Skin Condition Classification

Presented by Trinity Gahagen

Healthcare is “Skin-Deep”

- According to the American Academy of Dermatology, US healthcare data from 2013 shows that approximately 1 in 4 Americans were impacted by some sort of skin disease
- Skin conditions like acne, psoriasis, and rosacea affect millions of people
- How can we help patients feel more assured when dealing with the process of getting diagnosed?

Consider: Machine Learning

- Machine learning models that can predict what a skin condition could be based on just a picture
- A predictive model deployed onto a health portal can help both patients and healthcare professionals
  - Patients can feel reassured by not having to go for weeks without answers if they have a concern with diagnosing a potential skin condition
  - Can assist doctors to narrow down diagnoses/tests to order

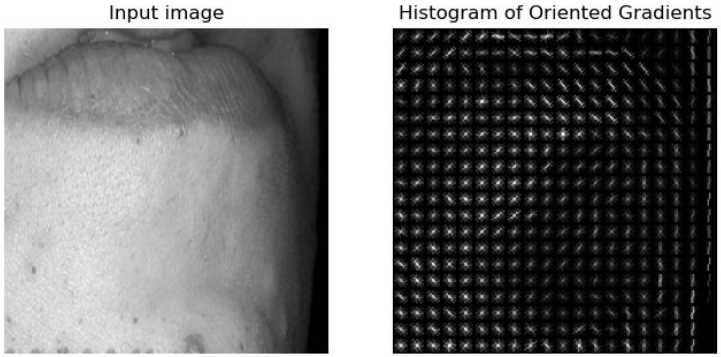
The Data

The data I used consisted of approximately 2.4k total images of 6 different skin conditions:

- Acne
- Carcinoma
- Eczema
- Keratosis
- Milia
- Rosacea

The First Approach: Classical ML with HOG Feature Extraction

- Transformed images using the Histogram of Oriented Gradients (HOG) feature extraction technique
- Initially “shotgunned” four different classifiers:
  - KNN
  - SVM
  - Gradient Boosted Trees
  - Random Forest



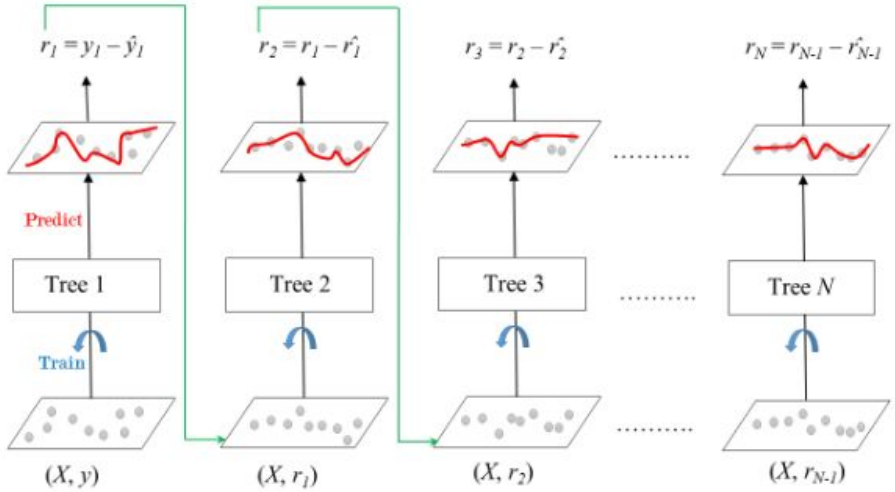


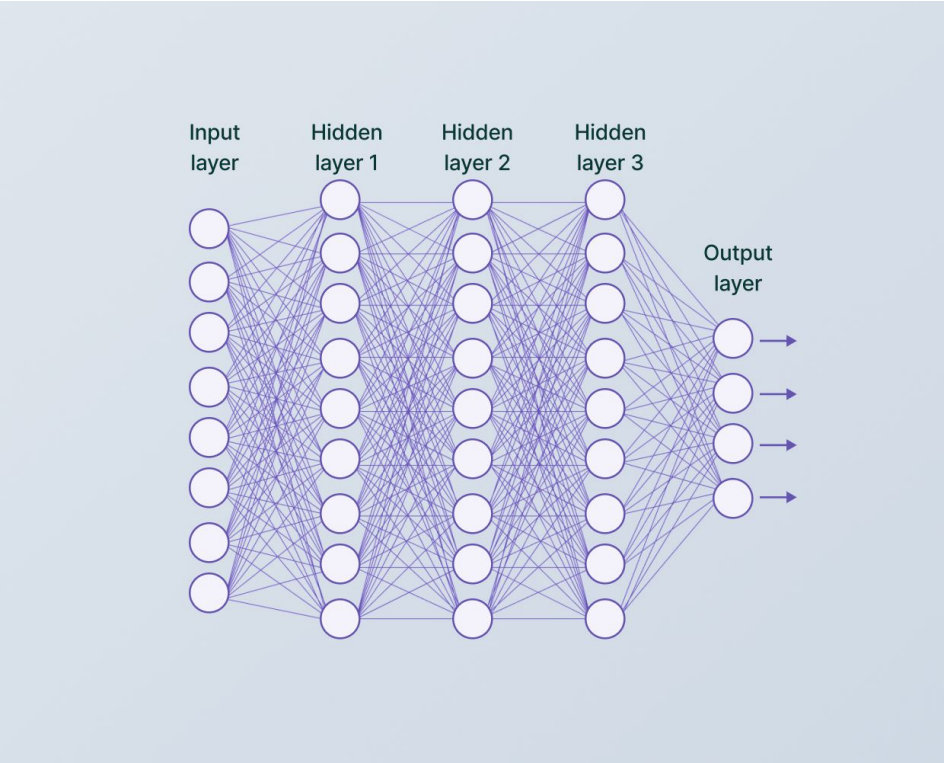
Image source: [Geeks for Geeks]

