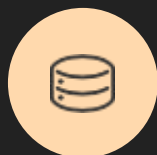# Medication Packaging Identification: Leveraging Deep Learning for Accurate and Reliable Medication Recognition

Leveraging deep learning to accurately classify medications from images of drug boxes and bottles
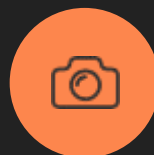
Data Collection and Curation

Custom data acquisition pipeline
Developed a dedicated Python-based scraper to interface with the DailyMed (NIH) REST API, an authoritative source for drug labeling and media

Diverse dataset
Images reflect real-world variations in packaging design, lighting conditions, and image quality
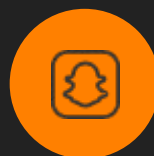
Targeted 20 commonly prescribed medications
Retrieved thousands of raw images associated with specific National Drug Codes (NDCs) and Structured Product Labeling (SPL) records

Heuristic image filtering
Implemented a computer vision-based engine using OpenCV to automatically discard non-informative samples, such as grayscale images, line drawings, barcodes, and document scans

The custom data curation pipeline enabled the collection of a diverse and high-quality dataset of medication packaging images, which is essential for training accurate and reliable deep learning models.

Model Architectures

ResNet-18 (Baseline CNN)

Lightweight model with 11.2M parameters, strong inductive bias for local texture and edge features.

ResNet-50 (Deeper CNN)

Deeper model with 23.5M parameters, allows learning more complex visual hierarchies.

ViT-B/16 (Vision Transformer)

85.8M parameters, tests global attention-based representations and long-range spatial relationships.

ViT-S/16 (Smaller Transformer)

21.7M parameters, investigates whether transformers can generalize well on a medium-sized, noisy real-world dataset.

Training and Evaluation

- **Loss Function**

  Categorical Cross-Entropy Loss

- **Optimizer**

  AdamW with learning rate 1e-4 and weight decay 1e-4

- **Learning Rate Scheduler**

  ReduceLROnPlateau, with factor 0.5 and patience 3 epochs

- **Training Configuration**

  15 epochs, batch size 32, on CPU/GPU/Apple MPS

- **Data Preprocessing and Augmentation**

  Input Normalization: Mean [0.485, 0.456, 0.406], Std [0.229, 0.224, 0.225]. Training Augmentations: Resize, Random Resized Crop, Random Horizontal Flip, Color Jitter. Evaluation Preprocessing: Resize, Center Crop, Normalize

- **Evaluation Metrics**

  Primary Metric: Top-1 Accuracy (Train/Validation/Test). Per-Class Metrics: Precision, Recall, F1-Score. Confusion Matrices to analyze systematic misclassifications

