Skin Disease Prediction through Image Analysis on Diverse Skin Tones

Amanda Walsh

Introduction

- **Objective:** To offer a triage solution through predictive modeling to a hospital executive board that is experiencing long wait times for appointment scheduling in Dermatology

- **Goal:** To show how AI can be one solution for triaging patients and making sure prompt care is provided to those who need it.

  Educate the board on the importance of using ethical models that are representative of diverse skin tones, to ensure equality in care across all groups.

Bias in Dermatology

- Lack of diversity in diagnostic images
- Doctor training across the board
- Limited, dated medical research

Warning- Medical Images

Data Collection

- Stanford University "Diverse Dermatology Images" Dataset
- Selected at Stanford Clinics from 2010-2020
- 656 Images of biopsy proven skin diseases
- Skin tone labeled and cross-referenced using Fitzpatrick Scale
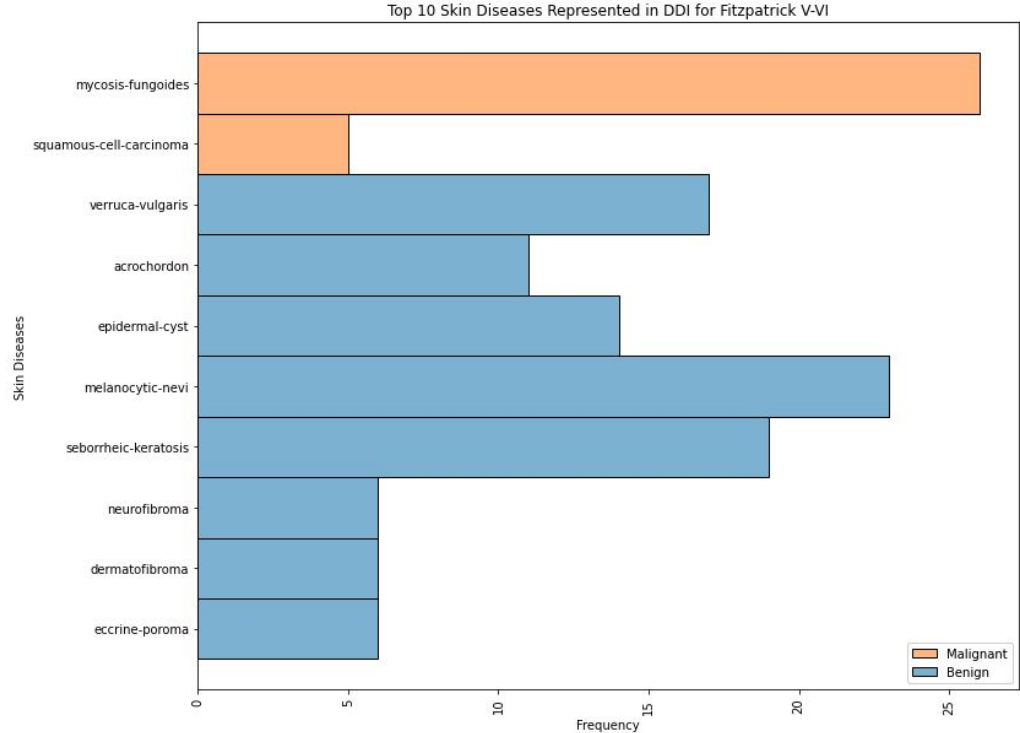
What is the Fitzpatrick Scale?

