Ocular Disease Prediction with CNNs

...

Lou Hines
July 14, 2023

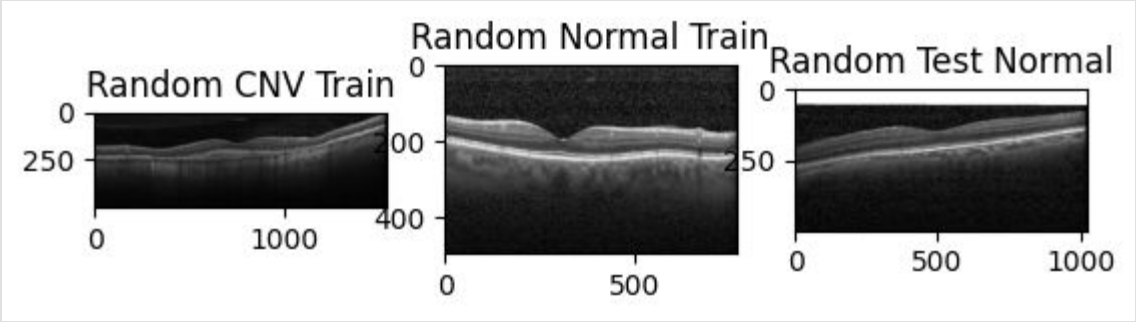
Optimizing Models to Predict Ocular Disease

AI really shines in medical profession, where recent technologies are able to create advanced mRNA vaccines, predict diabetes well ahead of other techniques, and catch missing heart attack diagnoses.

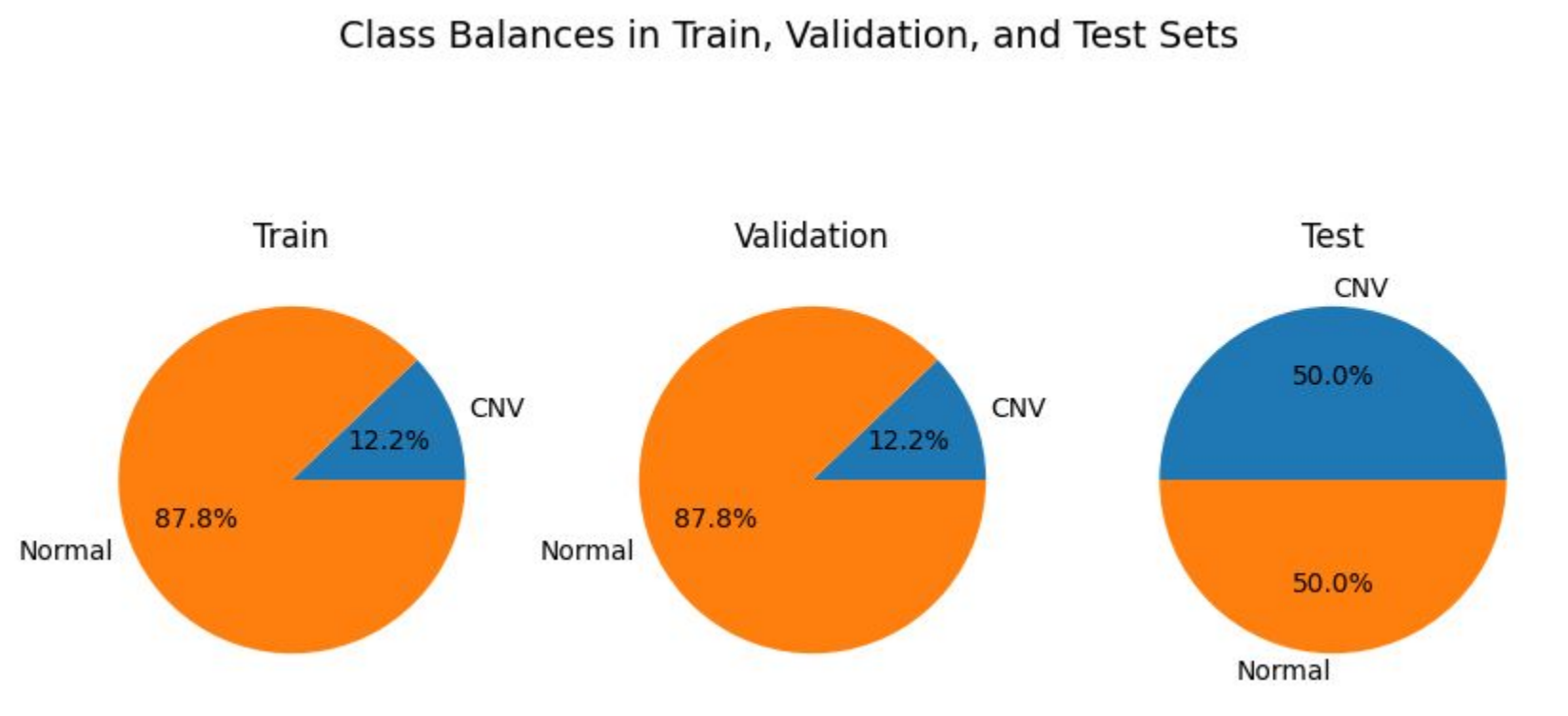
In this spirit, we optimized models using OCT retinal scans to detect CNV, (Inflammatory Choroidal Neovascularization), a degenerative ocular disorder commonly resulting in blindness, using a subset of data from *Cell*.

