Deep Learning for Skin Cancer Classification

Dermatoscopic Image using Vision Transformer

November 2025

By: Binh Diep,
Rashmi Thimmaraju
Kartavya Mandora
Kirtan Patel
Date: 11/07/2025
AI700-001_BK_FALL2025

LONG ISLAND UNIVERSITY

Outlines

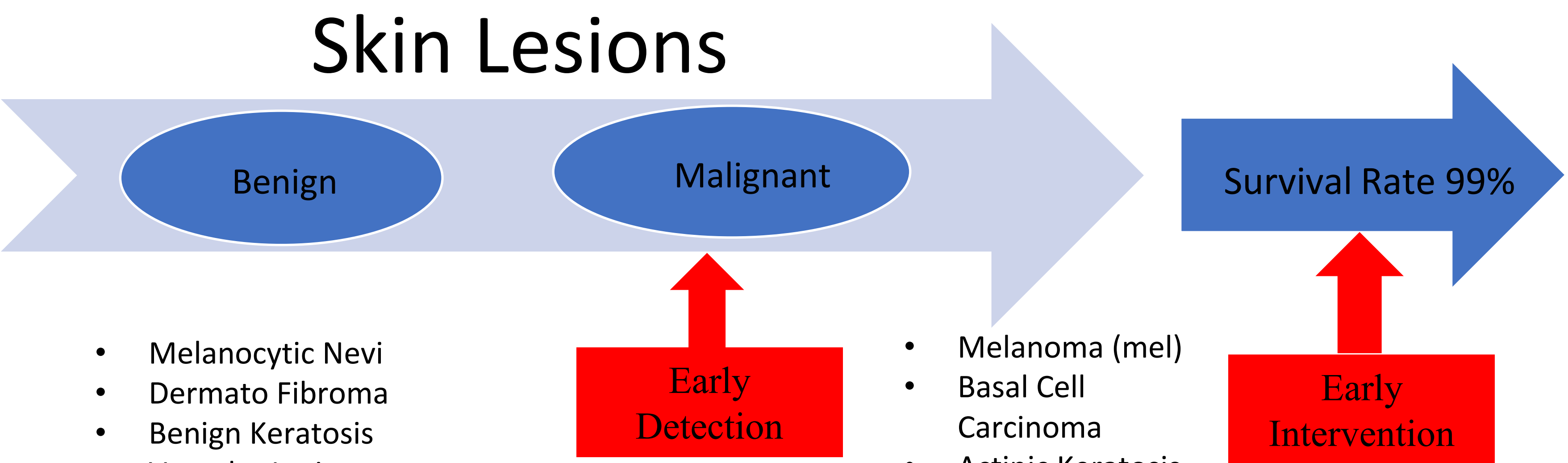
LIU

TABLE OF CONTENTS

LIU

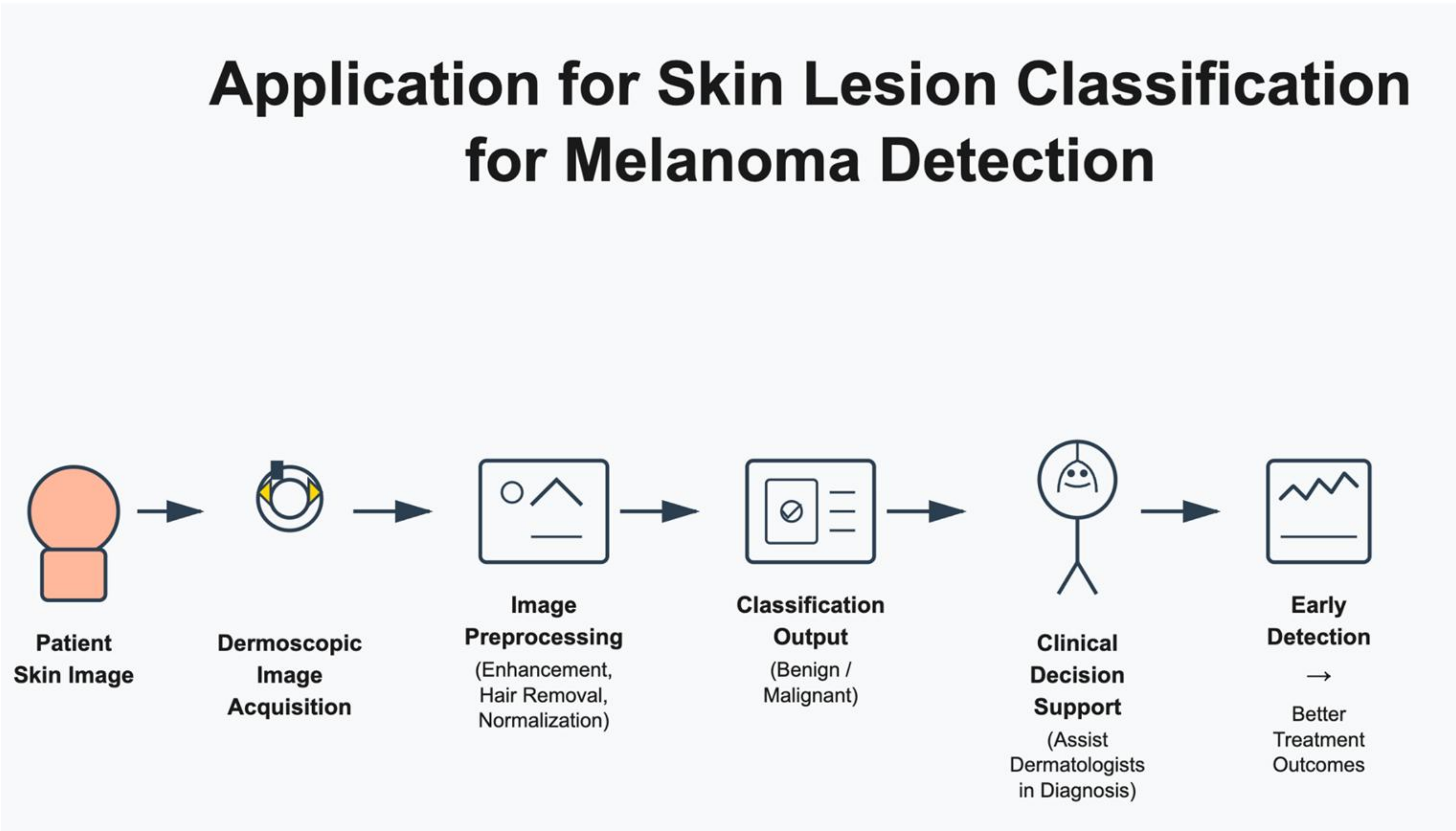
Prevalence

1. The most common cancer in the United States
2. 1 in 5 American will develop skin cancer by the age of 70
3. Greater than 200,000 people will be diagnosed with melanoma in 2025



- Melanocytic Nevi
- Dermato Fibroma
- Benign Keratosis
- Vascular Lesions

- Melanoma (mel)
- Basal Cell Carcinoma
- Actinic Keratosis


Application for Skin Lesion Classification for Melanoma Detection
Patient Skin Image
Dermoscopic Image Acquisition
Image Preprocessing
(Enhancement, Hair Removal, Normalization)
Classification Output
(Benign / Malignant)
Clinical Decision Support
(Assist Dermatologists in Diagnosis)
Early Detection
→
Better Treatment Outcomes

LITERATURE REVIEW

LIU

- *Transformers in Skin Lesion Classification a Diagnosis (Aldealnabi et al, 2024)*
  - Traditional CNN’s capture local features
  - Vision Transformers (ViTs) add global context and attention
  - Achieved 96% accuracy on HAM10000 dataset
- *Skin Lesion Analysis for Melanoma Detection Using Fuzzy GC-SCNN (Bhimavarapu et al, 2022)*
  - Combined fuzzy logic, GrabCut segmentation, and stacked CNNs
  - Achieved 99.75% accuracy and 100% sensitivity
  - Handles uncertain lesion boundaries
- *A Deep Learning Framework for Automated early Diagnosis and Classification of Skin Cancer Lesions (Al-Waisy et al 2025)*
  - Hybrid pipeline with preprocessing, segmentation, and ensemble classification
  - Used AGCWD, Mask R-CNN + GrabCut, HRNet+ attention
  - Achieved 100% accuracy and 99.9% AUC

HAM 10000 DATASET

- Human against machine
- 10,015 images
- 7 categories

Lesion_id	Image_id	dx	dx_type	age	sex	localization
HAM_0000118	ISIC_0027419	bkl	histo	80	Male	scalp
HAM_0000118	ISIC_0025030	Bkl	histo	80	Male	scalp
HAM_0002730	ISIC_0026769	Bkl	histo	80	Male	Scalp
HAM_0002730	ISIC_0125661	Bkl	histo	80	Male	Scalp
HAM_0001466	ISIC_0031633	bkl	histo	75	Male	Ear

