CSE 472 : Machine Learning

Skin Lesion Classification

Presented By :
Abrar Mahmud 1905097
Shahriar Raj 1905105

Problem Definition

Skin lesions, including skin cancer, are increasing globally, with significant cases attributed to prolonged sun exposure. Traditional diagnostic methods like dermoscopy enhance detection accuracy, but human factors such as fatigue and mental state contribute to diagnostic errors, with up to 21% misdiagnoses by dermatologists. The growing need for objective, rapid, and reliable classification of skin lesions underscores the importance of computer-aided systems.

Base Paper Link: https://ieeexplore.ieee.org/document/9873756

We did extensive analysis on multiple models to get better result for Skin Lesion Classification than the base paper. We also tried our best models on the new and updated dataset.

Dataset & Analysis

SKin Disease Type	Number of Samples	Percentage
Pigmented Nevus	6707	66.9562%
Melanoma	1113	11.1111%
Benign Keratosis	1099	10.9714%
Basal Carcinoma	514	5.1313%
Actinic Keratosis	327	3.2644%
Vascular disease	142	1.4176%
Dermatofibroma	115	1.1480%

Total Samples: 10017
Total Classes: 7

Dataset LinK : https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

Data Preprocessing Steps

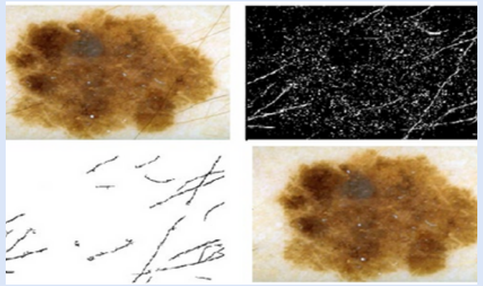
Crop Image

600 * 450

299 * 299

224 * 224

Removal of Hair



Amplification

Increasing /Decreasing/Unchanged

INPUT

Crop Image

- First Convert the 600 x 450 images into 299 x 299 images
- Now crop the images to make it 224 x 224

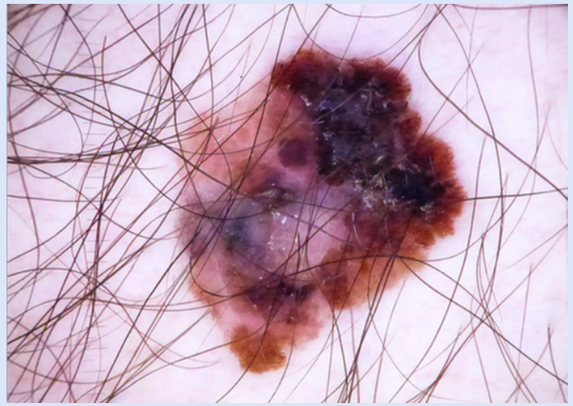
600 x 450 ⟶ 299 x 299 ⟶ 224 x 224

Removal of Hair Proceedure

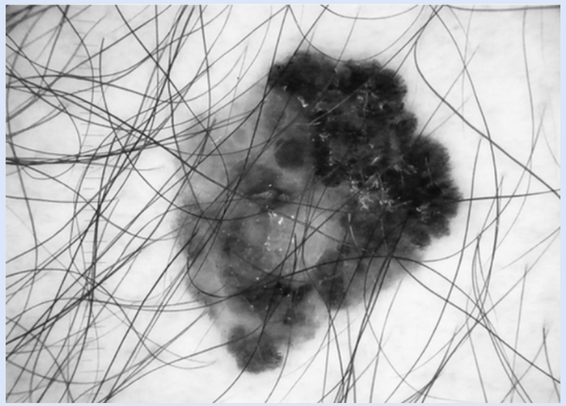
- After cropping the image, hair removal is done using dullrazor algorithm.
- The image is converted to grayscale as it reduces the image from three channels (RGB) to a single intensity channel.
- Black hat filter is applied to highlight dark structures (such as hair) on a lighter background, effectively isolating the hair patterns.
- A Gaussian blur is applied to the blackhat image to smooth the isolated hair patterns and reduce noise.
- Then a binary threshold (cv2.threshold) is applied to the blurred image, where pixel values greater than 10 are set to 255 (white) and others to 0 (black), generating a binary mask that isolates the hair patterns.
- The inpainting function is used to remove hair patterns from the original image using the binary mask.