***Especially given SARS-CoV-2,* early and quick detection of virally-induced disorders is complementary to this understanding.**

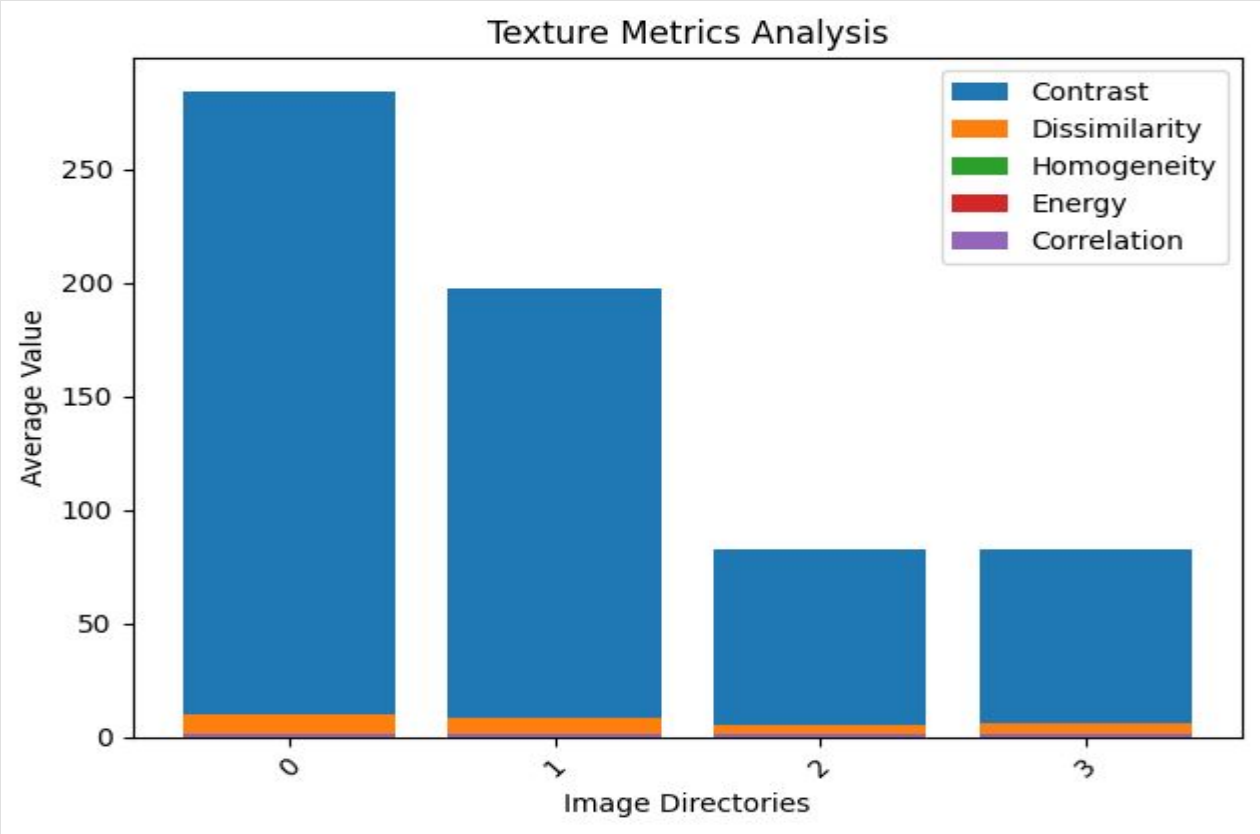
EDA:

