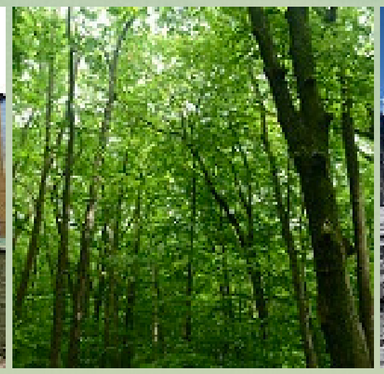
KLASIFIKASI PEMANDANGAN ALAM MENGGUNAKAN MOBILENETV3 DAN TEKNIK FINE-TUNING

Tugas Besar Pengolahan Citra Digital Lanjut

Muhammad Abiya Makruf (203012420034)

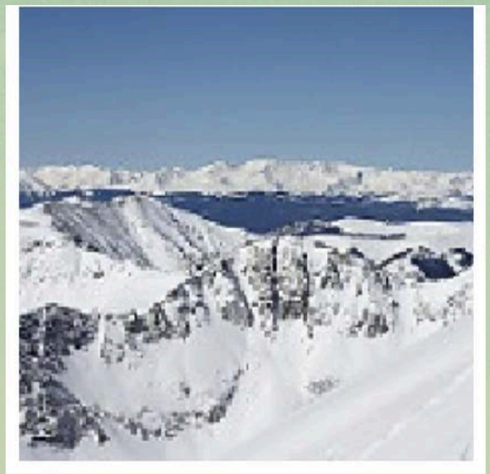
Pendahuluan

Pengenalan pemandangan alam adalah kemampuan inti dalam sistem persepsi visual, baik biologis maupun artifisial. Dengan meningkatnya volume data citra digital, kebutuhan akan sistem otomatis yang dapat memahami dan mengkategorikan konten visual dari pemandangan alam menjadi semakin penting.



Mengapa Klasifikasi Pemandangan Alam?

- Pentingnya pengenalan citra otomatis di era digital.
- Aplikasi luas: navigasi otonom, pemantauan lingkungan, pengelolaan konten visual, pariwisata, dll.
- Tantangan: Variasi intra-kelas yang tinggi, kemiripan antar-kelas, kondisi pencahayaan dan sudut pandang yang beragam.



(a)

(b)

Tujuan

- Mengimplementasikan sistem klasifikasi pemandangan alam untuk 6 kategori menggunakan dataset Intel Image Classification.
- Mengeksplorasi dan membandingkan performa arsitektur MobileNetV3 Small dan MobileNetV3 Large.
- Menerapkan teknik transfer learning untuk meningkatkan akurasi.
- Mengevaluasi model menggunakan metrik standar (akurasi, presisi, recall, F1-score) dan analisis kualitatif (Grad-CAM).
- Mengembangkan aplikasi web interaktif untuk demonstrasi.

Dataset & Persiapan

- Nama: Intel Image Classification (Kaggle)
- Jumlah Kelas: 6 (Bangunan, Hutan, Gletser, Gunung, Laut, Jalan)
- Ukuran: ~14k gambar latih, ~3k gambar uji.
- Resolusi Asli: 150x150 piksel.

- Pembagian Data: Latih (90% dari seg_train), Validasi (10% dari seg_train), Uji (seg_test).
- Resize Gambar: Ke 224x224 piksel (input MobileNetV3).
- Normalisasi: Menggunakan preprocess_input MobileNetV3.
- Augmentasi Data (pada data latih): Rotasi, pergeseran, zoom, flip horizontal.

Metodologi

- Mengapa MobileNetV3?
  - Efisiensi komputasional tinggi (cocok untuk aplikasi real-time/perangkat terbatas).
  - Performa akurasi yang kompetitif.
- Varian yang Digunakan:
  - MobileNetV3 Small
  - MobileNetV3 Large
- Transfer Learning:
  - Memanfaatkan bobot pre-trained dari ImageNet untuk pembelajaran fitur visual dasar.

