INTERNATIONAL UNIVERSITY - VNU HCMC
FINAL PROJECT (2025-2026,S1)
IT159IU Artificial Intelligence

# Self-Supervised Domain Adaptation for Skin Lesion Classification


Student name: [Trần Nam Anh] Student ID: [ITDSIU23030]
[Nguyễn Đức Hải] [ITDSIU23006]

Group [10 ] Place: [LA1.604]


Date:[24 December,2025]

# Self-Supervised Domain Adaptation for Skin Lesion Classification

Course: Artificial Intelligence

-IT159IU-

# Table of Content

# Introduction

- Real-world deployment is **limited** by two key challenges:
    - **Large, costly** labeled datasets
    - Domain shift → **degradation** clinical settings.

**Project Objective:**

- Develop a Self-Supervised Domain Adaptation **(SSDA)** framework
    - Self-Supervised Learning **(SSL)**: exploit unlabeled data
    - Domain Adaptation **(DA)**: improve cross-domain robustness
- Evaluate SSDA on **HAM10000** against strong supervised baselines

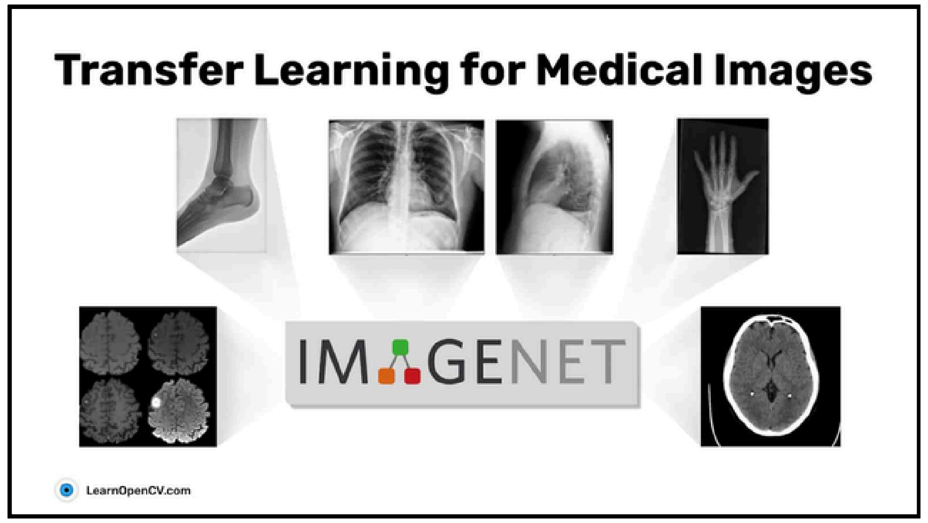