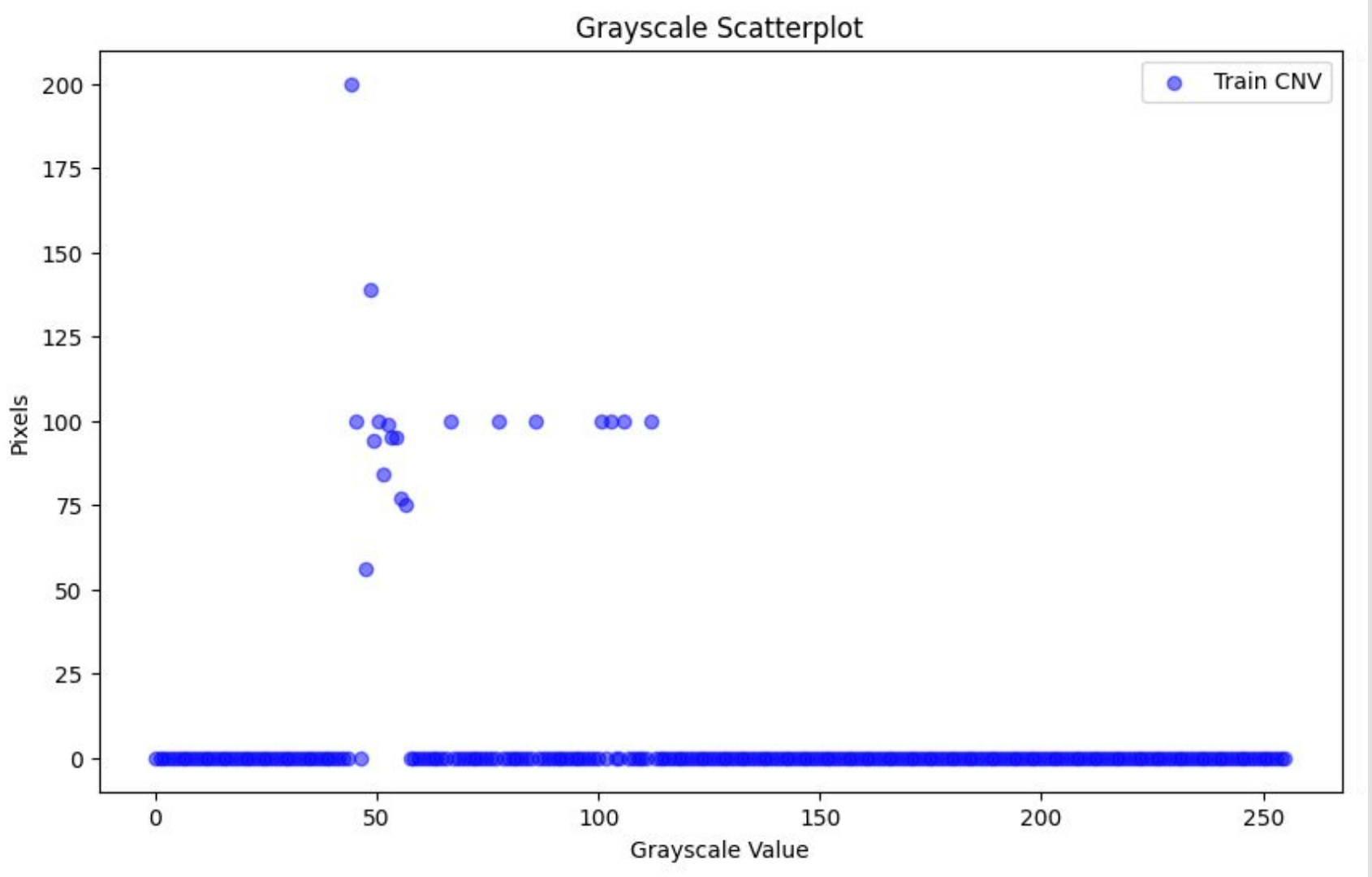

Random CNV Train
0
250
0
1000
Random Normal Train
0
00
400
0
500
Random Test Normal
0
250
0
500
1000

