

# SKIN CANCER CLASSIFICATION

---

MACHINE LEARNING WITH TENSORFLOW

AUTHORS: ANTHONY AKO SOWAH, STEPHAN PRIES,  
NATNAEL GUESH HAGOS & HAMMAM RAFAT MOHAMED HRAESHA



# AGENDA

- **Introduction**
- **Literature Review**
- **Dataset Characteristics**
- **Baseline Model**
- **Model Definition and Evaluation**
- **Results**
- **Challenges and Errors**
- **Discussion, Conclusion & Future Work**
- **Q & A**

# INTRODUCTION

---

- Skin cancer is one of the most common and dangerous types of cancer.
- This project uses deep learning to automatically classify skin lesion images into different cancer types using the **HAM10000** dataset.
- We apply a **MobileNetV2**, **EfficientNetBx** models to help improve early detection and support medical diagnosis.



## LITERATURE REVIEW

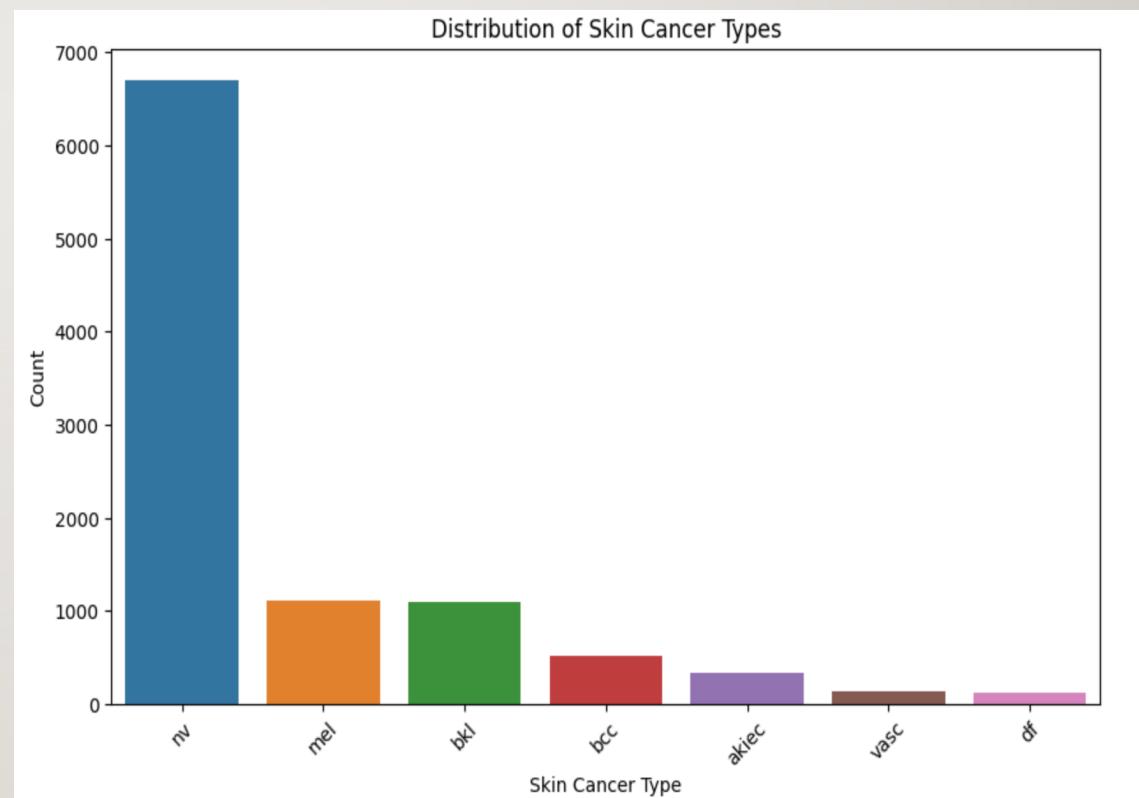
---

- Studies show that deep learning models like MobileNetV2 and EfficientNetBx, especially with transfer learning, perform well on HAM10000.
- Adding patient info (age, gender, lesion site) improves accuracy in tough cases.
- Ensemble methods and attention tools (like Grad-CAM) boost both performance and interpretability.
- Overall, combining image data with clinical context delivers the best results.