ALGORITHM/APPROACH/METHODOLOGY

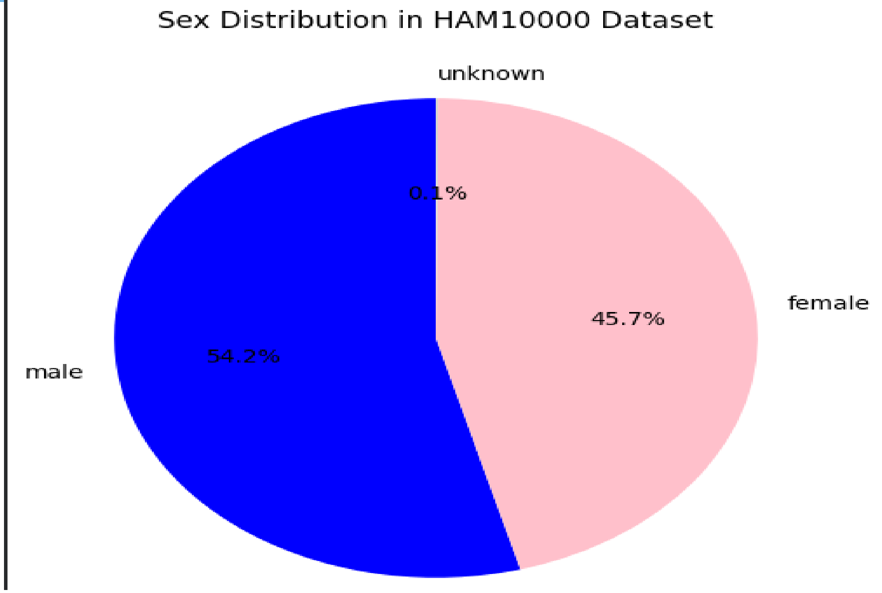
LIU

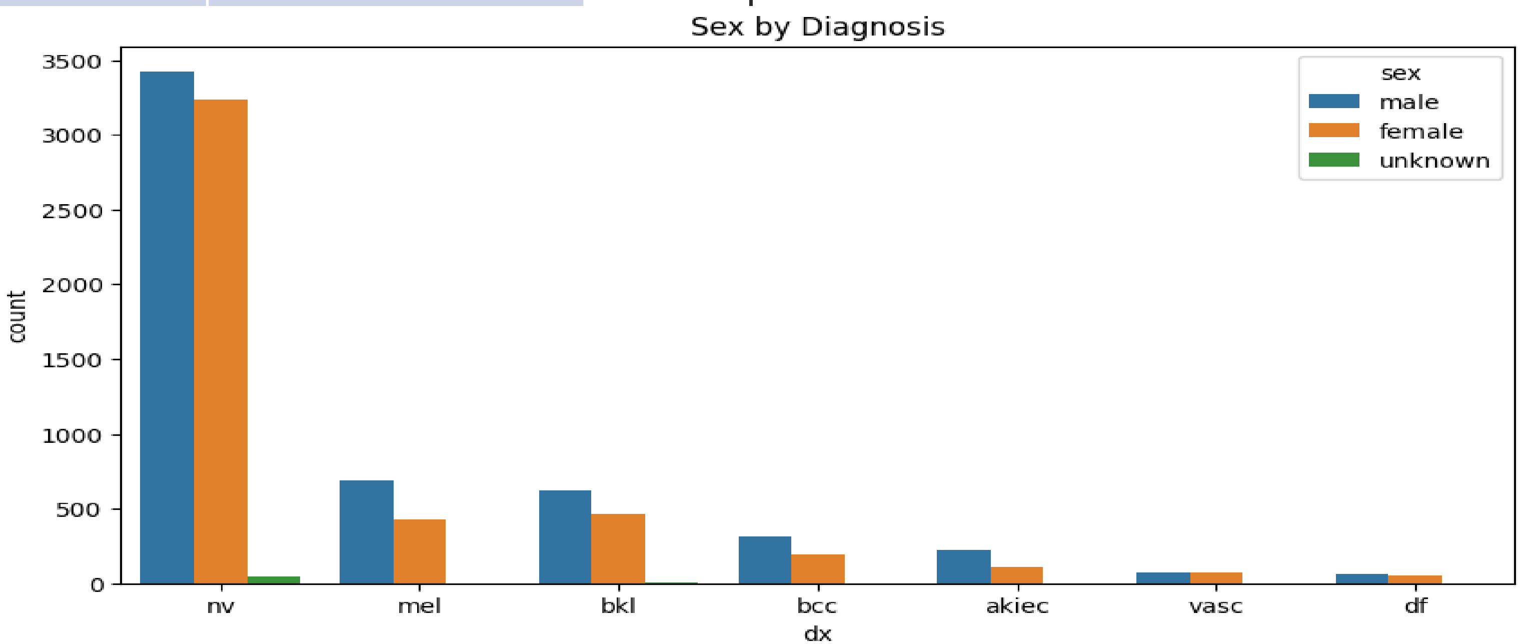
Step 1	**Dataset Collection:** HAM10000 (includes benign & malignant lesions).
Step 2	**Preprocessing:** Image resizing, normalization, data balancing, augmentation.
Step 3	**Feature Extraction:** Convolutional Neural Network for feature extraction and classification.
Step 4	**Training:** Backpropagation to minimize loss.
Step 5	**Model Evaluation:** Performance is assessed on unseen validation data using key metrics .
Step 6	**Result Analysis** –The final results are compared with **existing literature** .

Data Exploration

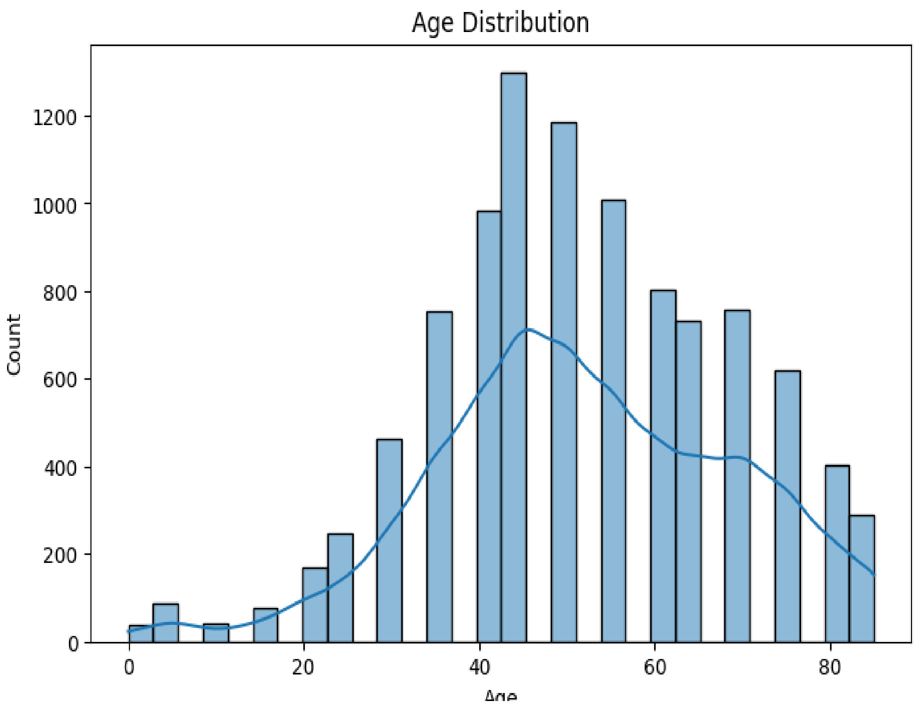
LIU

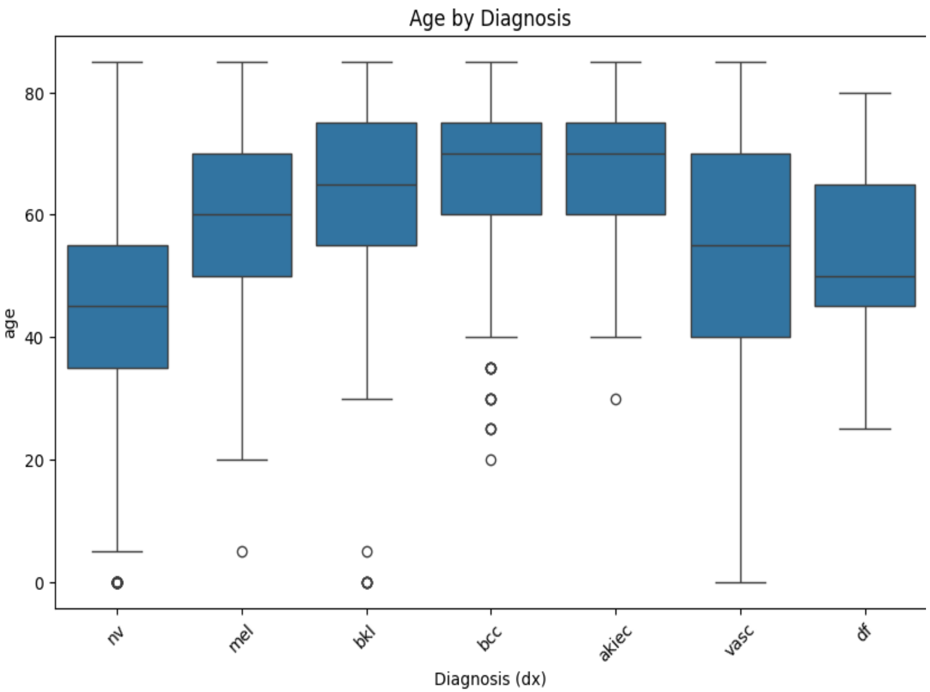
Skin Lesions	Abbreviation
Melanocytic nevi	Nv
Melanoma	Mel
Benign keratosis	Bkl
Basal Cell carcinoma	Bcc
Actinic Keratoses	Akiec
Vascular lesions	vasc
dermatofibroma	df





