Pharmaceutical Drugs & Vitamins Image Classification

Scan me!

Group 1
2702343302 - Audrey Theodora Phang
2702256630 - Eirene Michella Tjhan
2602240644 - Fernando Farrell
2702208575 - Michelle Nathania
2702320425 - Nadja Nayara

20,000

prescription drug products
approved for marketing

Can you indentify them?

Ibuprofen

Hydroxychloroquine

Could you tho?

Why a deep learning model?

- Automatically learns features from raw images
- Pixel-level patterns so we can gain spatial & visual feature (texture, shape, and color)
- Allow transfer learning

It could

Identification using phone camera

Automated sorting in iventory systems

Quick identification

Detect packaging error

Data Overview

Pharmaceutical Drugs and Vitamins Synthetic Images

This dataset contains images of 10 common drugs and vitamins in the Philippines, intended for educational use in drug image classification tasks.

Input

An image of a pharmaceutical drug or vitamin (like a photo of a pill or tablet)

Output

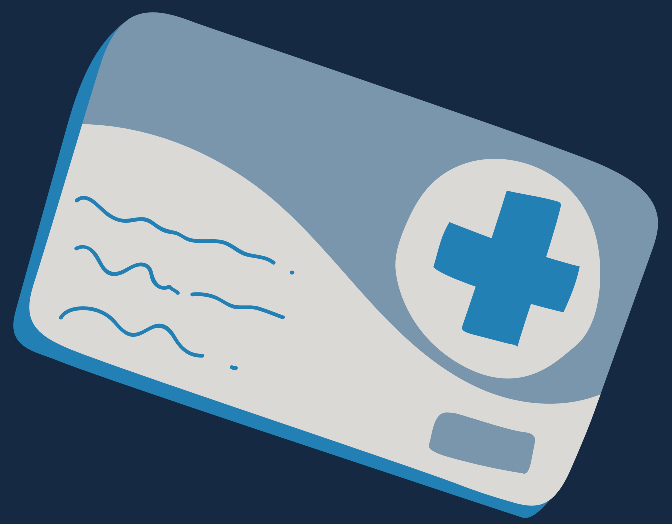
The predicted class label indicating which drug the image belongs to.

