500 images per class making up the testing dataset. We used cross entropy loss as our loss function, with L2 regularization to combat potential overfitting. The lambda value chosen for L2 regularization was 0.001. We used Adam optimizer with learning rate 0.0001 and a weight decay of 0.00001 to also combat overfitting. Standard accuracy is used to compare results. The purpose of our experiments is to better understand how a model using an ensemble of pretrained feature extractors compares to a model using only a single pretrained feature extractor in a few-shot setting. In a few-shot setting, small differences in the amount of data can make a very large difference in results, so we decided on testing 500-shot, 100-shot, 50-shot, and 10-shot datasets. For each of these datasets the ensemble of pretrained feature extractors and each individual pretained feature extractor are used to extract features to train our MLP and calculate testing accuracy. Using these accuracy values, we can compare each feature extractor to determine which yields the best accuracy in the few-shot setting, and how these relative accuracies change when the few-shot setting becomes more extreme.

## Results

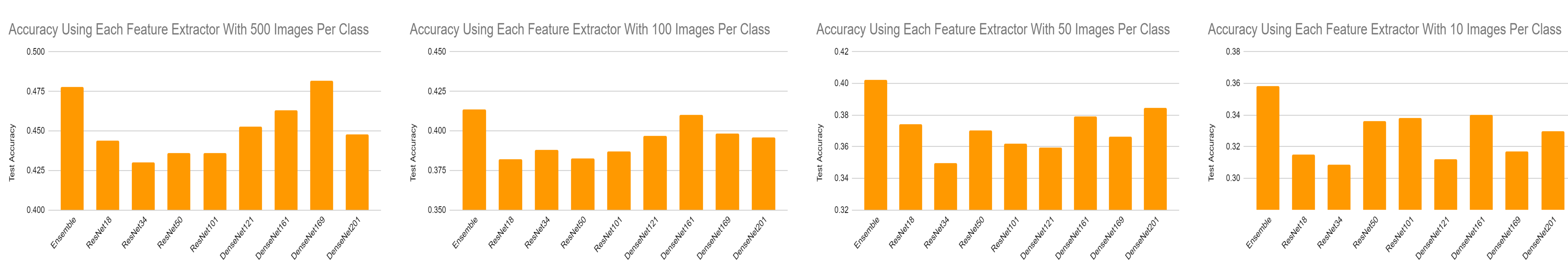


Figure 2 | **Results using 500-shot, 100-shot, 50-shot, and 10-shot datasets**

## Conclusion

As expected, overall accuracy declines when the few-shot setting becomes more extreme. An interesting observation is that in the 500-shot setting, DenseNet169 performed best among all feature extractors in accuracy, even surpassing the ensemble of feature extractors. This may be due to a larger number of overall features in the ensemble resulting in overfitting. However, as the few-shot situations become more extreme, the ensemble of feature extractors begins to strongly outperform all single feature extractors, with DenseNet169 especially regressing. This points to the possibility that with a large amount of data, DenseNet169 may perform well in diabetic retinopathy classification, but in a true few-shot setting, the ensemble of pretrained feature extractors performed better than any single feature extractor.

## Future Work

Future work concerns testing more deep learning classifiers for feature extraction and applying the proposed approach to different few-shot problems. Some of the ideas include:

- Incorporating VGGs, Inception_V3, and other pretrained feature extractors into our feature extractors.
- Using different datasets to test our model to evaluate performance.
- Looking into different eye conditions and testing whether the proposed approach consistently outperforms models using one network for feature extraction for few-shot problems.
- Research into applying DaSNN to binary classification few-shot problems

## Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1949972. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## References

1. Kaggle. Diabetic Retinopathy Detection (Data); 2015. Available from: https://www.kaggle.com/c/diabetic-retinopathy-detection/data
2. Wu L, Fernandez-Loaiza P, Sauma J, Hernandez-Bogantes E, Masis M. Classification of diabetic retinopathy and diabetic macular edema. World J Diabetes. 2013 Dec 15;4(6):290-4. doi: 10.4239/wjd.v4.i6.290. PMID: 24379919; PMCID: PMC3874488.
3. A. Chowdhury, M. Jiang, S. Chaudhuri and C. Jermaine, "Few-shot Image Classification: Just Use a Library of Pre-trained Feature Extractors and a Simple Classifier," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 9425-9434, doi: 10.1109/ICCV48922.2021.00931.