Data Exploration





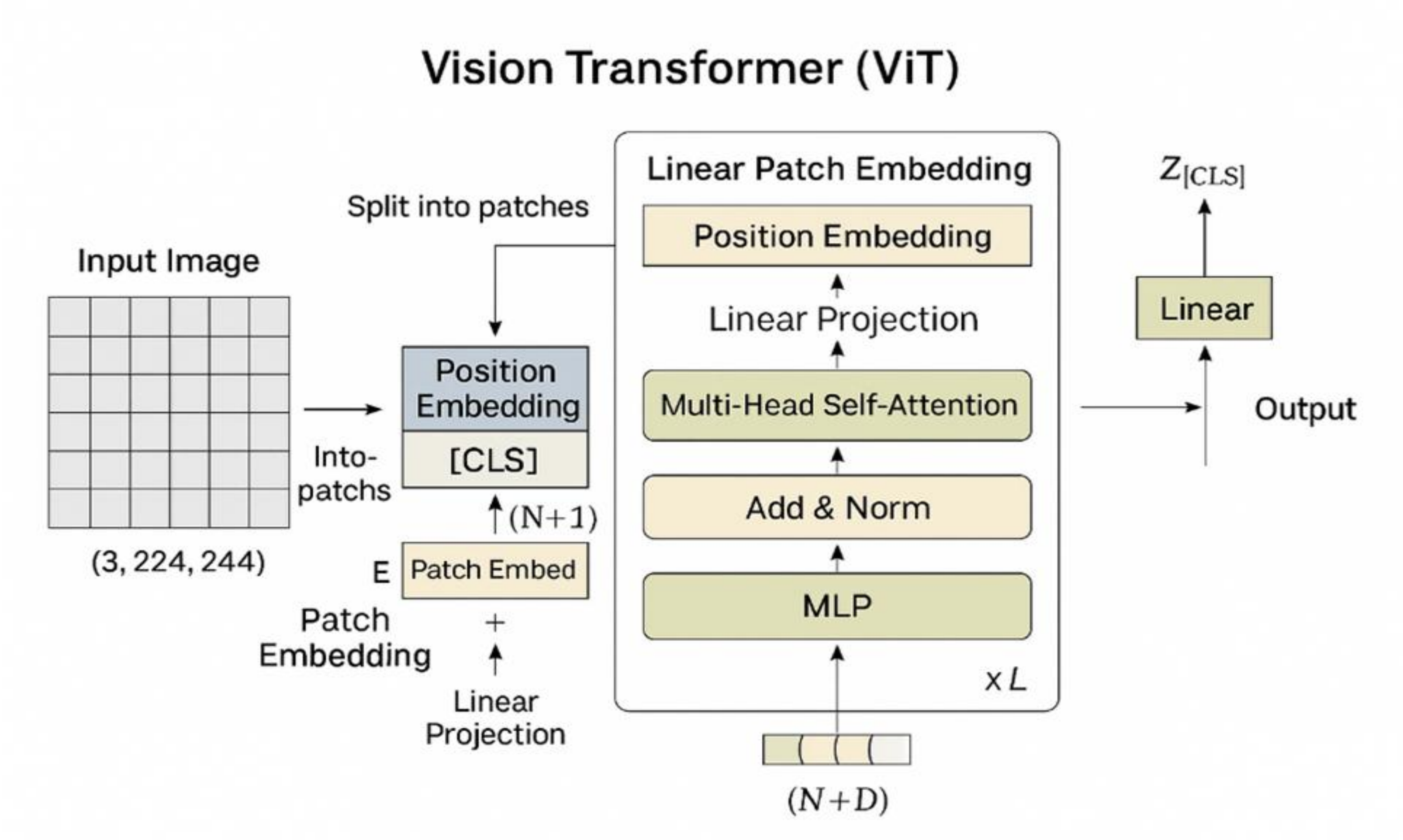
Data Pre-processing

Image Standardization
Resized and Normalized

Augmentation Techniques
Rotation, Flipping, Contrast adjustment

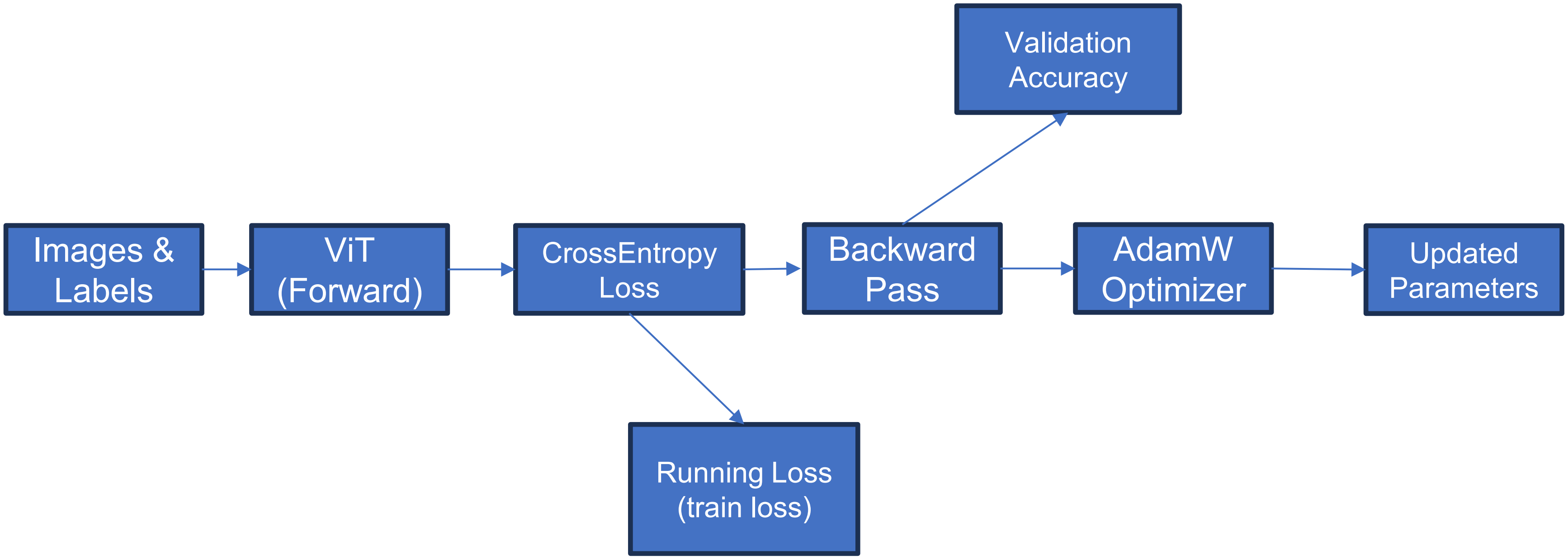
Image Enhancement
Hair removal

Transformer Model



Adebiyi, A. et. Al.(2024) *Heath Informatics*

Training Loop for ViT-Based Patch15-224



Training Components

Codes	Description
Model=timm.create_model (vit-base_patch16-224)	Loads a pre-trained Vision Transformer model with a classification head
Criterion = nn.CrossEntropyLoss	Defines the loss function used to measure how far predictions are from true labels
Optimizer = optim.AdamW (model.parameter())	Uses AdamW optimizer to update model weights efficiently while preventing overfitting

LIU

Training Breakdown

Codes	Description
Model.train()	Sets the model to training mode
For images, labels in train_loader	Iterates through mini-batches of data
Optimizer.zero_grad()	Clears old gradients before the next update
Outputs = model(images)	Forward pass through the ViT model
Loss = criterion(outputs, labels)	Computes classification loss
Loss.backward()	Performs backpropagation, computing gradients for all model parameters
Optimizer.step()	Updates model weights based on gradients