EDA:


Class Balances in Train, Validation, and Test Sets
Train
CNV
12.2%
87.8%
Normal
Validation
CNV
12.2%
87.8%
Normal
Test
CNV
50.0%
50.0%
Normal

EDA:


Texture Metrics Analysis
Contrast
Dissimilarity
Homogeneity
Energy
Correlation
Average Value
0
50
100
150
200
250
0
1
2
3
Image Directories


Grayscale Scatterplot
Train CNV
Pixels
Grayscale Value
200
175
150
125
100
75
50
25
0
0
50
100
150
200
250

Modelling

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_14 (Conv2D)          (None, None, None, 32)    320

 max_pooling2d_10 (MaxPooli  (None, None, None, 32)    0
 ng2D)

 conv2d_15 (Conv2D)          (None, None, None, 64)    18496

 max_pooling2d_11 (MaxPooli  (None, None, None, 64)    0
 ng2D)

 conv2d_16 (Conv2D)          (None, None, None, 64)    36928

 flatten_4 (Flatten)         (None, None)              0

 dense_8 (Dense)             (None, 64)                2560064

 dense_9 (Dense)             (None, 2)                 130

=================================================================
Total params: 2615938 (9.98 MB)
Trainable params: 2615938 (9.98 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fairly vanilla downsampled CNN with standard filters/kernals, optimization(adam), layer types and sequences, etc. 5 epochs. Added padding to input layer only. Utilized categorical cross-entropy and softmax and ran 50 epochs.

Other Models Tried:

CNN without MaxPool

Replaced MaxPool with striding, as per the Springenberg et al paper:

https://arxiv.org/pdf/1412.6806.pdf

Fairly strong results but not an overall improvement compared to our more vanilla model.

VGG16

Extremely time-intensive, but literature (if not our early results) seems promising. Uses pre-trained datasets.

More Models Tried:

CNN with Increased Filters

Doubled filter sizes in input and hidden layers in different models. Did not see improvement.