The automated medication packaging identification system leverages deep learning to accurately classify medicines, addressing the critical problem of look-alike, sound-alike (LASA) medication errors. By empowering patients, caregivers, and healthcare professionals with reliable medication identification, this system can improve patient safety and reduce medication-related incidents.
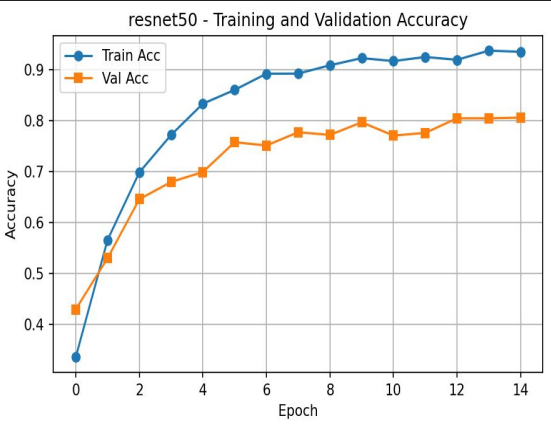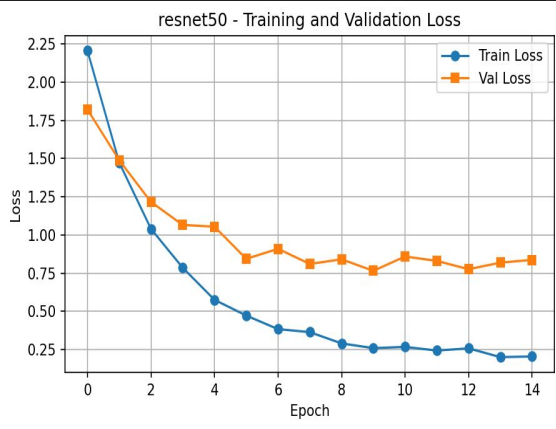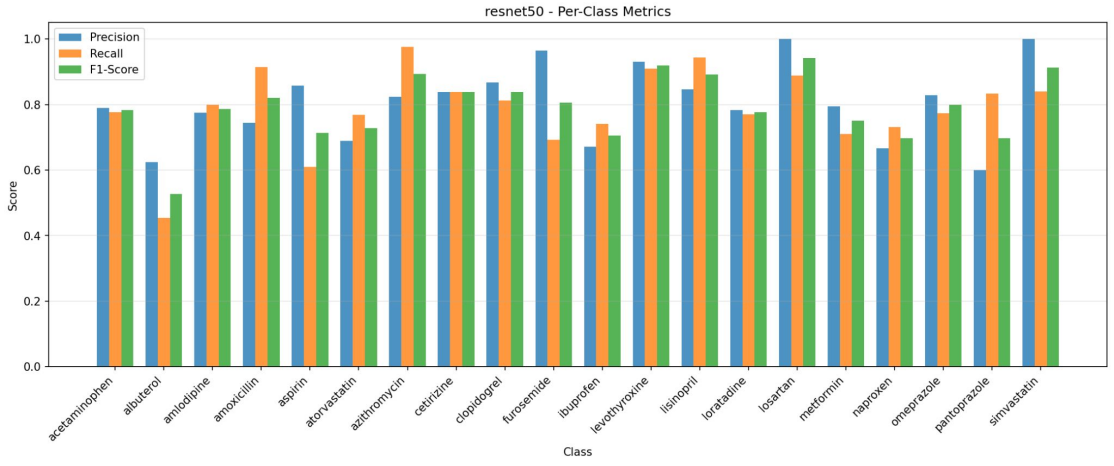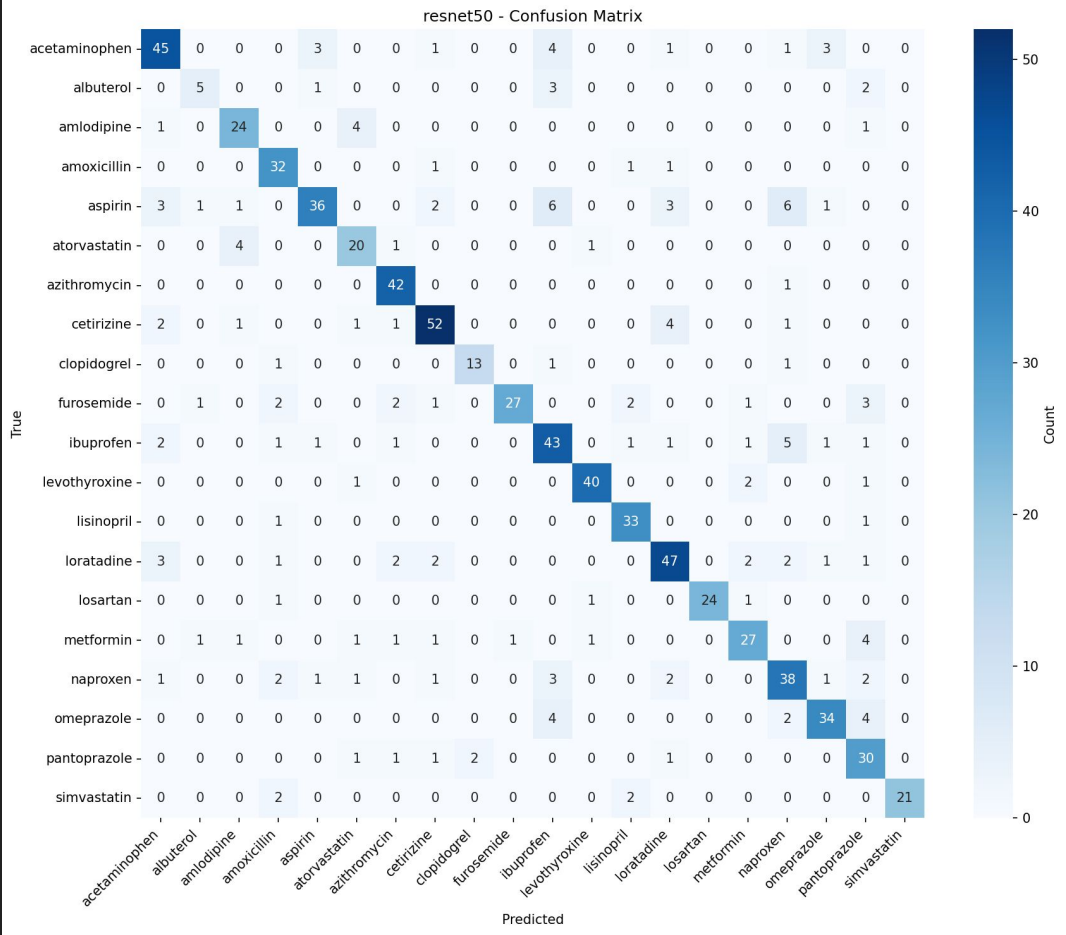
Results

