

Nama : Dody Adi Sancoko

NIM : 1301223071

TUGAS 1: Observasi Berbagai Arsitektur Cnn dalam Klasifikasi Objek

Dataset yang digunakan dalam tugas ini adalah Vegetable Image Dataset dari Kaggle, yang berisi 15.000 gambar yang telah terbagi secara seimbang ke dalam 15 kelas sayuran berbeda. Dataset ini memenuhi kriteria tugas 1 karena memiliki lebih dari 10 kelas dan jumlah total citra melebihi 1000. Dataset dibagi ke dalam tiga bagian: train, validation, dan test.

Saya menggunakan tiga arsitektur CNN: **VGG16**, **ResNet50**, dan **MobileNetV2**. Ketiganya digunakan sebagai *feature extractor* (dengan *pretrained weights* dari ImageNet), lalu ditambahkan beberapa lapisan klasifikasi di atasnya untuk mengenali 15 jenis sayuran.

A. Arsitektur yang Digunakan

1. VGG16

VGG16 adalah arsitektur yang terdiri dari 16 trained layers, dengan pola konvolusi kecil berukuran 3x3 dan max pooling 2x2 untuk mengecilkan dimensi. Setelah melalui konvolusi, fitur visual dari gambar diteruskan ke tiga lapisan fully connected yang menghasilkan prediksi akhir.

Cara Kerja: Dalam dataset ini, VGG16 belajar mengenali ciri visual dari sayuran seperti pola daun, kontur buah, atau warna khas (misalnya hijau pada brokoli, ungu pada brinjal).

2. ResNet50

ResNet50 memiliki 50 layer dan memperkenalkan *residual blocks*, yaitu *shortcut* yang melewati sebagian informasi dari layer sebelumnya ke layer selanjutnya. Tujuannya adalah mencegah degradasi performa pada jaringan dalam (deep).

Cara Kerja: Karena dataset memiliki banyak kelas dengan perbedaan visual yang halus, ResNet dapat menangkap fitur kompleks seperti pola urat pada daun atau bentuk unik dari masing-masing sayuran. *Shortcut connection* membantu model tetap stabil walaupun sangat dalam.

3. MobileNetV2

MobileNetV2 dirancang agar ringan dan cepat, menggunakan *depthwise separable convolution* dan *inverted residual blocks*. Arsitektur ini mengurangi jumlah komputasi tanpa terlalu mengorbankan akurasi.

Cara Kerja: MobileNetV2 fokus pada fitur penting sambil tetap efisien. Walau ukurannya kecil, ia bisa mengenali bentuk dan warna dasar dari sayuran.

B. Perbandingan Performa

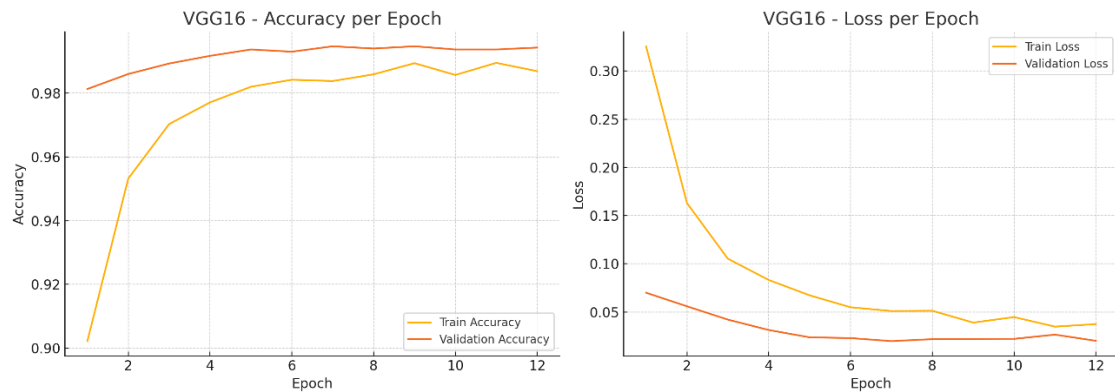
Arsitektur	Accuracy	Precision	Recall	F1-Score	Training time/epoch
VGG16	0.9930	0.9931	0.9930	0.9930	234.68 detik
ResNet50	0.9990	0.9990	0.9990	0.9990	203.90 detik
MobileNetV2	0.9983	0.9983	0.9983	0.9983	196.87 detik

Dari sisi **kinerja**, ResNet50 mencatatkan skor tertinggi pada semua metrik evaluasi—**Accuracy, Precision, Recall, dan F1-Score**—yakni sebesar **0.9990**, yang menunjukkan bahwa model ini sangat andal dan stabil dalam klasifikasi. Di sisi lain, MobileNetV2 menempati posisi kedua dengan metrik sekitar **0.9983**, hanya sedikit di bawah ResNet50, dan masih jauh lebih unggul dibandingkan VGG16 (0.9930). Ini menunjukkan bahwa meskipun ringan, MobileNetV2 tetap mampu mencapai akurasi yang sangat tinggi.