# DATASET CHARACTERISTICS

Benign Classes	Malignant / Precancerous Classes
Melanocytic nevi (nv)	Melanoma (mel)
Benign keratosis-like lesions (bkl)	Basal cell carcinoma (bcc)
Dermatofibroma (df)	Actinic keratoses and intraepithelial carcinoma (akiec)
Vascular lesions (vasc)	

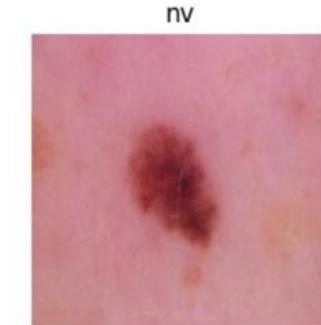
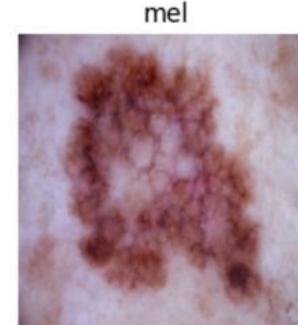


# DATASET CHARACTERISTICS

```
print(df.head())
```

	lesion_id	image_id	dx	dx_type	age	sex	localization
0	HAM_0000118	ISIC_0027419	bkl	histo	80.0	male	scalp
1	HAM_0000118	ISIC_0025030	bkl	histo	80.0	male	scalp
2	HAM_0002730	ISIC_0026769	bkl	histo	80.0	male	scalp
3	HAM_0002730	ISIC_0025661	bkl	histo	80.0	male	scalp
4	HAM_0001466	ISIC_0031633	bkl	histo	75.0	male	ear

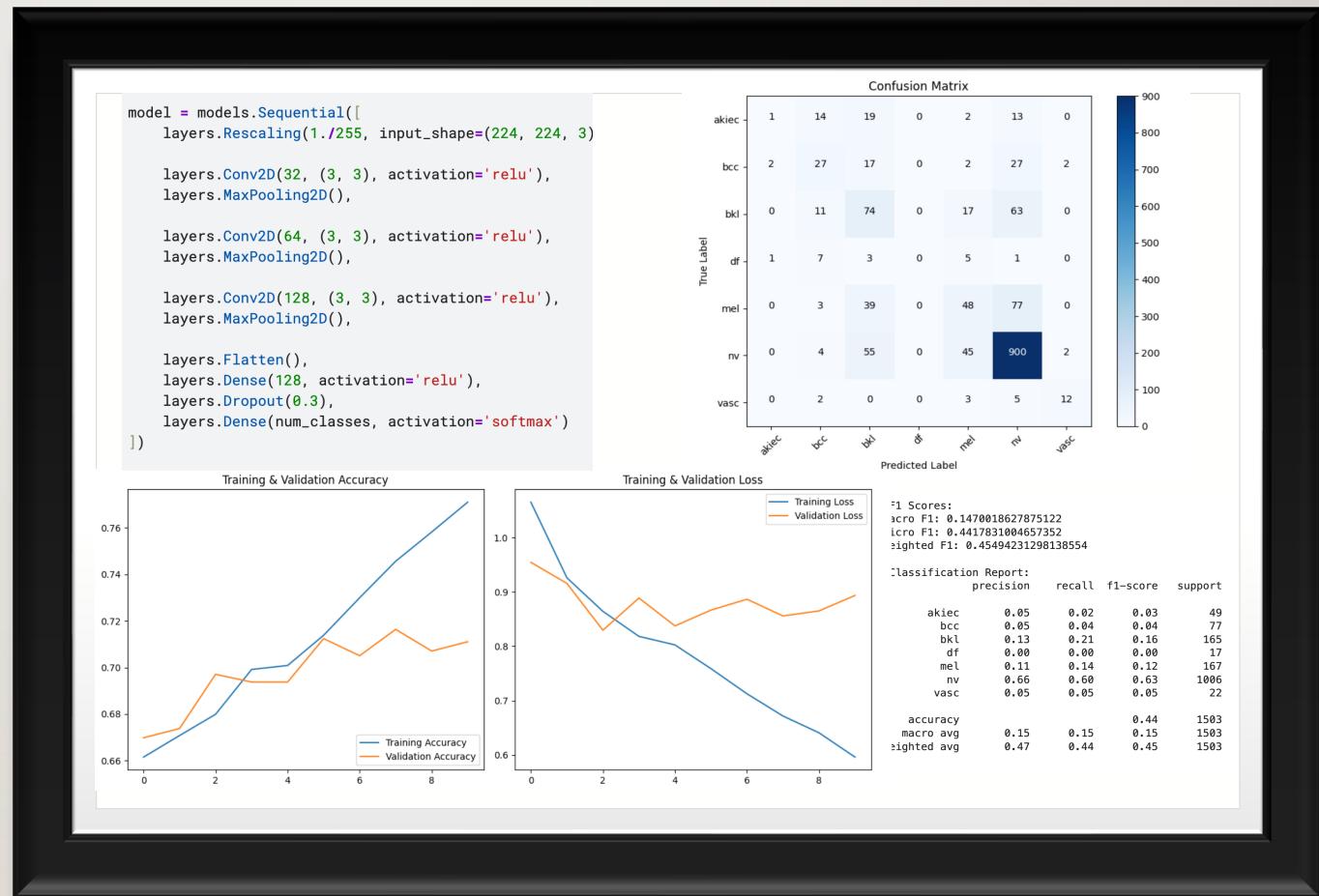
Classes: ['akiec', 'bcc', 'bkl', 'df', 'mel', 'nv', 'vasc']