Tracking Performance

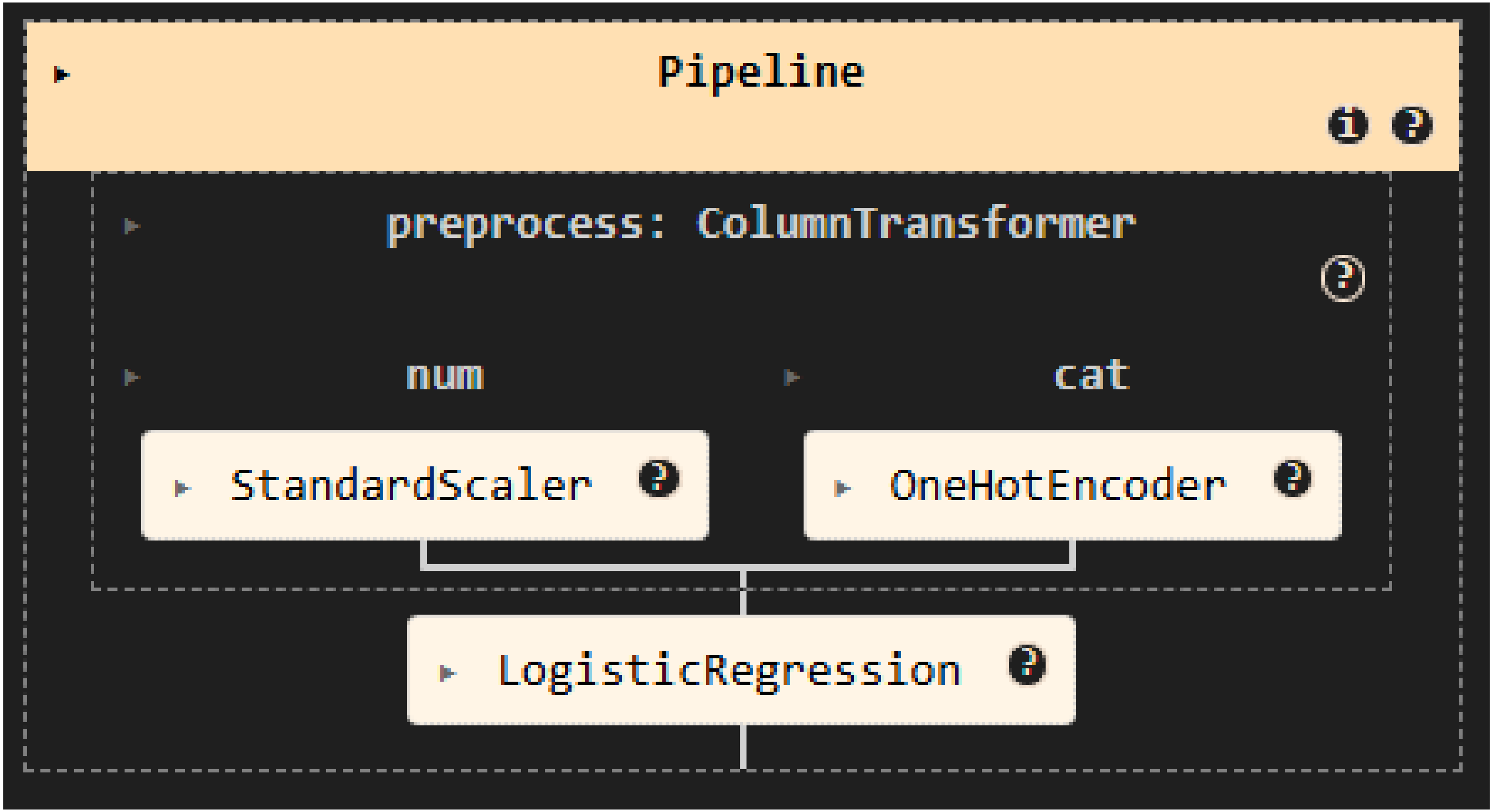
Codes	Description
Running_loss	Tracks average loss across batches for each epoch
Val_accuracies	Records accuracy on validation data to monitor progress
Best_acc	Keeps the highest validation accuracy for model checkpoints

MODEL EVALUATION- CNN MODEL

LIU

```
Columns: ['lesion_id', 'image_id', 'dx', 'dx_type', 'age', 'sex', 'localization']

Target distribution (dx):
dx
nv       6705
mel      1113
bkl      1099
bcc       514
akiec     327
vasc      142
df        115
```

MODEL EVALUATION-CNN MODEL



MODEL EVALUATION-CNN MODEL

LIU

Pipeline

```
Pipeline(steps=[('preprocess',
                 ColumnTransformer(transformers=[('num', StandardScaler(),
                                                  ['age']),
                                                 ('cat',
                                                  OneHotEncoder(handle_unknown='ignore'),
                                                  ['sex', 'localization'])])),
                ('clf',
                 LogisticRegression(class_weight={'akiec': np.float64(4.37305053025577),
                                                  'bcc': np.float64(2.7817460317460316),
                                                  'bkl': np.float64(1.3022478172023035),
                                                  'df': np.float64(12.36331569664903),
```

preprocess: ColumnTransformer

```
ColumnTransformer(transformers=[('num', StandardScaler(), ['age']),
                                ('cat', OneHotEncoder(handle_unknown='ignore'),
                                 ['sex', 'localization'])])
```

num

```
['age']
```

StandardScaler

```
StandardScaler()
```

cat

```
['sex', 'localization']
```

OneHotEncoder

```
OneHotEncoder(handle_unknown='ignore')
```

LogisticRegression

```
LogisticRegression(class_weight={'akiec': np.float64(4.37305053025577),
                                 'bcc': np.float64(2.7817460317460316),
                                 'bkl': np.float64(1.3022478172023035),
                                 'df': np.float64(12.36331569664903),
                                 'mel': np.float64(1.285530900421786),
                                 'nv': np.float64(0.21338772031292808),
                                 'vasc': np.float64(10.115440115440116)},
                   max_iter=2000, multi_class='multinomial')
```

MODEL EVALUATION-CNN MODEL

LIU

Validation report:

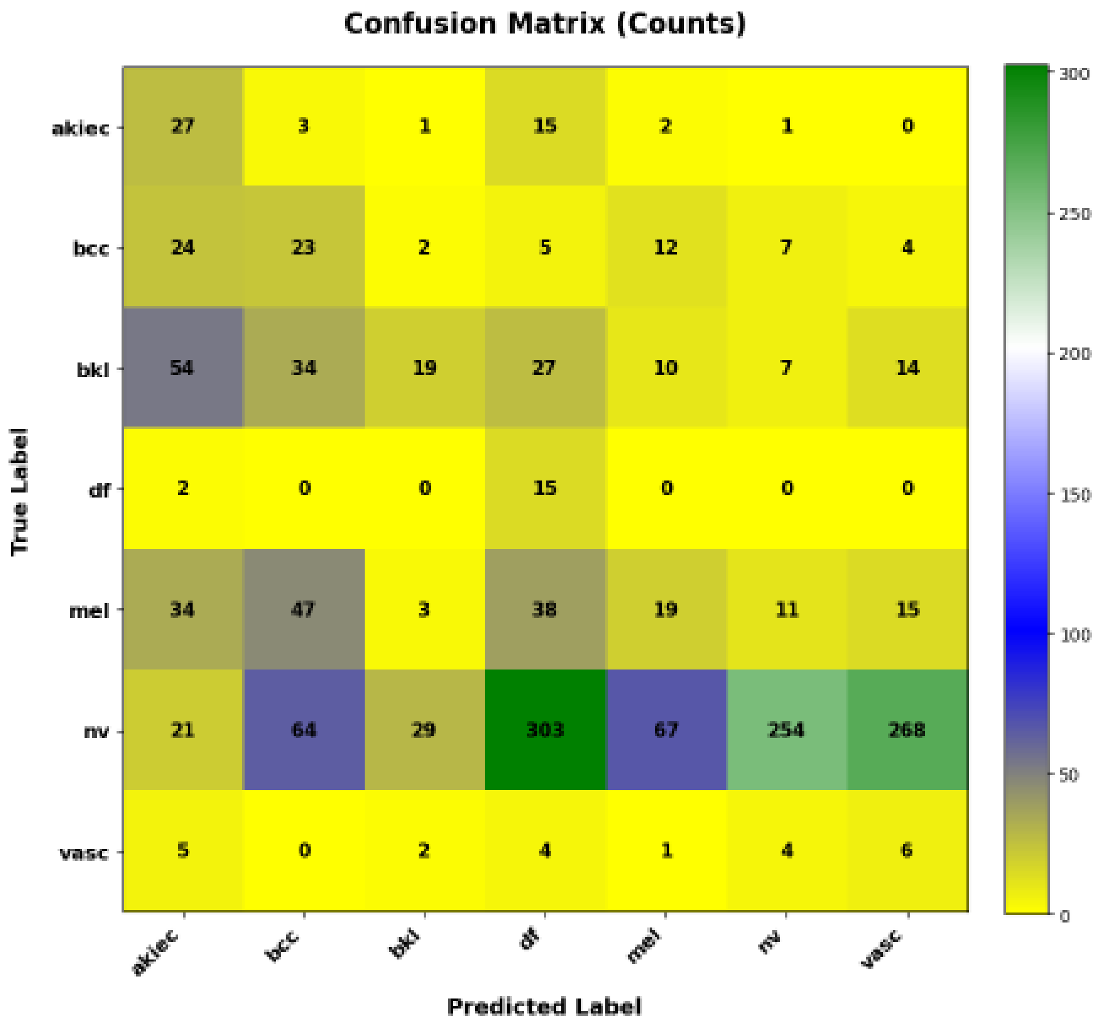
	precision	recall	f1-score	support
akiec	0.1377	0.4694	0.2130	49
bcc	0.1875	0.3896	0.2532	77
bkl	0.3091	0.1030	0.1545	165
df	0.0323	0.8235	0.0621	17
mel	0.1682	0.1078	0.1314	167
nv	0.8687	0.2237	0.3557	1006
vasc	0.0281	0.4286	0.0528	21
accuracy			0.2237	1502
macro avg	0.2474	0.3637	0.1747	1502
weighted avg	0.6494	0.2237	0.2912	1502