Removal of Hair

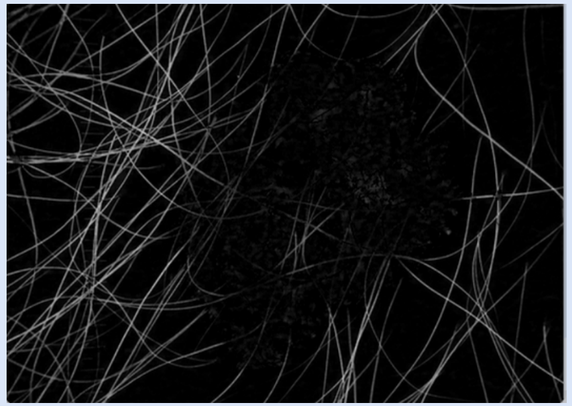
Cropped Image

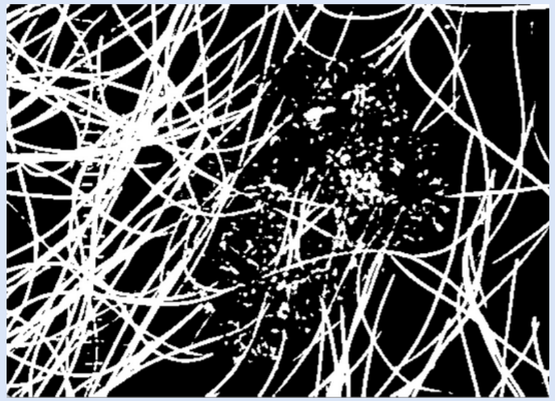


Grayscale

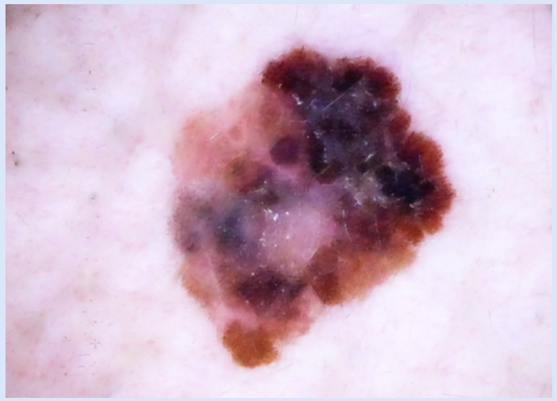


Blackhat





Binary Mask

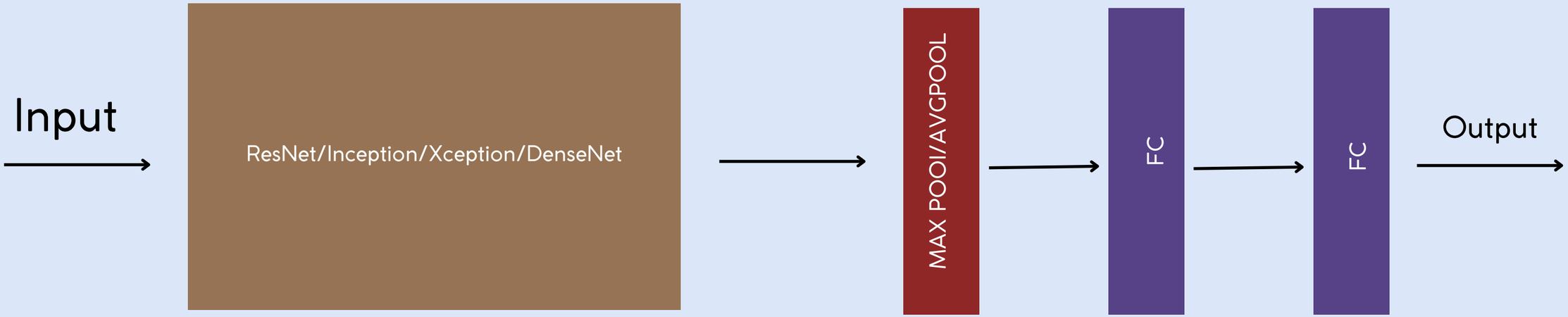


Clean Image

Amplification of Dataset

SKin Disease Type	Number of Samples	Adjusted Ratio	Adjusted Samples
Pigmented Nevus	6705	0.15	1005
Melanoma	1113	1	1113
Benign Keratosis	1099	1	1099
Basal Carcinoma	514	2	1028
Actinic Keratosis	327	3	981
Vascular disease	142	8	1136
Dermatofibroma	115	8	920

Model Architecture



Base Paper Performance Metrics

	Network	Categories	accuracy
Reference 1	VGGNet	2	78.66%
Reference 2	Res-152+SVM	2	86.28%
Reference 3	VGG19	3	71.34%
Reference 4	VGGNet-19	7	81.93%
Reference 4	Inception-V3	7	81.01%
Reference 4	ResNet-50	7	83.26%
This paper	InceptionV3	7	85.80%
This paper	ResNet50	7	86.69%
This paper	DenseNet201	7	86.91%