IMG\_SIZE 224x224

# BASELINE MODEL

- test accuracy: 0.7066
- poor f1-score per class
- nv - relatively high accuracy
- nv true positives at 900
- nv misclassifications: 55 as bkl, 45 as mel, etc.
- mel heavily confused with 77 nv
- mel true positive 48



# MODEL DEFINITION AND EVALAUATION

---

## Transfer learning

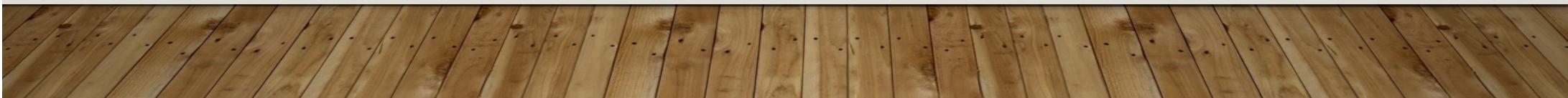
- MobileNetV2
- EfficientNetBx
- Ensemble  
(MobileNetV2/EfficientNetB1)

## Training steps

- Data augmentation
- Unfreezing of layers
- Focal-Loss with class weight
- DropOut
- Learning rate variation

## Accuracy Benchmark:

73 – 79 % Kaggle Challenges with specific dataset  
88 – 92 % Literature Optimum for Ensemble Models with other datasets



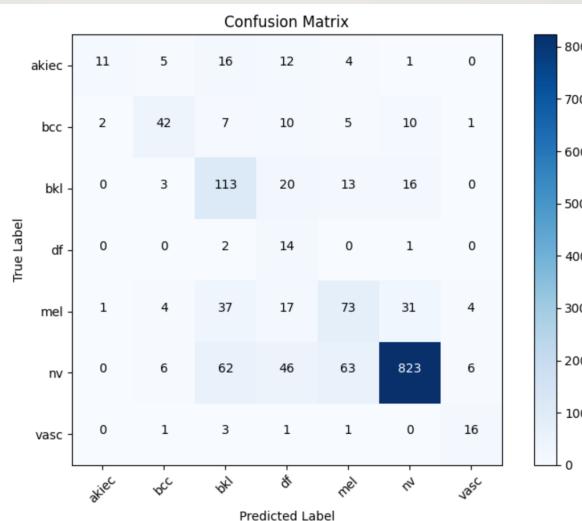
# TRANSFER LEARNING MOBILENETV2

Epochs	60
dynamic LR	1.e-4
reduced gamma	1.5

F1 Scores:  
Macro F1: 0.5275774677009114  
Micro F1: 0.7265469061876247  
Weighted F1: 0.7489959594416075

Classification Report:

	precision	recall	f1-score	support
akiec	0.79	0.22	0.35	49
bcc	0.69	0.55	0.61	77
bkl	0.47	0.68	0.56	165
df	0.12	0.82	0.20	17
mel	0.46	0.44	0.45	167
nv	0.93	0.82	0.87	1006
vasc	0.59	0.73	0.65	22
accuracy			0.73	1503
macro avg	0.58	0.61	0.53	1503
weighted avg	0.80	0.73	0.75	1503



Unfreeze deeper layers

Reduce LR

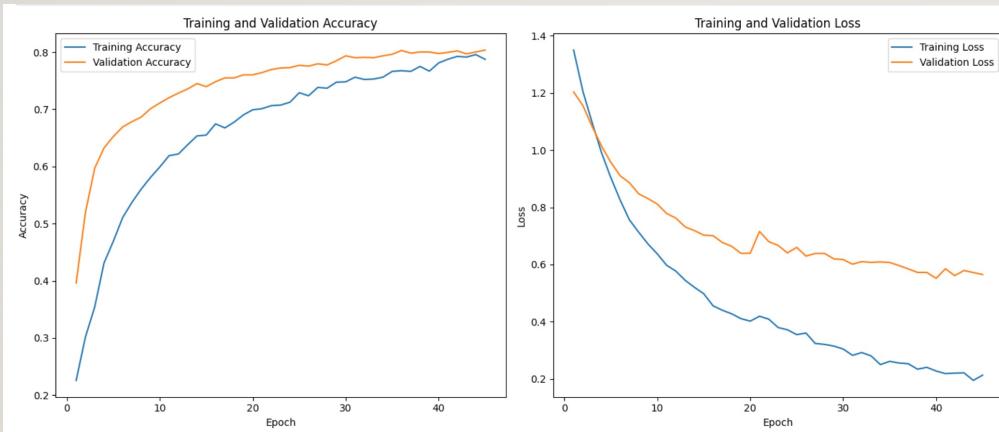
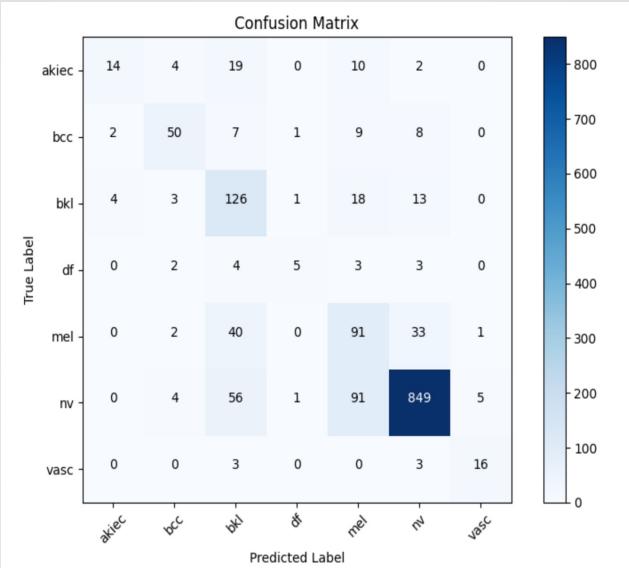
Early Stopping