# RELATED WORK

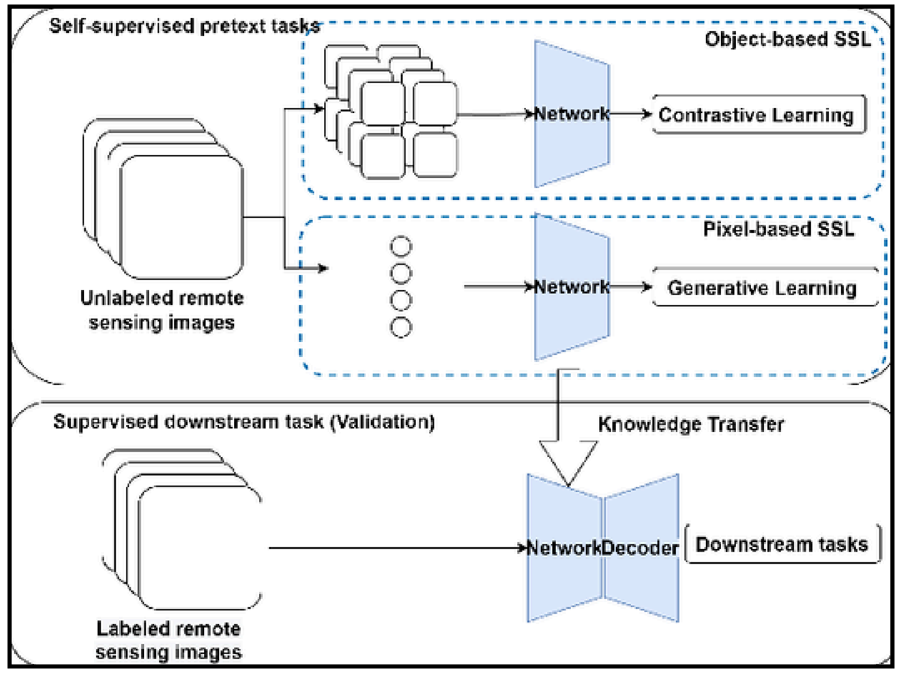


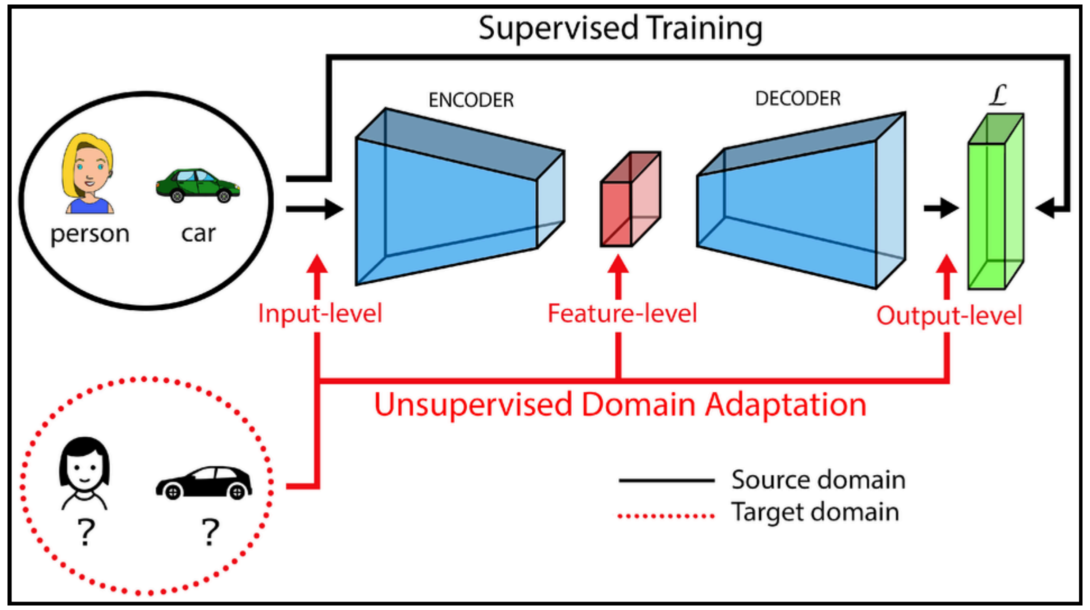
**SUPERVISED TRANSFER LEARNING IN MEDICAL IMAGING**

**SELF-SUPERVISED LEARNING (SSL)**

**UNSUPERVISED DOMAIN ADAPTATION (DA)**



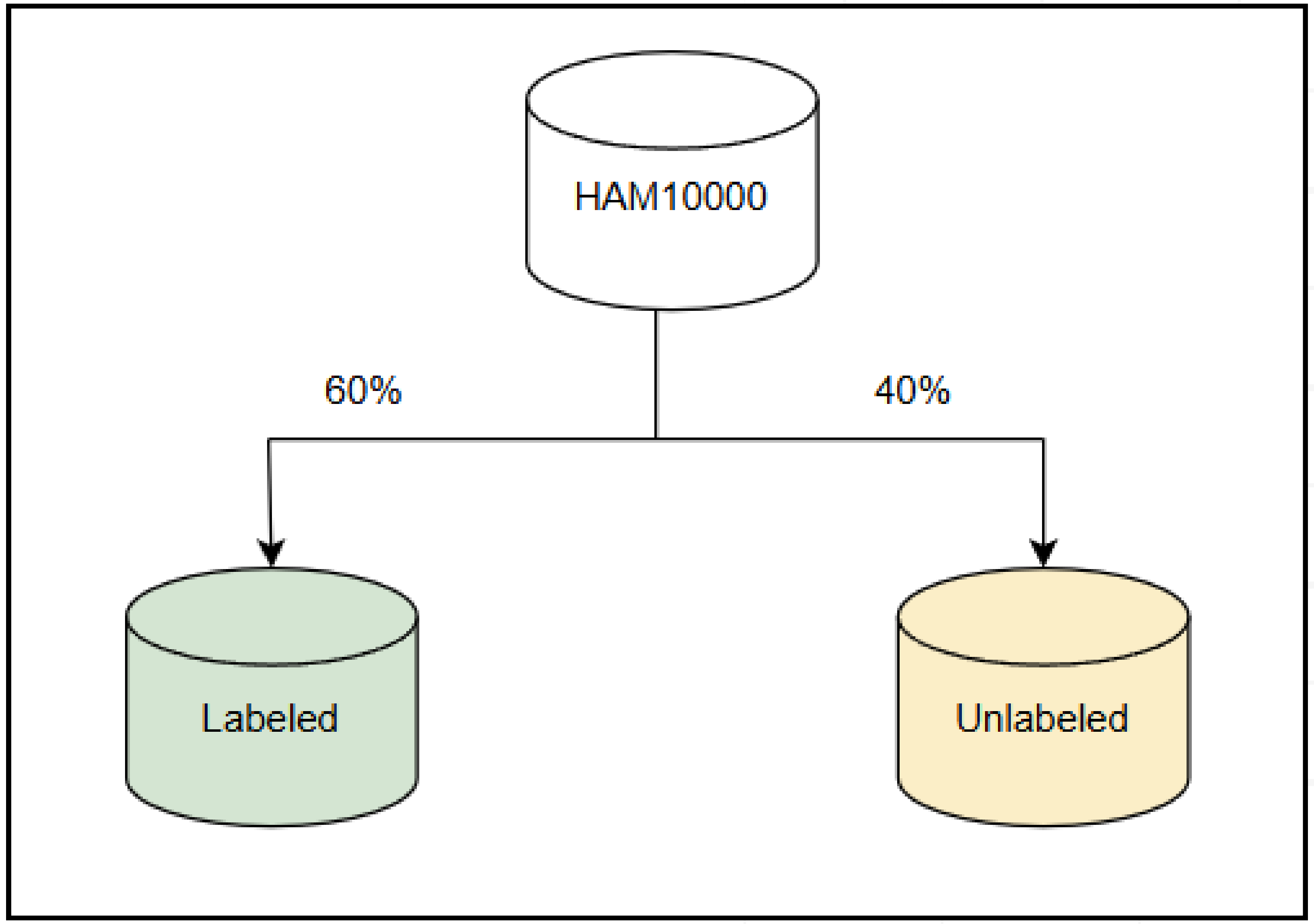



# METHODOLOGY

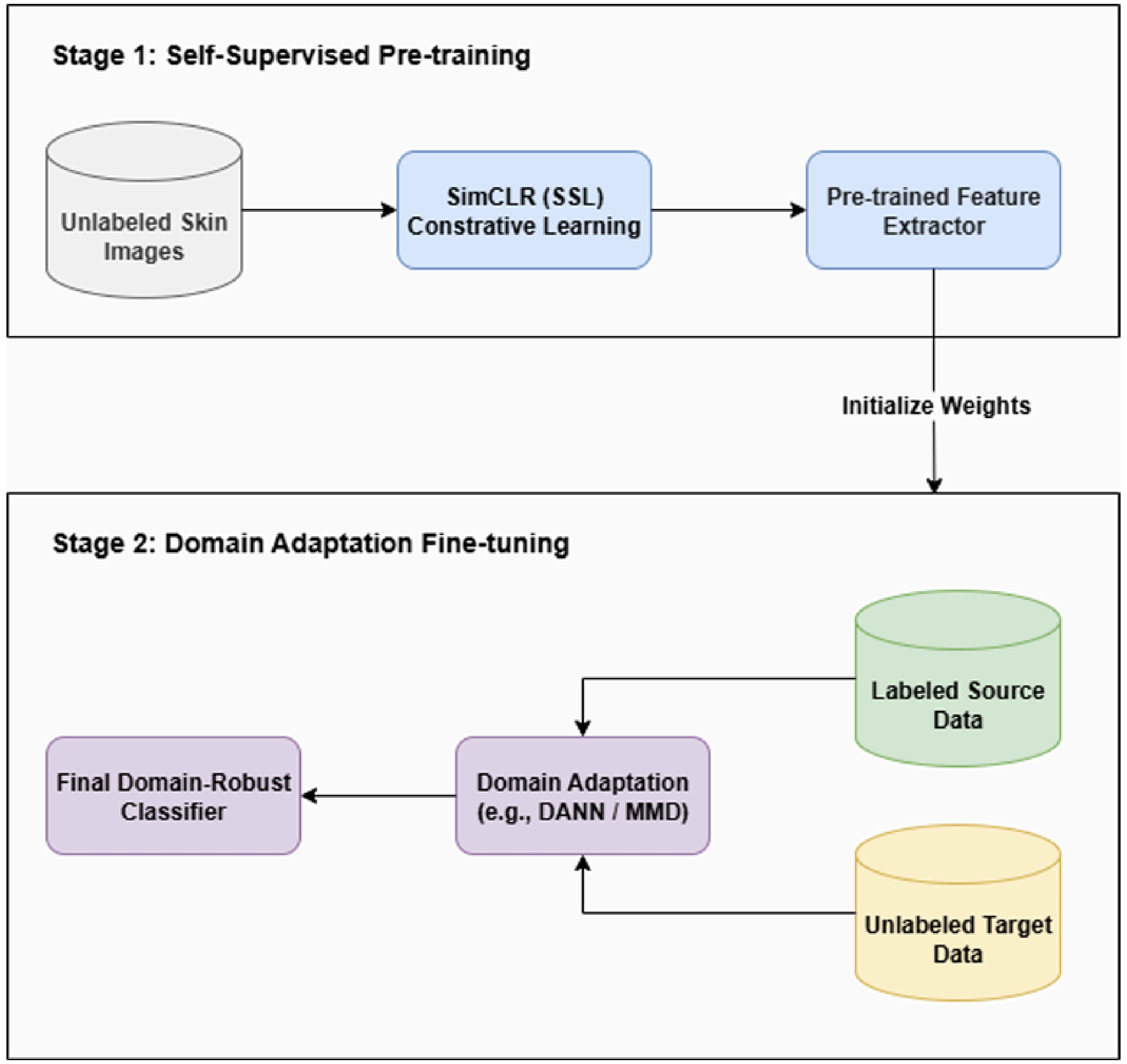
## 1.DATA PREPARATION AND SPLITTING

- HAM10000 dataset
- Golden Data Split
  - 60% labeled images (source domain) → supervised learning
  - 40% unlabeled images → evaluate model