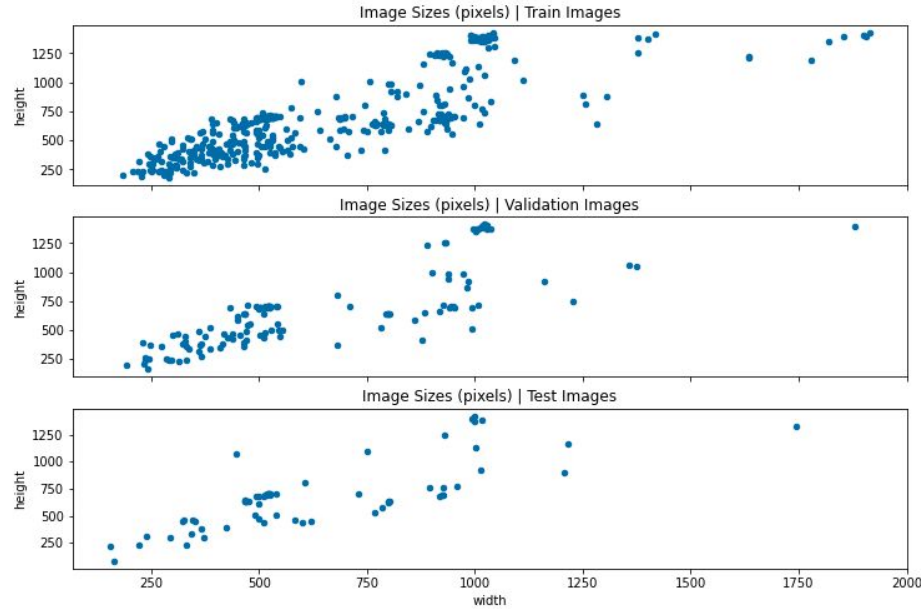| Score	Description	Female	Male
0–6 / Type I	**Pale white skin** — Extremely sensitive skin, always burns, never tans. *Example: red hair with freckles*		
7–13 / Type II	**White skin** — Very sensitive skin, burns easily, tans minimally. *Example: fair skinned, fair haired Caucasians, northern Asians*		
14–20 / Type III	**Light brown skin** — Sensitive skin, sometimes burns, slowly tans to light brown. *Example: darker Caucasians, some Asians*		
21–27 / Type IV	**Moderate brown skin** — Mildly sensitive, burns minimally, always tans to moderate brown. *Example: Mediterranean and Middle Eastern Caucasians, southern Asians*		
28–34 / Type V	**Dark brown skin** — Resistant skin, rarely burns, tans well. *Example: some Hispanics, some Africans*		
35+ / Type VI	**Deeply pigmented dark brown to black skin** — Very resistant skin, never burns, deeply pigmented. *Example: darker Africans, Indigenous Australians*		

- Fitzpatrick (FST) I-VI
- Score 1-35
- Describes skin's reaction to sunlight exposure commonly used by medical providers to describe patient skin color
- **Limitations**

