STAT 405 FINAL PROJECT

SKIN CANCER ANALYSIS: DNN

Kaiwei Tu, Joshua Hopwood, Brandon Wolf

Background

5.4 MM

Cases of skin cancer diagnosed per year

20%

Of Americans will develop skin cancer by 70

99%

survival rate with early detection

Lit Review

84%

sensitivity and specificity by dermatologists for BCC

15%

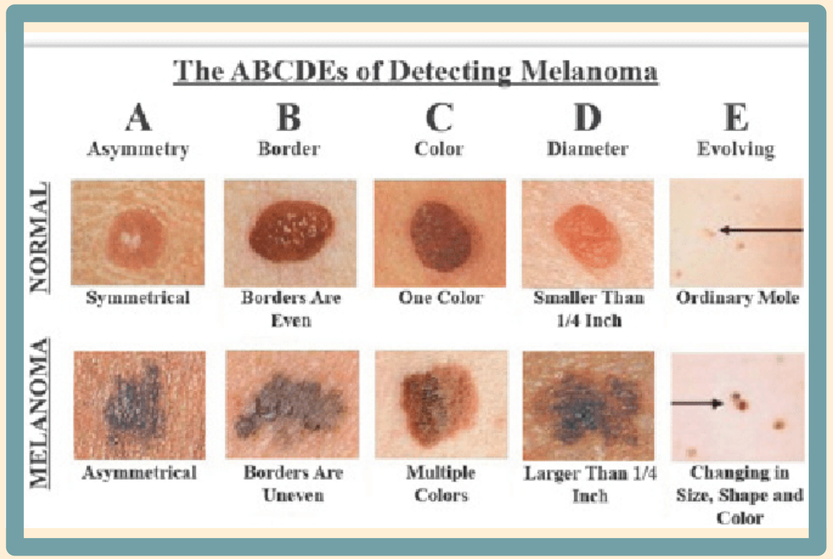
increase in accuracy when doctors use AI

10%

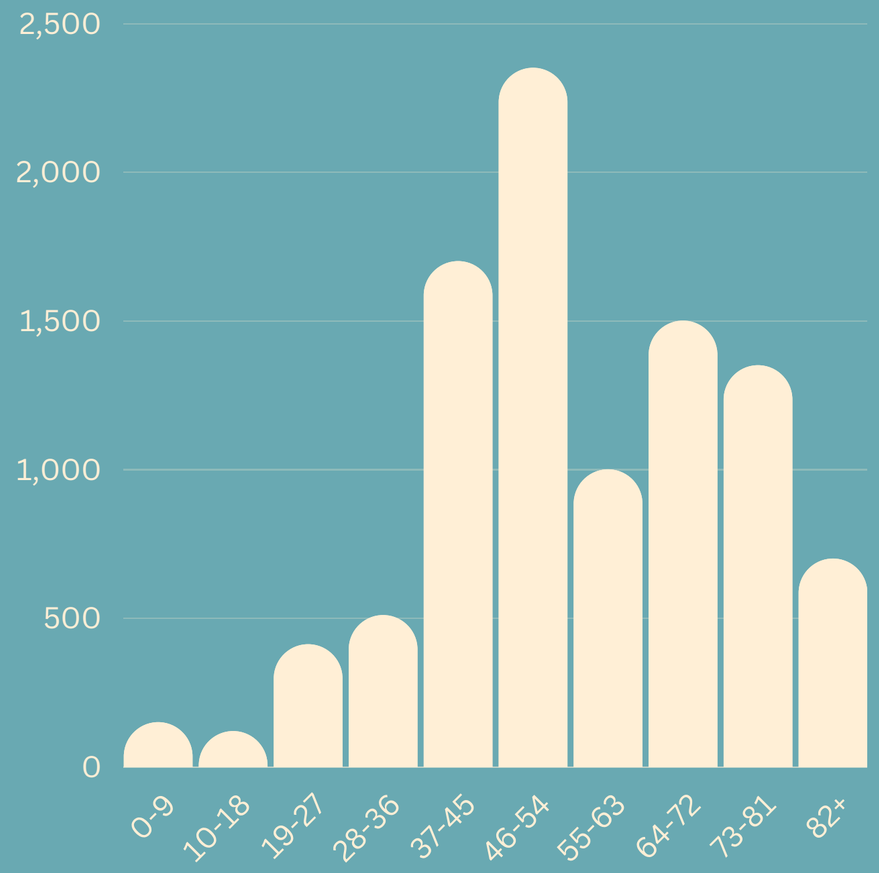
average image failure according to the 2018 classification study