Dari segi **efisiensi**, MobileNetV2 adalah yang **paling cepat** dengan rata-rata waktu pelatihan **196.87 detik/epoch**, disusul oleh ResNet50 (**203.90 detik**) dan terakhir VGG16 yang paling lambat (**234.68 detik/epoch**). Jadi, jika dilihat dari perpaduan kinerja tinggi dan efisiensi waktu, **MobileNetV2 merupakan model paling seimbang dan ideal**. Namun, jika fokus utama hanya pada akurasi maksimal tanpa memperhitungkan waktu pelatihan, maka **ResNet50 adalah pilihan terbaik**.

C. Analisis Dampak terhadap Kinerja dan Efisiensi

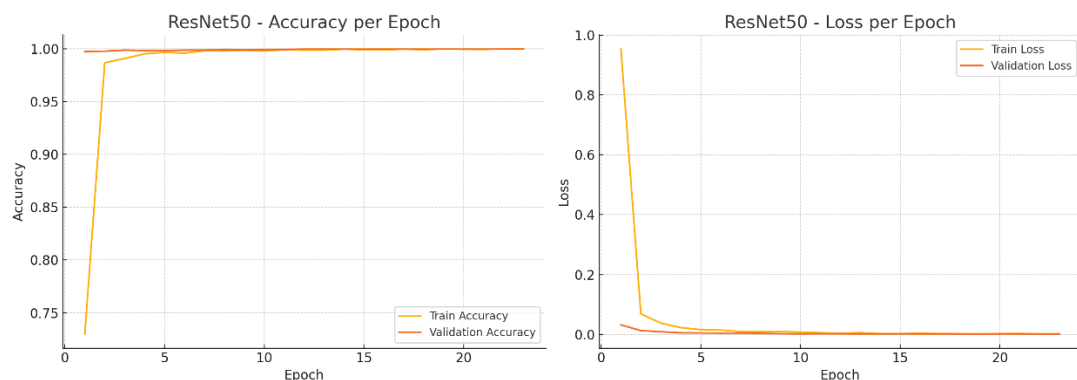
1. VGG16



Model VGG16 menunjukkan performa yang sangat baik dalam klasifikasi gambar sayuran. Sejak awal pelatihan, model sudah memberikan akurasi yang tinggi, yakni 90,22% untuk data training dan 98,13% untuk data validasi pada epoch pertama. Nilai **loss menurun drastis** dari **0.3257 ke 0.0237** hingga sekitar **epoch ke-5**, yang menjadi titik transisi sebelum kurva mulai melandai. Setelahnya, penurunan loss berjalan lebih lambat dan stabil di kisaran **0.02**, menunjukkan bahwa model telah masuk ke fase konvergensi. Dengan diterapkannya *early stopping*, proses pelatihan otomatis dihentikan di epoch ke-12, karena tidak ada lagi peningkatan signifikan pada *validation loss*. *Validation accuracy* terbaik tercapai pada **epoch ke-7 dan ke-9** yaitu sebesar **99,47%**, menunjukkan model telah belajar dengan sangat baik tanpa *overfitting* yang berarti.

Dari segi arsitektur, VGG16 unggul dalam kesederhanaan struktur, menggunakan layer konvolusi kecil (3x3) dan max pooling secara bertahap, membuatnya sangat cocok untuk mengekstraksi fitur visual dari gambar sayuran. Kelebihannya adalah **akurasi tinggi** dan **stabilitas performa**, namun arsitektur ini juga memiliki **jumlah parameter yang besar**, membuat proses pelatihan memakan waktu dan sumber daya komputasi yang lebih banyak dibanding model lain seperti MobileNet. Selain itu, tidak adanya *shortcut connection* membuat VGG16 kurang fleksibel dalam menjaga gradien saat training berlangsung lama. Meski demikian, untuk tugas klasifikasi ini, VGG16 tetap tampil solid dan efektif, memberikan performa tinggi dengan proses training yang stabil dan efisien.

