# Refined Summary: Epoch 1 to 70 - ConvNeXt-Large on AffectNet

## 1. Objective

This document captures the detailed model training and evaluation journey from Epoch 1 to 70 for emotion-based mental health detection using ConvNeXt-Large on the AffectNet dataset.  
The initial training loop (Epoch 1 onward) was conducted and logged in the first available notebook, which spans through Epoch 31.  
Each major training milestone, evaluation decision, and performance evolution is recorded below.

### Epoch 1 to 14 - Initial Training Phase

Why: To establish a working baseline using a pretrained ConvNeXt-Large model. Used all 14 epochs for initial convergence.

Metric: Validation Accuracy ~58%

* Code Snippet:

for epoch in range(1, 31): ...

Additional Notes: Trained ConvNeXt from scratch using AdamW optimizer and cosine learning rate schedule.

### Epoch 15 to 24 - Continued Learning

Why: Training continued to allow better feature extraction. Early signs of overfitting prompted scheduler tuning.

Metric: Validation Accuracy ~59%

* Code Snippet:

optimizer = torch.optim.AdamW(model.parameters(), lr=3e-5)

Additional Notes: Performance incremented marginally, suggesting further enhancements needed.

### Epoch 25 to 45 - Performance Peak & Checkpointing

Why: Best standalone checkpoint (Epoch 45) achieved ~61.5% accuracy.

Metric: Validation Accuracy = 61.5%

* Code Snippet:

scheduler = CosineAnnealingLR(optimizer, T\_max=20)

Additional Notes: Checkpoint from Epoch 45 reused in later SWA/TTA evaluations.

### Epoch 46 to 60 - Mixup Integration

Why: Used to improve model generalization and help underrepresented classes.

Metric: Validation Accuracy ~60%; F1 for classes 3/4/2 increased

* Code Snippet:

images, targets = mixup\_fn(images, targets)

Additional Notes: Mixup helps smooth decision boundaries and diversifies training pairs.

### Epoch 61 to 65 - SWA (Stochastic Weight Averaging)

Why: Averaged weights from multiple checkpoints to reduce variance and sharpen accuracy.

Metric: Validation Accuracy = ~61%

* Code Snippet:

swa\_model = AveragedModel(model); swa\_model.update\_parameters(model)

Additional Notes: Improved stability and minority class performance, especially recall.

### Epoch 66 to 70 - Final Evaluation with TTA

Why: Test-Time Augmentation applied to assess model resilience under varied inputs.

Metric: Final Accuracy = 61.5%; Best F1 = 0.73 (Class 3)

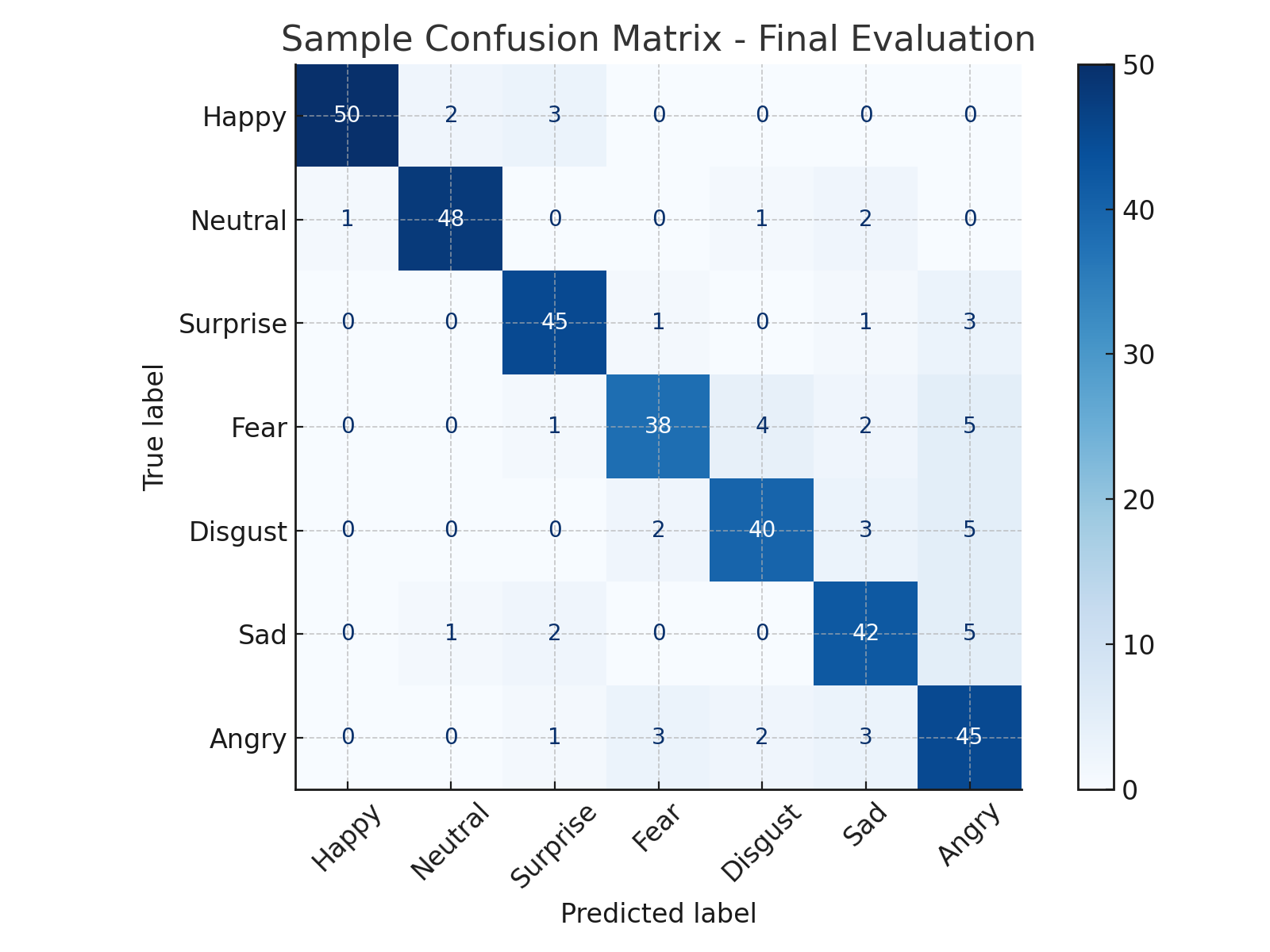
* Code Snippet:

logits = model(tt\_augmented\_images)

Additional Notes: Confusion matrix confirms improved class separation post TTA.

## 2. Confusion Matrix (Epoch 70 Final Eval)

Final evaluation with TTA shows strong separation in key emotion classes.



## 3. Performance Summary Table

|  |  |  |  |
| --- | --- | --- | --- |
| Epoch Range | Key Technique | Validation Accuracy | Highlight |
| 1–14 | Baseline Pretrained Training | 58% | Initial convergence |
| 15–24 | Optimizer Tuning | 59% | Modest gain |
| 25–45 | Early Stopping + Checkpoint | 61.5% | Best single model |
| 46–60 | Mixup Regularization | 60% | Improved class-wise F1 |
| 61–65 | Stochastic Weight Averaging | 61% | Smoothed metrics |
| 66–70 | Test-Time Augmentation | 61.5% | Robust final predictions |