Laporan Training - Macro NAS + Fusion untuk Face Anti-Spoofing (Training Kedua)

Eksekutif Summary

Eksperimen ini melanjutkan Neural Architecture Search (NAS) untuk mencari backbone terbaik pada setiap modalitas (RGB, IR, Depth), kemudian mengimplementasikan concatenation fusion untuk menggabungkan ketiga modalitas. Hasil menunjukkan bahwa individual modality (terutama Depth) memberikan performa superior dibanding fusion model.

---

1. Pemilihan Backbone (Macro NAS)

1.1 Metodologi

- **Backbone yang Diuji**: ResNet18, ResNet34, ResNet50, EfficientNet-B0, MobileNet-V3-Large
- **Training**: 15 epochs dengan early stopping (patience=5)
- **Evaluasi**: Validation AUC sebagai metric utama
- **Optimisasi**: AdamW dengan learning rate berbeda untuk backbone vs classifier

1.2 Hasil Per Modalitas

RGB Modality

Backbone	Val AUC	ACER	Params	Status
ResNet18	0.8908	0.3499	11.44M	
ResNet34	0.9016	0.3073	21.55M	
ResNet50	0.8755	0.2819	24.56M	
EfficientNet-B0	0.9051	0.3209	4.66M	
MobileNet-V3	0.9195	0.3148	3.46M	⭐ SELECTED

Pemenang RGB: MobileNet-V3-Large

- Alasan: AUC tertinggi (0.9195) dengan parameter paling efisien (3.46M)
- Trade-off: Balance optimal antara performance dan efficiency

IR Modality

Backbone	Val AUC	ACER	Params	Status
ResNet18	0.9713	0.2340	11.44M	
ResNet34	0.9839	0.2073	21.55M	⭐ SELECTED
ResNet50	0.9832	0.3966	24.56M	Close
EfficientNet-B0	0.9714	0.1563	4.66M	Good ACER
MobileNet-V3	0.9623	0.2583	3.46M	

Pemenang IR: ResNet34

- Alasan: AUC tertinggi (0.9839) dengan ACER terbaik (0.2073)
- Karakteristik: IR Modality IR menunjukkan performa konsisten tinggi di semua backbone

Depth Modality

Backbone	Val AUC	ACER	Params	Status
ResNet18	0.9999	0.0415	11.44M	⭐ SELECTED
ResNet34	0.9959	0.0961	21.55M	Excellent
ResNet50	0.9994	0.0391	24.56M	Near-perfect
EfficientNet-B0	0.9993	0.0189	4.66M	Best ACER
MobileNet-V3	0.9851	0.1188	3.46M	Good

Pemenang Depth: ResNet18

- Alasan: AUC near-perfect (0.9999) dengan ACER sangat rendah (0.0415)
- Insight: Depth information memberikan discriminative power tertinggi untuk face anti-spoofing

---

2. Evaluasi Test Set (Individual Models)

2.1 Performance Individual Terbaik

Modalitas	Backbone	Test AUC	Test ACER	APCER	BPCER	Accuracy
RGB	MobileNet-V3	0.8375	0.4082	0.8161	0.0003	43.07%
IR	ResNet34	0.9812	0.1897	0.3794	0.0000	73.54%
DEPTH	ResNet18	0.9301	0.2065	0.4118	0.0013	71.24%

Key Findings:

- Depth modality memberikan performance terbaik (AUC: 0.9301)
- IR modality sangat strong (AUC: 0.9812) dengan BPCER = 0
- RGB modality paling challenging dengan APCER tinggi (0.8161)

3. Concatenation Fusion Experiment

3.1 Arsitektur Fusion

Fusion Model Architecture:

- RGB Backbone: MobileNet-V3 (960 features) [FROZEN]
- IR Backbone: ResNet34 (512 features) [FROZEN]
- Depth Backbone: ResNet18 (512 features) [FROZEN]
- Fusion Classifier: Concatenated features (1984 dim)

Input: Concatenated features (1984 dim)
Hidden: 1024 → 512 → 1 (with dropout & batch norm)
Total Params: 39.51M (2.56M trainable)

3.2 Two-Stage Training Strategy

Stage 1: Frozen Backbone Training

- **Duration**: 20 epochs
- **Strategy**: Train fusion classifier only, freeze all backbones
- **Learning Rate**: 1e-3 (AdamW)
- **Result**: Val AUC = 0.8623, Val ACER = 0.1409