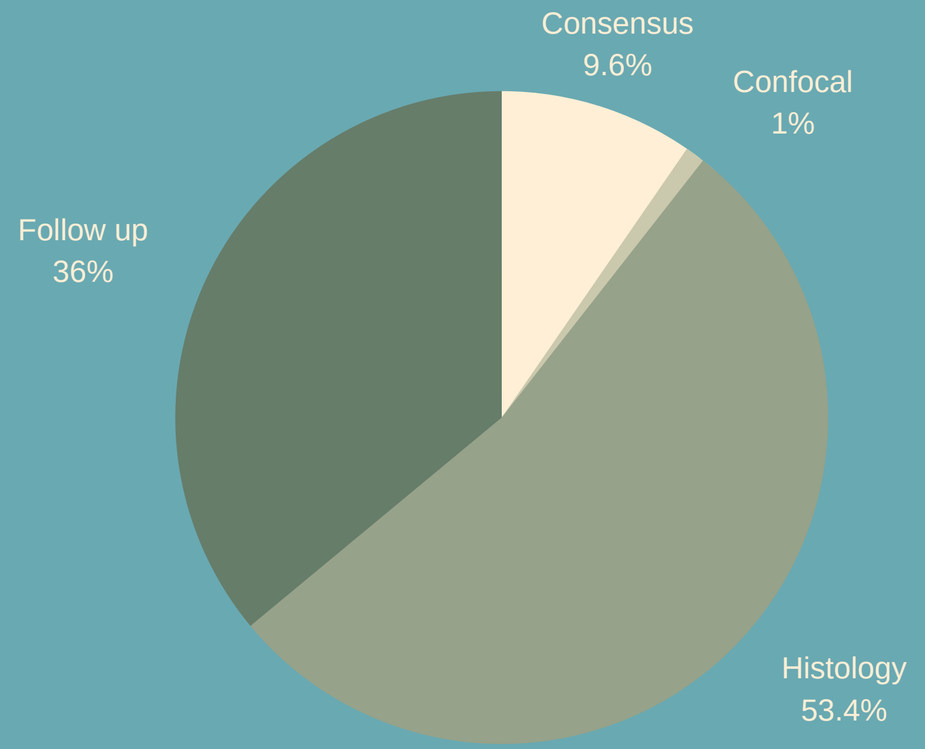
Description

- 10000 training Jpegs
  - 300kb each
- 1500 test
- Metadata: Age, Locus, Sex, Histology
- Competition - to compare our accuracies and efficacy



Distribution of Images by Age





Diagnosis Denotation

Packages

PyTorch with a conda environment

Preprocessing

Need to homogenize the location of the object in the center of frame.