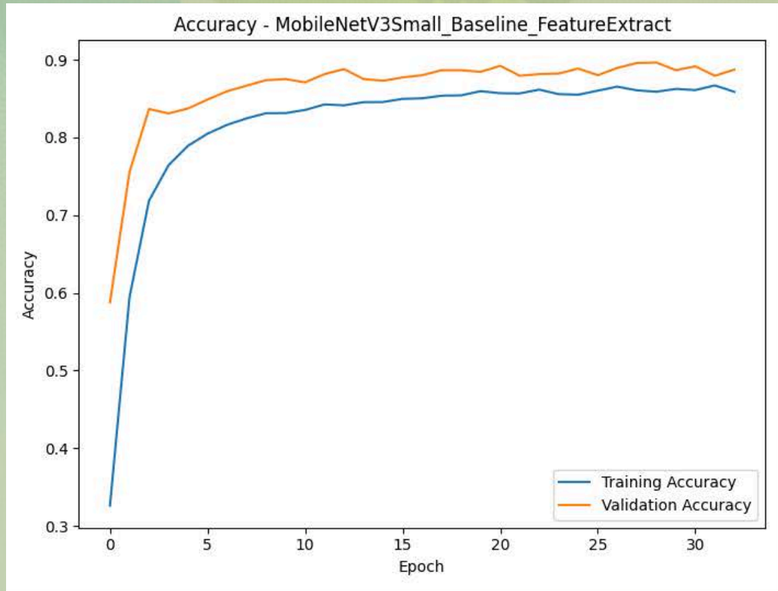
Metodologi - Strategi Pelatihan

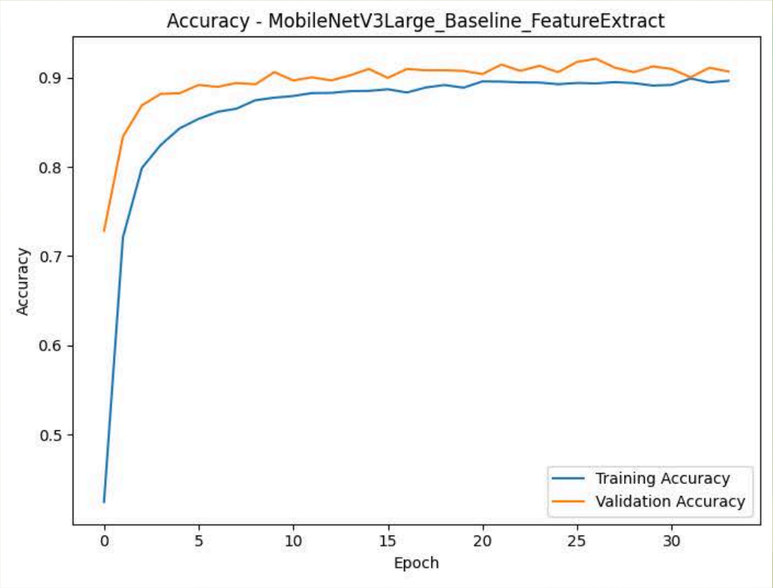
- Tahap 1: Baseline)
  - Base model (MobileNetV3) di-freeze.
  - Hanya head classifier baru (untuk 6 kelas) yang dilatih.
  - Tujuan: Mengadaptasi output layer ke dataset baru.
- Tahap 2: Fine-tuning
  - Sebagian/seluruh base model di-unfreeze.
  - Pelatihan dilanjutkan dengan learning rate yang lebih kecil.
  - Tujuan: Menyesuaikan fitur yang lebih dalam agar spesifik ke dataset pemandangan alam.
- Detail Pelatihan:
  - Optimizer: Adam
  - Loss: Categorical Crossentropy
  - Callbacks: ModelCheckpoint, EarlyStopping, ReduceLROnPlateau.

Hasil Eksperimen - Baseline

- Akurasi (Test Set):
  - MBV3S 89.27%
  - MBV3L 91.53%

- Observasi Singkat:
  - Performa awal yang baik menunjukkan efektivitas fitur pre-trained.
  - MobileNetV3 Large sedikit lebih unggul.