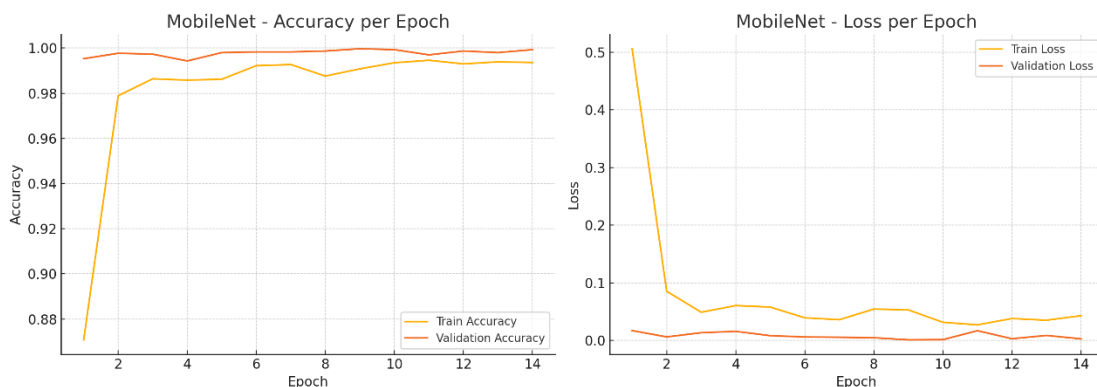
2. ResNet50



ResNet50 menunjukkan kinerja yang sangat baik dalam melakukan klasifikasi gambar pada dataset sayuran. Model ini berhasil mencapai *validation accuracy* sebesar **99,97%** hanya dalam beberapa epoch awal, dengan akurasi training yang juga sangat tinggi. Penurunan loss yang signifikan terjadi pada **epoch 1 hingga sekitar epoch 6**, di mana nilai loss turun drastis dari **0.95 menjadi sekitar 0.014**, menunjukkan proses pembelajaran fitur berjalan sangat efektif. Setelah itu, nilai loss mulai melandai, terutama mulai epoch 9 hingga akhir, dengan nilai *validation loss* yang sangat rendah dan stabil di kisaran **0.0012**. Proses pelatihan otomatis dihentikan lebih awal oleh mekanisme *early stopping* dengan *patience*=5 berdasarkan *validation loss*, sehingga training berakhir di epoch ke-23, saat model sudah dianggap tidak lagi mengalami peningkatan signifikan.

Dari segi arsitektur, kekuatan utama ResNet50 terletak pada penggunaan *residual connections* yang memudahkan jaringan dalam mempelajari fitur mendalam tanpa mengalami degradasi performa. Hal ini memungkinkan model untuk bekerja sangat baik dalam mengenali berbagai fitur visual seperti warna, bentuk, dan tekstur pada gambar sayuran yang cukup beragam. Namun, kelemahan dari arsitektur ini adalah kebutuhan akan waktu komputasi dan sumber daya yang tinggi, karena jumlah parameter yang besar dan kedalaman jaringan yang kompleks. Dibandingkan arsitektur ringan seperti MobileNetV2, ResNet50 memang lebih lambat, namun tetap menjadi pilihan ideal jika **prioritas utama adalah akurasi tinggi**, terutama untuk tugas klasifikasi yang kompleks dan variatif seperti ini.

3. MobileNetV2



Model MobileNet menunjukkan performa yang sangat baik dalam proses training selama 14 epoch. Pada epoch pertama, akurasi training sudah cukup tinggi yaitu 87%, dan langsung melonjak menjadi hampir 99% hanya dalam beberapa epoch. Hal ini mencerminkan efisiensi MobileNet sebagai model ringan namun efektif. Selama proses training, nilai loss juga turun secara konsisten, menunjukkan bahwa model berhasil belajar dengan cepat. Validasi akurasi bahkan mencapai angka sempurna (99.97%) pada epoch ke-9, yang menandakan generalisasi model ke data validasi sangat baik.

Namun, dibandingkan dengan model seperti ResNet dan VGG16, performa MobileNet sedikit lebih fluktuatif dalam hal **validation loss**. Terlihat bahwa pada beberapa epoch seperti ke-3, ke-4, dan ke-11, val_loss sempat meningkat sebelum akhirnya turun lagi. Ini bisa jadi karena karakteristik model yang ringan, yang membuatnya sedikit lebih sensitif terhadap variasi data. Meskipun begitu, tren keseluruhan menunjukkan bahwa model tetap stabil dan tidak mengalami overfitting serius. MobileNet cocok untuk deployment di perangkat dengan keterbatasan komputasi karena ukurannya kecil dan kecepatannya tinggi, sambil tetap mempertahankan akurasi tinggi yang kompetitif.