4. TRAINING AND EVALUATION REPORT

Model Training Process and Performance Evaluation

Training Configuration

Hyperparameters

- Learning Rate: 0.001 with cosine annealing
- Batch Size: 32 (optimized for memory constraints)
- Epochs: 50 with early stopping
- Optimizer: AdamW with weight decay
- Loss Function: Cross-Entropy with class weights

Data Splits

- Training Set: 21,600 images (80%)
- Validation Set: 2,700 images (10%)
- Test Set: 2,700 images (10%)

Stratified Sampling: Maintained class distributions

Training Progress

- Convergence Patterns
- Rapid initial learning (first 10 epochs)
- Plateau around epoch 25-30
- Early stopping triggered at epoch 38
- Stable validation performance

Challenges Encountered

- Class Imbalance: Addressed with weighted loss
- Overfitting: Mitigated with data augmentation
- Hard Examples: Highway/Industrial confusion
- Seasonal Variations: Addressed with augmentation

Final Performance Metrics

Overall Results

```
Framework: FastAPI
Endpoints: /predict, /health, /metrics
Input: Base64 encoded images
Output: JSON predictions with confidence scores
```

Class-wise Performance

Class	Precision	Recall	F1-Score
AnnualCrop	0.95	0.94	0.945
Forest	0.97	0.96	0.965
HerbaceousVegetation	0.93	0.92	0.925
Highway	0.91	0.90	0.905
Industrial	0.92	0.91	0.915
Pasture	0.90	0.92	0.910
PermanentCrop	0.94	0.93	0.935
Residential	0.93	0.94	0.935
River	0.96	0.95	0.955
SeaLake	0.97	0.96	0.965

Precision Recall F1-Score

Key Insights

- Best Performing Classes: Forest, SeaLake (high visual distinctiveness)
- Most Challenging Classes: Highway, Industrial (structural similarities)
- Confusion Patterns: Urban classes show highest inter-class confusion
- Generalization: Strong performance on unseen test data