Encoding

Used one hot encoding for metadata to be added to a later layer

Network

We used res net with some unfrozen layers, and later an ensemble

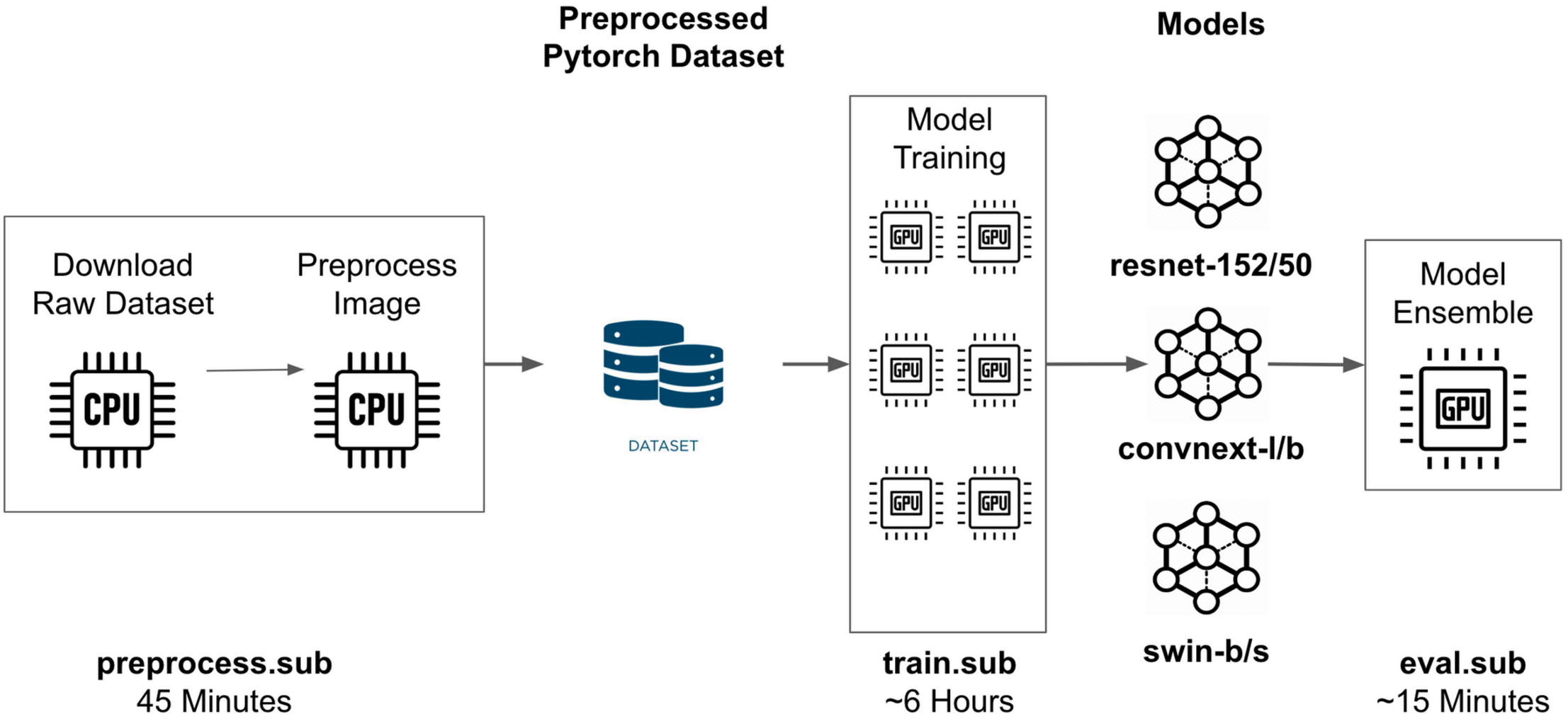
Metadata

Added this encoded data as a layer after running the images through the first few layers

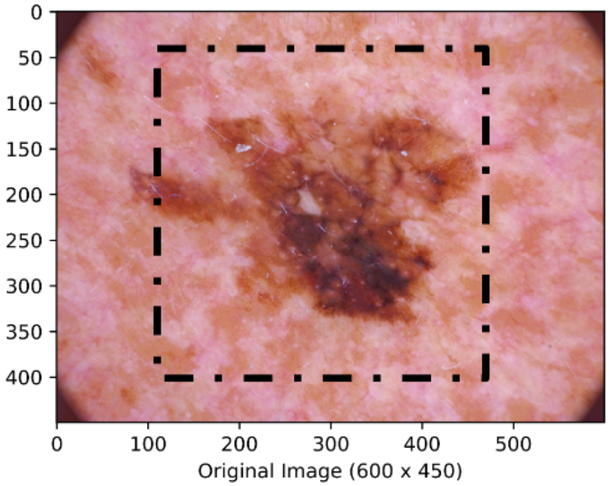
Results

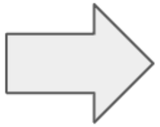
Differs from other methodologies. Parallel training for ensemble.

