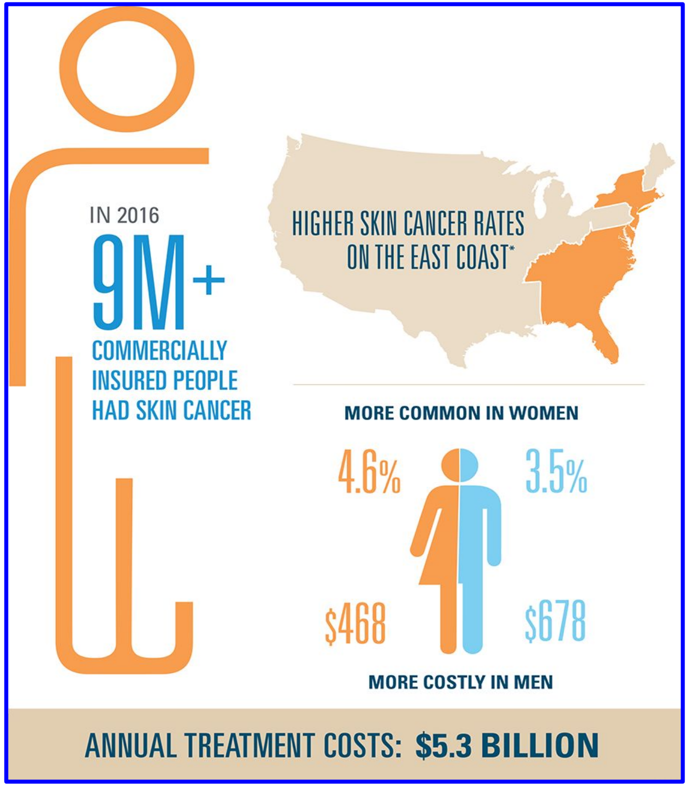
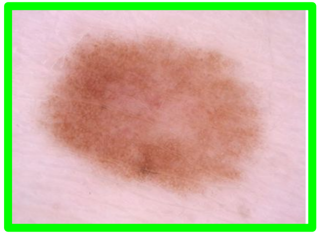
Background - Skin Cancer

- 1 in 3 cancer patients have a form of skin cancer
- Over 9,000 patients are diagnosed on a daily basis in the US.
- Skin cancer hits rural areas most with lack of access to professional healthcare workers and equipment.
- Many cases go undetected, and proper care can't be administered in time.