Test Accuracy:7265

Macro F1: 0.5987029290253126  
Micro F1: 0.7658017298735862  
Weighted F1: 0.7751940376592819

Classification Report:

	precision	recall	f1-score	support
akiec	0.70	0.29	0.41	49
bcc	0.77	0.65	0.70	77
bkl	0.49	0.76	0.60	165
df	0.62	0.29	0.40	17
mel	0.41	0.54	0.47	167
nv	0.93	0.84	0.89	1006
vasc	0.73	0.73	0.73	22
accuracy			0.77	1503
macro avg	0.67	0.59	0.60	1503
weighted avg	0.80	0.77	0.78	1503



# TRANSFER LEARNING EFFICIENTNETB1

Epochs	70
LR	0.0001
Unfreeze	100
Gamma	1.5
Callback	55

Continue Training

Adjusting Parameters

Test Accuracy:  
0.7658

# TRANSFER LEARNING ENSEMBLE MOBILENETV2 AND EFFICIENTNETB1

---

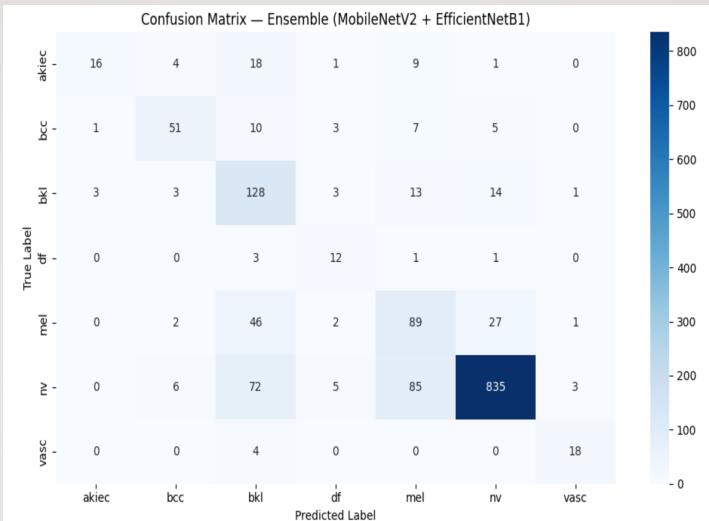
Ensemble Classification Report:

	precision	recall	f1-score	support
akiec	0.80	0.33	0.46	49
bcc	0.77	0.66	0.71	77
bkl	0.46	0.78	0.57	165
df	0.46	0.71	0.56	17
mel	0.44	0.53	0.48	167
nv	0.95	0.83	0.88	1006
vasc	0.78	0.82	0.80	22
accuracy			0.76	1503
macro avg	0.66	0.66	0.64	1503
weighted avg	0.81	0.76	0.78	1503