- CNN-based models (ResNet-18, ResNet-50) outperformed ViTs on this dataset, despite lower parameter counts.
- This suggests:
  - Strong local texture cues dominate medication packaging recognition
  - Transformers require larger datasets to fully leverage global attention
- ResNet-50 achieved the best overall test accuracy (~79%), demonstrating a clear depth-performance trade-off.
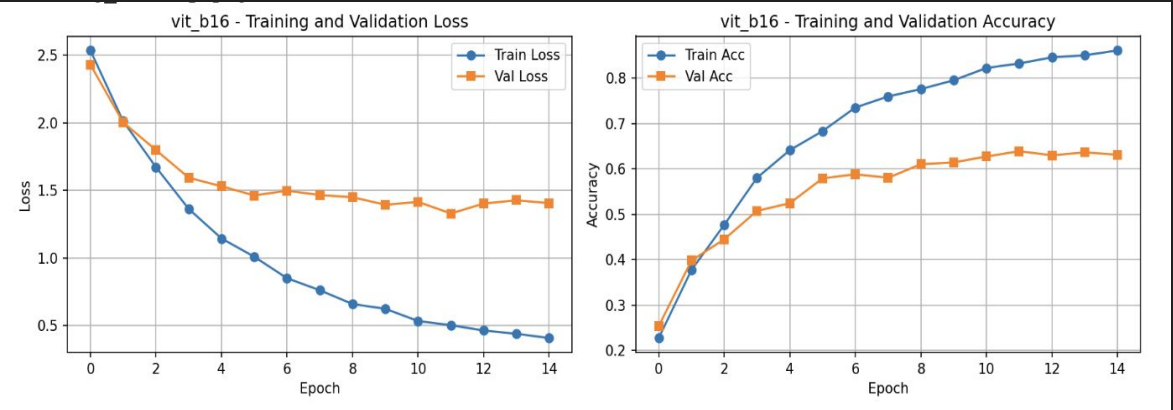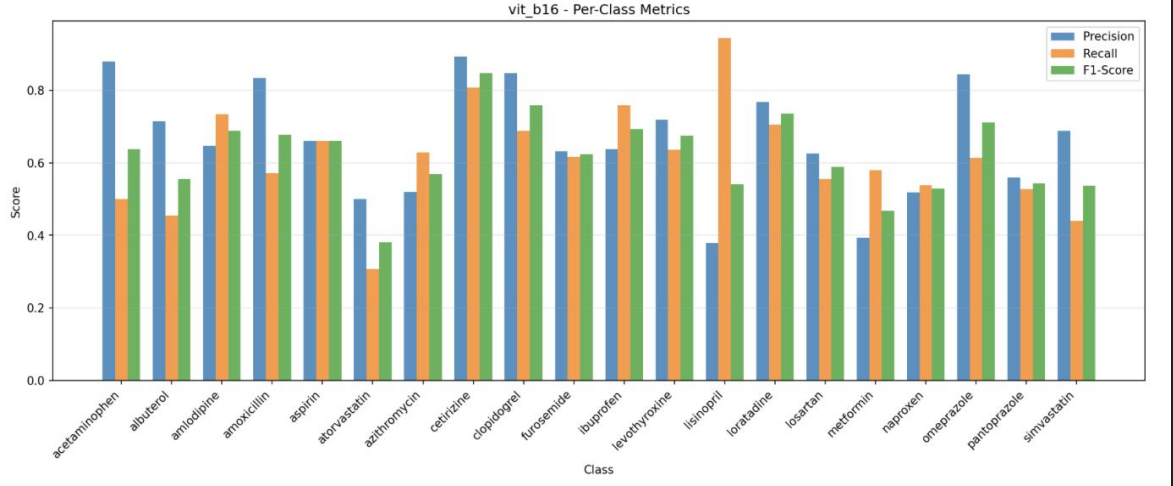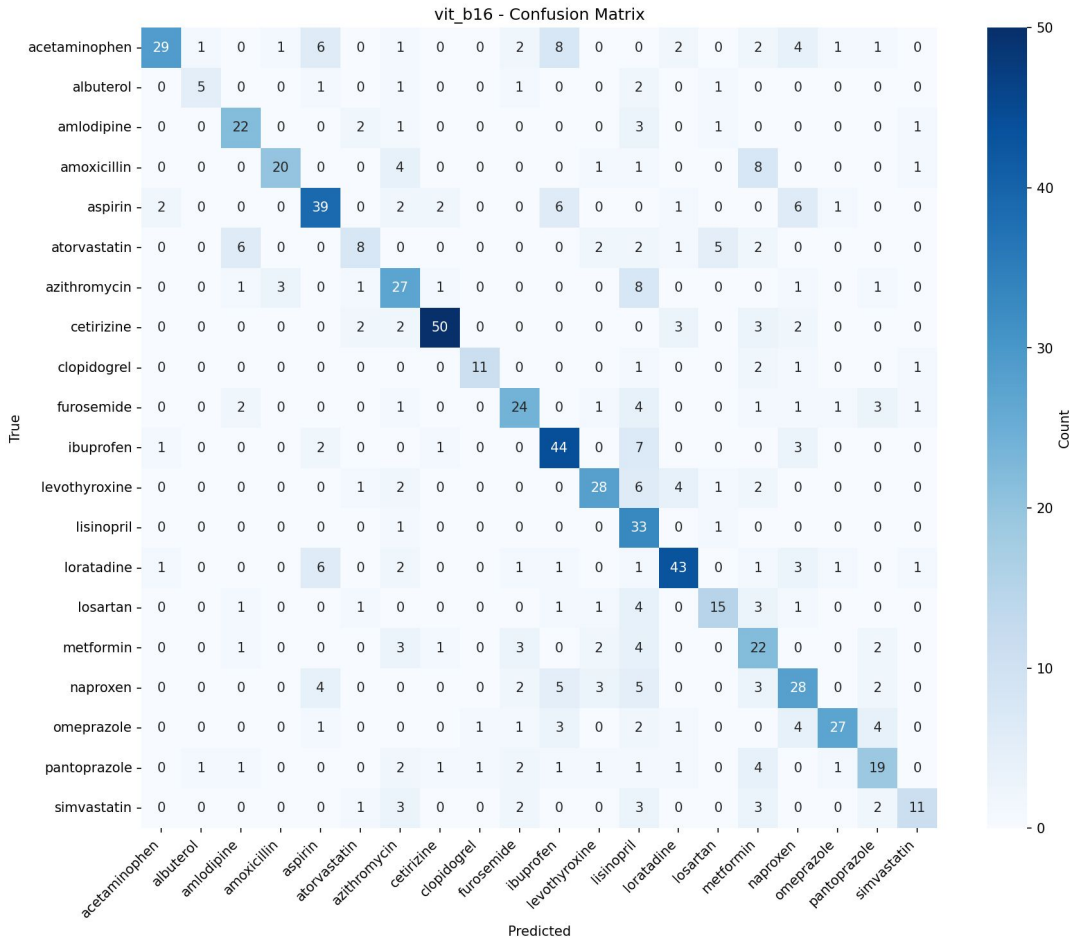