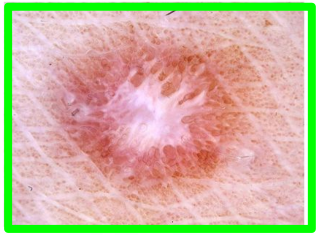


Research

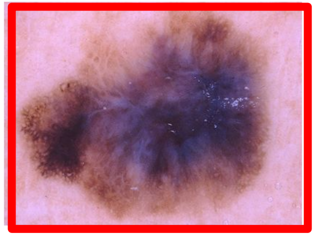
Nevus



Dermatofibroma



Melanoma

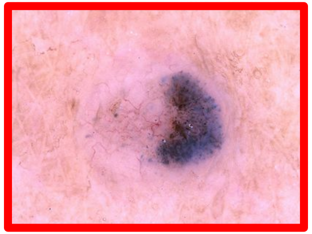


Vascular

Pigmented Bowen’s

Pigmented Benign Keratoses

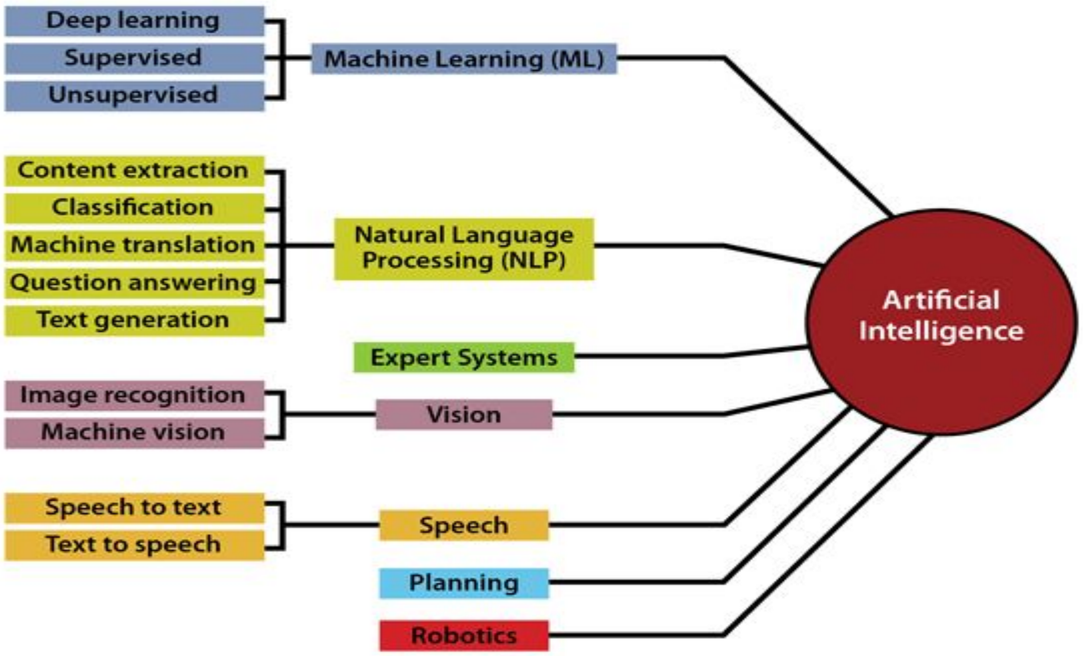
Basal Cell Carcinoma



Non-Cancerous

Cancerous

How can Artificial Intelligence (AI) help?



Artificial Intelligence (AI) technology goal is to mimic and learn the way humans see, read, listen, talk and interpret things

Machine Learning (ML) is a sub branch of AI that automatically learns from the past data without feeding logic.

Trained ML models can read skin lesion images and predict classification

Hypothesis

The purpose of this project is how AI/ML can be used to predict skin cancer by optimizing the parameters that improve the efficiency.

This project involved four different but related hypotheses to optimize efficiency of the machine learning model.

1) Deep learning ML methods like Convolutional Neural Networks (CNN) are more efficient than traditional ML methods.

2) ML models trained on higher image size perform better than small size.

3) ML models trained on less blurred images perform better than more blurred images

4) Higher the batch size, number of iterations increase the efficiency of the ML model

ML Model Performance Measurement:

1) Accuracy - Higher the accuracy better the performance

2) ROC AUC - Higher the AUC score better the performance

Data Collection & Procedure

1. Set up the environment: Install python packages (Python, Pandas, Tensorflow, sklearn, OpenCV)

2. Prepare the data: Download the ISIC skin cancer images (~10,000) and metadata with classification. Augment the data as needed. Split the data for training and testing the model

3. Identify performance measures: In this project two measures are considered (a) Efficiency of the ML model (b) Area Under the Curve (AUC) of Receiving Operating Characteristic (ROC) curve

4. Train & Test: Train the ML models with training dataset and measure accuracy & ROC with test dataset

- **Test Hypothesis#1:** Analyze which ML model performs better K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), Random Forest Classifier (RFC), Decision Tree Classifier (DTC) that will perform the detection of skin cancer.
- **Test Hypothesis#2:** Analyze five different skin lesion resolution sizes in pixels (25x25, 50x50, 75x75, 100x100, 125x125) and measure the ML model performance for each size to determine which resolution performs better.
- **Test Hypothesis#3:** Analyze five different blur sizes (5x5, 10x10, 15x15, 20x20, 25x25) and measure the accuracy and AUC to determine the blur impact on ML model performance.
- **Test Hypothesis#4:** Analyze model specific parameters like epochs(10,20,30, 40, 50) and batch size (10,20,30,40,50) at different intervals and measure performance.

4. Conclusion: Document observations and finalize conclusions.

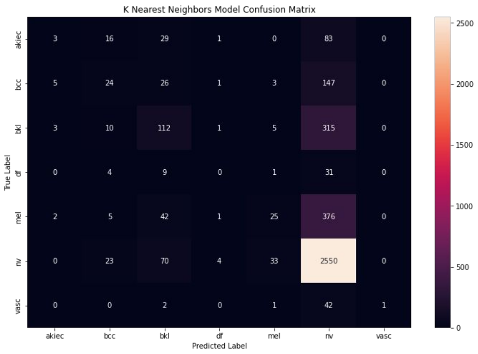
Hypothesis #1

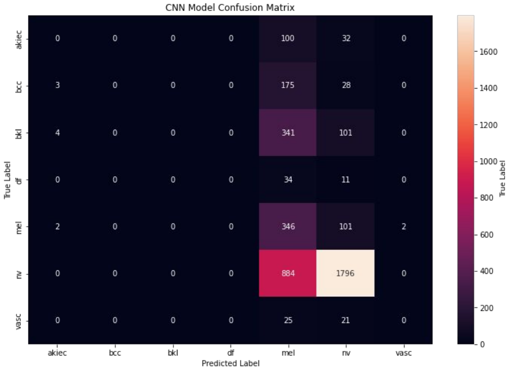
Deep learning ML methods like Convolutional Neural Networks (CNN) are more efficient than traditional ML methods