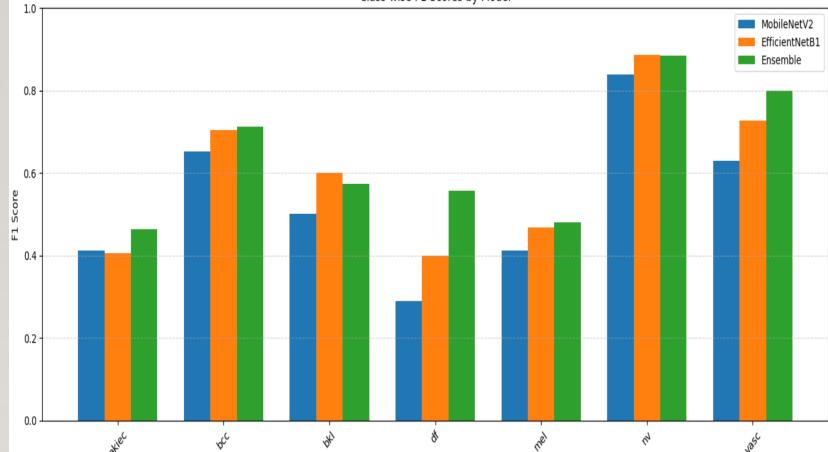
Macro F1: 0.6390050578966966

Micro F1: 0.7644710578842314

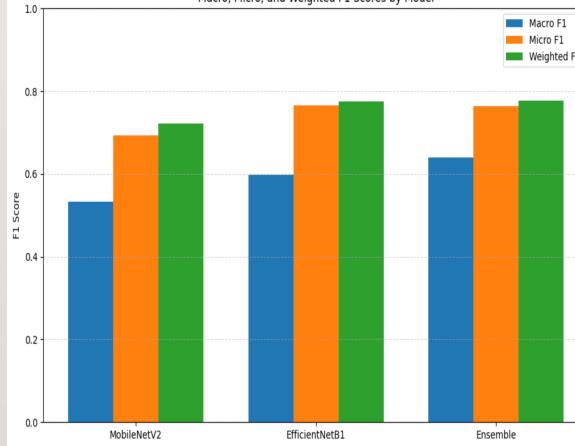
Weighted F1: 0.7777369482046602

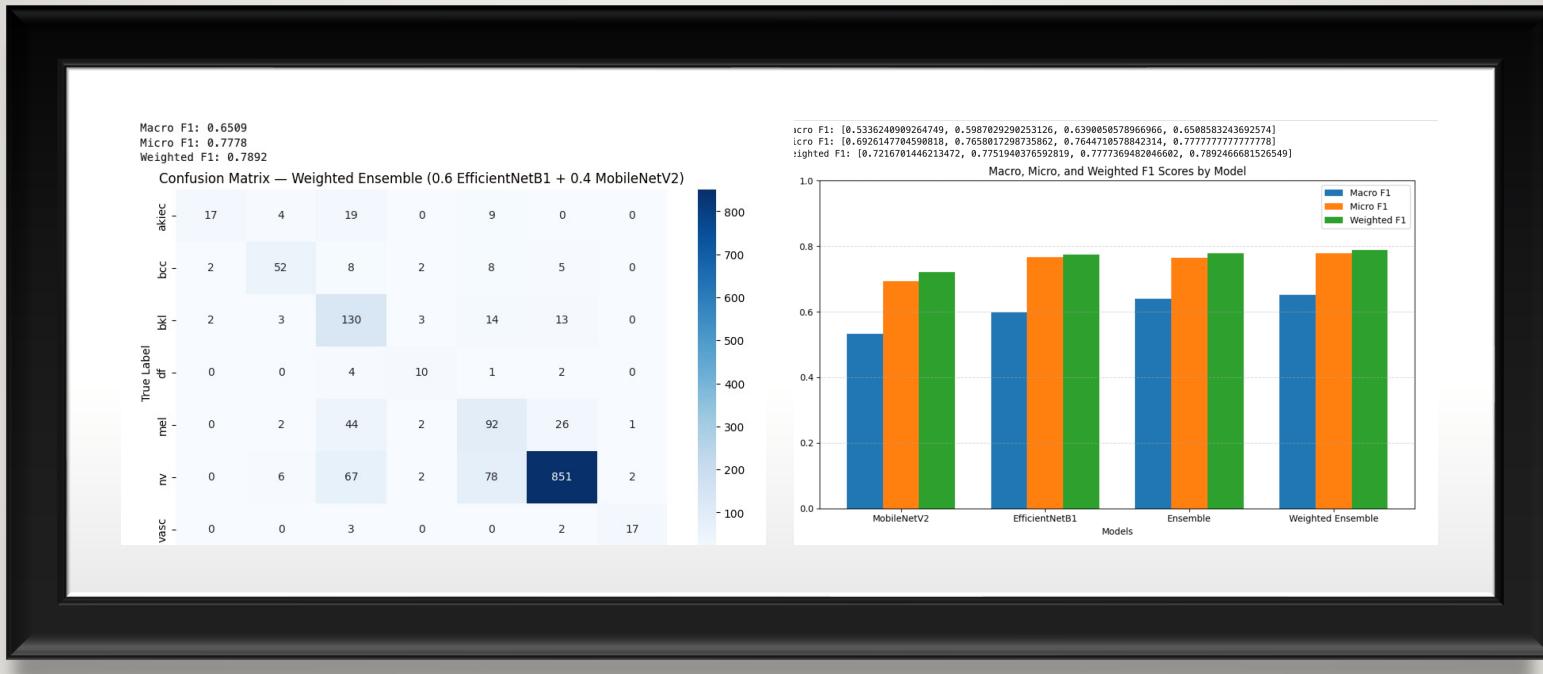


Class-wise F1 Scores by Model



Macro, Micro, and Weighted F1 Scores by Model





## ENSEMBLE WITH WEIGHT (0.6 AND 0.4)

# CHALLENGES

---



Running out of RAM



Dataset imbalance



Image resolution



Short time frame

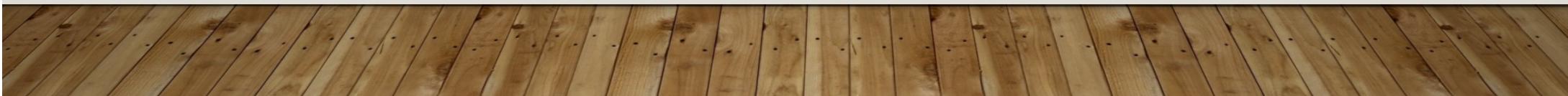
## **DISCUSSION, CONCLUSION AND FUTURE WORKS**

---

- our approach during training was not to misclassify malignant lesions but to optimize f1 and precision