Results - Resnet 18
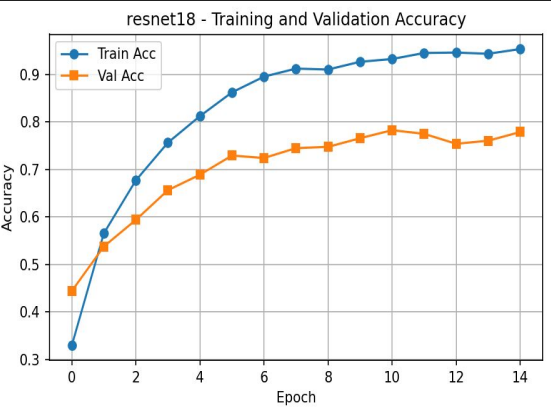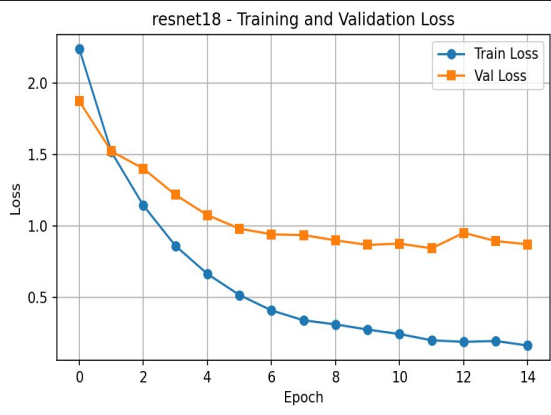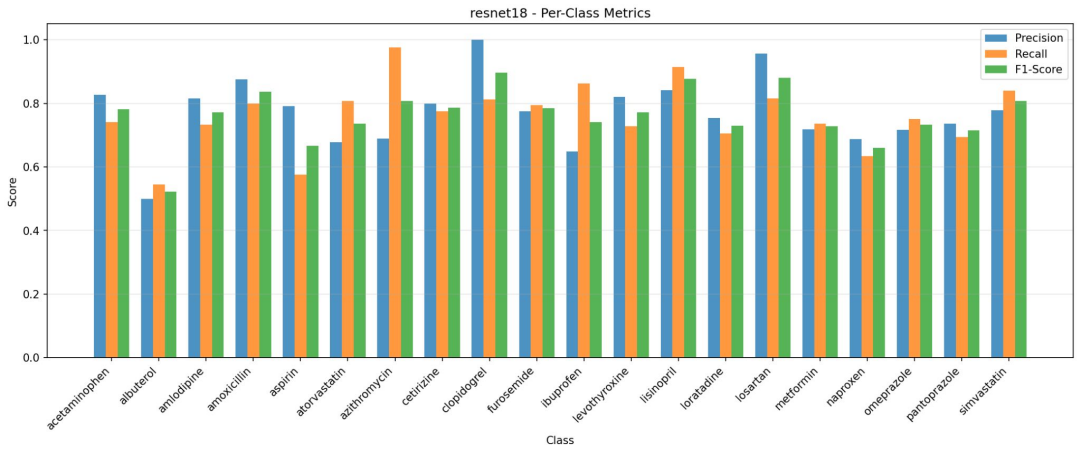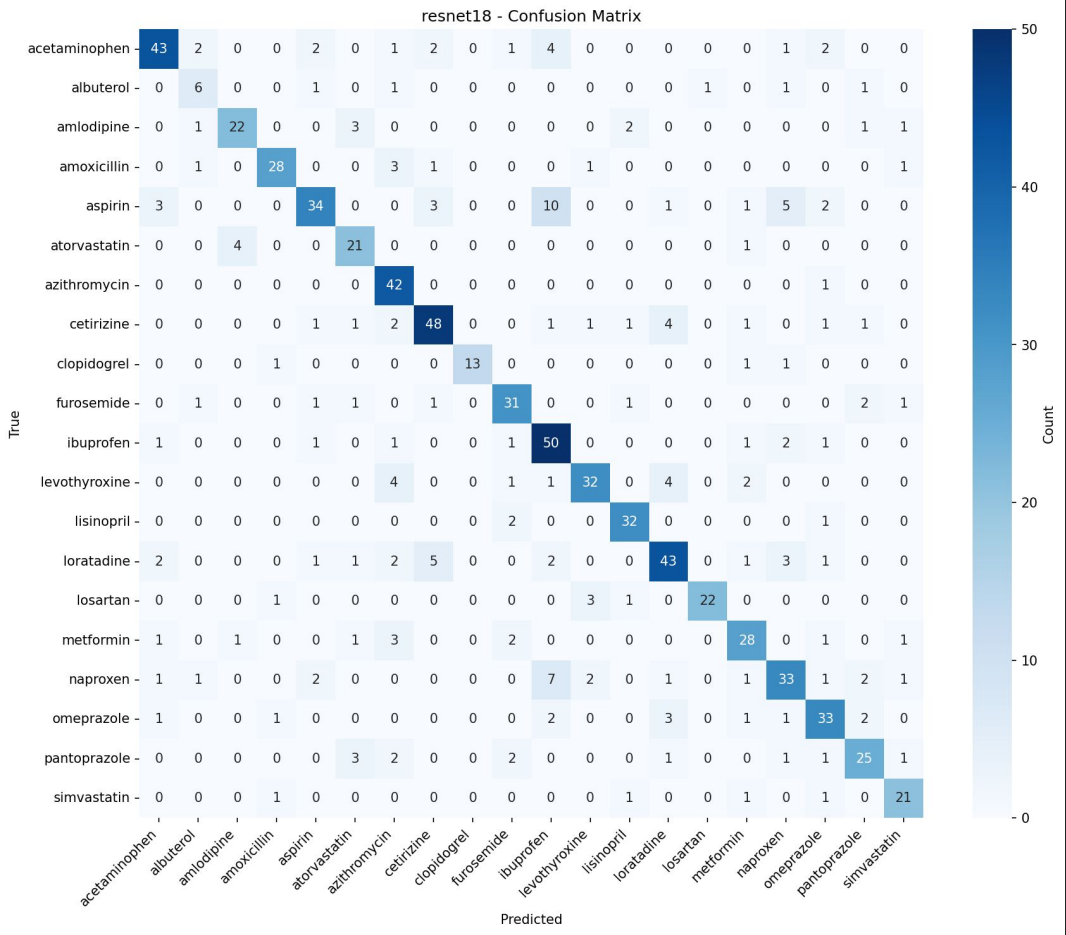
Results - Resnet 50

Results - VIT-B16

Results - ViT-S/16