

LAPORAN UTS ANALISA VISUAL TASK

DATASET KLASIFIKASI IKAN



Disusun oleh :

Rafif Zain Hibatullah

(1103213163)

PRODI S1 TEKNIK KOMPUTER

FAKULTAS TEKNIK ELEKTRO

UNIVERSITAS TELKOM BANDUNG

2025

1. Arsitektur CNN dengan X lapisan konvolusi menghasilkan akurasi training 98% tetapi akurasi validasi 62%. Jelaskan fenomena vanishing gradient yang mungkin terjadi pada lapisan awal, dan bagaimana cara memitigasinya! Mengapa penambahan Batch Normalization setelah lapisan konvolusi ke-Y justru memperburuk generalisasi, serta strategi alternatif untuk menstabilkan pembelajaran?

a. Fenomena Vanishing gradient terjadi ketika gradien (nilai perubahan bobot selama backpropagation) menjadi sangat kecil, terutama pada lapisan awal. Ini menyebabkan model sulit untuk belajar, karena update bobot pada lapisan-lapisan awal menjadi sangat kecil, sehingga pelatihan tidak dapat dilanjutkan dengan efektif. Ini sering terjadi pada model yang sangat dalam dengan aktivasi seperti sigmoid atau tanh, yang menghasilkan gradien kecil saat nilai input terlalu besar atau kecil.

Mitigasi:

- Aktivasi ReLU dapat mengatasi masalah vanishing gradient karena memberikan gradien tetap (nilai tetap jika input positif). Hal ini mempercepat pelatihan dan mengurangi risiko vanishing gradient.
- Menggunakan metode Xavier atau He initialization untuk mencegah bobot yang terlalu kecil atau terlalu besar, yang dapat menyebabkan masalah pada lapisan awal.
- Menambahkan Batch Normalization (BN) dapat mengurangi masalah vanishing gradient dengan menormalkan input untuk setiap lapisan dan mempercepat konvergensi.

b. Penambahan Batch Normalization setelah beberapa lapisan konvolusi dapat memperburuk generalisasi jika tidak dipasang dengan hati-hati, terutama jika layer konvolusi terlalu banyak atau jika model sudah mengalami overfitting pada data pelatihan. BN mengurangi internal covariate shift, namun jika digunakan berlebihan atau di tempat yang tidak tepat, dapat menyebabkan regularisasi berlebih yang justru membatasi kemampuan model untuk belajar dengan optimal.

Strategi Alternatif:

- Data Augmentation: Teknik ini membantu memperkenalkan variasi pada data pelatihan untuk meningkatkan kemampuan model dalam generalisasi.
- Dropout: Dropout dapat diterapkan setelah lapisan konvolusi untuk mencegah overfitting dan meningkatkan generalisasi.

2. Ketika melatih CNN dari nol, loss training stagnan di nilai tinggi setelah XXX(3 digit epoch) epoch. Identifikasi tiga penyebab potensial terkait laju pembelajaran (learning rate), inisialisasi berat, atau kompleksitas model! Mengapa penggunaan Cyclic Learning Rate dapat membantu model keluar dari local minima, dan bagaimana momentum pada optimizer SGD memengaruhi konvergensi?

a. Penyebab Potensial:

1. Learning Rate Terlalu Tinggi

Jika learning rate terlalu tinggi, model bisa melewati minima dan tidak konvergen dengan baik, yang menyebabkan stagnasi pada nilai loss yang tinggi.

2. Inisialisasi Berat yang Buruk

Inisialisasi bobot yang tidak tepat (misalnya, terlalu kecil atau besar) bisa membuat model kesulitan untuk belajar pada awal pelatihan.

3. Kompleksitas Model yang Terlalu Tinggi

Model yang terlalu kompleks (terlalu banyak lapisan atau neuron) dapat menyebabkan overfitting, yang bisa mengarah pada stagnasi pada nilai loss yang tinggi.

b. Cyclic Learning Rate (CLR) mengubah learning rate secara siklik (bervariasi) dalam interval tertentu, yang memungkinkan model keluar dari local minima dan menemukan minima global. Dengan CLR, learning rate akan naik dan turun, yang membantu model menghindari terjebak di local minima dan mempercepat konvergensi.

c. Momentum pada Optimizer SGD membantu mempercepat pelatihan dengan mempertahankan arah gradien sebelumnya, mengurangi fluktuasi dalam pembaruan bobot, dan membantu model bergerak lebih cepat ke arah yang lebih baik. Momentum dapat menstabilkan pelatihan, membantu model keluar dari local minima, dan mengurangi waktu yang dibutuhkan untuk konvergensi.

3. Pada klasifikasi spesies ikan menggunakan CNN, penggunaan fungsi aktivasi ReLU tidak menunjukkan peningkatan akurasi setelah 50 epoch, meskipun learning rate telah dioptimasi. Jelaskan fenomena dying ReLU yang mungkin terjadi dan bagaimana hal ini mengganggu aliran gradien selama backpropagation!

Fenomena Dying ReLU terjadi ketika beberapa unit ReLU (Rectified Linear Unit) pada jaringan tidak aktif sepanjang pelatihan dan selalu mengeluarkan nilai 0. Hal ini bisa terjadi jika nilai input pada unit ReLU sangat besar atau sangat kecil, yang menyebabkan gradien menjadi 0 selama backpropagation, sehingga pembaruan bobot menjadi tidak ada.

Dampak pada Aliran Gradien: Jika unit ReLU mati, gradien tidak dapat mengalir melalui unit tersebut, yang menyebabkan backpropagation terhambat dan mencegah model belajar. Ini akan membuat model tidak dapat memperbarui bobot dan meningkatkan akurasi.

5. Pada arsitektur CNN untuk klasifikasi ikan, peningkatan kompleksitas model justru menyebabkan penurunan akurasi validasi dari 85% ke 65%, meskipun akurasi training mencapai 98%. Jelaskan fenomena overfitting yang terjadi, dan mengapa penambahan kapasitas model tidak selalu meningkatkan generalisasi! Identifikasi 3 kesalahan desain arsitektur yang memicu degradasi performa.

a. Fenomena Overfitting terjadi ketika model sangat baik dalam mempelajari data pelatihan, tetapi gagal untuk generalize pada data validasi atau uji. Dengan model yang lebih kompleks, model bisa menangkap noise atau detail yang tidak relevan dalam data pelatihan, yang menyebabkan performa buruk pada data validasi.

b. Meskipun model dengan kapasitas yang lebih besar (lebih banyak lapisan dan neuron) bisa menangkap hubungan yang lebih kompleks, model yang terlalu kompleks lebih cenderung overfit pada data pelatihan, terutama jika jumlah data pelatihan terbatas.

c. Kesalahan Desain Arsitektur yang Memicu Degradasi Performa:

1. Terlalu banyak neuron atau lapisan dapat meningkatkan risiko overfitting karena model dapat "menghafal" data pelatihan tanpa belajar generalisasi.
2. Tidak menggunakan teknik regularisasi seperti dropout, batch normalization, atau data augmentation bisa menyebabkan model mengandalkan fitur yang tidak penting dalam data pelatihan.
3. Memiliki model yang terlalu dalam (terlalu banyak lapisan konvolusi atau fully connected) tanpa cukup data untuk melatihnya akan menyebabkan overfitting.