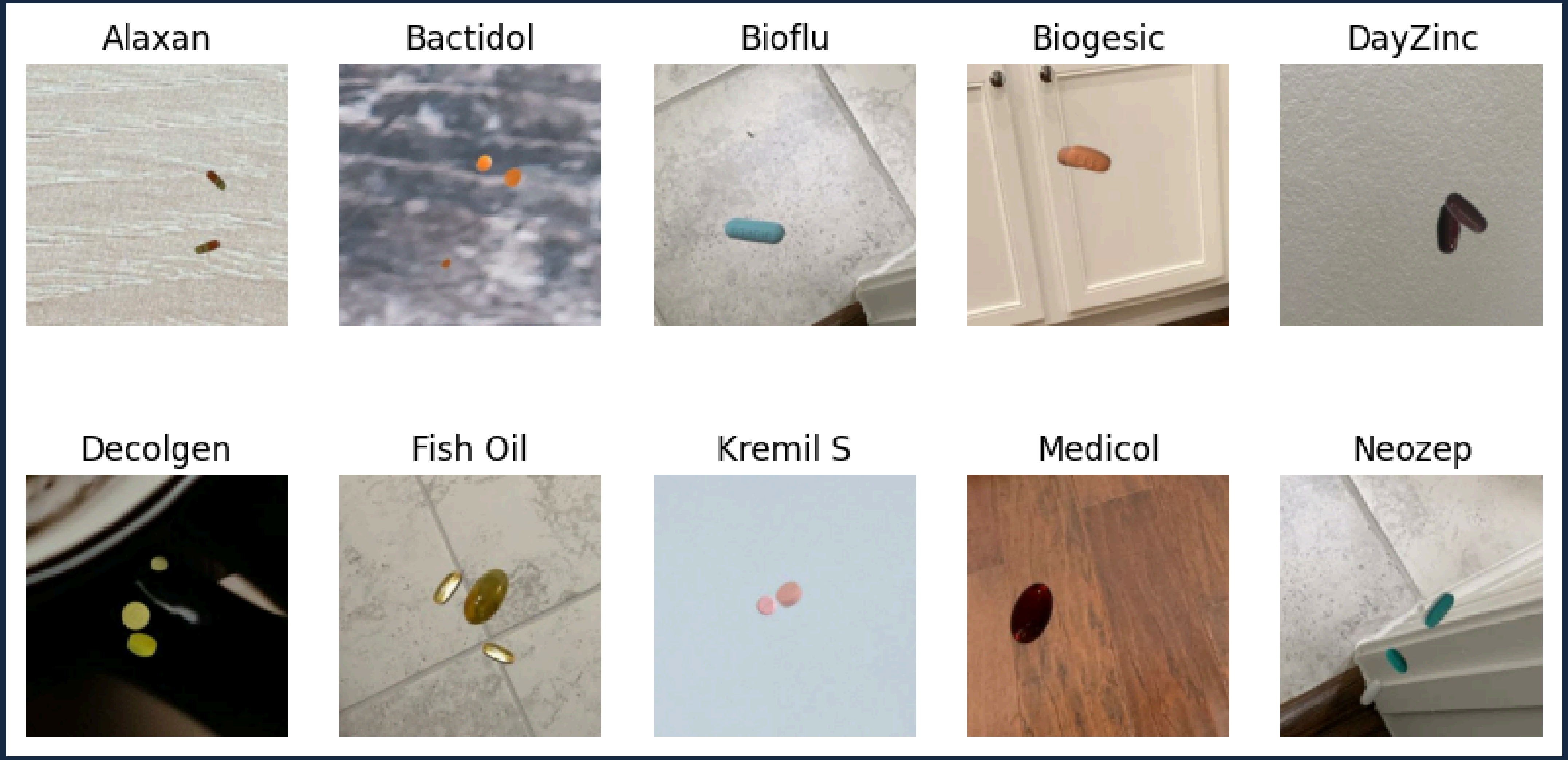
10,000

Images

10

Classes

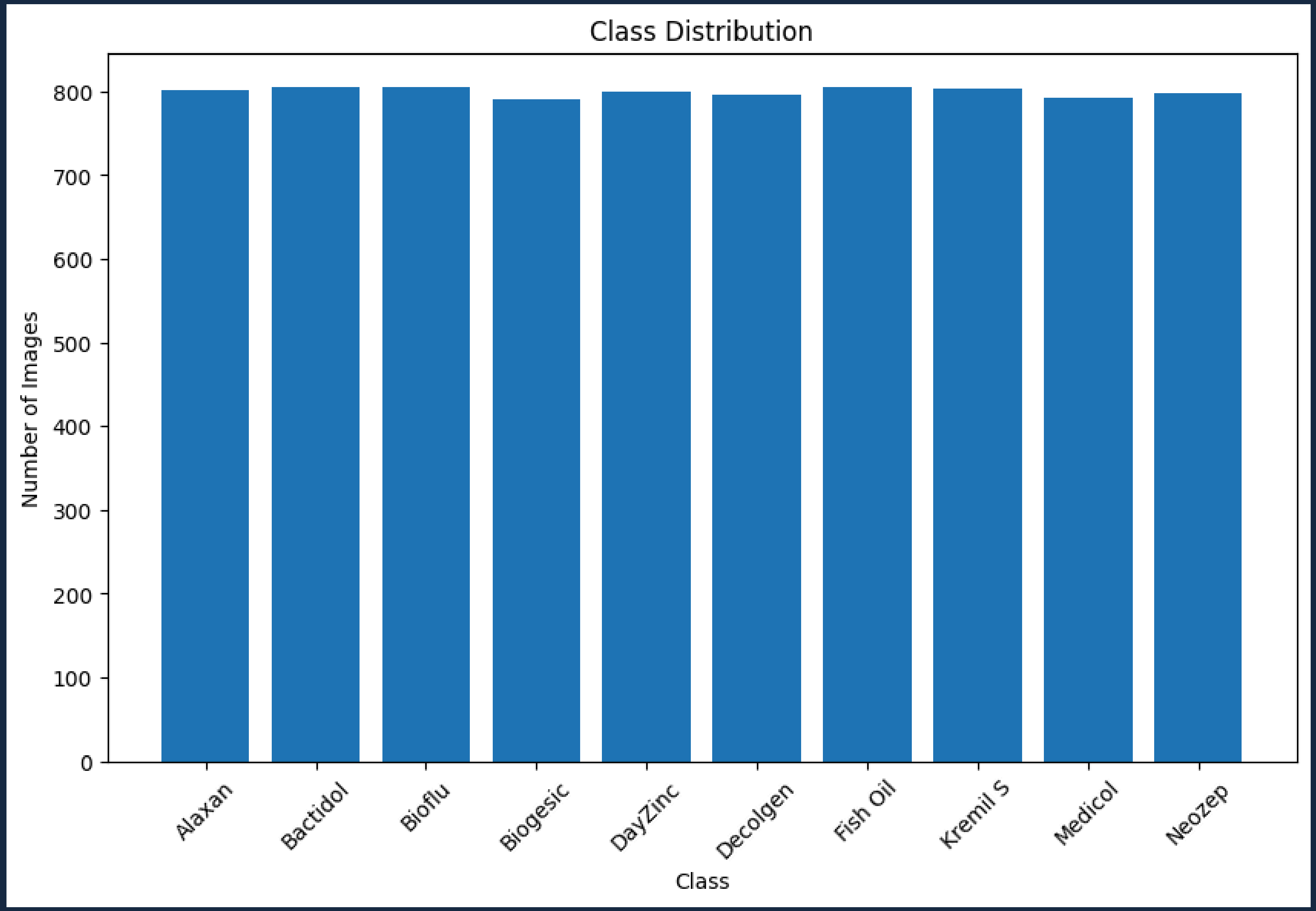




Precisely 1,000 images per class

Exploratory Data Analysis

Split data: 80% Test 10% Val 10% Test



Class distribution in train_ds:

Alaxan: 802

Bactidol: 806

Bioflu: 806

Biogesic: 791

DayZinc: 799

Decolgen: 797

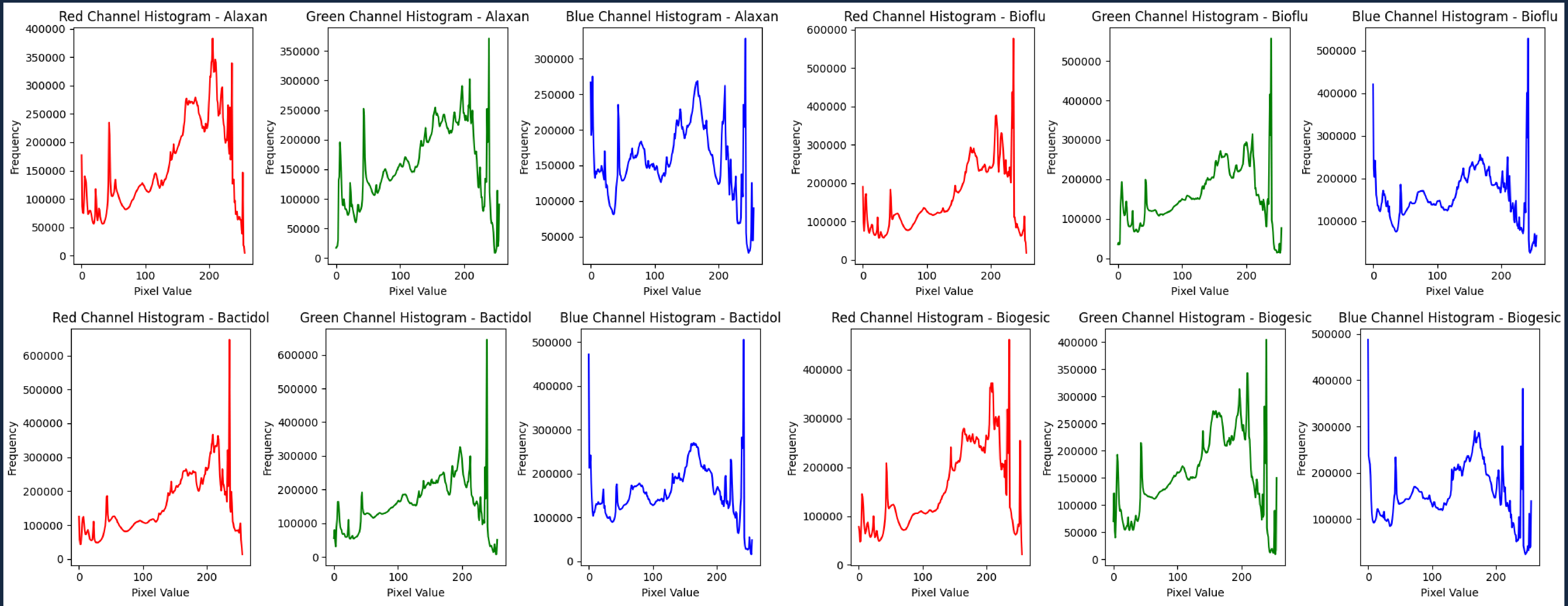
Fish Oil: 806

Kremil S: 803

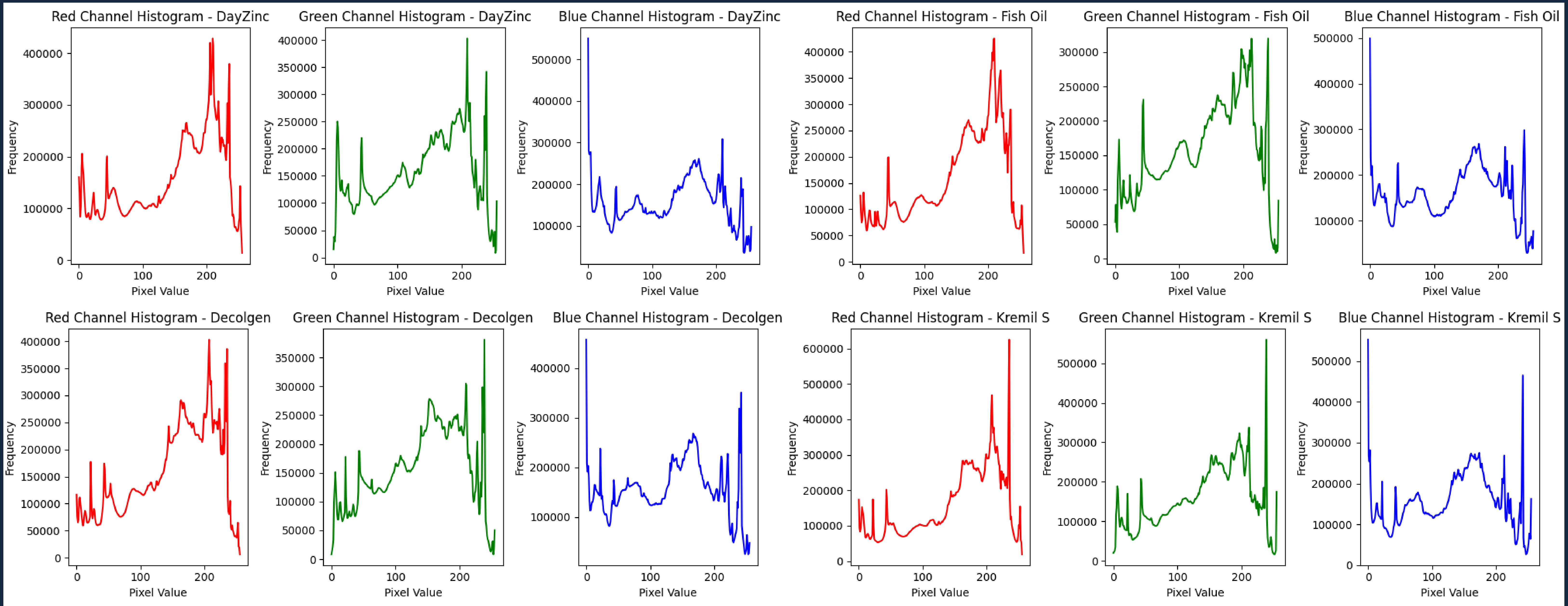
Medicol: 792

Neozep: 798

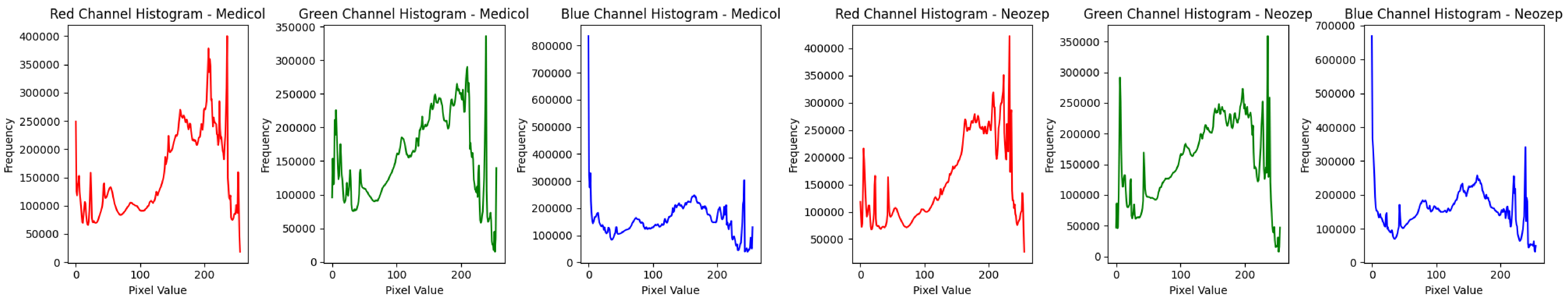
RGB Distribution Among Clases



RGB Distribution Among Clases



RGB Distribution Among Clases



All images are generally bright, with RGB histograms peaking near 255. However, the most valuable insight lies in the blue channel:

- Alaxan, Bactidol, Bioflu: High blue peaks → cooler tone.
- Biogesic, Fish Oil, Kremil S: Lower blue peaks → warmer tone.
- DayZinc, Decolgen, Medicol, Neozep: Lowest blue peaks → warmest tone.

Also, since the RGB color patterns across the classes look very similar, the model might have a hard time telling the classes apart just by looking at pixel colors.

Preprocessing on Train Data

Augmentation helps the model avoid memorizing images.

1. Standardization
2. Random Flip (left-right)
3. Random Zoom (crop)
4. Resize to 224x224x3
5. Random Contrast
6. Random Brightness

Preprocessing on Val & Test Data

No Augmentation is applied because they're used to measure true model performance.

We only **Standardized** to match the training data scale

Modeling

For Modeling we use 2 model such as EfficientNetB0 and ResNet50V2

Structure

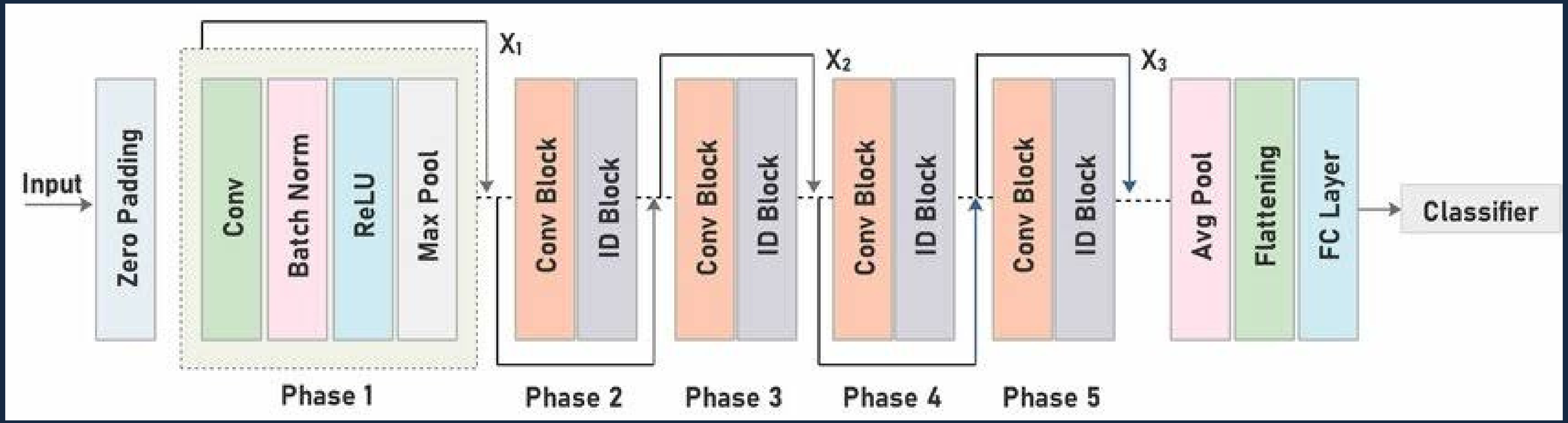
- GlobalAveragePooling2D
- Dense layers (512 neurons, ReLU)
- Dropout 0.3
- Ouput layers (10 neurons, softmax for classification)

Training

- 100 epochs
- Adam optimizer
- Learning rate 1e-4
- Early stopping with patience 10
- Loss using sparse Categorical Cross-Entropy

ResNet50V2 Base Model

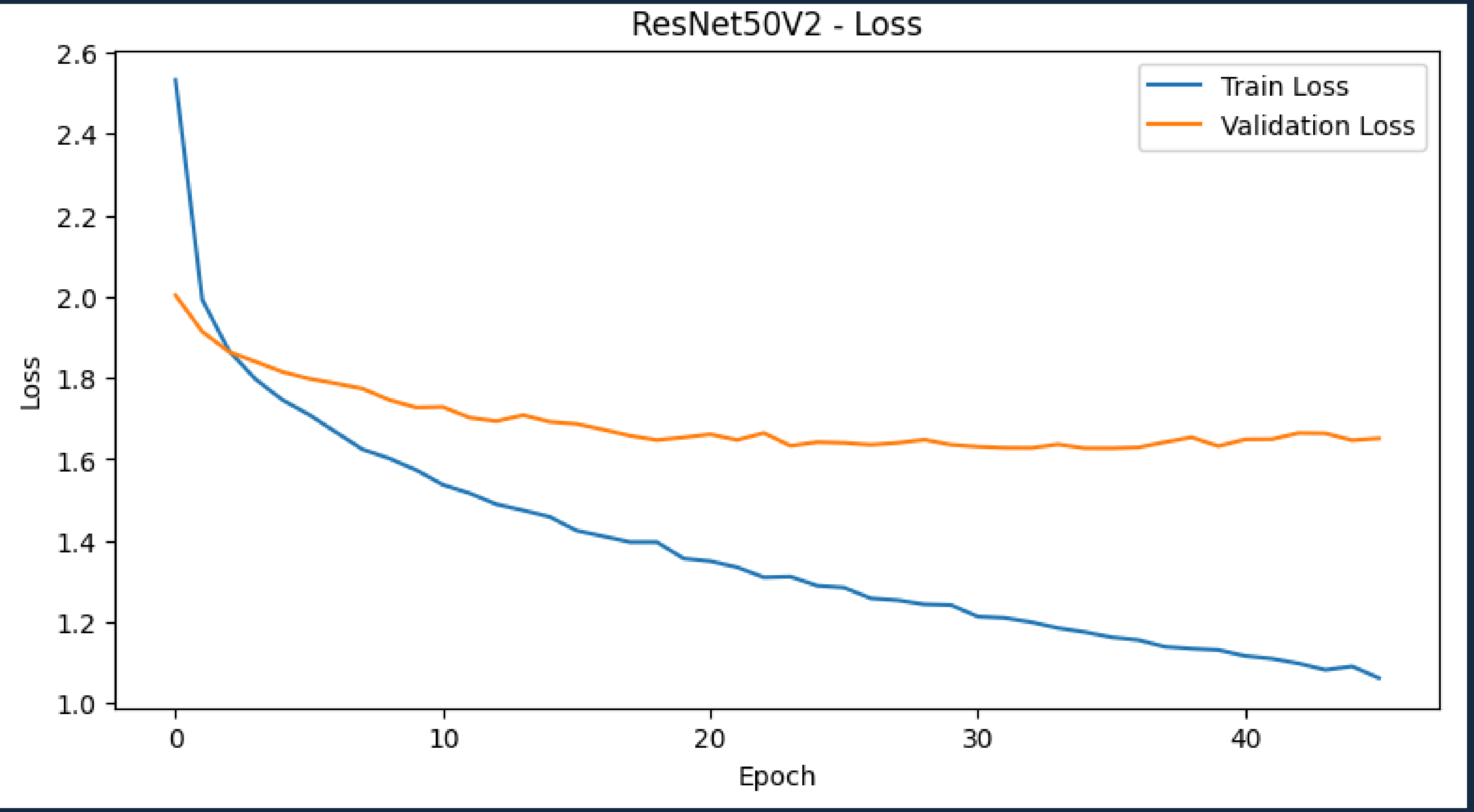
Uses **ResNet50V2 with all layers frozen**, serving only as a fixed feature extractor **without updating weights** during training.



