Report

07_densenet121_384_final_best. ipynb

DenseNet121 Osteoarthritis Severity Classification

---

1. Objective of This Notebook

The purpose of this notebook is to implement and optimize a transfer learning pipeline using DenseNet121 for automatic Knee Osteoarthritis (KOA) grading (KL 0–4).

This notebook contains two full experimental configurations:

1. **DenseNet121 @320 (Baseline with advanced tricks)**
2. **DenseNet121 @384 (Final Best Optimized Model)**

The goal was to systematically improve performance while keeping the architecture fixed (DenseNet121).

---

2. Dataset Handling (Common to Both Codes)

Dataset:

Knee Osteoarthritis dataset with KL grades (0–4)

Automatic Drive Loading

- Google Drive mounted
- ZIP file automatically unzipped if not already extracted
- Auto-detection of train, val, and test folders

Dataset Sizes:

- Train: 5,778 images
- Validation: 826 images
- Test: 1,656 images

Class Distribution:

Highly imbalanced (KL4 smallest class)

Imbalance Handling:

- **WeightedRandomSampler** used during training

---

3. Code Block 1

DenseNet121 @320 – Advanced Training Tricks Baseline

3.1 Configuration

- Image size: 320 × 320
- Batch size: 16
- Epochs: 15
- Optimizer: AdamW
- Learning rate: 1e-4
- Weight decay: 1e-4
- Cosine LR scheduler with warm-up
- Automatic Mixed Precision (AMP)
- Exponential Moving Average (EMA)
- MixUp (constant)
- No label smoothing

3.2 Data Augmentation

- Resize
- Random Horizontal Flip
- Small Rotation (±7°)
- ImageNet normalization

3.3 Model Architecture

- DenseNet121 pretrained on ImageNet
- Final classifier replaced with 5-class linear layer

3.4 Results – 320 Model

Test Performance:

- Test Accuracy: **67.39%**
- Test Macro-F1: **0.6914**

Observations:

- Strong performance on KL0, KL3, KL4
- KL1 most confused class
- Errors mostly between adjacent grades

This model serves as a strong baseline.

---

4. Code Block 2

DenseNet121 @384 – Final Best Configuration

This is the improved configuration over the 320 baseline.

4.1 Updated Configuration

- Image size: **384 × 384**
- Batch size: 8 (due to larger resolution)
- Epochs: 18
- AdamW optimizer
- Cosine LR + warm-up
- AMP enabled
- EMA enabled

Additional Improvements:

1. **Label Smoothing = 0.07**
2. **Scheduled MixUp**
    - Enabled during first half of training
    - Disabled during second half

4.2 Why These Improvements?

Increased Resolution (384)

- Preserves fine joint space details
- Improves subtle osteophyte detection

Label Smoothing

- Reduces overconfidence
- Helps KL1 borderline cases

Scheduled MixUp

- Early regularization
- Later sharp convergence

EMA

- Stabilizes evaluation
- Reduces noisy weight fluctuations

4.3 Results – 384 Final Model

Final Test Performance:

- Test Accuracy: **70.89%**
- Test Macro-F1: **0.7085**

---

5. Confusion Matrix Analysis (384 Model)

Strong Performance:

- KL0 (Normal)
- KL3 (Moderate)
- KL4 (Severe)

Most Challenging:

- KL1 (Mild OA)

Most misclassifications occur between:

- KL0 ↔ KL1
- KL1 ↔ KL2

This pattern is clinically expected due to subtle radiographic differences in early OA.

---

6. Comparison Between the Two Codes

Feature	DenseNet121 @320	DenseNet121 @384
Image Size	320	384
Epochs	15	18
Label Smoothing	No	Yes (0.07)
MixUp	Always On	Scheduled
EMA	Yes	Yes
Test Accuracy	67.39%	**70.89%**
Test Macro-F1	0.6914	**0.7085**

Improvement:

+3.5% absolute accuracy gain
+1.7% macro-F1 improvement

---

7. Final Conclusion of This Notebook

The notebook demonstrates a structured optimization process:

1. Baseline with strong training tricks (320 resolution)
2. Systematic improvement via:
    - Higher resolution
    - Label smoothing
    - MixUp scheduling
    - EMA stabilization

The final DenseNet121 @384 configuration achieved:

70.89% test accuracy
0.7085 macro-F1

This represents the best-performing single DenseNet model implemented in this project.