Performance Report

Best Model

We got the best result from Inception ResNetV2. The accuracy from the proposed architecture was 89.44%

Best Model From the Base Paper

The Base paper recorded 86.91% accuracy with the help of model DenseNet201

Performances of Different Models

ResNet50

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	99.61%	84.37%
Average	64	ReLU	0.5	97.66%	84.03%
Average	128	ReLU	0.2	98.95% (Epoch-30)	83.62%
Average	32	ReLU	0.2	97.99%	82.45%
Max	64	ReLU	0.2	98.90% (Epoch-30)	85.20%
Max	64	Sigmoid	0.2	90.04%	79.85%

ResNet152

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	91.96%	81.84%
Average	64	ReLU	0.5	94.20%	81.15%
Average	32	ReLU	0.2	87.36%	78.55%
Max	64	ReLU	0.2	91.95%	77.86%
Max	64	Sigmoid	0.2	86.60%	80.19%

InceptionV3

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	99.98%	86.77%
Average	64	ReLU	0.5	99.79%	87.32%
Average	128	ReLU	0.2	100.00%	87.80%
Average	128	ReLU	0.5	99.97%	88.35%
Average	256	ReLU	0.5	100.00%	87.59%
Average	32	ReLU	0.5	97.25%	87.39%
Max	64	ReLU	0.2	100.00%	86.50%
Max	64	Sigmoid	0.2	99.74%	86.98%

Inception ResNet V2

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	99.85%	87.39%
Average	64	ReLU	0.5	99.91%	87.80%
Average	128	ReLU	0.2	99.98%	88.62%
Average	128	ReLU	0.5	100.00%	89.44%
Average	32	ReLU	0.5	99.62%	86.98%
Max	128	ReLU	0.5	99.38%	85.95%
Max	64	ReLU	0.2	100.00%	88.62%
Max	64	Sigmoid	0.2	100.00%	86.91%

Xception

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	100.00%	89.03%
Average	64	ReLU	0.5	100.00%	88.97%
Average	128	ReLU	0.2	100.00%	88.90%
Average	32	ReLU	0.2	100.00%	88.97%
Max	64	ReLU	0.2	100.00%	89.03%
Max	64	Sigmoid	0.2	100.00%	88.97%
Average	64	Sigmoid	0.2	100.00%	88.62%

DenseNet121

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	98.88%	87.87%
Average	64	ReLU	0.5	98.69%	86.50%
Average	128	ReLU	0.2	99.85%	88.49%
Average	256	ReLU	0.5	98.06%	86.29%
Max	64	ReLU	0.2	99.67%	87.53%
Max	64	Sigmoid	0.2	98.06%	86.98%