# METHODOLOGY

## 2. METHOD 1: THE PROPOSED SSDA FRAMEWORK (SIMCLR + DANN)

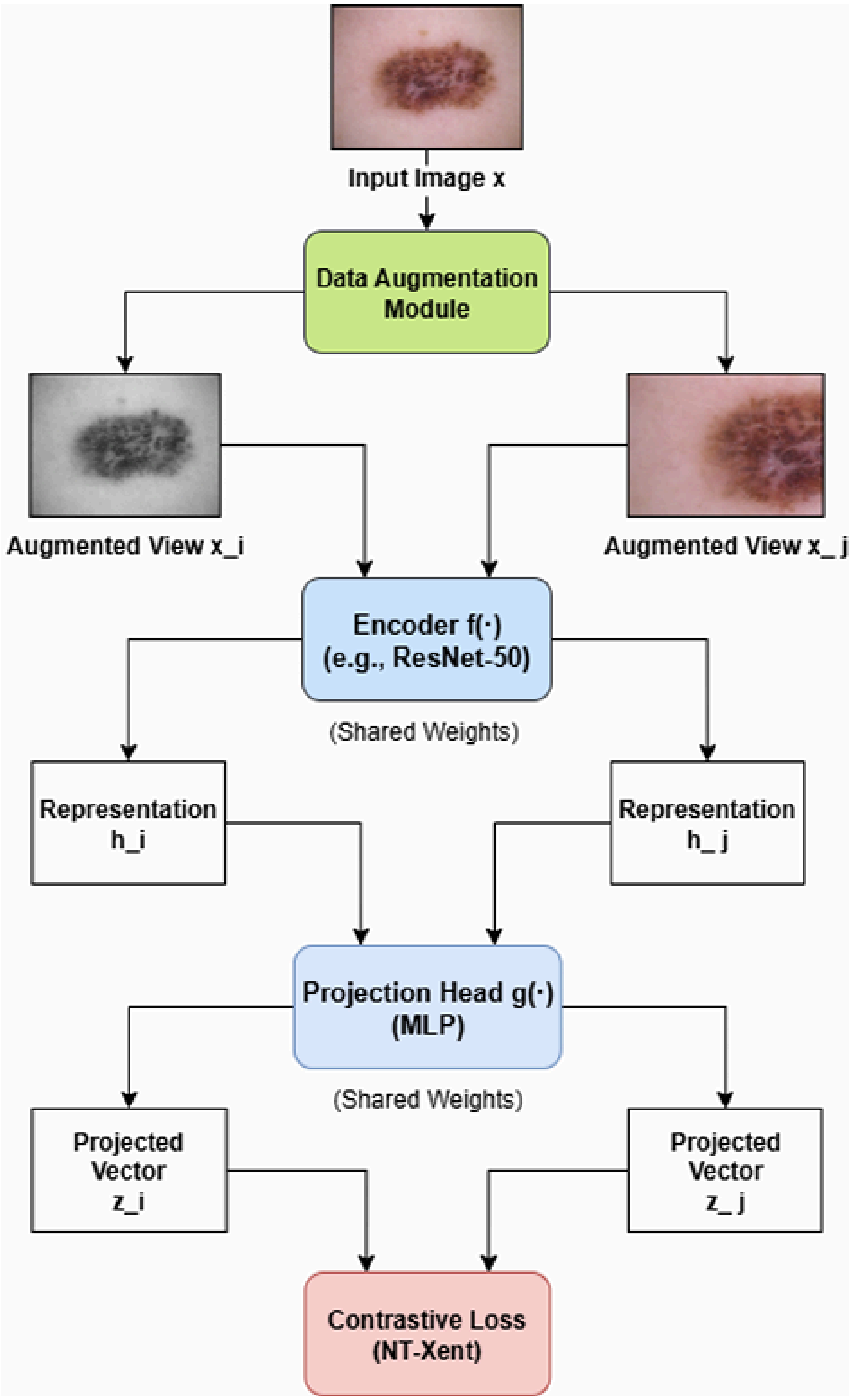


# METHODOLOGY

### 2.1. STAGE 1: SELF-SUPERVISED PRE-TRAINING (SSL)

- **Encoder** f(·): ResNet-50 extracts features from augmented images.
- **Projection Head** g(·): 2-layer MLP maps features to the contrastive space.
- **NT-Xent Loss**: Maximizes similarity between two views of the same image and minimizes similarity to other images in the batch.