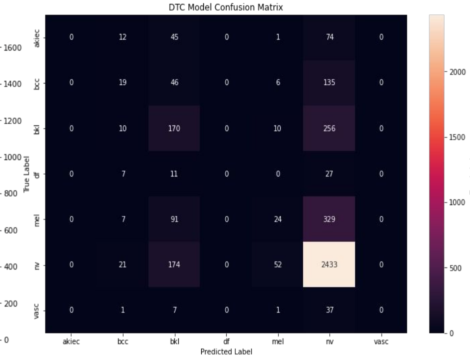
Use 4 ML (machine-learning) models: K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), Random Forest Classifier (RFC), Decision Tree Classifier (DTC) that will perform the detection of skin cancer. Measure Accuracy and the ROC AUC score.

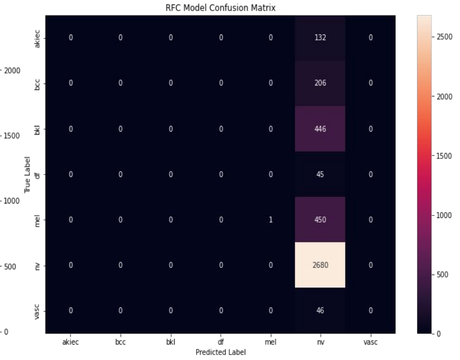
	Accuracy	ROC %
KNN	0.67599	0.62361
CNN	**0.68148**	0.55767
DTC	0.66475	0.64432
RFC	0.66925	0.64987

Confusion Matrix (CM) for KNN, CNN, DTC, RFC







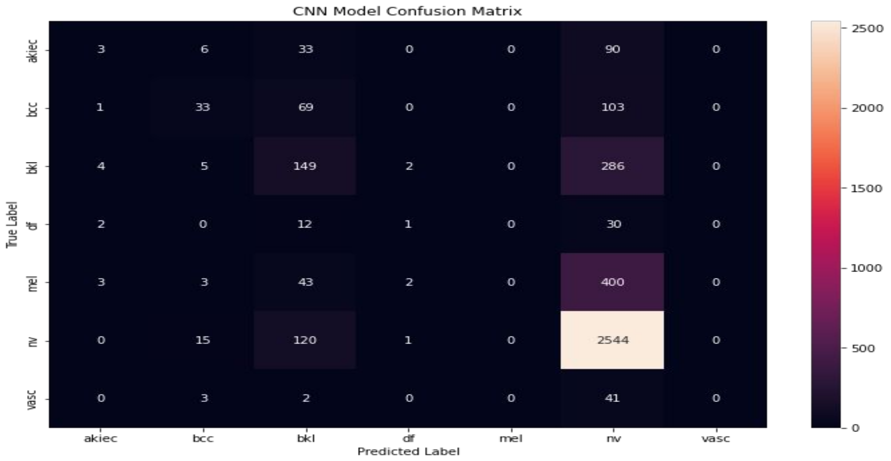


Hypothesis #2

ML models trained on higher image size perform better than small size

Using OpenCV (Open Computer Vision) software resize the images to five different resolution sizes in pixels (25x25, 50x50, 75x75, 100x100, 125x125) and measure the ML model performance for each size to determine which resolution performs better

Confusion Matrix of the best performance resolution (75x75)



ML Model Performance Data Collection

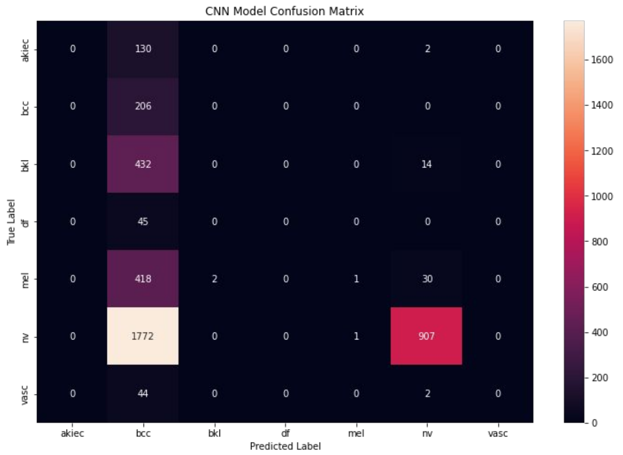
Resolution	Accuracy	ROC %
Image 100x75	0.6707	0.5065
Image 25X25	0.1707	0.6434
Image 50x50	0.6705	0.5268
Image 75x75	**0.6815**	0.6011
Image 100x100	0.0345	0.5021
Image 125x125	0.1510	0.5110