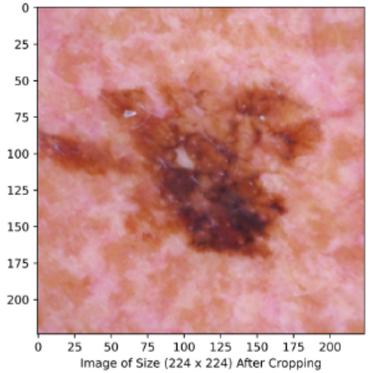
Computation Pipeline on CHTC

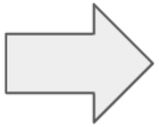


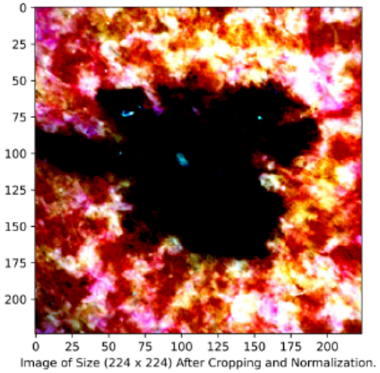
Preprocess









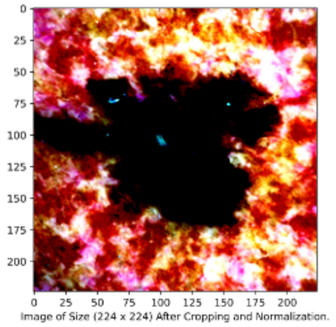


Original Image
(600 x 450)

Resized and Cropped Image
(224 x 224)

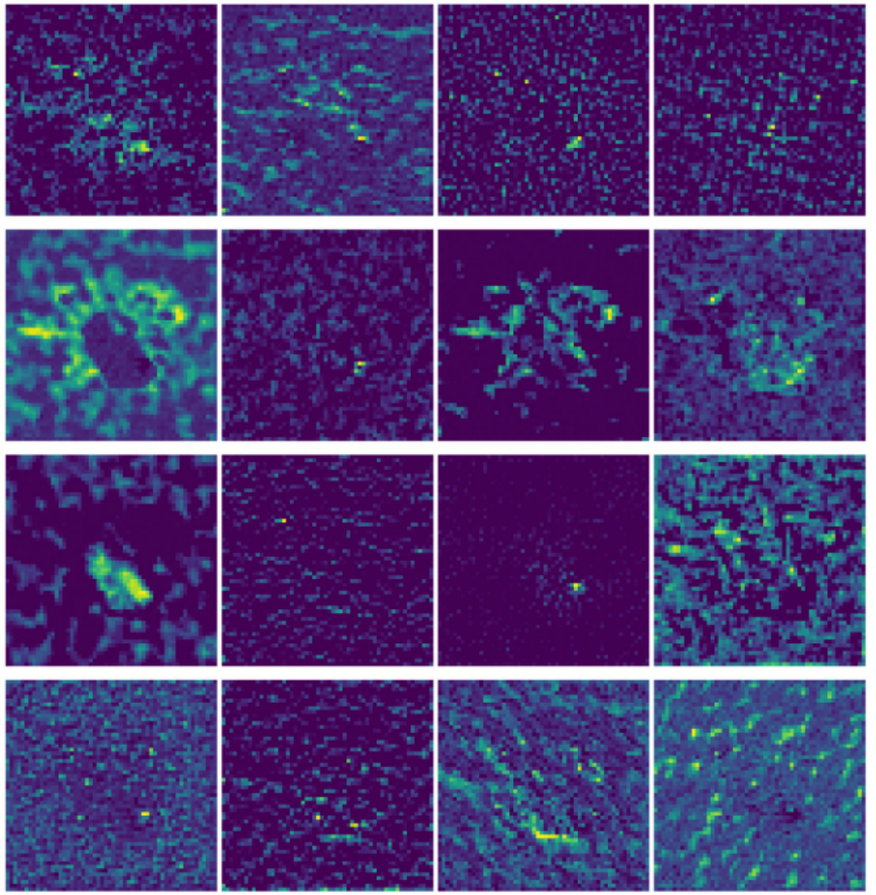
Normalized Image
(224 x 224)