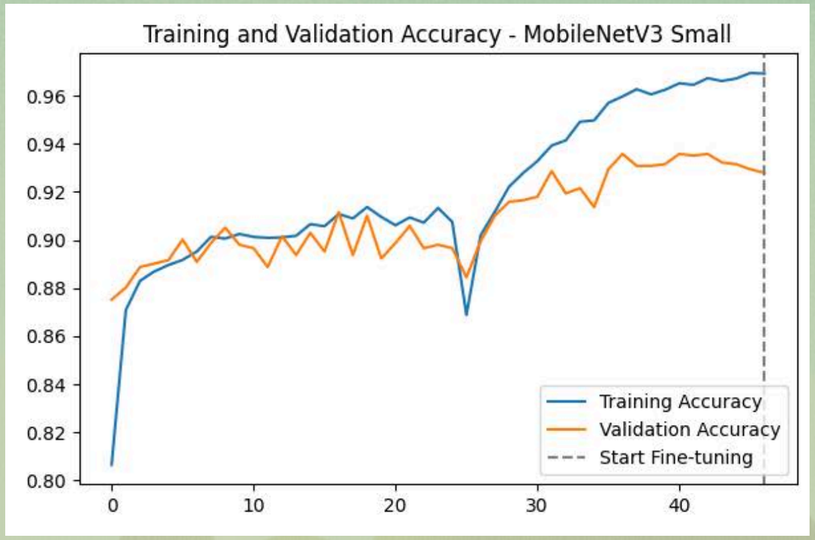


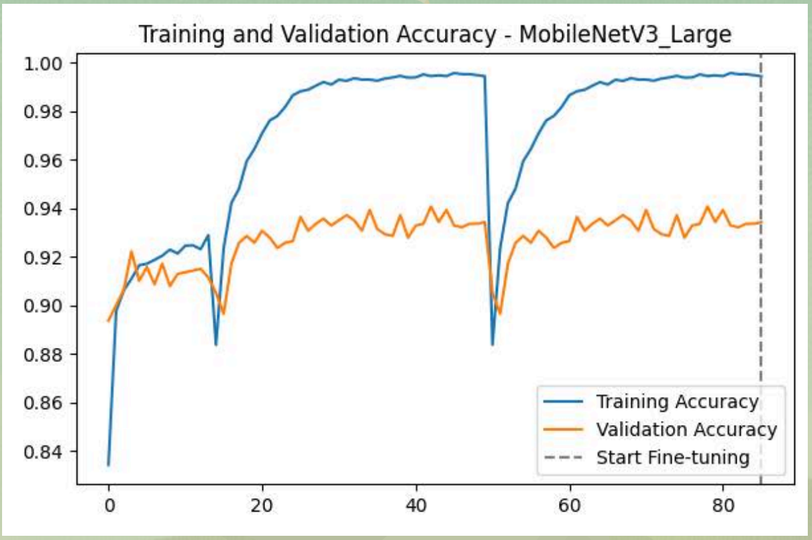


Hasil Eksperimen - Finetuning

- Akurasi (Test Set):
  - MBV3S 92.93%
  - MBV3L 93.63%

- Observasi Singkat:
  - Peningkatan akurasi yang signifikan setelah fine-tuning untuk kedua model.
  - MobileNetV3 Large (Fine-tuned) mencapai performa terbaik.