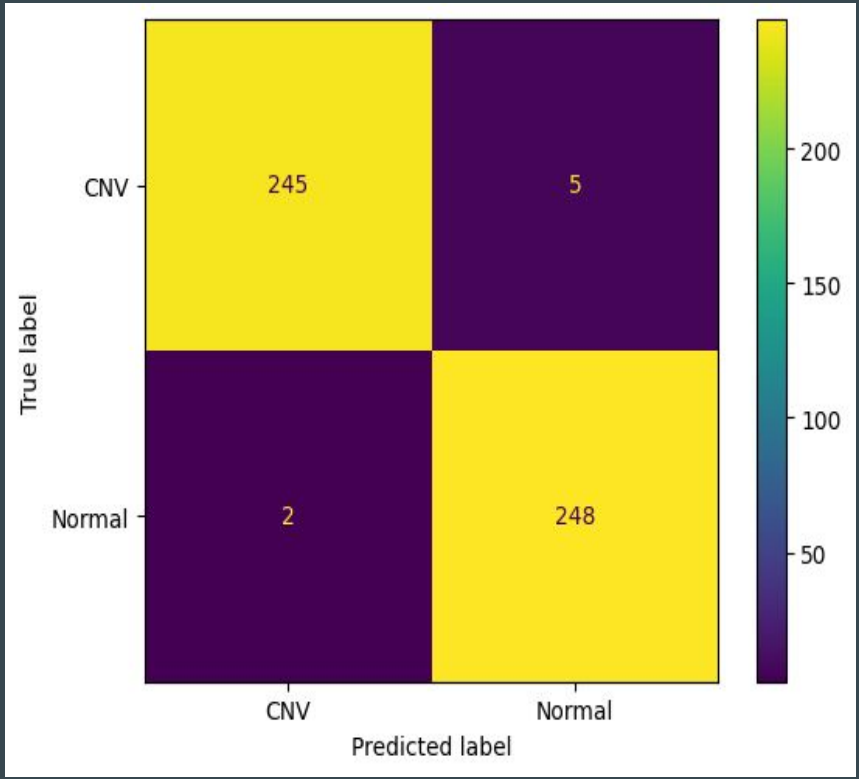
CNN with Increased Learning Rate

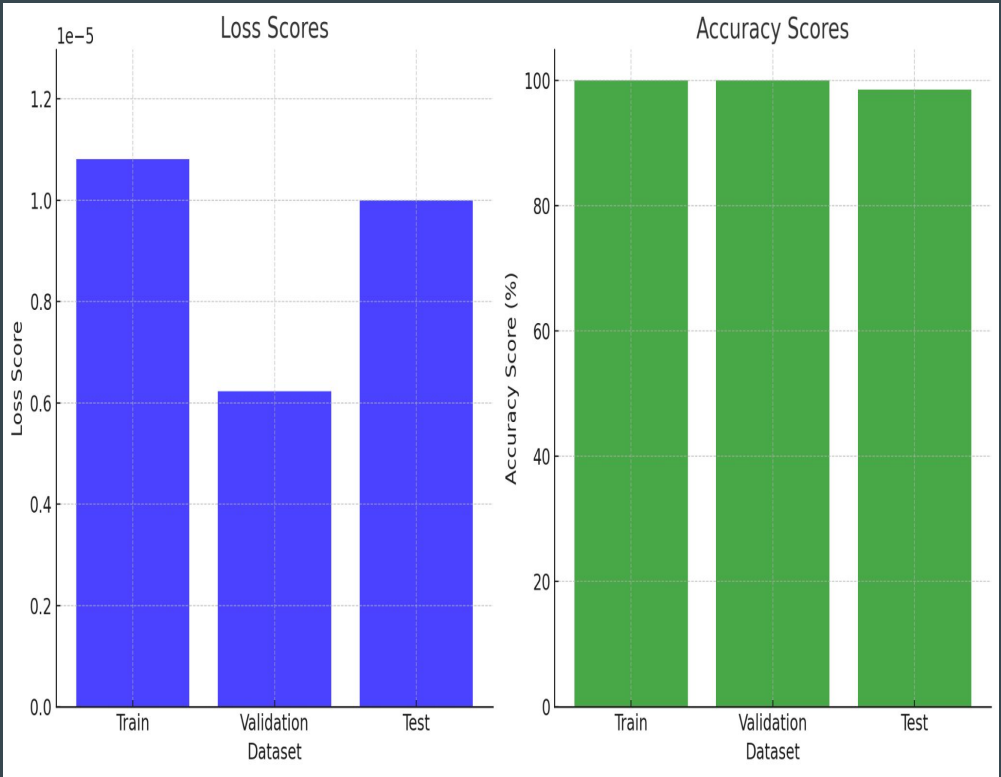
Adjusted from default .001 to .01. Major drop-off in scores.

CNN without Downsampling or Striding (1 Epoch)

Achieved strong results (average test accuracy of 98.5% over 100 evaluations) but was too time-consuming and we needed more epochs to feel confident in our outcome.

Results


CNV
Normal
True label
245
5
2
248
CNV
Normal
Predicted label
200
150
100
50


Loss Scores
1e−5
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Loss Score
Train
Validation
Test
Dataset
Accuracy Scores
100
80
60
40
20
0
Accuracy Score (%)
Train
Validation
Test
Dataset

Suggestions

Save time!	• Batch process hundreds of images in seconds! • Thousands in minutes!
Save money!	• Less time reviewing test results = more patients can be seen
Improve diagnostic accuracy	• Accurately diagnose disease, saving on expensive resources using multiple diagnostics

For Future:

Work more with VGG16: seems promising, but must optimize computation time.

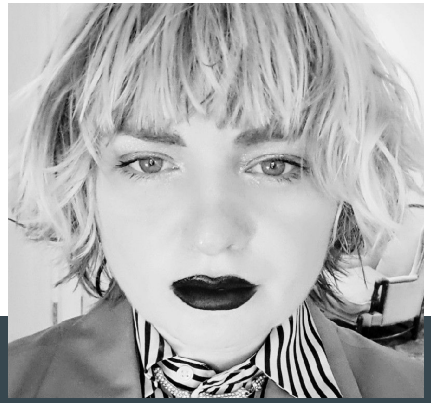
Visualize activation layers for better model understanding.

Optimize for image noise, and make contrast more consistent.

Run more epochs (eg, 100) to improve false negatives.

Questions?

Contact Info



Lou Hines, Data Scientist

lou.m.hines@gmail.com

https://www.linkedin.com/in/lou-hines-data-scientist/