Hypothesis #3

ML models trained on less blurred images performed better than more blurred images

Using OpenCV (Open Computer Vision) software to blur the images to five different kernel sizes in pixels (5x5, 10x10, 15x15, 75x75, 25x25) kept under the best performing image resolution (75x75), measuring the ML model performance for each size to determine which blur performs better

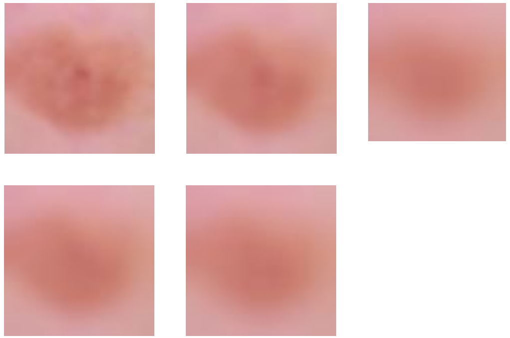
Confusion Matrix of 10x10 blur kernel



ML Model Performance Data

Blur Kernel Size	Accuracy	ROC %
Blur 5x5	0.2781	0.5493
Blur 10x10	**0.5634**	0.6296
Blur 15x15	0.2594	0.5397
Blur 20x20	0.1700	0.5235
Blur 25x25	0.4666	0.5944

Blurred skin lesion images



Hypothesis #4

Higher the batch size and number of iterations increases the efficiency of the ML model

Modified specific parameters like epochs (10,20,30,40,50) and batch size (10,20,30,40,50), which were tested at different intervals. Performance measures will be recorded.

ML Model Performance Data (Epochs)

Epochs	Accuracy	ROC %
epochs 10	0.6762	0.5487
epochs 20	0.6815	0.6011
epochs 30	0.6695	0.5033
epochs 40	0.3617	0.6198
epochs 50	0.4054	0.5935

ML Model Performance Data (Batch-Size)

Batch-Size	Accuracy	ROC %
batch 10	0.6727	0.5263
batch 20	0.6815	0.6011
batch 30	0.6707	0.5264
batch 40	0.4725	0.6406
batch 50	0.3802	0.6168

Results and Conclusion

- The CNN model performed better accuracy than the other 3 models

- The 75x75 was the image resolution (A lower resolution gave a better performance than the original resolution). This is an important observation because with less computing power the model can perform better.

- Blurring the images resulted in a significant negative shift in the accuracy. Further study can be conducted by augmenting the dataset with more blurred images to train the ML model.

- Additional parameters such as the epochs and batch-size unfortunately did not have any contribution to a positive shift in accuracy

Further Study

This research project can be further continued with

- Compare model performance with transfer learning models (AlexNet, ImageNet, GoogleNet)
- Build a mobile application for skin cancer lesion classification
- Train the ML model with more data augmentation
- Increase the ML model accuracy from ~70% to ~90%

Works Cited

Morris, K. L., PhD. (2018, August 1). Decision Tree Model vs Traditional Measures to Identify Patterns of Sun-Protective Behaviors and Sun. Kasey L. Morris. https://jamanetwork.com/journals/jamadermatology/fullarticle/2684788

Classification Models for Skin Tumor Detection Using Texture Analysis in Medical Images. (2020, June 19). MDPI. https://www.mdpi.com/2313-433X/6/6/51

An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. (2021, September 15). ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2666827021000177

What Is Deep Learning? | How It Works, Techniques & Applications. (1994). MATLAB & Simulink. https://www.mathworks.com/discovery/deep-learning.html

NCBI. (2021, December). Deep Learning-Based Transfer Learning for Classification of Skin Cancer. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8662405/

Singh, L. (2021, September 30). A Deep Learning-Based Transfer Learning Framework for the Early Detection and Classification of Dermoscopic Images of Melanoma. Biomedical and Pharmacology Journal. https://biomedpharmajournal.org/vol14no3/a-deep-learning-based-transfer-learning-framework-for-the-early-detection-and-classification-of-dermoscopic-images-of-melanoma/