- accuracy from base model of 0.71 increased up to 0.79 which comes close to Kaggle Challenges but is far away from literature

### **Future work**

- Optimize f1 and precision etc.
- Add metadata for training
- Run model with different datasets (probably with higher resolution)
- Employ Grid search to ideentify best optimized values for the two ensemble models





# **BACKUP**

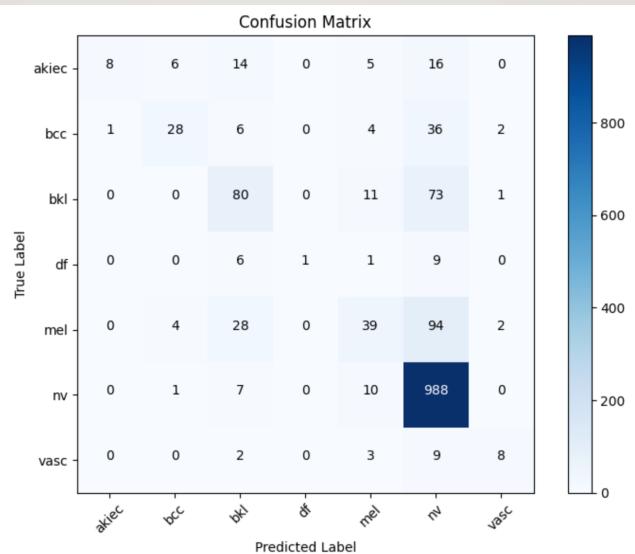
---



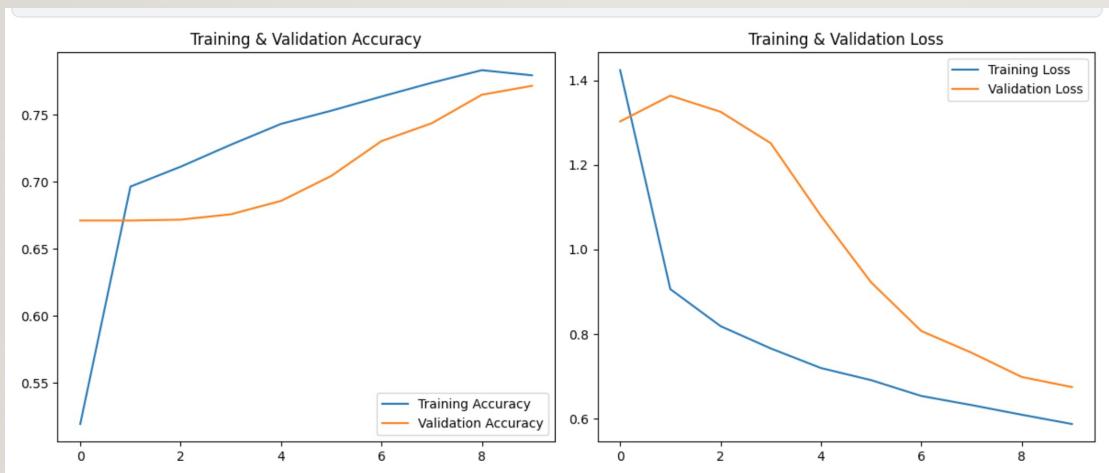
F1 Scores:  
Macro F1: 0.43672237950671267  
Micro F1: 0.7664670658682635  
Weighted F1: 0.7276384967831031