Image Distribution in Data

- 656 images
- 208 FST I-II Images (Fair Skin Tones)
  - 159 benign, 49 malignant
- 241 FST III-IV Images (Medium Skin Tones)
  - 167 benign, 74 malignant
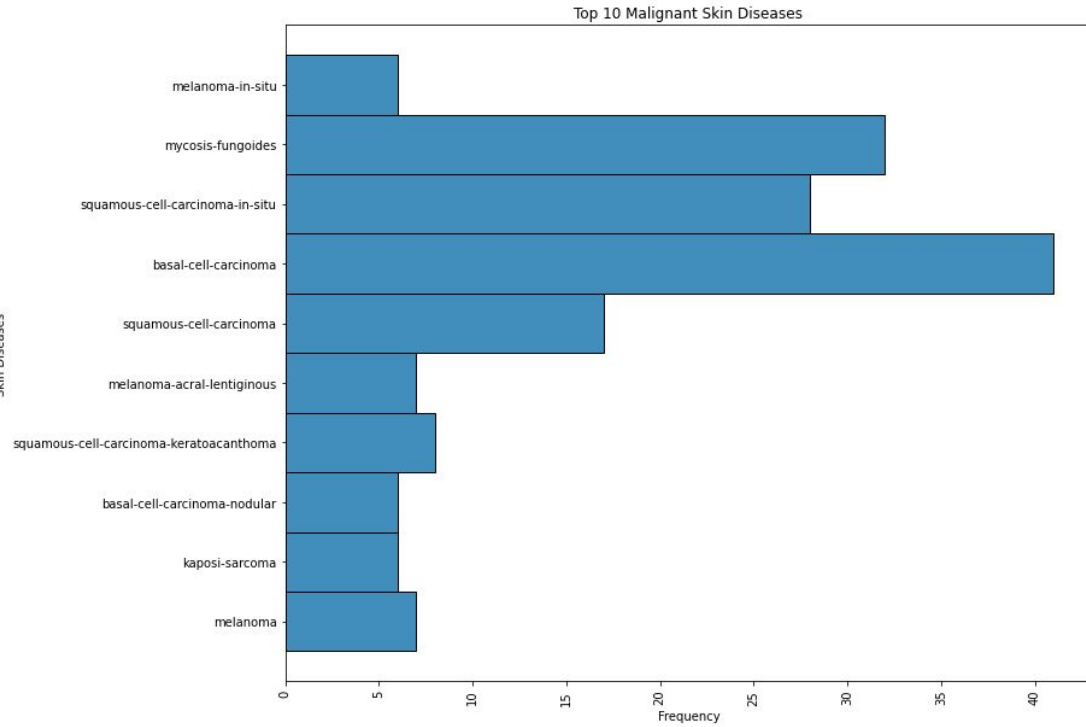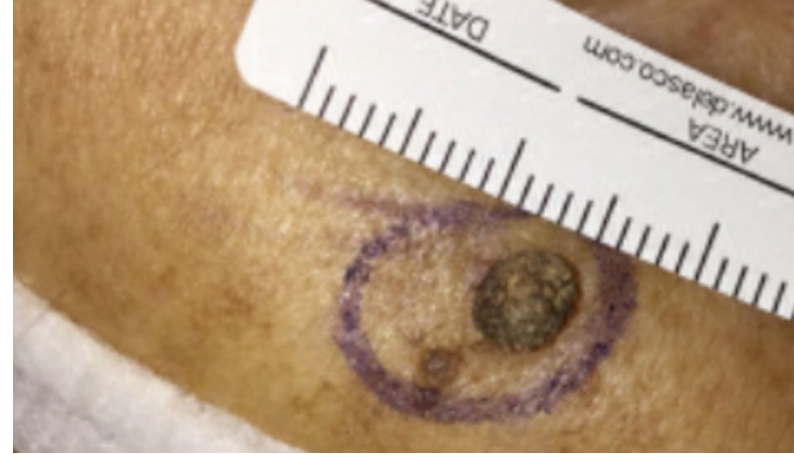- 207 FST V-VI Images (Dark Skin Tones)
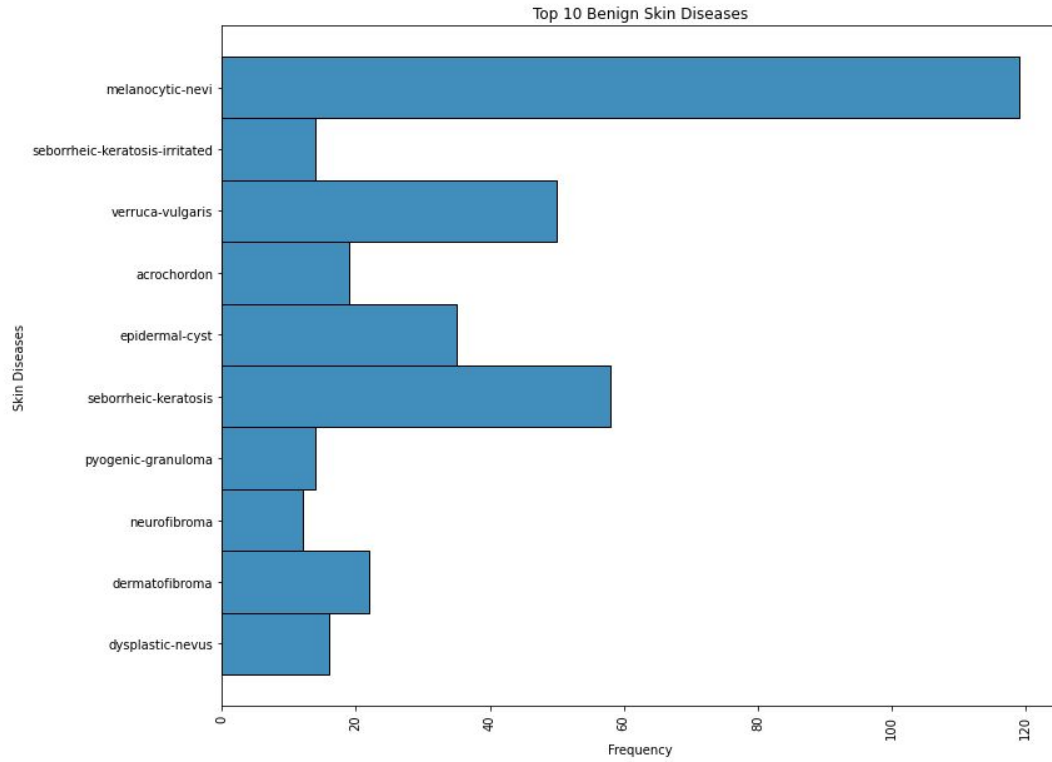 - 159 benign, 48 malignant