Feature Extraction



Input Image
(224 x 224)

Deep
Neural
Network



Extracted Features
(28 x 28)

Deep
Neural
Network

[
0.14,
0.33,
0.12,
.
.
.
0.08,
0.19
]

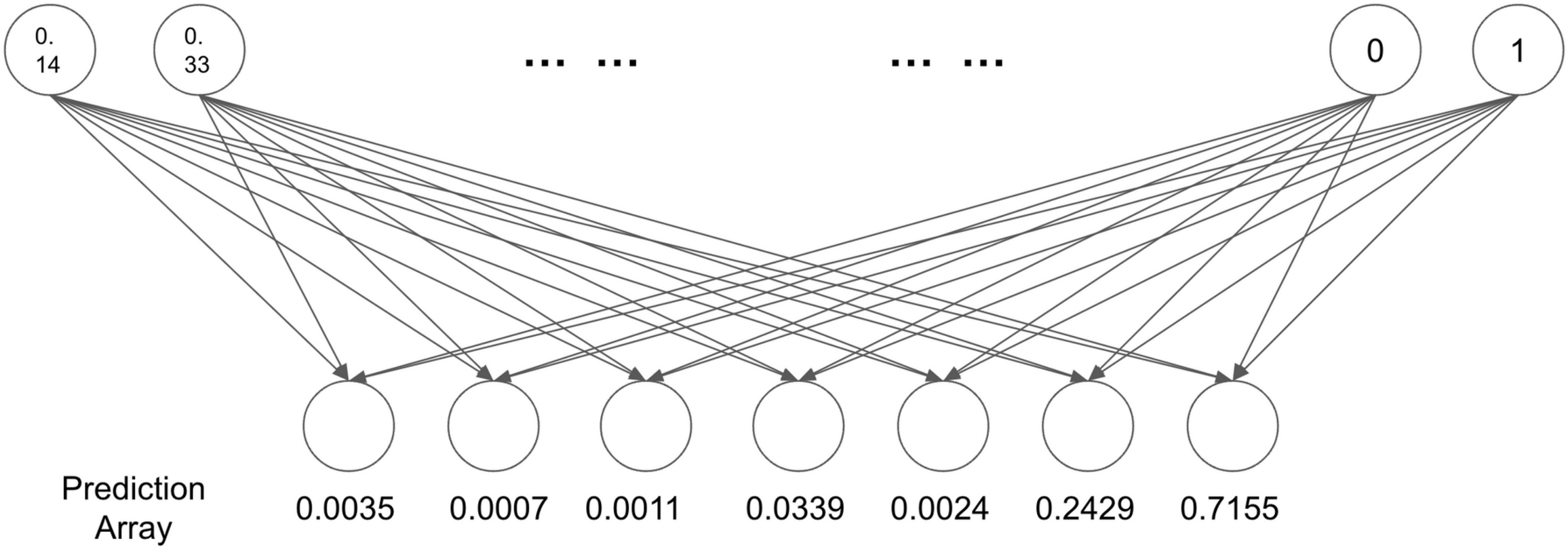
Features Vector
1 x 2048

Prediction

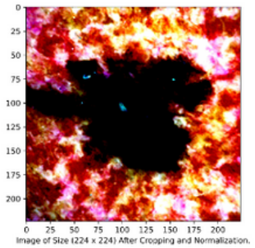
[0.14, 0.33, … , 0.08, 0.19] + **[0, 1, …, 0, 1]**