Evaluation Metrics

Test Loss: 1.6451

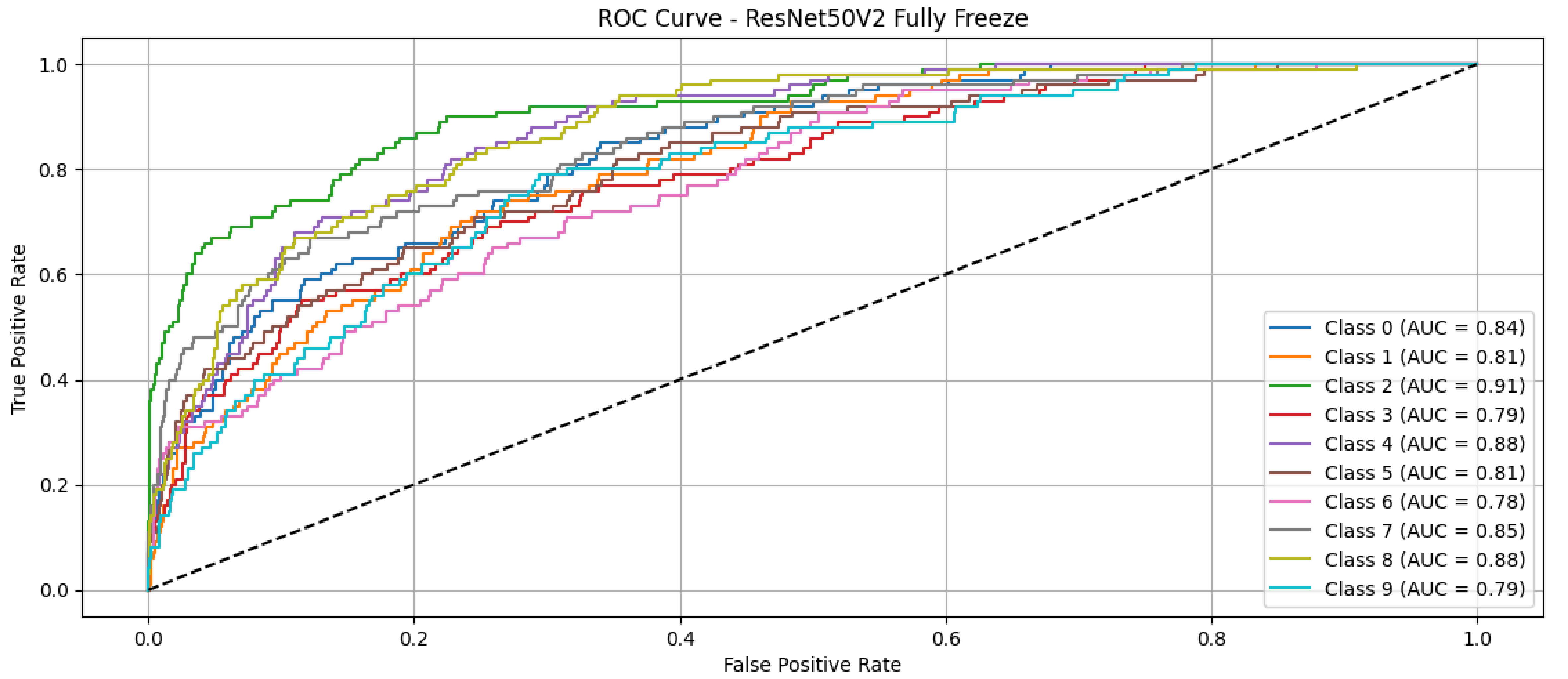
Test Accuracy: 0.4390

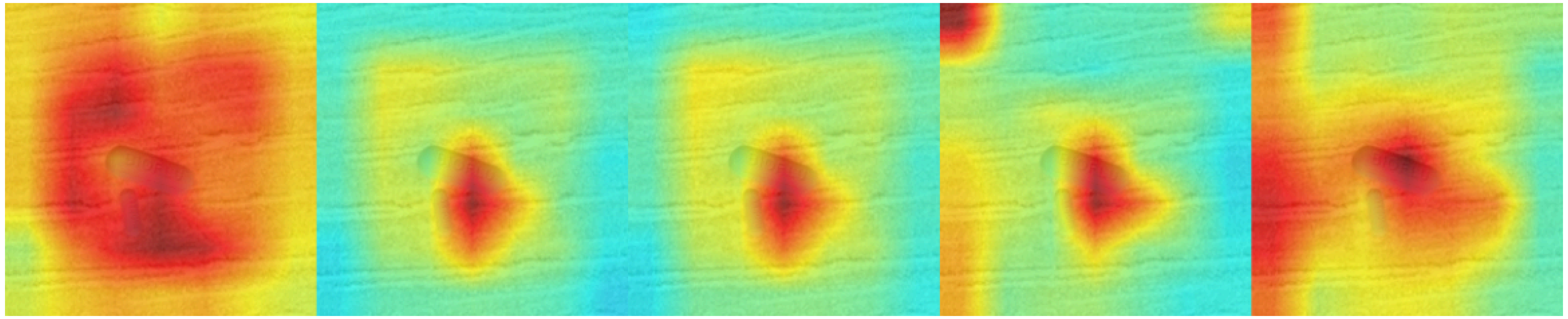


Classification Report ResNet50 V2:

	precision	recall	f1-score	support
0	0.36	0.45	0.40	100
1	0.43	0.33	0.38	100
2	0.54	0.62	0.58	100
3	0.46	0.37	0.41	100
4	0.39	0.55	0.46	100
5	0.47	0.40	0.43	100
6	0.52	0.33	0.40	100
7	0.55	0.46	0.50	100
8	0.41	0.55	0.47	100
9	0.34	0.33	0.34	100
accuracy			0.44	1000
macro avg	0.45	0.44	0.44	1000
weighted avg	0.45	0.44	0.44	1000