Top 10 Skin Diseases Represented in DDI for Fitzpatrick V-VI

Data Science Workflow

Image Sizes (pixels) | Train Images

Image Sizes (pixels) | Validation Images

Image Sizes (pixels) | Test Images

- Cleaned Data
- Exploratory Data Analysis

Top 10 Benign Skin Diseases

Data Science Workflow

- Preprocessed images
- Multiple Convolutional Neural Network Models
  - Recall Optimization (Minimizing False Negatives)
- Best Model Results Report:

	precision	recall	f1-score	support
benign	0.89	0.88	0.88	48
malignant	0.67	0.71	0.69	17
accuracy			0.83	65
macro avg	0.78	0.79	0.78	65
weighted avg	0.83	0.83	0.83	65

benign 100% (benign)

malignant 100% (malignant)

benign 89% (benign)

malignant 100% (malignant)

benign 100% (benign)

malignant 100% (malignant)

benign 100% (malignant)

benign 64% (malignant)

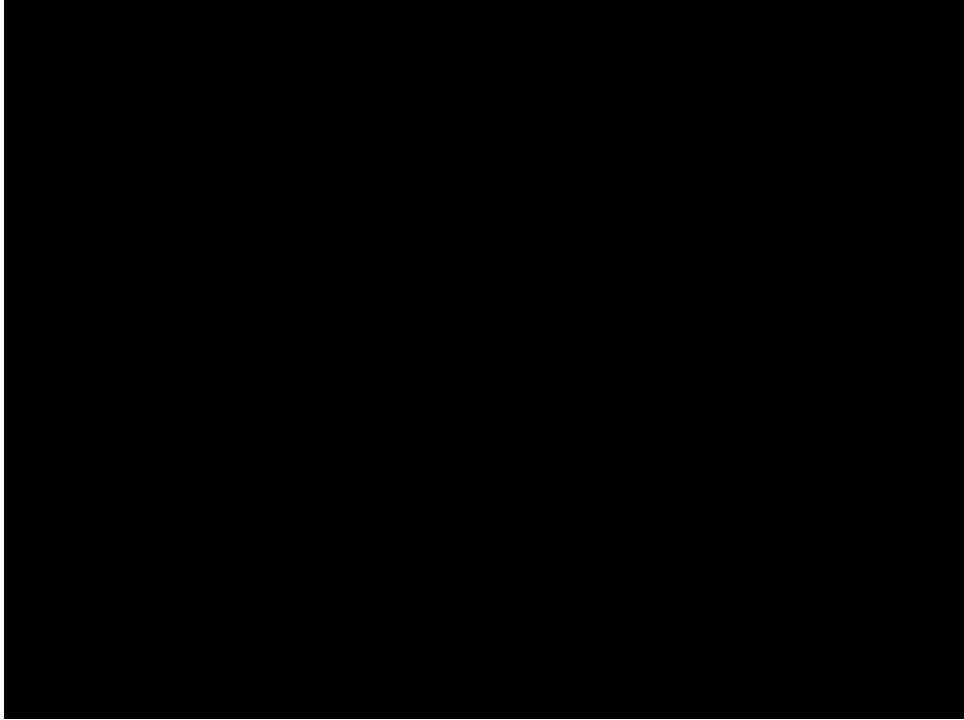benign 100% (malignant)

malignant 100% (benign) malignant 100% (benign) malignant 100% (benign)

Conclusions and Recommendations

- Training models on diverse populations
- Knowing model limitations
- Making sure models are up to date and ethical

Final Thoughts:

- The Diverse Dermatology Images Dataset needs more images, class imbalance should be addressed
- Model and application improvement

Thank you!