The Best Model: Gradient Boosted Trees

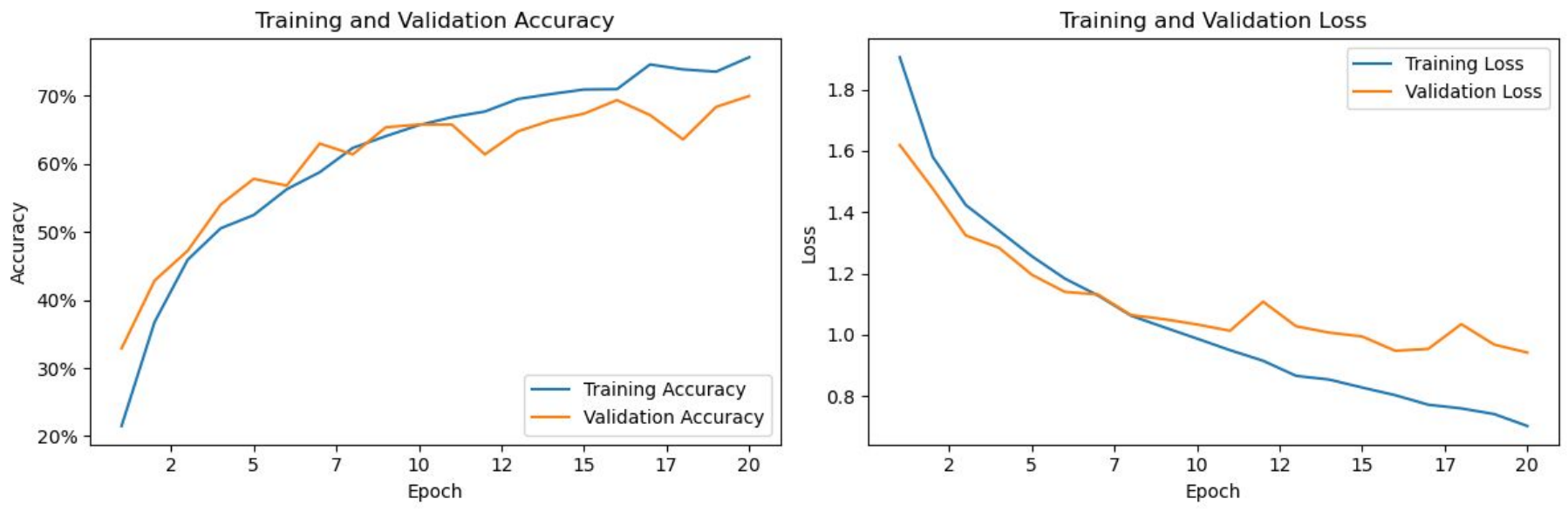
- The best performing model out of the four classifiers listed on the previous slide: **Gradient Boosting Classifier**
- Achieved a peak model accuracy of only **60%**
- Classical machine learning did not do well on this kind of image data
- A more powerful approach is necessary to tackle this task

The Second Approach: Convolutional Neural Networks

- A custom convolutional neural network with the following layers:
  - 2 augmentation layers (horizontal flipping, noise)
  - 2 Convolutional 2D layers (ReLU activation)
  - 1 Max Pooling layer
  - 2 dropout layers
  - 1 fully connected Dense hidden layer (leaky ReLU activation)
  - 1 output layer (softmax activation)