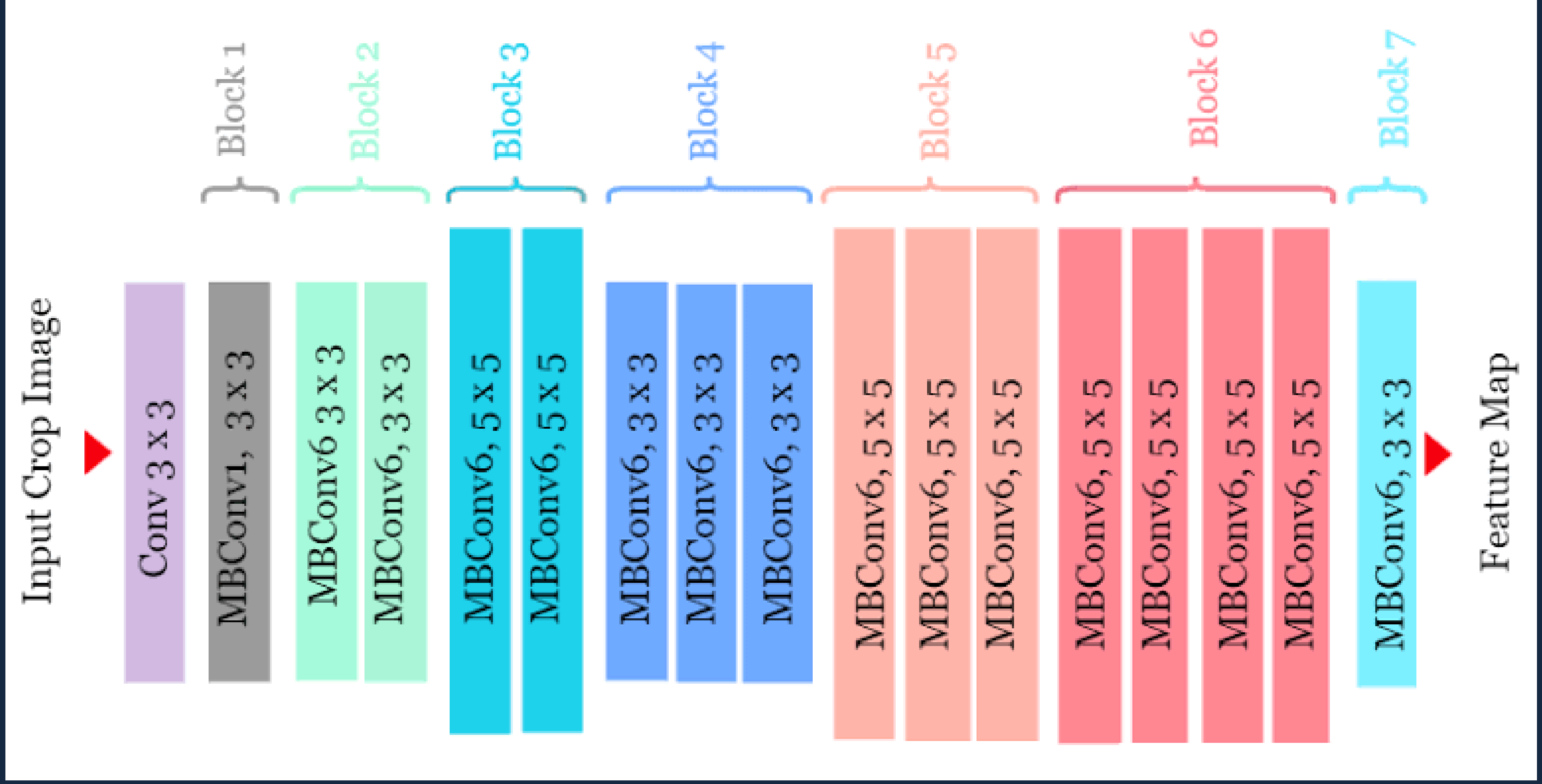
Evaluation Metrics





EfficientNetB0 Base Model

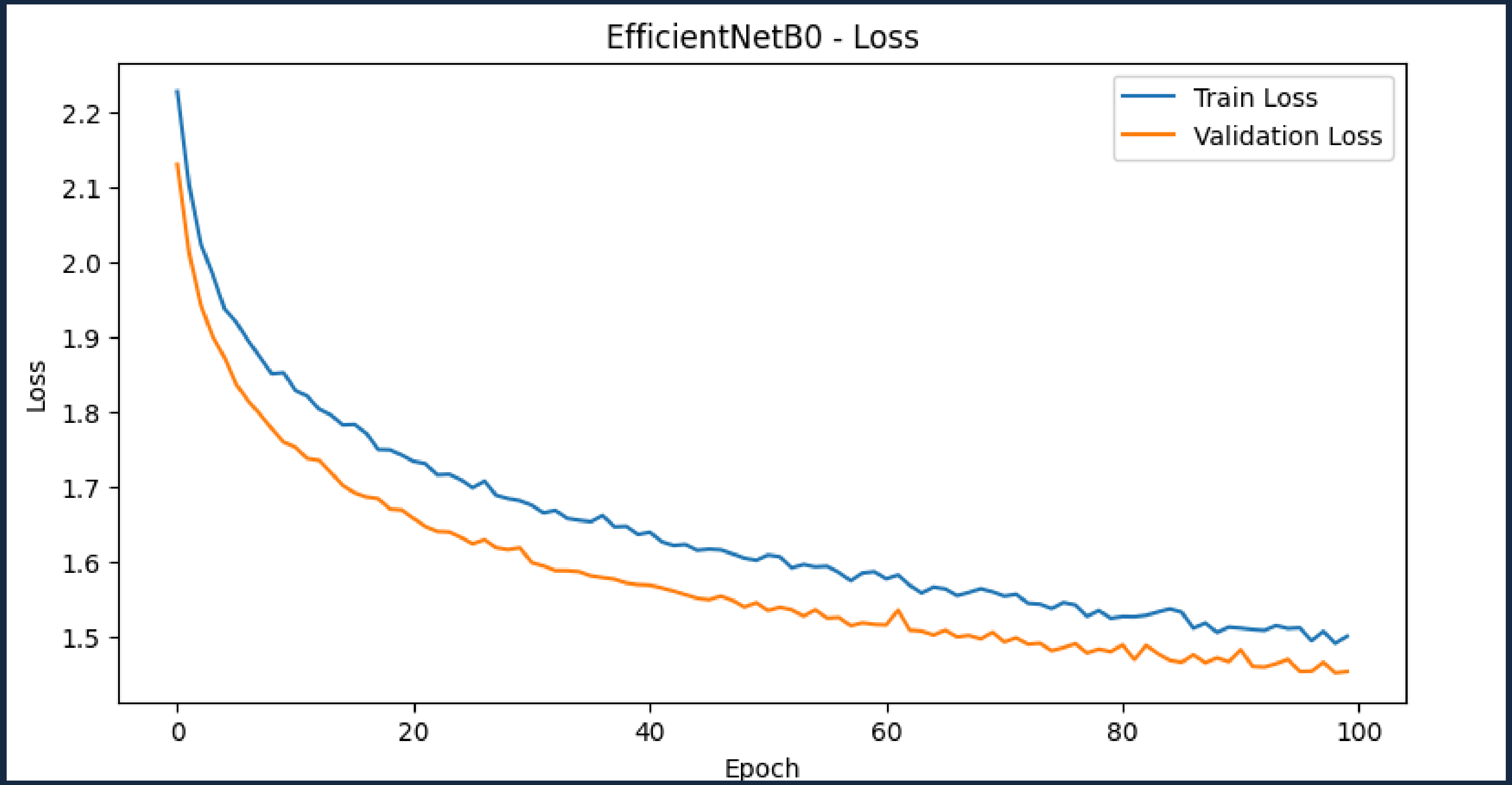
EfficientNet is pretrained with ImageNet weights. Setting trainable = False freezes all layers, so they act only as feature extractors without updating during training.