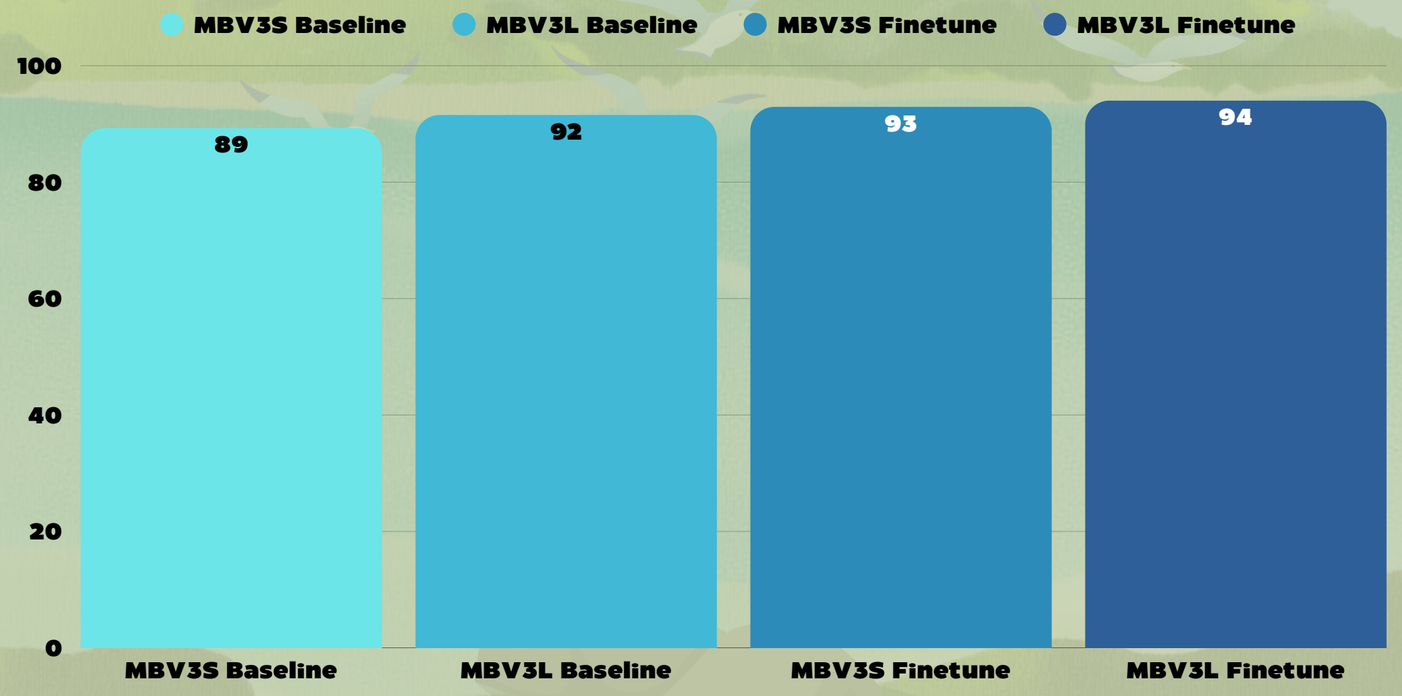




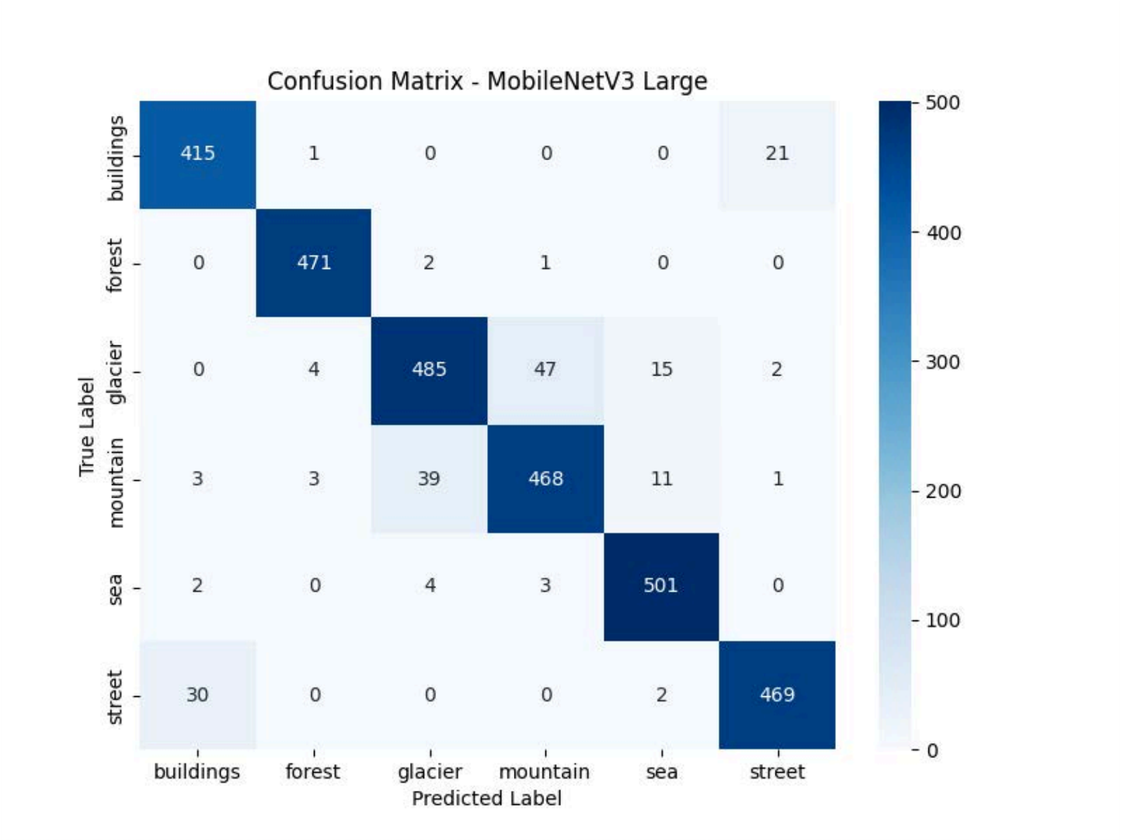
Analisis Perbandingan Model

- Jumlah Parameter
- Perbandingan MobileNetV3 Small vs Large (trade-off akurasi vs efisiensi).