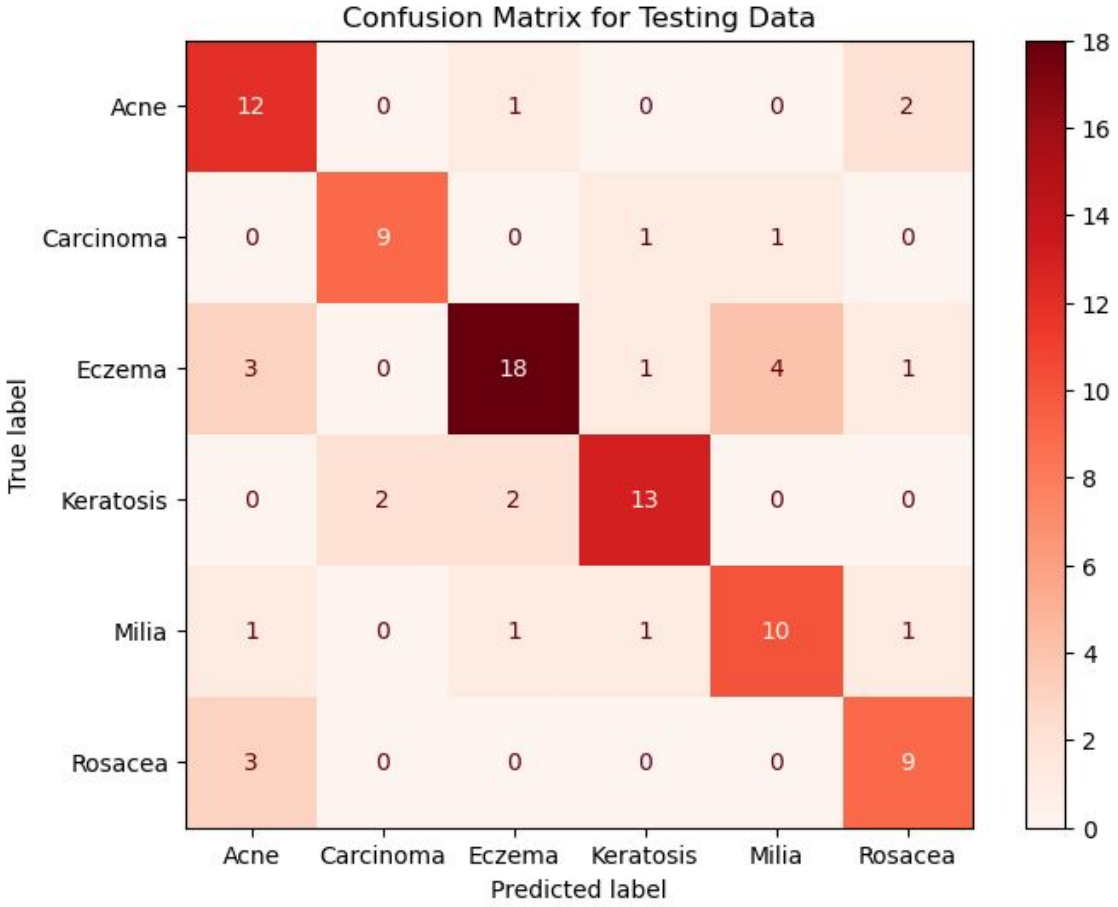


Training/Validation Accuracy and Loss



73%

Testing Accuracy



Results

- While the convolutional neural net did not achieve a very high accuracy, it *improved* the accuracy score from the classical machine learning approach by approximately **13%**
- With more time and research, I would try to implement a pre-trained neural network architecture to increase predictive power

Thank You!

External Resources/Appendix

External Resources: Domain

https://www.healthline.com/health/skin-disorders

https://www.mayoclinic.org/diseases-conditions/rosacea/symptoms-causes/syc-20353815

https://www.aad.org/media/stats-numbers

External Resources: Engineering

https://selfdriving5.github.io/udacity/Self-Driving%20Car%20Engineer%20v5.0.0(us)/Part%2003-Module%2001-Lesson%2001_Object%20Detection/20.%20scikit-image%20HOG.html

https://machinelearningmastery.com/opencv_hog/

https://scikit-image.org/docs/stable/api/skimage.feature.html#skimage.feature.hog

https://medium.com/swlh/histogram-of-oriented-gradients-hog-for-multiclass-image-classification-and-image-recommendation-cf0ea2caaae8

https://customers.pyimagesearch.com/lesson-sample-histogram-of-oriented-gradients-and-car-logo-recognition/

https://learnopencv.com/histogram-of-oriented-gradients/

https://github.com/henrhoi/image-classification/blob/master/feature_extraction_and_exploratory_data_analysis.ipynb

https://medium.com/@sehjadkhoja0/title-exploring-and-analyzing-image-data-with-python-79a7f72f4d2b

Motivation

I actually have an undiagnosed skin condition myself and was motivated to do this type of project because of my complete lack of knowledge of what it could be. After completing the project, I ran an image of my own skin through the model and this is the result:

Rosacea is included in my family history, so I thought this was a really interesting result!

I had to reshape the image and normalize before running it through the model.