Evaluation Metrics

Test Loss: 1.4543

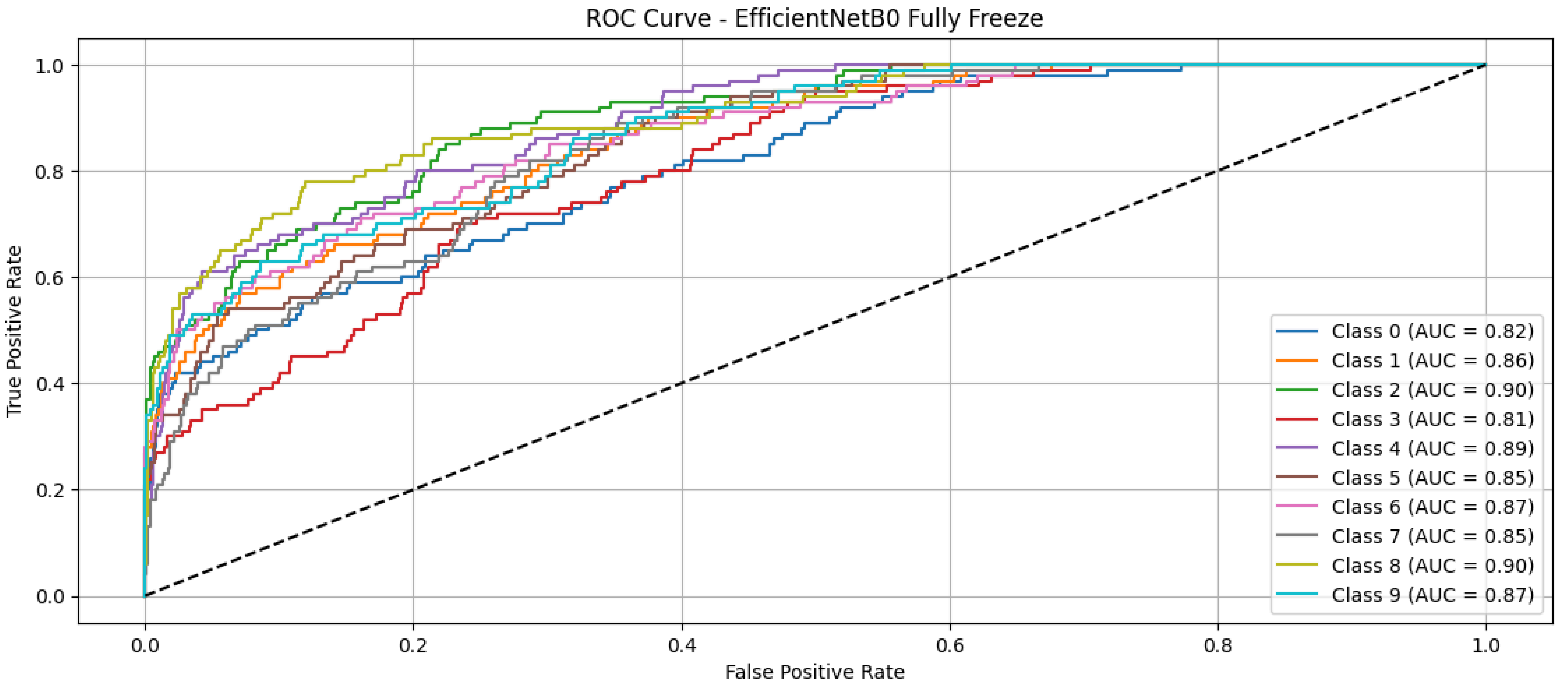
Test Accuracy: 0.4910

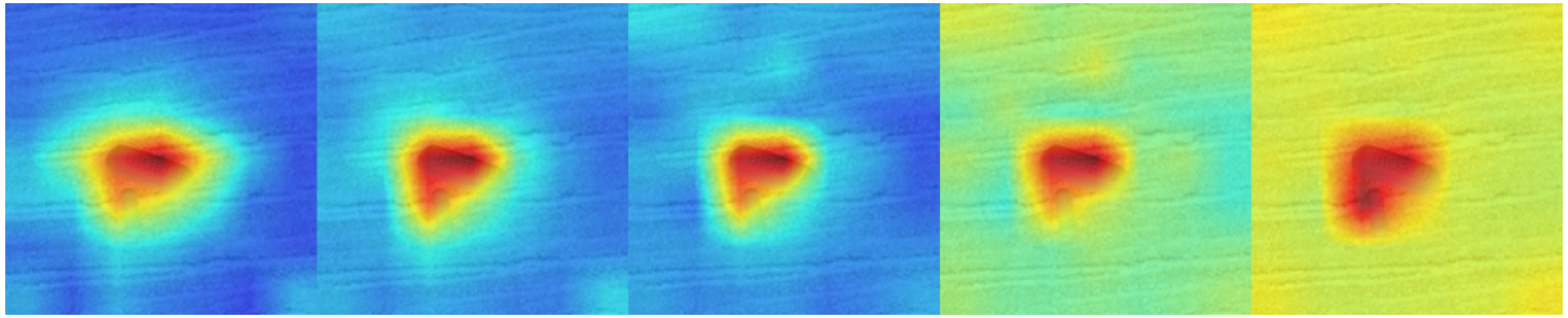


Classification Report EfficientNetB0:

	precision	recall	f1-score	support
0	0.63	0.43	0.51	100
1	0.59	0.41	0.49	100
2	0.48	0.62	0.54	100
3	0.46	0.33	0.38	100
4	0.53	0.54	0.53	100
5	0.28	0.48	0.35	100
6	0.62	0.49	0.55	100
7	0.40	0.43	0.42	100
8	0.57	0.64	0.60	100
9	0.61	0.54	0.57	100
accuracy			0.49	1000
macro avg	0.52	0.49	0.49	1000
weighted avg	0.52	0.49	0.49	1000

Evaluation Metrics





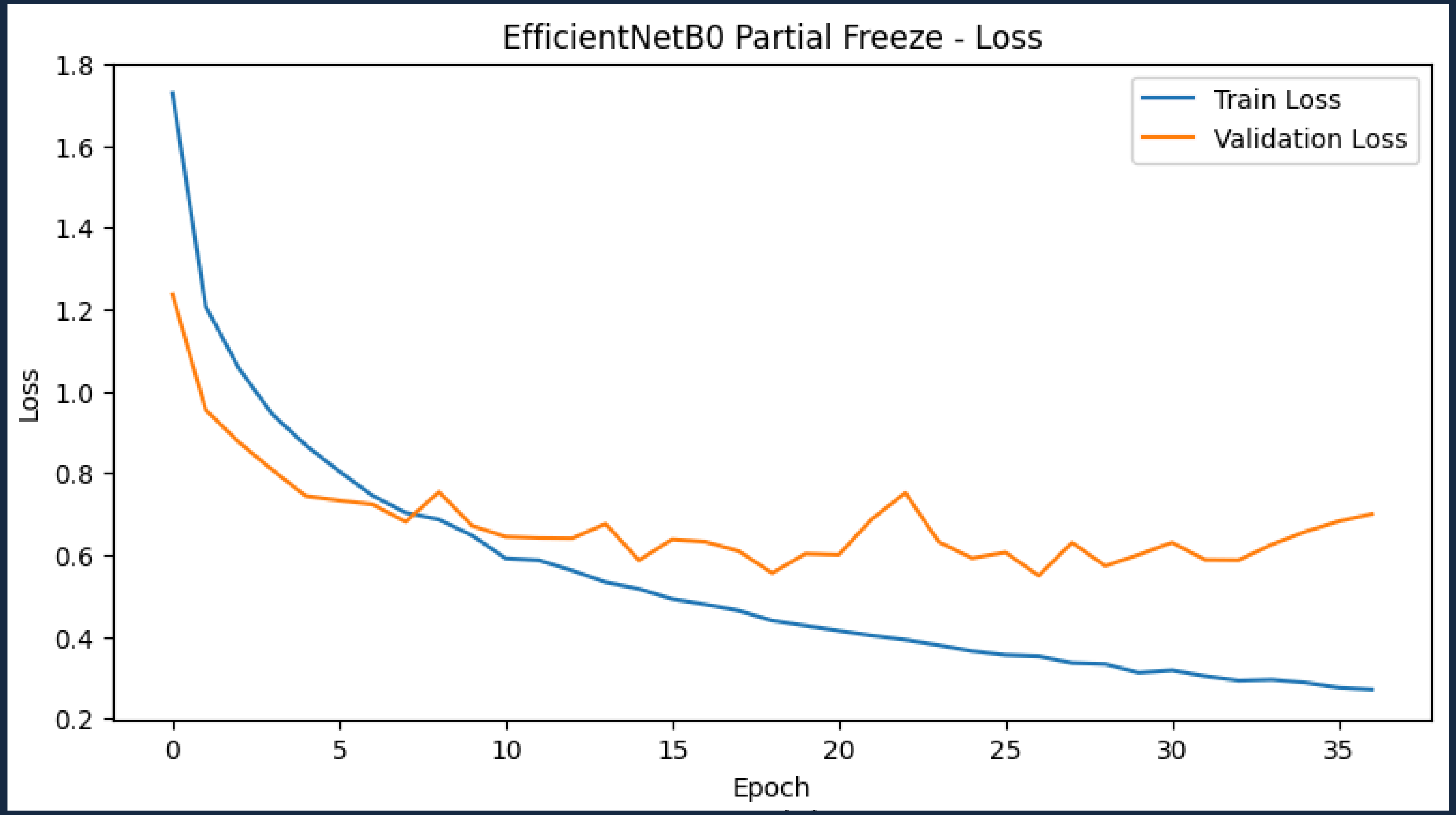
EfficientNetB0 Modified Model

Uses EfficientNetB0 but unfreezes Block6 and Block7 layers to fine-tune high-level features, aiming to improve performance by allowing partial weight updates during training.