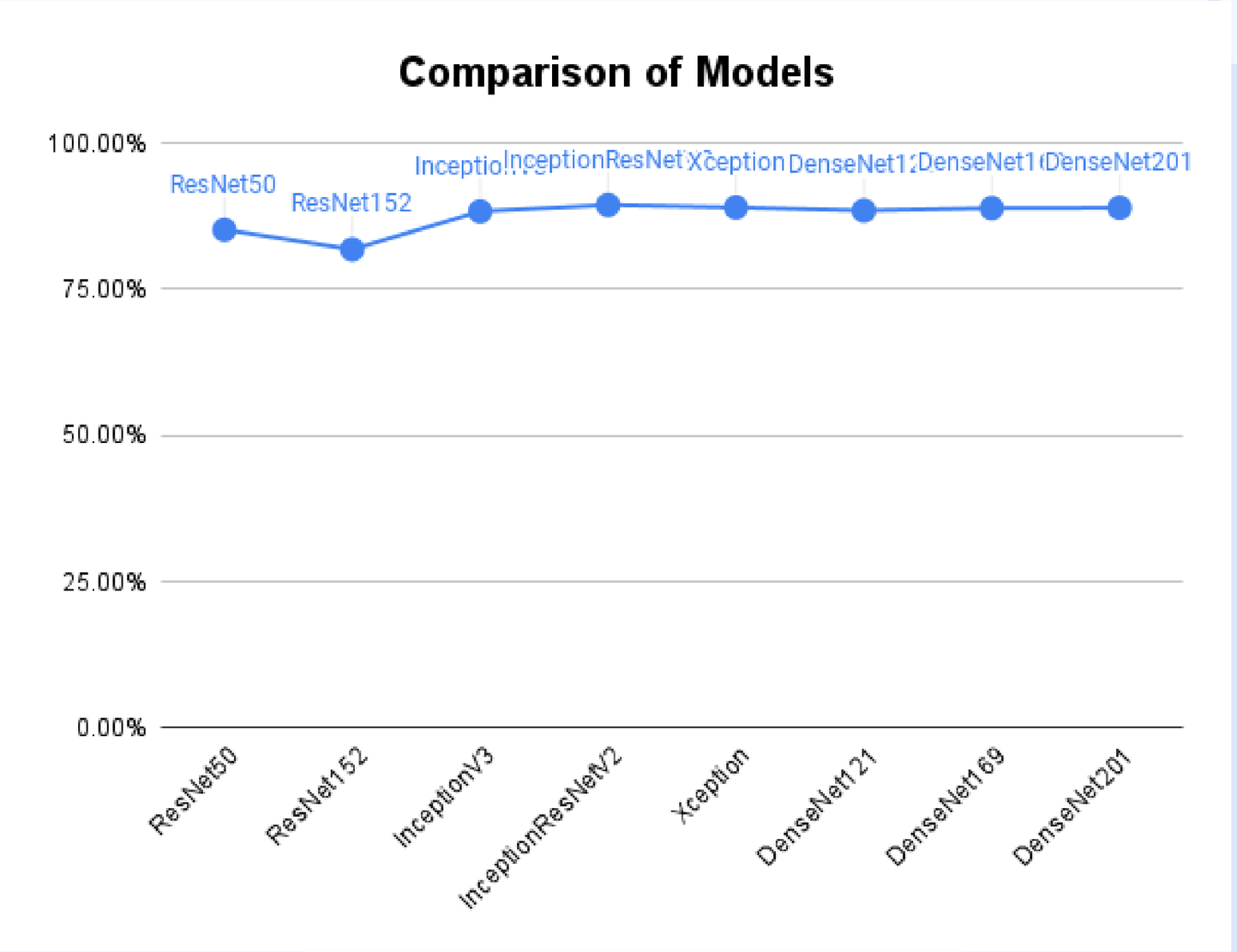
DenseNet169

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	99.42%	87.53%
Average	64	ReLU	0.5	97.30%	86.50%
Average	128	ReLU	0.2	99.86%	88.90%
Average	32	ReLU	0.2	97.63%	87.39%
Max	64	ReLU	0.2	98.97%	86.63%
Max	64	Sigmoid	0.2	96.72%	86.63%

DenseNet201

Pooling	Batch Size	Activation	Dropout	Training Accuracy	Testing Accuracy
Average	64	ReLU	0.2	99.50%	88.28%
Average	64	ReLU	0.5	97.78%	87.80%
Average	128	ReLU	0.2	99.98%	88.97%
Average	32	ReLU	0.2	97.42%	87.11%
Max	64	ReLU	0.2	99.88%	87.87%
Max	64	Sigmoid	0.2	98.21%	87.39%

Comparison of Models



New Dataset & Analysis

Skin Disease Type	Number of Samples	Percentage
Pigmented Nevus	12875	50.82704986%
Melanoma	4522	17.85164423%
Benign Keratosis	2624	10.35884884%
Basal Carcinoma	3323	13.11831353%
Actinic Keratosis	867	3.422683668%
Vascular disease	253	0.9987762031%
Dermatofibroma	239	0.943507954 7%
Squamous Cell Carcinoma	628	2.4791%

Total Samples: 25331
Total Classes: 8

Dataset LinK : https://www.kaggle.com/datasets/andrewmvd/isic-2019/data

Amplification on New Dataset

SKin Disease Type	Number of Samples	Adjusted Ratio	Adjusted Samples
Pigmented Nevus	12875	0.14	1802
Melanoma	4522	0.4	1808
Benign Keratosis	2624	0.7	1836
Basal Carcinoma	3323	0.54	1794
Actinic Keratosis	867	2	1734
Vascular disease	253	7	1771
Dermatofibroma	239	7	1673
Squamous Cell Carcinoma	628	3	1884

Performance on New Dataset

Model	Training Accuracy	Testing Accuracy
Inception ResNetV2	100.00%	83.17%
Xception	100.00%	84.46%
DenseNet169	99.52%	82.93%

Challenges

As training CNN models need a lot of computing power, we had to use free GPU of kaggle which had a time limit. Preprocessing the dataset was also a challenge

Discussion

Result

We almost got a 3% of improvement from the base paper.

Other Experiments

We used models like visual transformer but gave accuracy below 80%

Thank you