Classification Report:

	precision	recall	f1-score	support
akiec	0.89	0.16	0.28	49
bcc	0.72	0.36	0.48	77
bkl	0.56	0.48	0.52	165
df	1.00	0.06	0.11	17
mel	0.53	0.23	0.33	167
nv	0.81	0.98	0.89	1006
vasc	0.62	0.36	0.46	22
accuracy			0.77	1503
macro avg	0.73	0.38	0.44	1503
weighted avg	0.75	0.77	0.73	1503



# TRANSFER LEARNING MOBILENETV2



Epochs	10
LR	0.0001

1<sup>st</sup> 10 epochs terrible

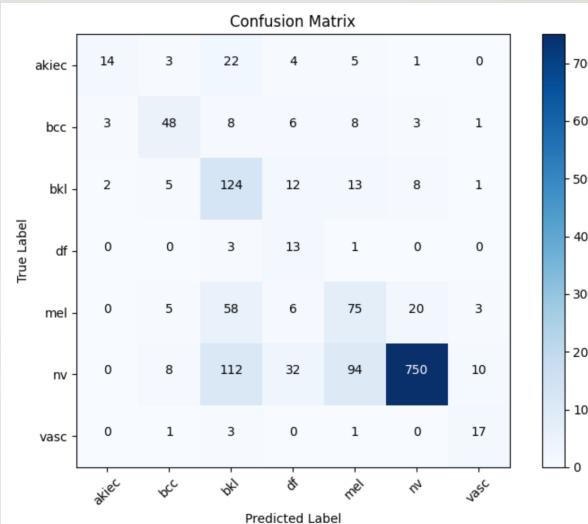
Unfreez layers

Test Accuracy:  
0.7665

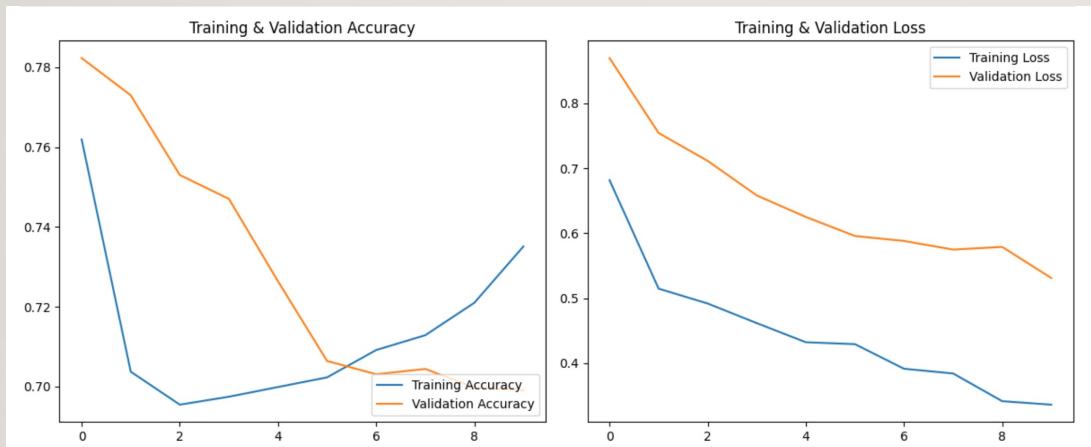
F1 Scores:  
Macro F1: 0.5336240909264749  
Micro F1: 0.6926147704590818  
Weighted F1: 0.7216701446213472

Classification Report:

	precision	recall	f1-score	support
akiec	0.74	0.29	0.41	49
bcc	0.69	0.62	0.65	77
bkl	0.38	0.75	0.50	165
df	0.18	0.76	0.29	17
mel	0.38	0.45	0.41	167
nv	0.96	0.75	0.84	1006
vasc	0.53	0.77	0.63	22
accuracy			0.69	1503
macro avg	0.55	0.63	0.53	1503
weighted avg	0.79	0.69	0.72	1503



# TRANSFER LEARNING MOBILENETV2



Epochs	10
LR	0.0001

Compute Class weight

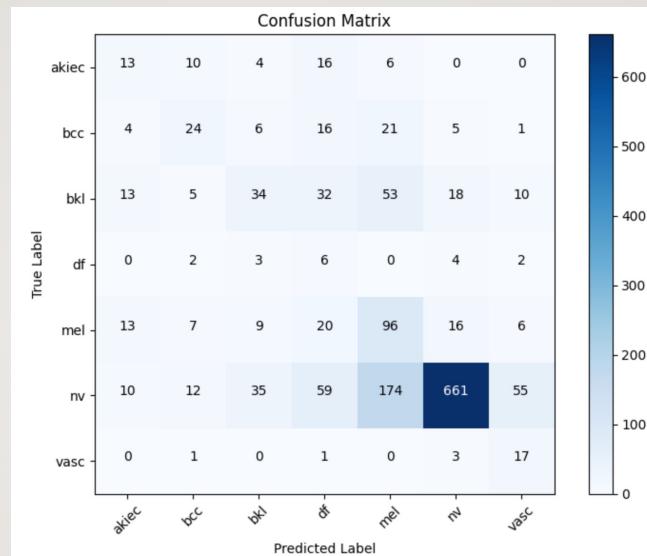
Focal Loss

Test Accuracy:  
0.6926

F1 Scores:  
Macro F1: 0.34060186774788775  
Micro F1: 0.5662009314703925  
Weighted F1: 0.6183215214070296