Evaluation Metrics

Test Loss: 0.6390

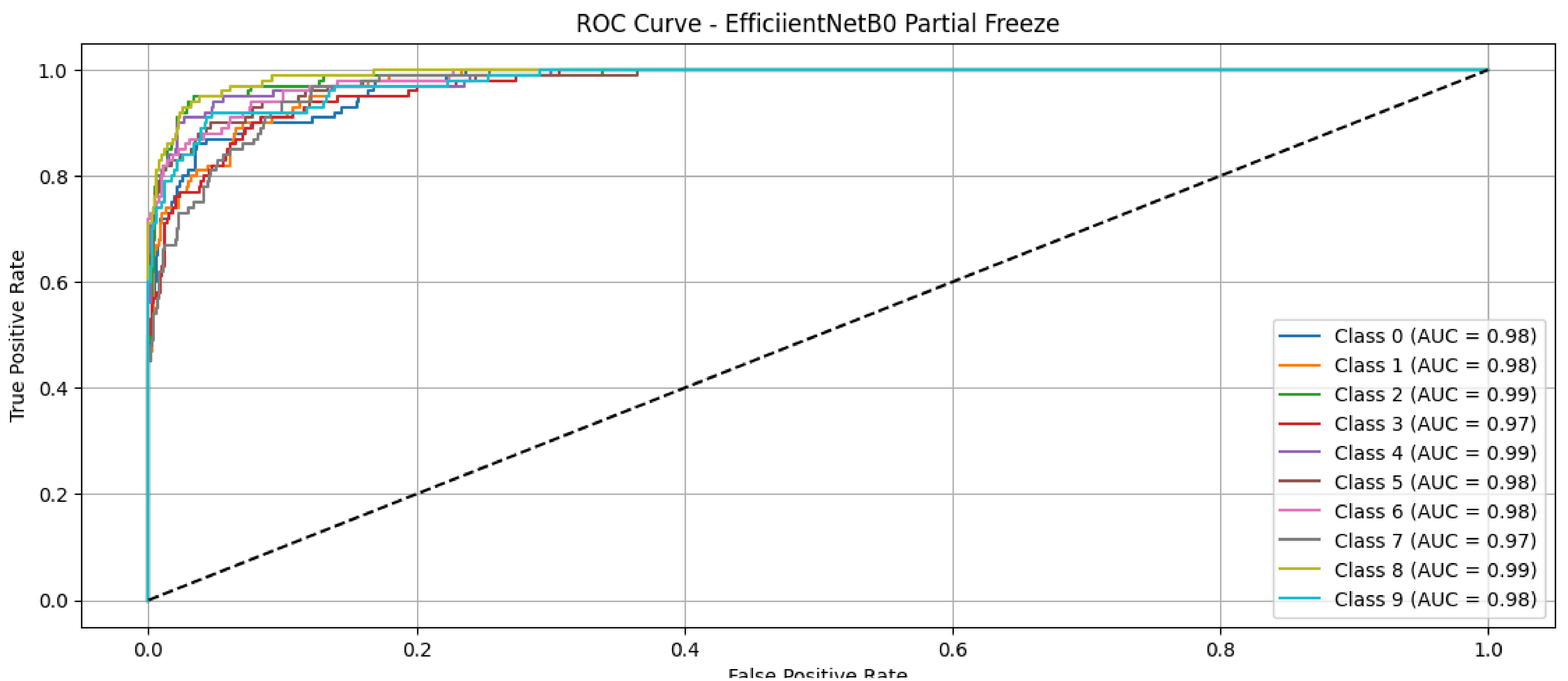
Test Accuracy: 0.7970

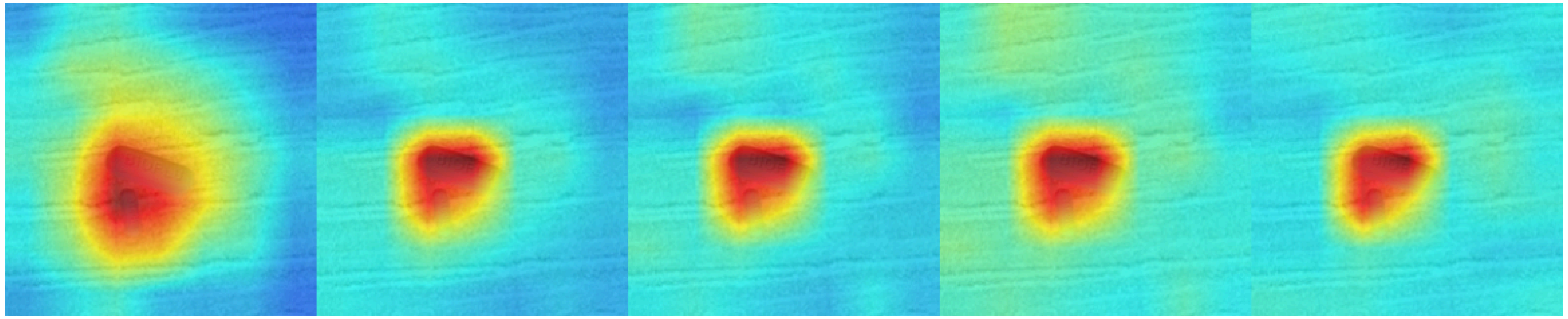


Classification Report EfficientNetB0 Partial Freeze:

	precision	recall	f1-score	support
0	0.78	0.77	0.77	100
1	0.72	0.78	0.75	100
2	0.86	0.85	0.85	100
3	0.71	0.79	0.75	100
4	0.86	0.83	0.85	100
5	0.86	0.79	0.82	100
6	0.93	0.74	0.82	100
7	0.67	0.77	0.72	100
8	0.85	0.86	0.86	100
9	0.81	0.79	0.80	100
accuracy			0.80	1000
macro avg	0.80	0.80	0.80	1000
weighted avg	0.80	0.80	0.80	1000

Evaluation Metrics





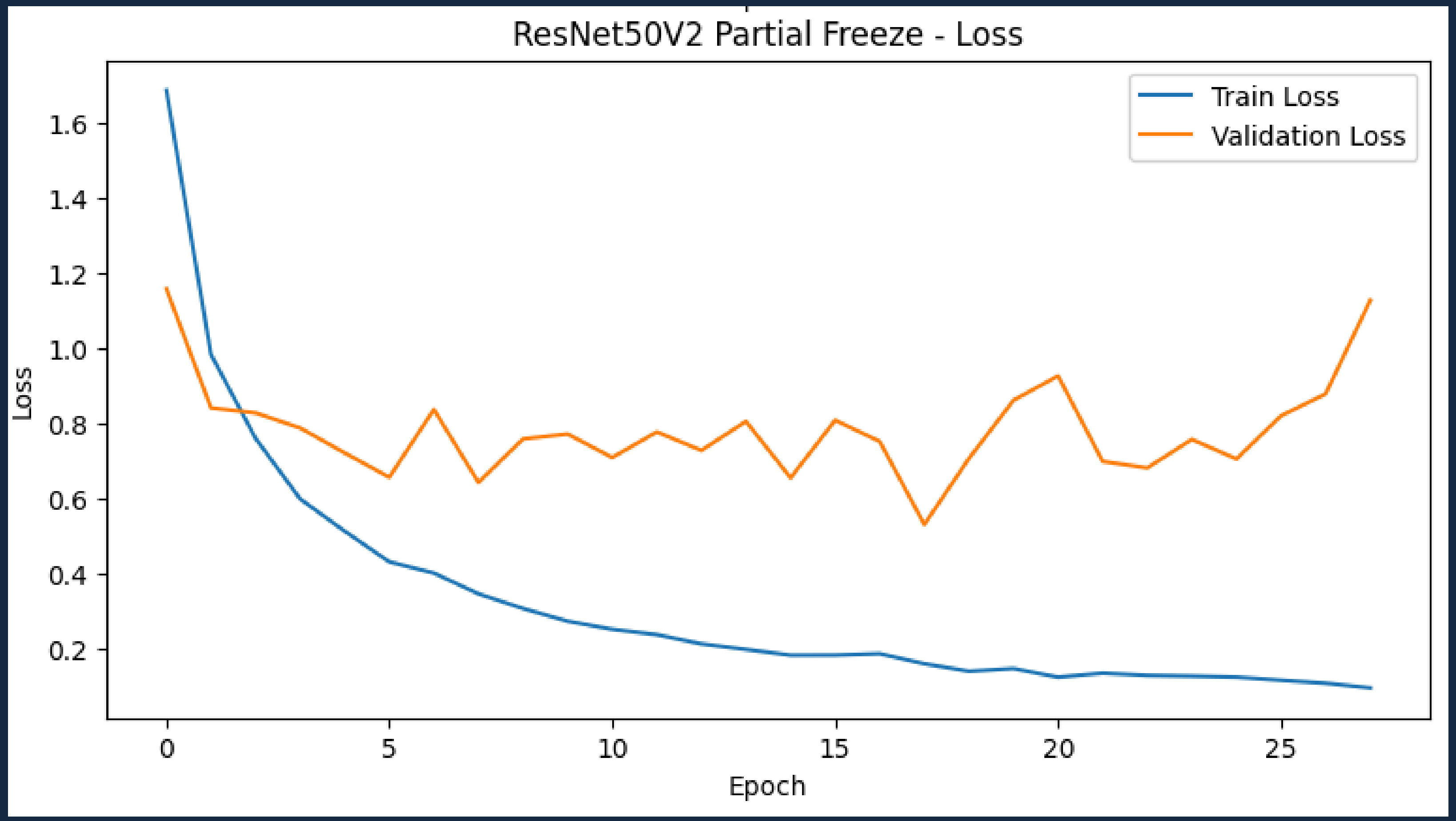
ResNet50V2 Modified Model

Uses ResNet50V2 but unfreezes Conv4 and Conv5 layers to fine-tune high-level features, aiming to improve performance by allowing partial weight updates during training.