Feature Vector

Metadata Encoding



Ensemble Prediction



Resnet-152/50

Swin-b/s

Convnext-l/b

0.00, 0.00, 0.00, 0.04, 0.00, 0.24, 0.72

0.00, 0.00, 0.00, 0.08, 0.00, 0.20, 0.72

0.00, 0.00, 0.00, 0.03, 0.00, 0.24, 0.73

0.00, 0.03, 0.00, 0.00, 0.00, 0.24, 0.73

0.00, 0.00, 0.00, 0.06, 0.00, 0.14, 0.80

0.00, 0.00, 0.03, 0.06, 0.00, 0.11, 0.80

Average

[
0.00,
0.00,
0.00,
.
.
.
0.20,
0.75
]

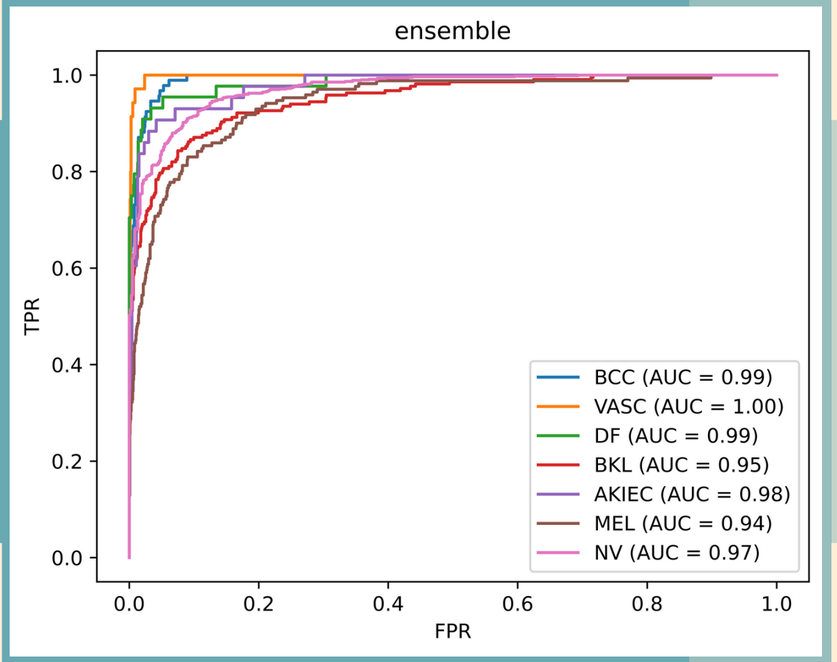
Input Image
(224 x 224)

Deep Neural
Network

Prediction
Array

Ensembled
Prediction
Array

RESULTS



The resnet/swin/convnext ensemble

Overall accuracy of 84%
BAC = 81.4

We ran separate jobs for parallel preprocessing and for the networks (3 gpu)

Memory utilized: 15GB
Disk space required: 40GB
NVIDIA A100-SXM4-80GB

	BCC	VASC	DF	BKL	**AKIEC**	**MEL**	NV	Mean
All human readers	86	**100**	77.3	80.6	48.8	73.1	88.8	79.2
Expert readers	**89.2**	**100**	**86.4**	**84.3**	51.2	67.8	**89.3**	81.2
Our Algorithm	88.2	88.6	79.5	76.5	**74.4**	**75.4**	87.3	**81.4**