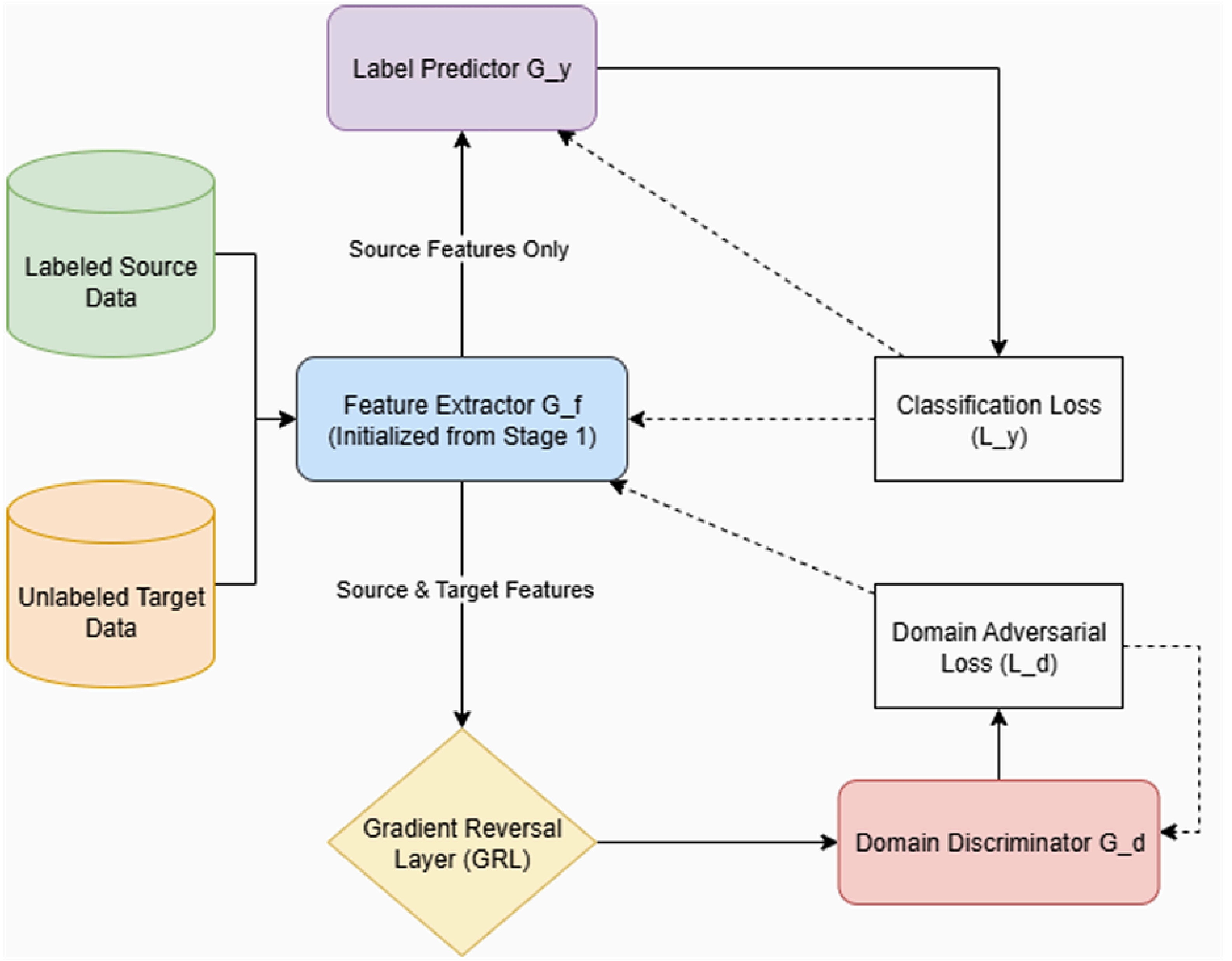
→ After 50 pre-training epochs, the projection head g(.) is removed, and the encoder f(.) is used as the pre-trained feature extractor for the next stage.



# METHODOLOGY

### 2.2. STAGE 2: DOMAIN-ADVERSARIAL FINE-TUNING (DA)

- **Feature Extractor G_f**: encoder from Stage 1
- **Label Predictor G_y**: A classifier trained only on labeled source data, measured by:
  - Classification Loss (L_y): a weighted cross-entropy loss
- **Domain Discriminator G_d**: A classifier trained to distinguish source and target domains measured by:
  - Domain Adversarial Loss (L_d)
- **Gradient Reversal Layer (GRL)**: reverses the sign of the gradient from Domain Discriminator



# METHODOLOGY

## 3.1: METHOD 2: THE ABLATION STUDY MODEL (SSL-FT)

To isolate the effect of SSL pre-training:

**Ablation study:** **SSL-FT** model SSL-FT follows the baseline training procedure but initializes the encoder with **SimCLR pre-trained weights** instead of ImageNet weights (without using any DANN components)

# METHODOLOGY

## 4.1: METHOD 3: THE SUPERVISED BASELINE

To establish a strong baseline:

- we implemented a **supervised transfer learning** model with a **ResNet-50** encoder **pre-trained on ImageNet.**
- The model is fine-tuned end-to-end on **labeled source data** only and trained under the **same optimized settings** (epochs, SGD, learning-rate scheduler, weighted cross-entropy) to ensure fair comparison.

# Implementation & Results



## Summary of Performance

| Model | Pre-training Strategy | Fine-tuning Method | Final Target Accuracy (%) |
|---|---|---|---|
| **Supervised Baseline** | ImageNet Supervised | Supervised FT | **83.33%** |
| **SSDA (Proposed)** | ImageNet -> SimCLR | DANN | **75.67%** |
| **SSL-FT** | ImageNet -> SimCLR | Supervised FT | **70.73%** |