Model	Total Parameter
MobileNetV3Small	2,554,968
MobileNetV3Large	5,507,432

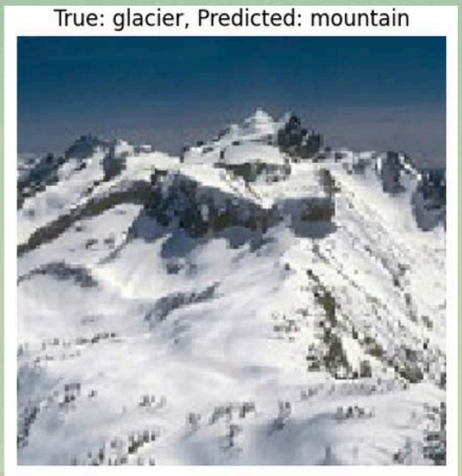


Analisis Kualitatif - Confusion Matrix



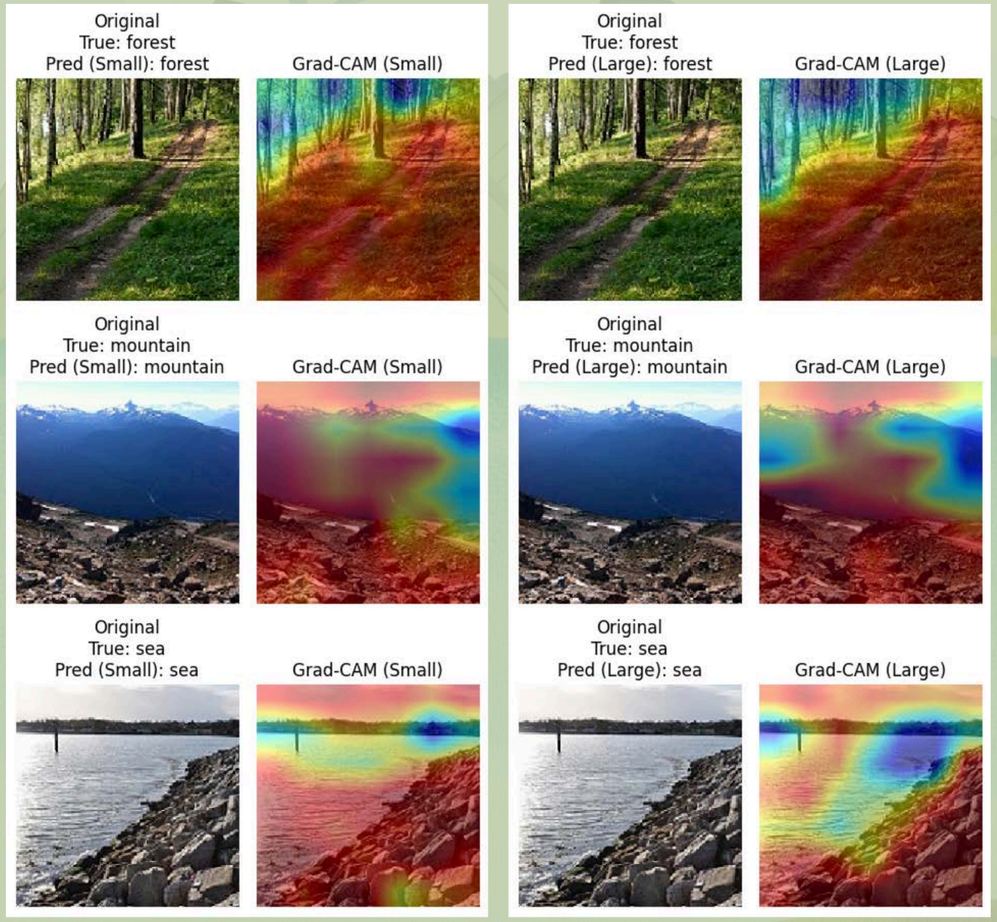
model terbaik yaitu MobileNetV3 Large (finetune), terdapat beberapa kelas yang salah klasifikasi seperti kelas glacier yang terprediksi mountain dan sebaliknya atau kelas buildings yang terprediksi street dan sebaliknya.

(a)



(b)