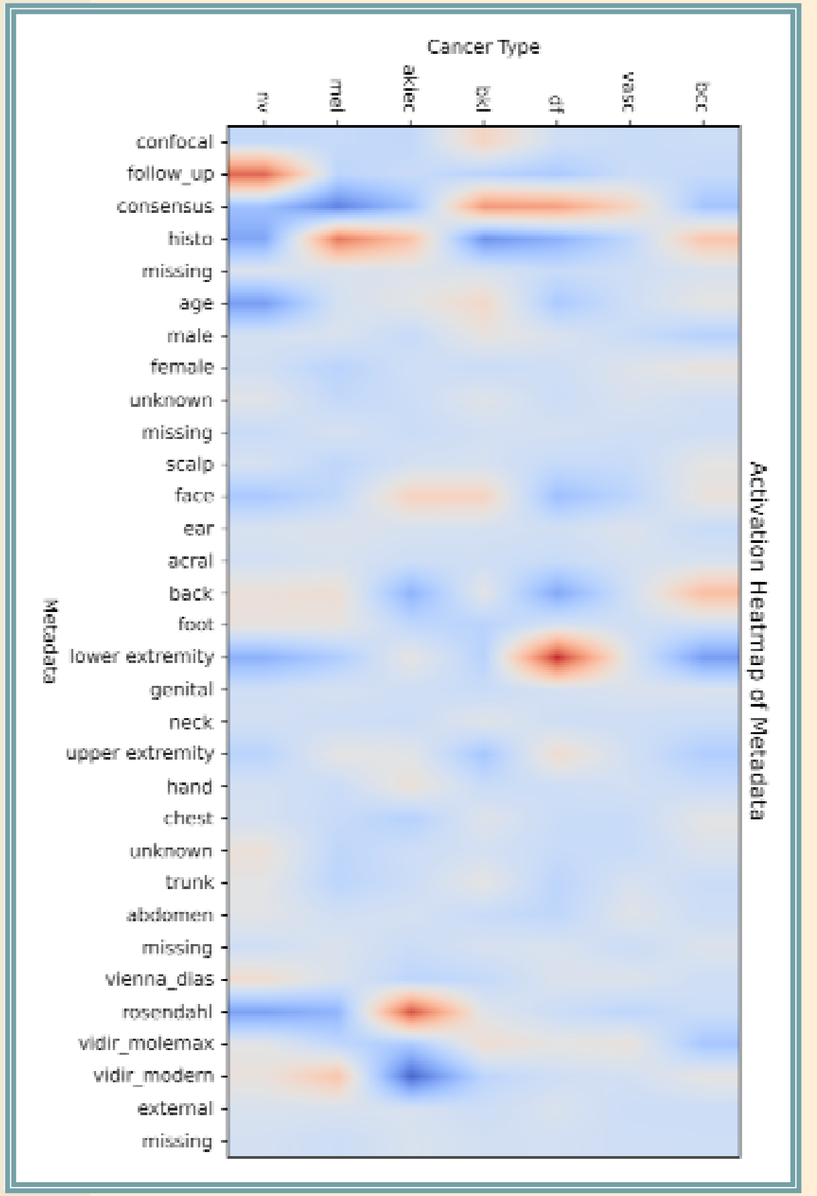
Table 1: Prediction Sensitivity: Human Readers vs. Algorithms.

	BCC	VASC	DF	BKL	AKIEC	MEL	NV
BCC	82	0	1	2	5	2	1
VASC	3	31	0	0	0	0	1
DF	4	0	35	2	1	0	2
BKL	6	0	3	166	10	20	12
AKIEC	3	0	1	4	32	3	0
MEL	3	2	1	10	3	129	23
NV	9	2	7	34	3	60	793

- Best = VASC @ 88.6
- Worst = AKIEC @ 74.4

	BCC	VASC	DF	BKL	AKIEC	MEL	NV	Mean Value
Accuracy	0.974	0.995	0.985	0.932	0.978	0.916	0.898	0.954
Specificity	0.98	0.997	0.991	0.96	0.985	0.937	0.935	0.969
Sensitivity	0.882	0.886	0.795	0.765	0.744	0.754	0.873	0.814
AUC	0.992	0.998	0.986	0.95	0.978	0.943	0.972	0.974

- Best AUC = .998 (VASC)
- Worst = .943 (MEL)



CONSIDERATIONS

- The most common physical indicator is a lump or localized swelling. Our data is trained on 2d images, which might limit efficacy.
- May have issues with lighting variability and lack of standardization
- Ways to improve:
  - Lighting
  - Dimensionality
  - More runs with different layers

SOURCES

https://pubmed.ncbi.nlm.nih.gov/9602875/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1495212/

https://www.sciencedirect.com/science/article/pii/S0959804919303491

https://www.ncbi.nlm.nih.gov/books/NBK68729/

https://challenge.isic-archive.com/landing/2018/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4513405/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7519424/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8237239/pdf/nihms-1716706.pdf