MODEL EVALUATION-CNN MODEL

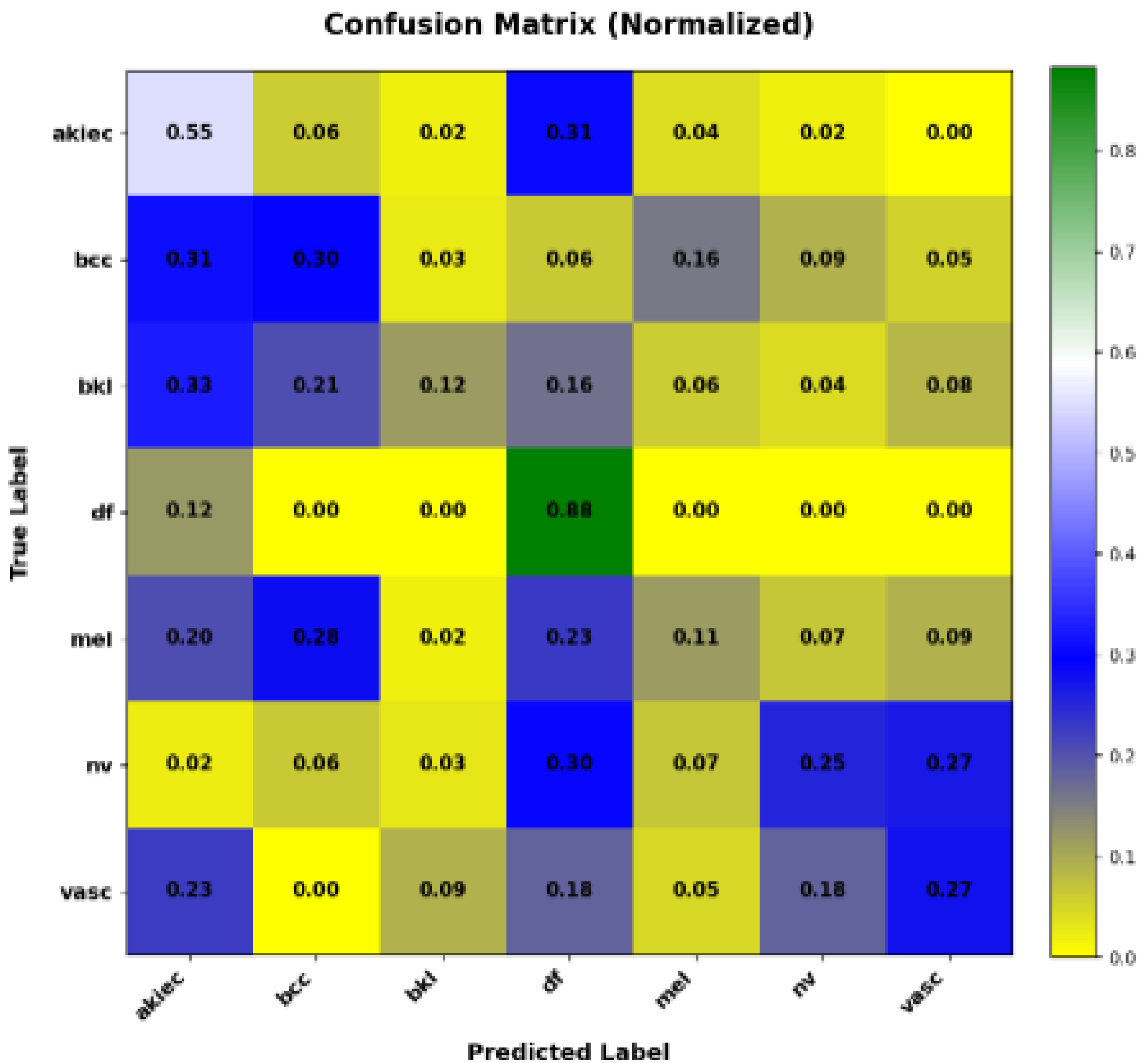
LIU

```
Test report:
              precision    recall  f1-score   support

       akiec     0.1617    0.5510    0.2500        49
         bcc     0.1345    0.2987    0.1855        77
         bkl     0.3393    0.1152    0.1719       165
          df     0.0369    0.8824    0.0708        17
         mel     0.1712    0.1138    0.1367       167
          nv     0.8944    0.2525    0.3938      1006
        vasc     0.0195    0.2727    0.0365        22

    accuracy                         0.2415      1503
   macro avg     0.2511    0.3552    0.1779      1503
weighted avg     0.6678    0.2415    0.3166      1503

array([[ 27,   3,   1,  15,   2,   1,   0],
       [ 24,  23,   2,   5,  12,   7,   4],
       [ 54,  34,  19,  27,  10,   7,  14],
       [  2,   0,   0,  15,   0,   0,   0],
       [ 34,  47,   3,  38,  19,  11,  15],
       [ 21,  64,  29, 303,  67, 254, 268],
       [  5,   0,   2,   4,   1,   4,   6]])
```