Evaluation Metrics

Test Loss: 0.5835

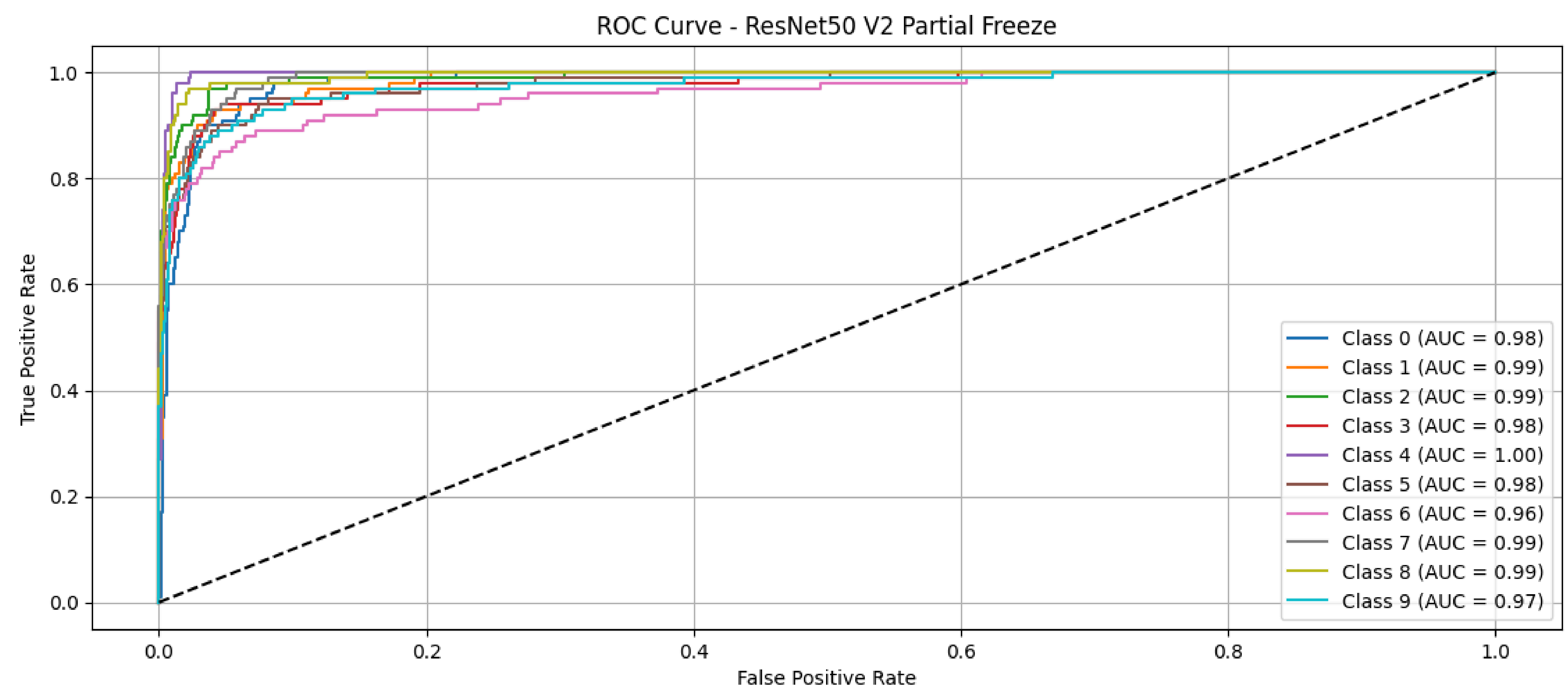
Test Accuracy: 0.8370

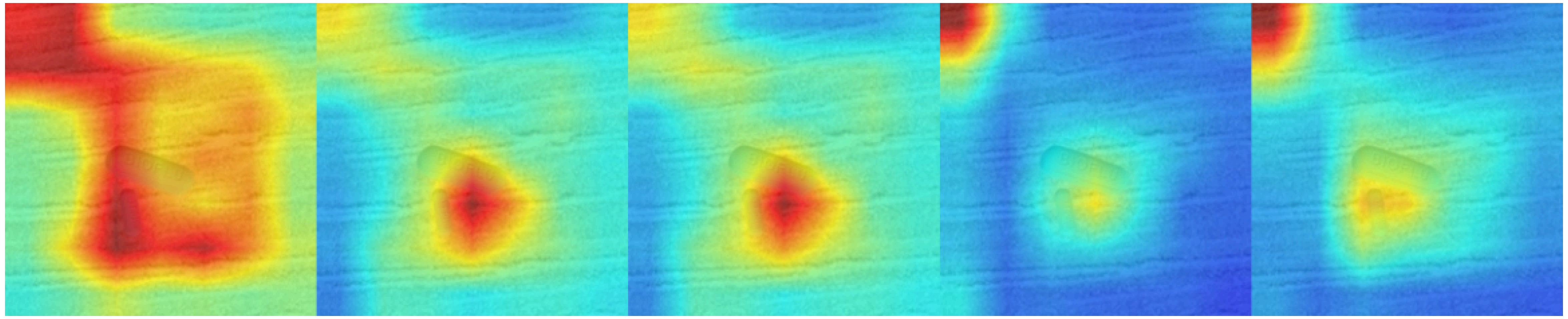


Classification Report ResNet50 V2 Partial Freeze:

	precision	recall	f1-score	support
0	0.78	0.86	0.82	100
1	0.84	0.83	0.83	100
2	0.89	0.83	0.86	100
3	0.76	0.87	0.81	100
4	0.95	0.89	0.92	100
5	0.80	0.83	0.81	100
6	0.88	0.74	0.80	100
7	0.84	0.78	0.81	100
8	0.85	0.93	0.89	100
9	0.81	0.81	0.81	100
accuracy			0.84	1000
macro avg	0.84	0.84	0.84	1000
weighted avg	0.84	0.84	0.84	1000

Evaluation Metrics





Conclusion

Partially unfrozen ResNet50V2 achieved the highest accuracy (84%) and strongest generalization to validation data.

Unfreezing selective layers improves the model’s ability to learn from domain-specific visual features.

Attention behavior confirmed model focus on primary drug objects, validating model's learning path.

Existing Challenges

Visual Ambiguity
Many drugs have similar visual appearances, which often leads to misclassification (example : Bioflu vs Neozep)

Limited Data Scope
The model currently recognizes only 10 drug types, limiting its practical deployment in broader use cases.

Noisy Real Data
Training data consists of synthetic images, limiting the model's robustness in real-world settings with complex backgrounds.

In some models (example : fully frozen ResNet), attention maps revealed focus outside the main object, reducing prediction precision.

Future Opportunities

Multimodal Learning

Combine image input with textual metadata (ex : brand name, dosage) for better disambiguation of lookalike drugs.

User-Centered Design

Enhance model output by displaying drug details (usage, dosage, warnings) for user confirmation and safety.

Commercial Integration

Partner with pharmacies or health tech platforms to enable real-time drug identification, with potential monetization models.

Dataset Expansion

Gradually include more diverse, real-world medication images across brands, forms (capsule, tablet), and lighting conditions.

Thankyou!!

Poster

Notebook

i asked my model if it was feeling okay.
it said…

i asked my model if it was feeling okay.
it said...

my val loss is HIGH and i'm having
an ✨existensial crisis✨

i asked my model if it was feeling okay.
it said...

my val loss is HIGH and i'm having
an ✨existensial crisis✨

whether i'm just a glorified machine
to classify 🐟fish oil🐟