Classification Report:

	precision	recall	f1-score	support
akiec	0.25	0.27	0.25	49
bcc	0.39	0.31	0.35	77
bkl	0.37	0.21	0.27	165
df	0.04	0.35	0.07	17
mel	0.27	0.57	0.37	167
nv	0.93	0.66	0.77	1006
vasc	0.19	0.77	0.30	22
accuracy			0.57	1503
macro avg	0.35	0.45	0.34	1503
weighted avg	0.73	0.57	0.62	1503



Tony:  
Check  
values and  
provide  
charts

# TRANSFER LEARNING EFFICIENTNETB0

Epochs 10  
LR 0.0001  
What else?

Frozen Layers

Test Accuracy:0.5662

F1 Scores:  
Macro F1: 0.5103708655992099  
Micro F1: 0.6560212907518297  
Weighted F1: 0.6887008713170641

Classification Report:  

	precision	recall	f1-score	support
akiec	0.46	0.27	0.34	49
bcc	0.56	0.58	0.57	77
bkl	0.37	0.64	0.47	165
df	0.16	0.47	0.24	17
mel	0.28	0.45	0.35	167
nv	0.94	0.72	0.81	1006
vasc	0.78	0.82	0.80	22
accuracy			0.66	1503
macro avg	0.51	0.56	0.51	1503
weighted avg	0.76	0.66	0.69	1503

Tony:  
Check  
values,  
provide  
charts and  
matrix

# TRANSFER LEARNING EFFICIENTNETB0

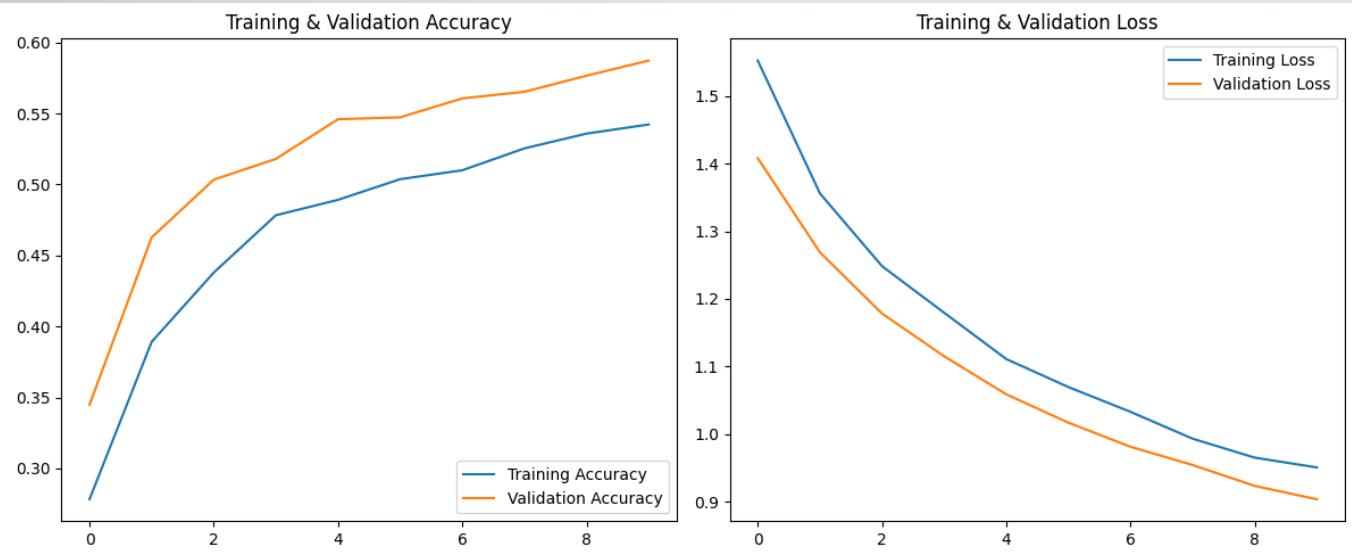
Epochs	30
LR	0.0001
Unfreeze	100
Gamma	1.5

Unfreeze  
layers

Weighted  
Focal Loss

Reduce  
gamma

Test  
Accuracy:  
0.6560



Same as  
previous  
slide

# TRANSFER LEARNING EFFICIENTNETB1

---

Epochs 10  
LR 0.00001  
early\_stopping patience=3  
results not  
focal loss + Class Weight