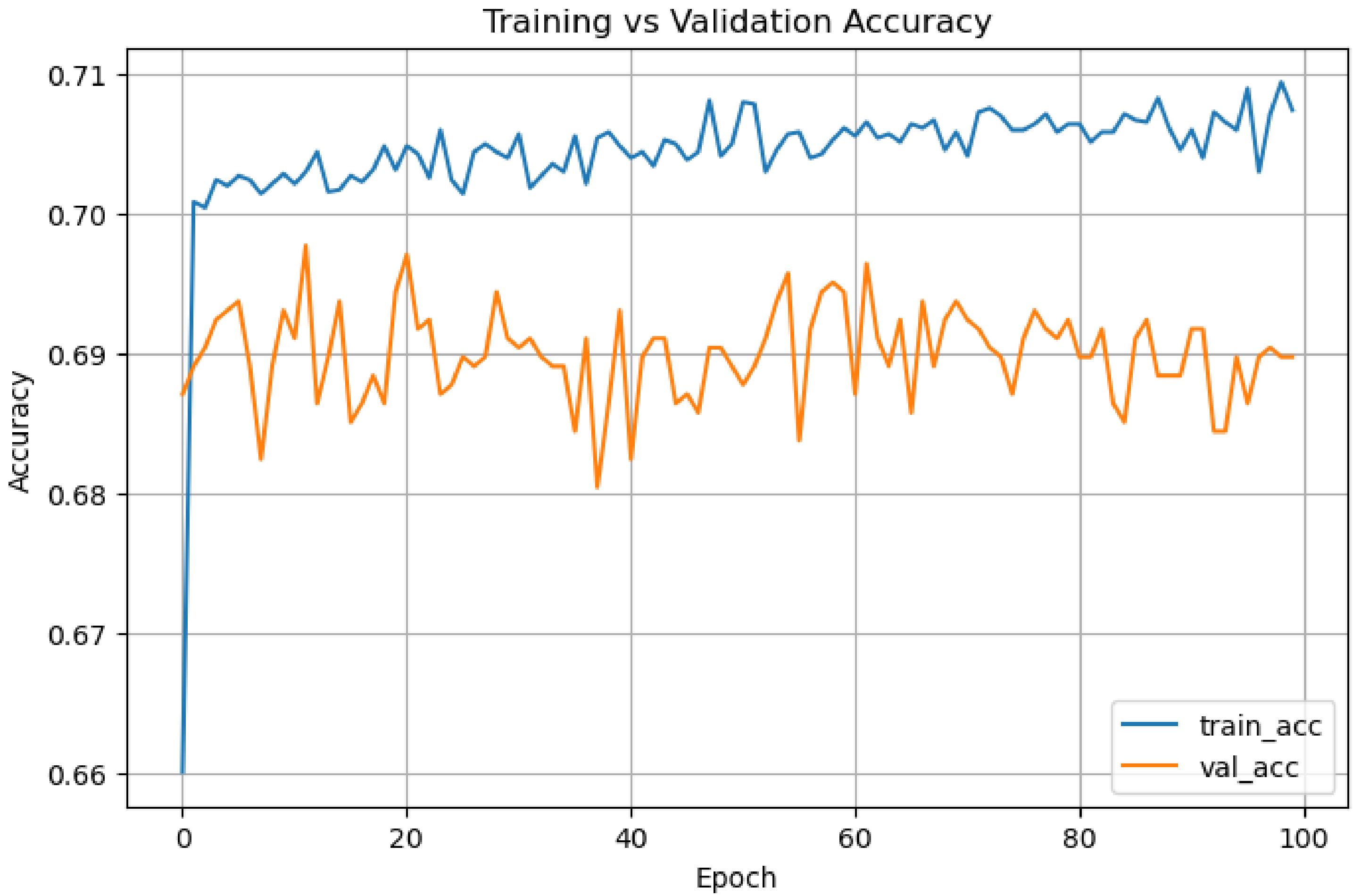
MODEL EVALUATION-CNN MODEL



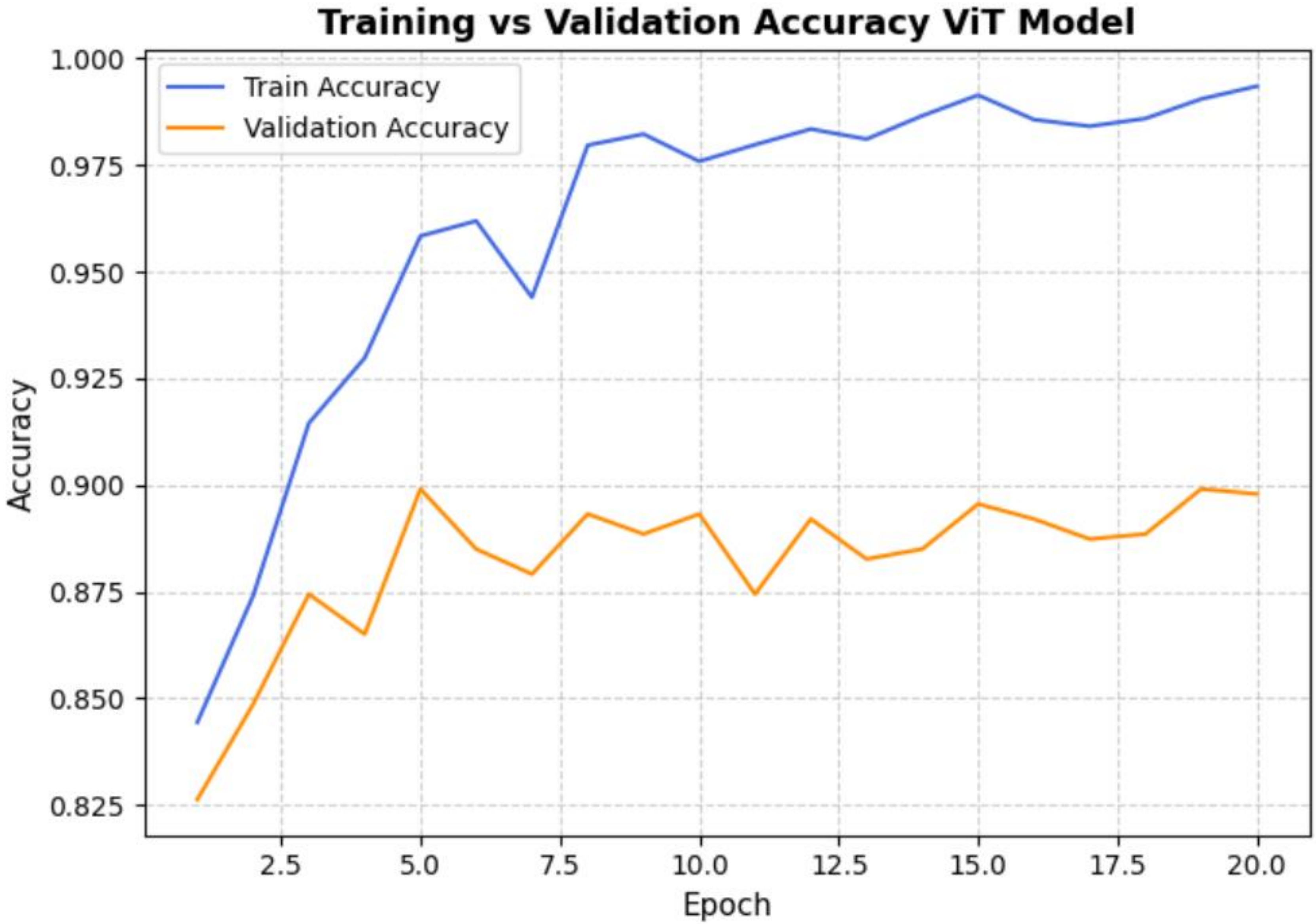
MODEL EVALUATION-CNN MODEL



MODEL EVALUATION-CNN MODEL

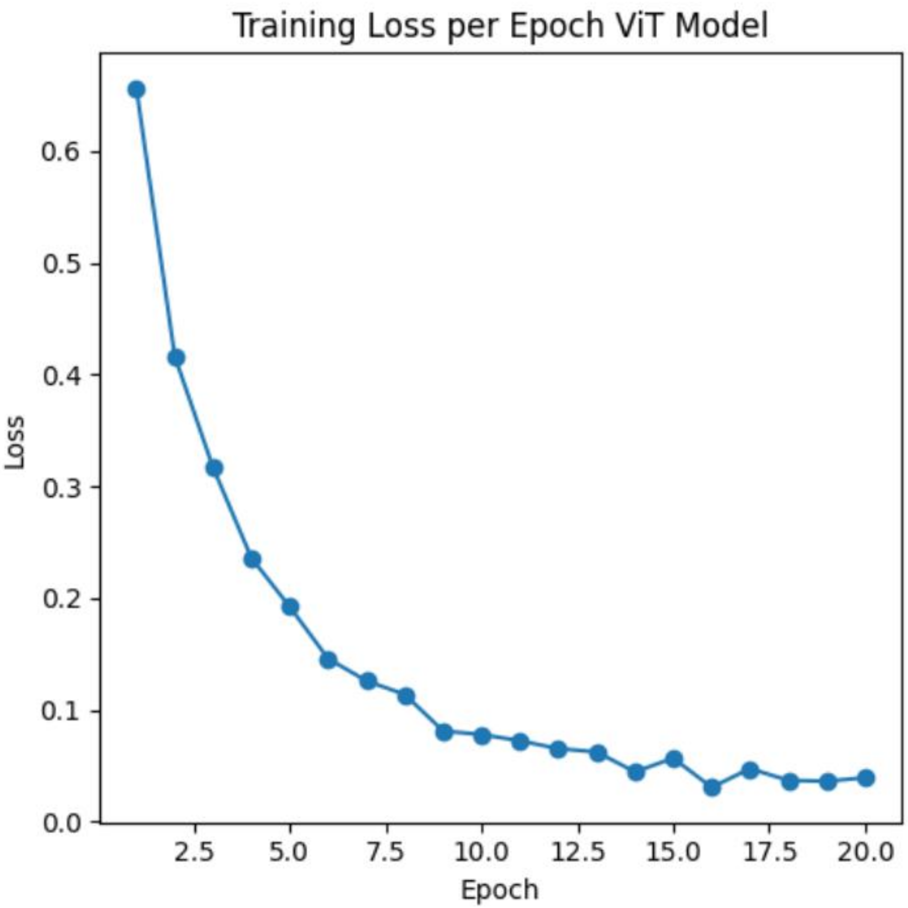


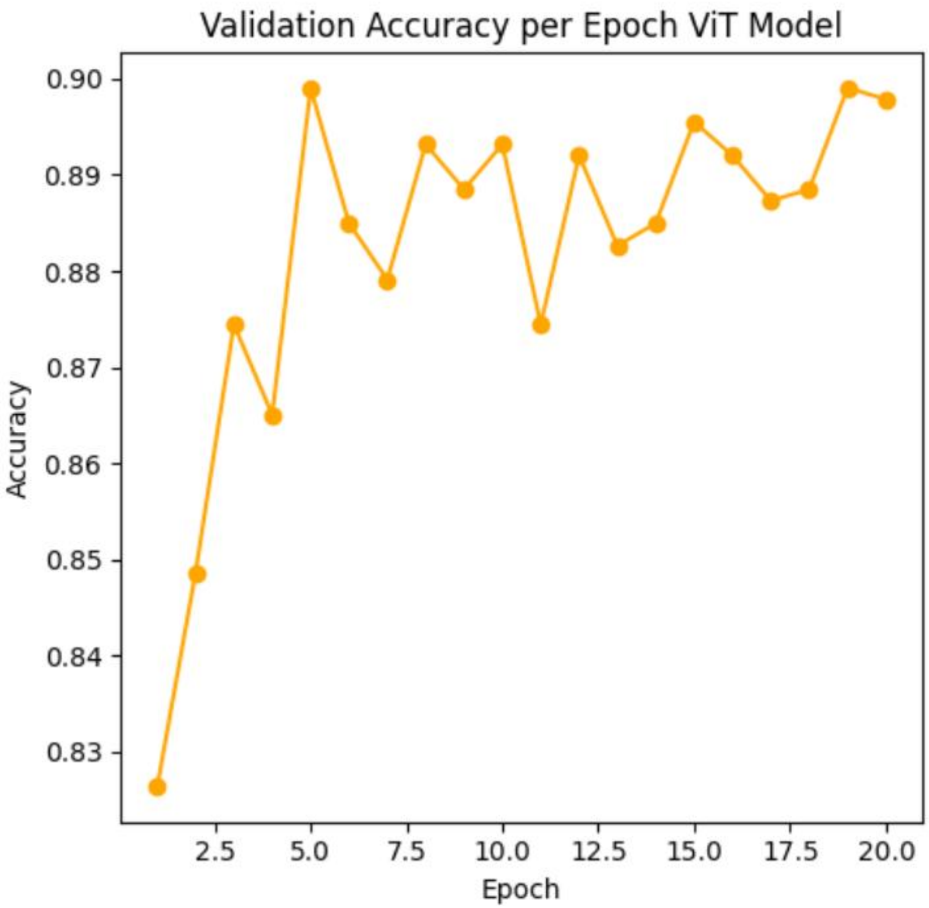
MODEL EVALUATION-VIT MODEL



MODEL EVALUATION-ViT MODEL

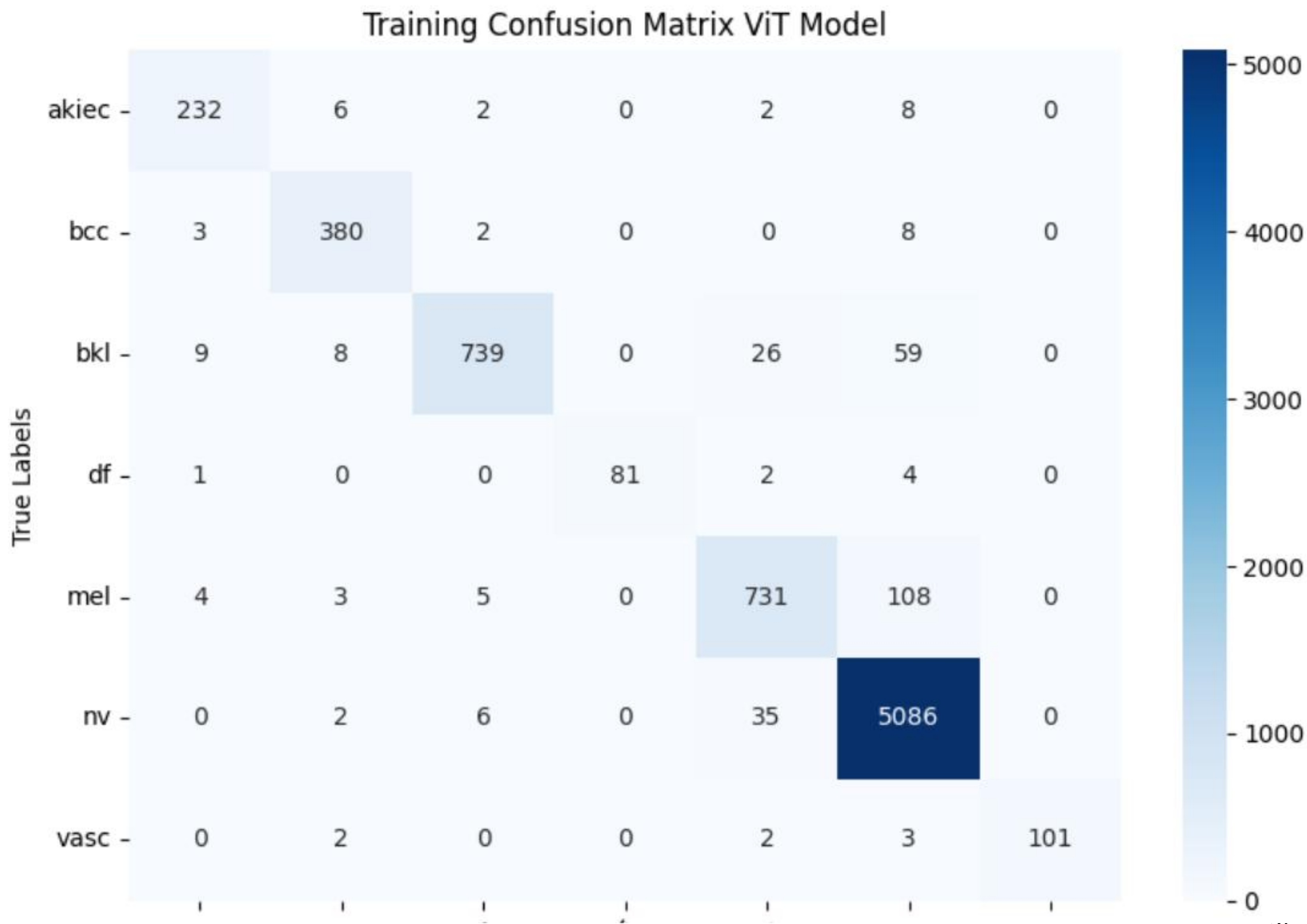
LIU



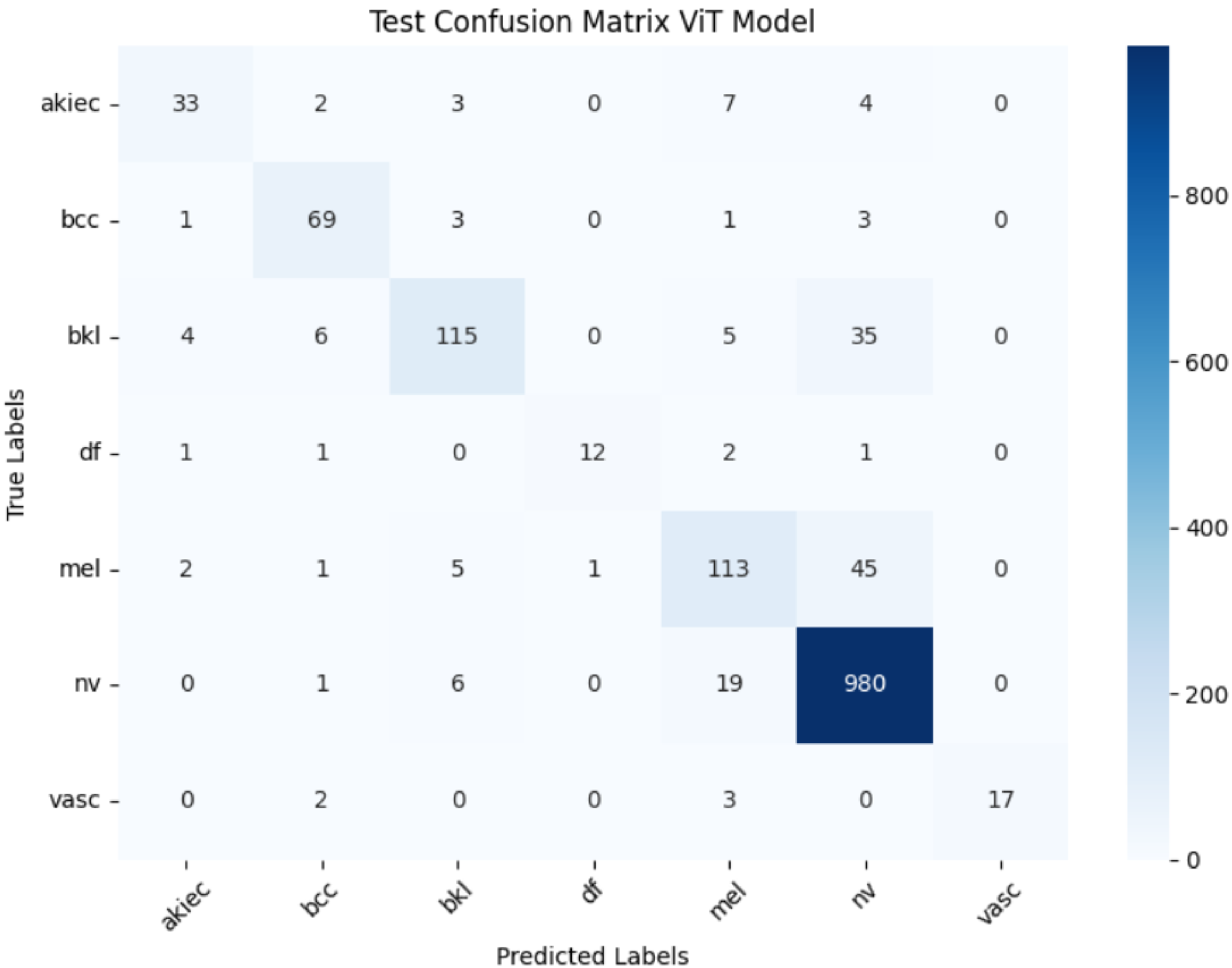


MODEL EVALUATION-ViT MODEL

LIU



MODEL EVALUATION-ViT MODEL



MODEL EVALUATION-ViT MODEL

LIU

Training Classification Report

Class	Precision	Recall	F1-Score	Support
Akiec	0.93	0.93	0.93	250
Bcc	0.95	0.97	0.96	393
Bkl	0.98	0.88	0.93	841
Df	1.00	0.92	0.96	88
Mel	0.92	0.86	0.89	851
Nv	0.96	0.99	0.98	5129
Vasc	1.00	0.94	0.97	108
Overall Accuracy			0.96	7660
Macro Avg	0.96	0.93	0.94	7660
Weighted Avg	0.96	0.96	0.96	7660