Stage 2: End-to-End Fine-tuning

- **Duration**: 15 epochs (early stopped at 14)
- **Strategy**: Unfreeze all parameters
- **Learning Rate**: 5e-4 (different LR for backbone vs fusion)
- **Result**: Val AUC = 0.8604, Val ACER = 0.1431

Selected Model: Stage 1 (Frozen) - better validation performance

3.3 Fusion Test Results

Model Type	AUC	ACER	APCER	BPCER	Accuracy	F1-Score
Best Individual (Depth)	0.9301	0.2065	0.4118	0.0013	71.24%	67.75%
Fusion Model	0.8596	0.1435	0.2869	0.0001	79.98%	75.14%
Performance Gap	-11.1%	+43.6%	+43.5%	-100%	+17.6%	+16.4%

4. Analisis Hasil

4.1 Temuan Utama

Positive Findings:

1. **Accuracy Improvement**: Fusion model meningkat +17.6% accuracy vs best individual
2. **Balanced Performance**: F1-score naik +16.4%, menunjukkan prediksi lebih balanced
3. **BPCER = 0**: Tidak ada false negative (spoof ter-deteksi sebagai real faces)
4. **APCER Reduction**: Attack classification error turun dari 0.4118 → 0.2869

Concerning Findings:

1. **AUC Degradation**: Fusion AUC turun -11.1% dibanding best individual (Depth)
2. **ACER Inconsistency**: ACER meningkat (+43.6%) tapi AUC menurun
3. **Model Complexity**: 39M parameter vs 11.44M (individual Depth model)

4.2 Root Cause Analysis

Mengapa Fusion Underperform?

1. **Depth Dominance**: • Depth modality sudah near-perfect (AUC: 0.9999 validation) • Menambahkan RGB/IR justru malah menghuruhkan performa • Concatenation tidak optimal karena komplexity information
2. **Training Imbalance**: • Frozen training mungkin insufficient untuk optimal fusion • End-to-end training overfitting (early stopped at 14) • Learning rate strategy belum optimal
3. **Feature Redundancy**: • RGB, IR, dan Depth mungkin capture similar discriminative features • Simple concatenation tidak eksploits complementary information • Perlu attention mechanism atau weighted fusion
4. **Dataset Characteristics**: • Face anti-spoofing dataset mungkin lebih suitable untuk single modality • Depth information sudah sufficient untuk classification task ini • Perlu attention mechanism atau weighted fusion

4.3 Recommendations untuk Improvement

Immediate Fixes:

1. **Weighted Fusion**: Replace concatenation dengan learned weighting
2. **Attention Mechanism**: Implementasi attention untuk selective feature fusion
3. **Progressive Training**: Fine-tune fusion layer dulu, baru unfreeze backbones
4. **Regularization**: Tambah L2 regularization dan stronger dropout

Alternative Strategies:

1. **Ensemble Voting**: Weighted average predictions instead of feature fusion
2. **Hierarchical Fusion**: RGB + IR → intermediate, lalu + Depth
3. **Adaptive Fusion**: Dynamic weighting based on input quality/confidence
4. **Knowledge Distillation**: Use best individual (Depth) as teacher

---

5. Kesimpulan

5.1 Macro NAS Success

- **NAS berhasil** menemukan backbone optimal per modalitas: • RGB: MobileNet-V3 (efficiency-focused) • IR: ResNet34 (performance-focused)
  • Depth: ResNet18 (simplicity wins)

5.2 Fusion Insights

- **Simple concatenation fusion tidak optimal** untuk dataset ini
- **Individual Depth model remains superior** untuk face anti-spoofing task
- **Fusion memberikan balanced prediction** tapi mengorbankan peak performance

5.3 Production Recommendation

For Deployment:

- **Use individual Depth Model (ResNet18)** untuk maximum AUC
- **Consider Fusion Model** jika butuh balanced prediction dan higher accuracy
- **Fallback Strategy**: Depth primary, RGB/IR sebagai backup jika depth unavailable

Model Files Tersimpan:

Individual Models: `/content/drive/MyDrive/Macro_NAS_Results/`

- `best_rgb_mobilenet_v3_large_20250928_152143.pth`
- `best_ir_resnet34_20250928_152143.pth`
- `best_depth_resnet18_20250928_152143.pth`

Fusion Model: `fusion_model_20250928_152144.pth`