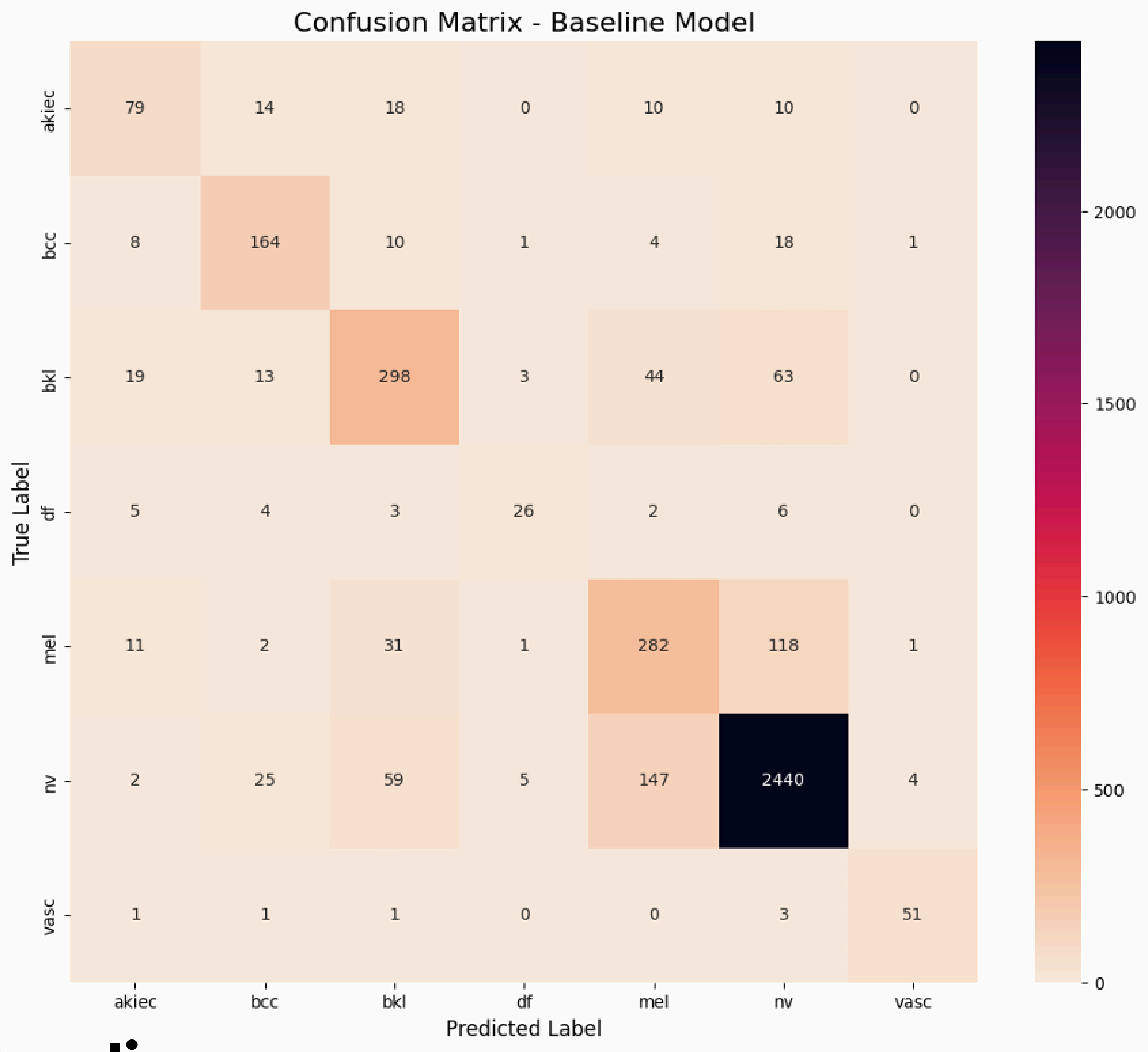
# Implementation & Results

```
==================================================
BASELINE MODEL RESULTS
   - Final Accuracy on Target: 83.33%
==================================================

Classification Report:
              precision    recall  f1-score   support

       akiec     0.6320    0.6031    0.6172       131
         bcc     0.7354    0.7961    0.7646       206
         bkl     0.7095    0.6773    0.6930       440
          df     0.7222    0.5652    0.6341        46
         mel     0.5767    0.6323    0.6032       446
          nv     0.9180    0.9098    0.9139      2682
        vasc     0.8947    0.8947    0.8947        57

    accuracy                         0.8333      4008
   macro avg     0.7412    0.7255    0.7315      4008
weighted avg     0.8358    0.8333    0.8342      4008
```