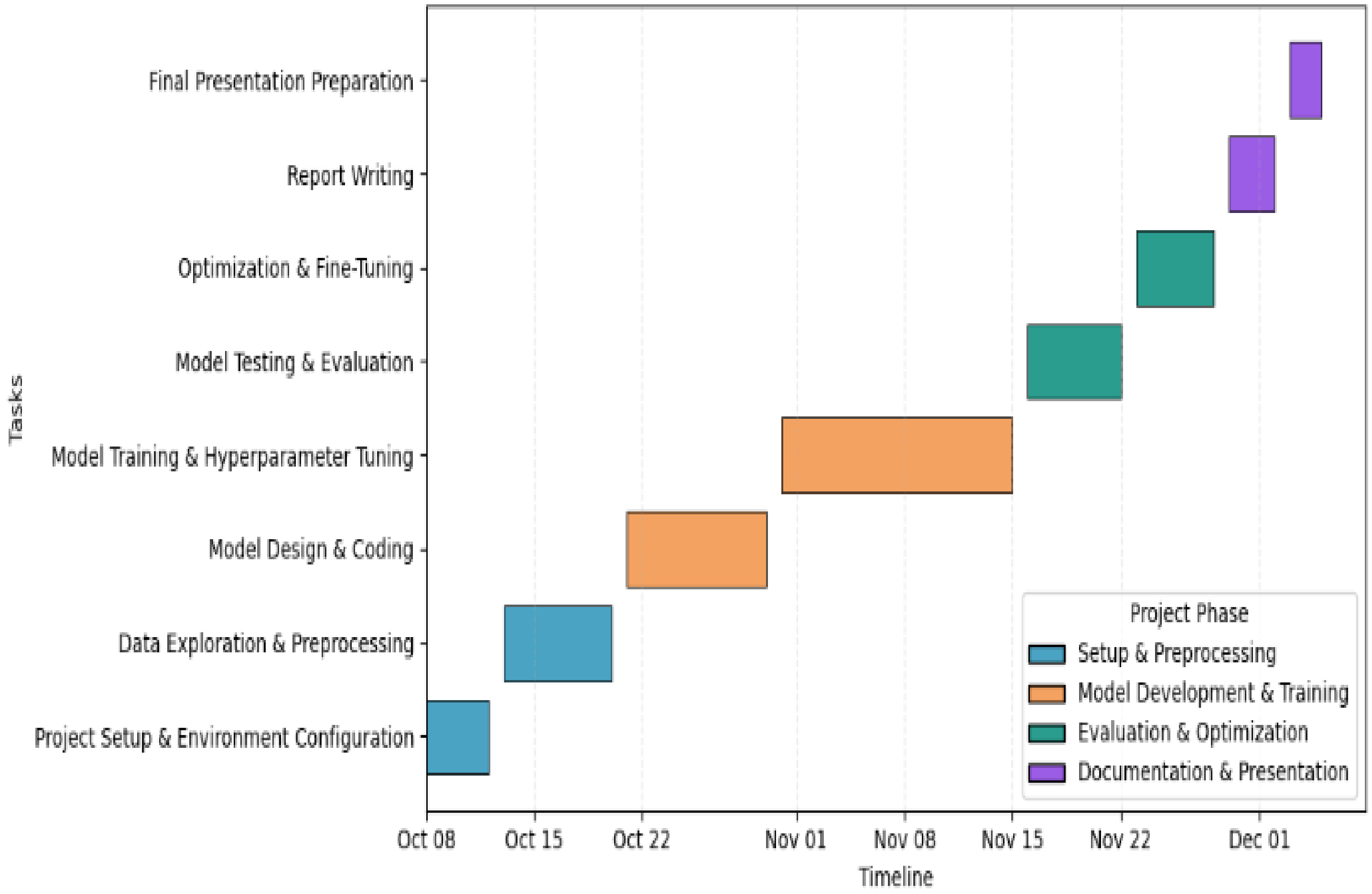
Testing Classification Report

Class	Precision	Recall	F1-Score	Support
Akiec	0.80	0.67	0.73	49
Bcc	0.84	0.90	0.87	77
Bkl	0.87	0.70	0.77	165
Df	0.92	0.71	0.80	17
Mel	0.75	0.68	0.71	167
Nv	0.92	0.97	0.95	1006
Vasc	1.00	0.77	0.77	22
Overall Accuracy			0.89	1503
Macro Avg	0.87	0.77	0.82	1503
Weighted Avg	0.89	0.89	0.89	1503

MODEL EVALUATION

CNN Model Accuracy = 70.53%
ViT Model Accuracy = 96%

TIMELINE


Final Presentation Preparation
Report Writing
Optimization & Fine-Tuning
Model Testing & Evaluation
Model Training & Hyperparameter Tuning
Model Design & Coding
Data Exploration & Preprocessing
Project Setup & Environment Configuration
Tasks
Oct 08
Oct 15
Oct 22
Nov 01
Nov 08
Nov 15
Nov 22
Dec 01
Timeline
Project Phase
Setup & Preprocessing
Model Development & Training
Evaluation & Optimization
Documentation & Presentation

MODEL COMPARISION: CNN, VIT MODEL, & RESEARCH

LIU

Aspect	CNN Model	VIT Model
Feature learning	Uses the convolutional layers to directly extract the spatial features	Uses the backpropogation, tranformer and feature extraction to deal with data
Accuracy	70.53%	91%
Effectiveness	More viable for less dataset sizes	More viable for large datasets sizes
Generalization	Works well for simple tasks like simple image classification	Works well for complex and complicated tasks like feature extraction and back propagation with hybrid data usage

Model	Training Accuracy	Loss
CNN	70.53%	1.20 (approx.)
VIT	96%	0.03 (approx.)
Research paper	Fuzzy GC-SCNN-99.75% Systemic review-96%	0.0005 (approx.) 0.04(approx.)

CONCLUSION

- Achieved 96% accuracy with the Vision Transformer(ViT)model.
- Captured both local details and global patterns in skin lesions.
- Outperformed traditional CNN models in reliability and precision.
- Shows strong potential for early and accurate melanoma detection.
- Supports AI-assisted diagnosis for faster clinical decisions.

FUTURE WORK

- Integrate the techniques from Skin Caner DeepNet such as adaptive contrast enhancement(AGCWD), hybrid Mask R-CNN + GrabCut segmentation.
- Validate model robustness on larger and more diverse datasets such as ISCI 2019 and PH2
- Employ visualization tools like Grad-CAM and attention heatmaps to interpret ViT attention regions and improve clinical trust
- Combine clinical metadata (e.g., patient history, lesion evolution) with image-based learning to enhance diagnostic accuracy
- Develop a web-based diagnostic tool for clinical testing.

REFERENCES

LIU

- U. Bhimavarapu, G Battineni, “Skin Lesion Analysis for Melanoma Detection Using the Novel Deep Learning Model Fuzzy GC-SCNN”, *Healthcare*, vol. 10, no. 5, p.962, May 2022
- American Academy of Dermatology, “Skin cancer,” *AAD*, June. 20, 2025. [Online]. Available: https://www.aad.org/media/stats-skin-cancer. [Accessed: Oct. 04, 2025].
- A. Adealnabi, S. Eduardo J, M. Becevic, E. H. Smith, P Rao, “Transformers in Skin Lesion Classification and Diagnosis: A systematic Review, Health Informatics, September 2024.
- A. S. Al-Waisy, S. AlFahdawi, M.I. Khalaf, M. A. Mohmmed, B. All-Attar, M.N. AL-Andoli, “ A Deep Learning Framework for Automated early Diagnosis and Classification of Skin Cancer Lesions in Dermoscopy Images”, Scientific Reports, vol. 15, no. 1, p.31234, August 2025

Questions