Analisis Kualitatif - Grad-CAM

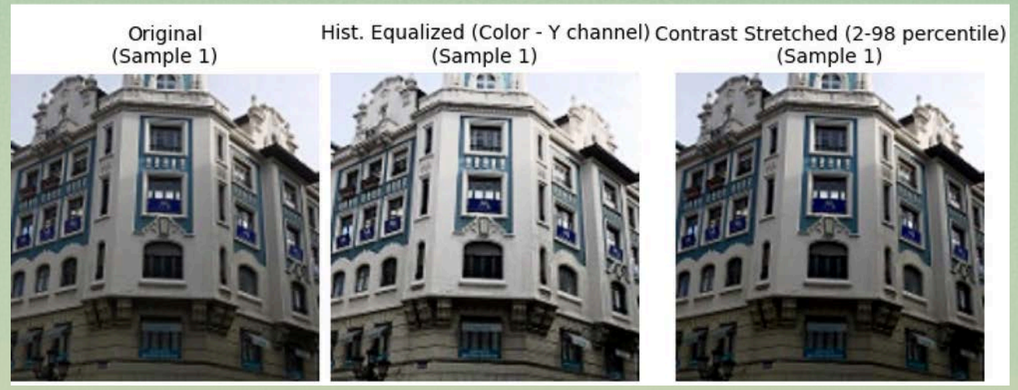


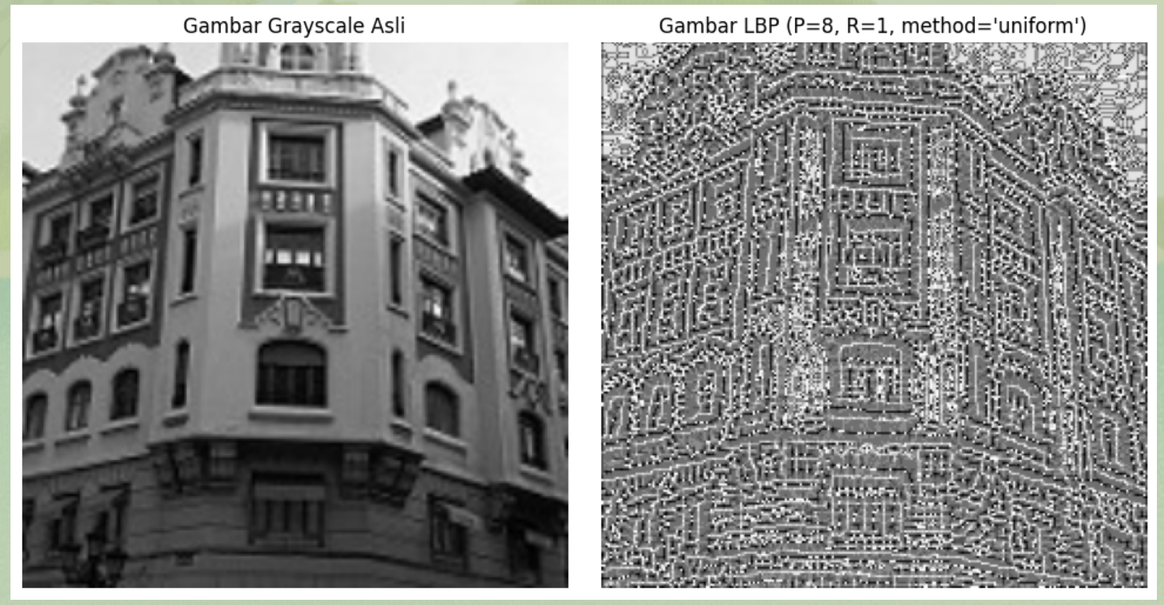
(a)

(b)

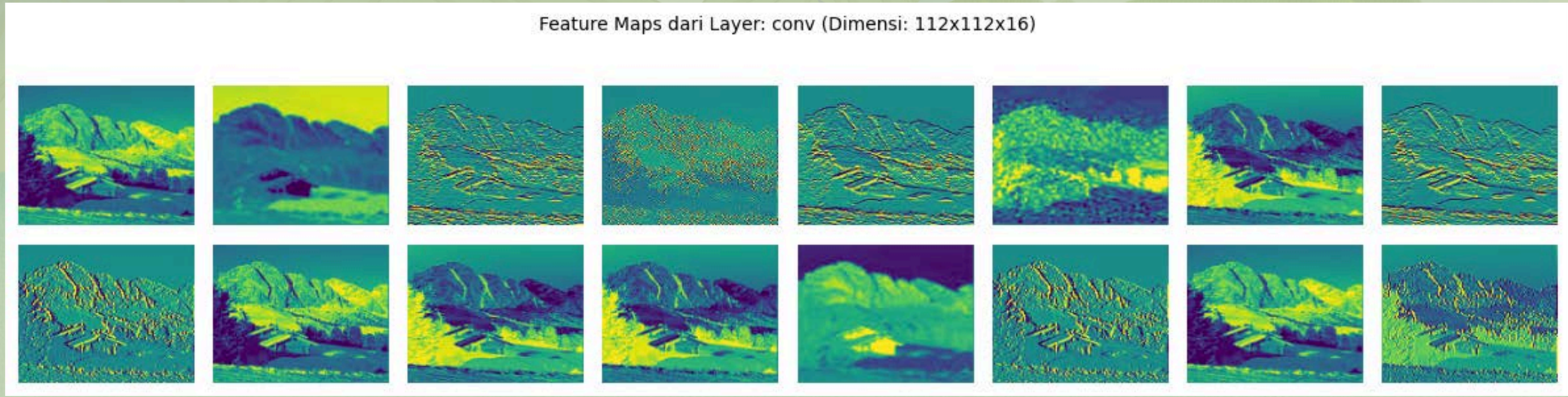
- Apa itu Grad-CAM? Penjelasan singkat: menyoroti area penting pada gambar yang digunakan model untuk prediksi.
- Manfaat: Memvalidasi keputusan model, membantu debugging kualitatif.