**Supervised Baseline**

# Implementation & Results

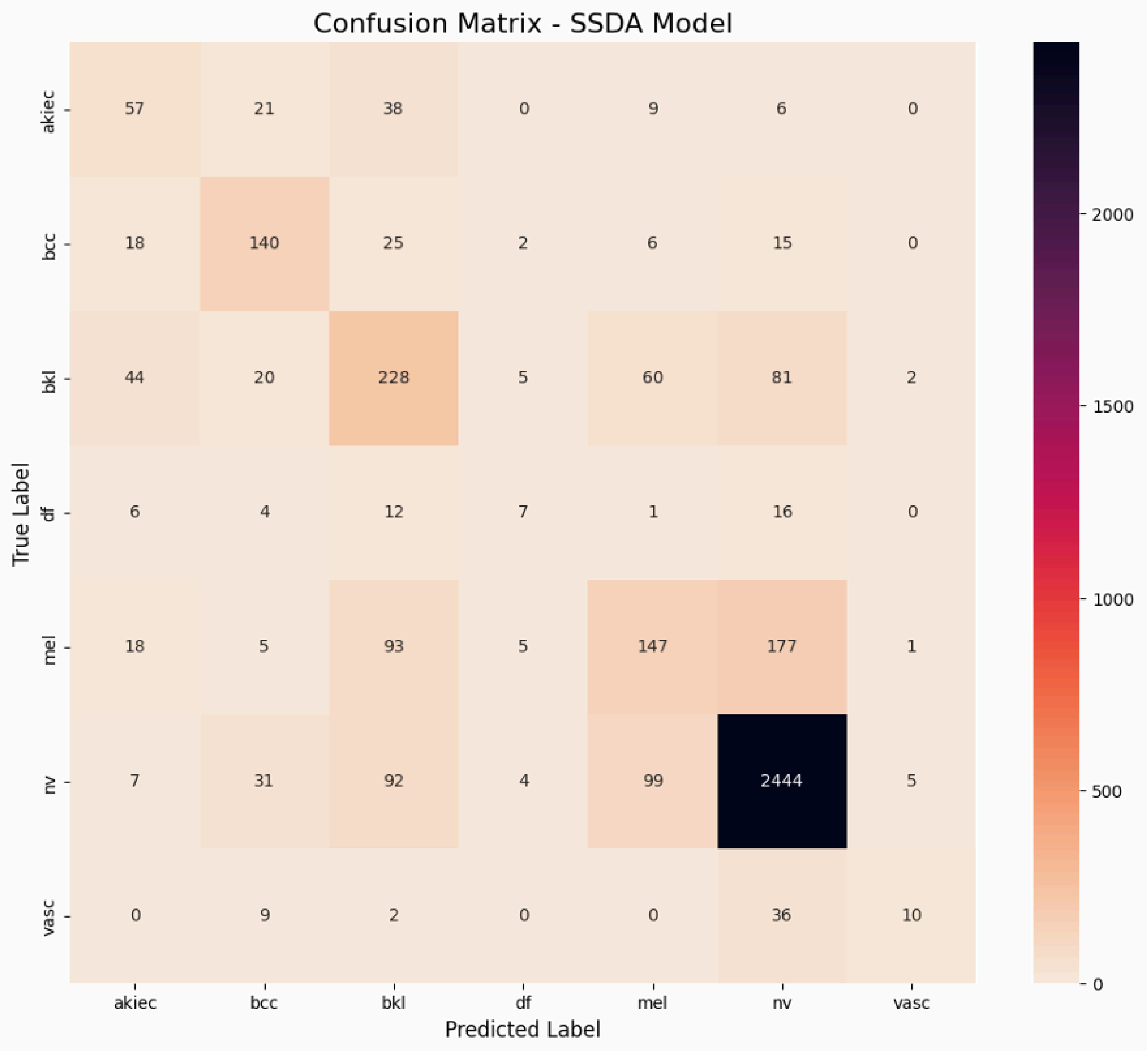


```
==================================================
SSDA MODEL RESULTS
  - Final Accuracy on Target: 75.67%
==================================================

Classification Report:
              precision    recall  f1-score   support

       akiec     0.3800    0.4351    0.4057       131
         bcc     0.6087    0.6796    0.6422       206
         bkl     0.4653    0.5182    0.4903       440
          df     0.3043    0.1522    0.2029        46
         mel     0.4565    0.3296    0.3828       446
          nv     0.8807    0.9113    0.8957      2682
        vasc     0.5556    0.1754    0.2667        57

    accuracy                         0.7567      4008
   macro avg     0.5216    0.4573    0.4695      4008
weighted avg     0.7463    0.7567    0.7482      4008
```




**SSDA Model (Proposed Method)**

# Implementation & Results

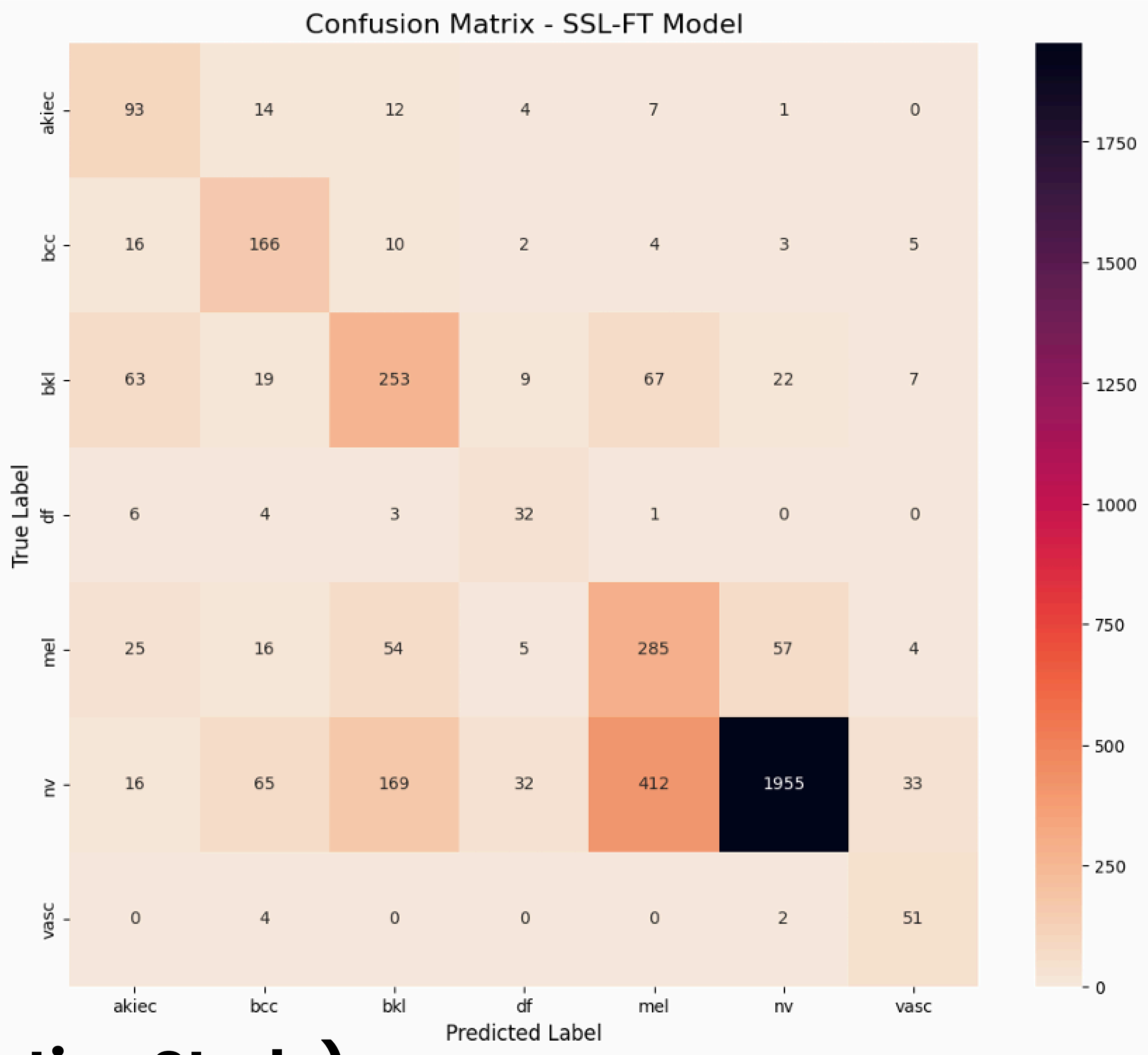


```
================================================================
PROPOSED METHOD (SSL-FT) RESULTS
   - Final Accuracy on Target: 70.73%
================================================================

Classification Report:
              precision    recall  f1-score   support

       akiec     0.4247    0.7099    0.5314       131
         bcc     0.5764    0.8058    0.6721       206
         bkl     0.5050    0.5750    0.5377       440
          df     0.3810    0.6957    0.4923        46
         mel     0.3673    0.6390    0.4664       446
          nv     0.9583    0.7289    0.8280      2682
        vasc     0.5100    0.8947    0.6497        57

    accuracy                         0.7073      4008
   macro avg     0.5318    0.7213    0.5968      4008
weighted avg     0.7927    0.7073    0.7318      4008
```




**SSL-FT Model (Ablation Study)**

# Evaluation & Discussion

**Finding 1:** The Limited but Evident Impact of Domain Adaptation

- **SSDA** (75.67% accuracy)
- **SSL-FT** (70.73% accuracy)

→ **increase** 5%

- **DANN** architecture success in align:
  - distributions
  - mitigating

  | negative effect of domain shift → ↑ model's generalization performance

- **The training logs**
  - show domain loss stabilizing ~ 0.63(theoretical optimum)

# Evaluation & Discussion

**Finding 2**: The Unexpected Superiority of the Supervised Baseline

- Achieving a final accuracy of 83.33% > 
  - **SSL-FT** (70.73%)
  - **SSDA** (75.67%)

- Two **hypotheses** to explain this phenomenon:
  - First, **regarding feature robustness**, the ResNet-50 model has learned rich and 20 general-purpose set of visual features.
  - Second, **regarding the adaptation process,** Fine-tuning the ImageNet-pretrained encoder with SimCLR on the small HAM10000 dataset may have caused "**catastrophic forgetting**", reducing general visual features needed to distinguish visually similar classes (e.g., bcc vs. nv).

→ A well-optimized, simpler baseline can be remarkably difficult to surpass, especially upon a strong pre-trained foundation like ImageNet.

# Conclusion

**SSDA** framework implemented and evaluated

SSL and DA effective individually, but not **jointly**

ImageNet pre-training is the **dominant** performance factor

SSL were **outweighed** by the potential loss

# Future Work

1. Exploring Higher Image Resolution
2. Investigating Alternative Backbones
3. Applying Pseudo-Labeling

# Thank You

## For Your Attention

# Reference

1. Ahmed, S., et al. (2023). Human Skin Diseases Detection and Classification using CNN.
2. Azizi, S., et al. (2021). Big Self-Supervised Models Advance Medical Image Classification. CVPR.
3. Ben-David, S., Blitzer, J., Crammer, K., & Pereira, F. (2010). A theory of learning from different domains. Machine Learning, 79(1-2), 151–175.
4. Berthelot, D., et al. (2021). Semi-Supervised Models are Strong Unsupervised Domain Adaptation Learners. NeurIPS.
5. Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations (SimCLR). ICML.
6. Ganin, Y., & Lempitsky, V. (2015). Unsupervised Domain Adaptation by Backpropagation (DANN). ICML.
7. He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum Contrast for Unsupervised Visual Representation Learning (MoCo). CVPR.
8. Li, J., et al. (2021). Speech SimCLR: Combining Contrastive and Reconstruction Objectives for Self-Supervised Speech Representation Learning.
9. Yadav, R., & Bhat, P. (2024). A Systematic Literature Survey on Skin Disease Detection and Classification using Machine Learning and Deep Learning.