Perbandingan Manual dan Otomatis

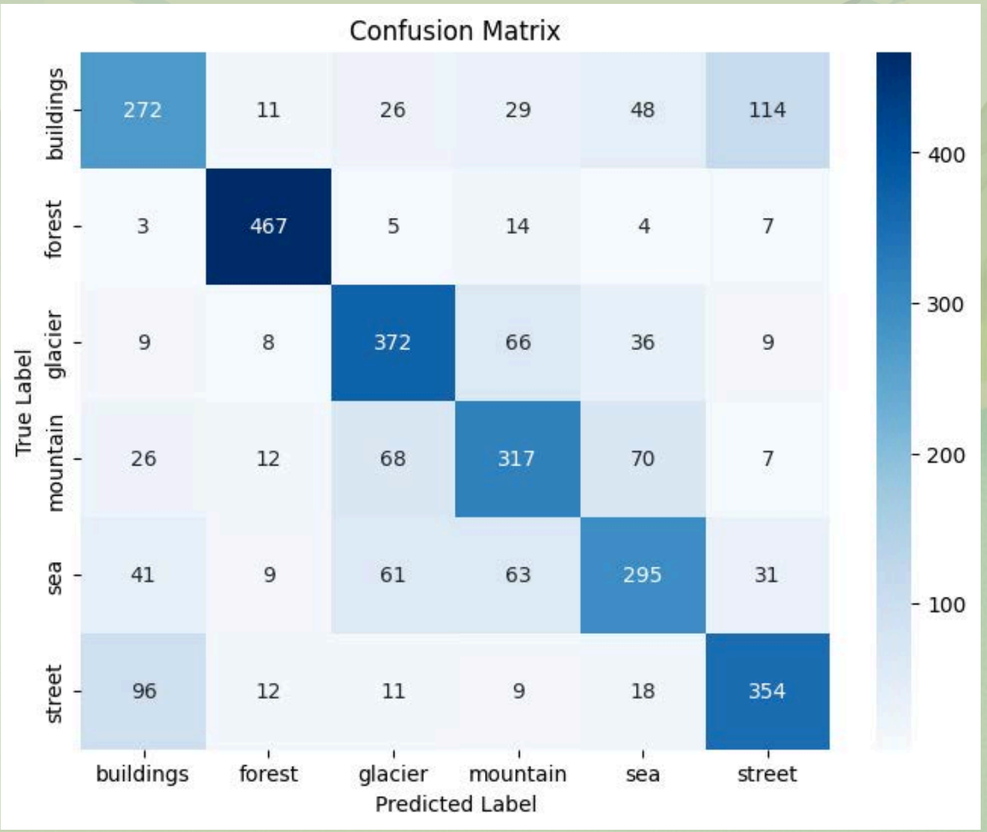




Perbandingan Manual dan Otomatis



Perbandingan Manual dan Otomatis




```
Mengevaluasi model...
Akurasi Model pada Test Set: 69.23%

Laporan Klasifikasi:
              precision    recall  f1-score   support

   buildings       0.61      0.54      0.57       500
      forest       0.90      0.93      0.92       500
     glacier       0.69      0.74      0.71       500
    mountain       0.64      0.63      0.64       500
         sea       0.63      0.59      0.61       500
      street       0.68      0.71      0.69       500

    accuracy                           0.69      3000
   macro avg       0.69      0.69      0.69      3000
weighted avg       0.69      0.69      0.69      3000
```

Demo Aplikasi Web (Streamlit)

Kesimpulan

- Berhasil mengimplementasikan sistem klasifikasi pemandangan alam dengan akurasi tinggi (93.63% dengan MobileNetV3 Large Fine-tuned).
- Demonstrasi efektivitas transfer learning dan fine-tuning.
- Analisis model menggunakan Grad-CAM memberikan wawasan.
- Aplikasi web fungsional berhasil dikembangkan.

Saran

- Eksplorasi arsitektur SOTA lain (EfficientNet, ViT,dll).
- Hyperparameter tuning yang lebih ekstensif.
- Teknik augmentasi data yang lebih canggih.
- Studi lebih mendalam tentang dampak pra-pemrosesan klasik.
- Optimalisasi model untuk deployment (TF Lite, ONNX).
- Penambahan fitur pada aplikasi web (misalnya, analisis batch, perbandingan model live).

Tautan Github

THANK YOU