

ResNet or DenseNet?

Muh. Sulthan Nazhim - Taufiq Alif
Universitas Hasanuddin

CONTENTS

I	Introduction	1
II	Problem Definition	1
III	Perbandingan Architecture	1
III-A	DenseNet	1
III-B	ResNet	1
IV	Research Methodology	2
V	Analysis and Interpretation	2
VI	Conclusions and Recommendations	2
	References	2

LIST OF FIGURES

1	Simulation Results	1
2	Simulation Results	1
3	Simulation Results	1
4	Simulation Results	2

LIST OF TABLES

ResNet or DenseNet?

Abstract—ResNet atau DenseNet? Saat ini, sebagian besar pendekatan berbasis pembelajaran mendalam diimplementasikan dengan jaringan tulang punggung seminal, di antaranya dua yang paling terkenal adalah ResNet dan DenseNet. Terlepas dari kinerja kompetitif dan popularitas mereka yang luar biasa, ada kelemahan yang melekat pada keduanya.

I. INTRODUCTION

Deep neural networks (DNNs) telah mencapai kinerja cang-gih dalam berbagai tugas visi komputer [14, 18, 11, 31, 12, 25, 50, 49, 9], dan interpretasi DNN juga telah diselidiki dari lensa visualisasi [46, 27, 28] serta kekokohan [39, 13, 3, 47, 48, 4, 2]. DenseNet telah terbukti memiliki efisiensi penggunaan fitur yang lebih baik, mengungguli ResNet dengan parameter yang lebih sedikit [18]. Meskipun demikian, DenseNet membutuhkan memori GPU yang berat karena operasi penggabungan.

II. PROBLEM DEFINITION

Secara kontra-intuitif, dengan menghubungkan cara ini, DenseNets membutuhkan lebih sedikit parameter daripada CNN tradisional yang setara, karena tidak perlu mempelajari peta fitur yang berlebihan. Selain itu, beberapa variasi ResNets telah membuktikan bahwa banyak lapisan hampir tidak berkontribusi dan dapat dijatuhkan. Faktanya, jumlah parameter ResNets besar karena setiap lapisan memiliki bobot untuk dipelajari. Sebaliknya, lapisan DenseNets sangat sempit (misalnya 12 filter), dan mereka hanya menambahkan satu set kecil peta fitur baru. Masalah lain dengan jaringan yang sangat dalam adalah masalah untuk melatih, karena aliran informasi dan gradien yang disebutkan. DenseNets memecahkan masalah ini karena setiap lapisan memiliki akses langsung ke gradien dari fungsi kerugian dan gambar input asli.

III. PERBANDINGAN ARCHITECTURE

A. DenseNet

DenseNet terdiri dari blok Dense seperti yang ditunjukkan di bawah ini. Di dalam blok-blok itu, lapisan-lapisan itu saling terhubung secara rapat: Setiap lapisan mendapat masukan dari peta fitur keluaran lapisan sebelumnya. Penggunaan kembali residu yang ekstrem ini menciptakan pengawasan yang mendalam karena setiap lapisan menerima lebih banyak pengawasan dari lapisan sebelumnya dan dengan demikian fungsi kerugian akan bereaksi sesuai dan karena metodologi ini, ini menjadikannya jaringan yang lebih kuat.

B. ResNet

Elemen dasar utama ResNet adalah blok sisa. Saat kita masuk lebih dalam ke jaringan dengan sejumlah besar lapisan, perhitungan menjadi lebih kompleks. Lapisan-lapisan ini diletakkan di atas satu sama lain dan setiap lapisan mencoba

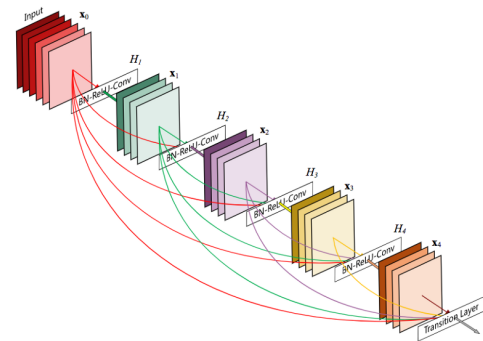


Fig. 1. Simulation Results

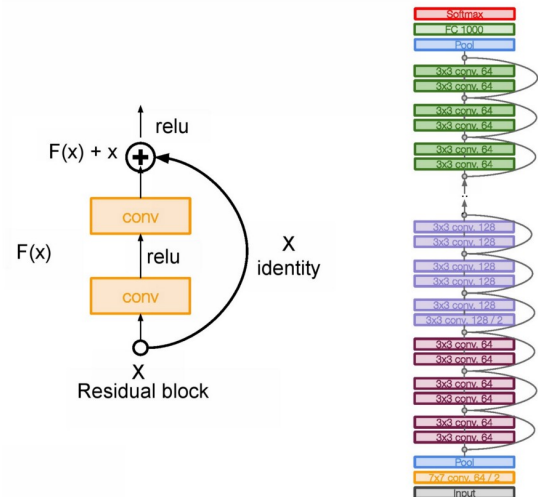


Fig. 2. Simulation Results

mempelajari beberapa pemetaan yang mendasari fungsi yang diinginkan dan alih-alih memiliki blok-blok ini, kami mencoba dan menyesuaikan pemetaan residu. Di sini, di sebelah kanan

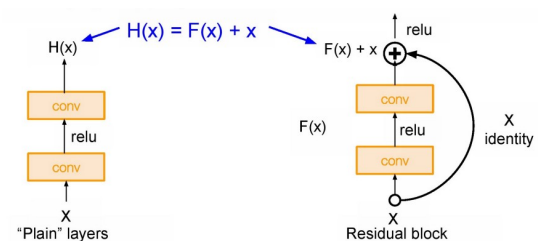


Fig. 3. Simulation Results

ini di mana input ke blok-blok ini hanyalah input yang masuk sedangkan di sisi lain, kita akan menggunakan lapisan kita untuk mencoba dan menyesuaikan beberapa sisa dari $H(X)$ - X kita alih-alih fungsi yang diinginkan $H(X)$ secara langsung. Jadi pada dasarnya, di akhir blok ini dibutuhkan koneksi lewat di sini, di mana ia hanya mengambil input dan meneruskannya sebagai identitas, jadi jika tidak ada lapisan bobot di antaranya, itu hanya akan menjadi identitas. Ini akan menjadi hal yang sama dengan output, tetapi sekarang kami menggunakan lapisan bobot tambahan untuk mempelajari beberapa delta, untuk beberapa sisa dari X kami.

IV. RESEARCH METHODOLOGY

Eksperimen CIFAR-10 Dalam technical report ini, kami menggunakan dataset CIFAR-10. CIFAR-10 merupakan dataset yang berisi kumpulan gambar yang terdiri dari 60000 gambar yang dibagi menjadi 10 kelas dimana masing-masing kelas terdiri dari 6000 gambar. Masing-masing data merupakan gambar berwarna 3 berukuran 32x32. Dataset ini terdiri dari 50000 data latih dan 10000 data uji. [5]. 10 kelas yang terdapat pada dataset CIFAR10 adalah airplane, automobile, cat, deer, dog, frog, horse, ship, truck. Sebelum dataset tersebut digunakan, maka perlu dilakukan normalisasi pada dataset tersebut menggunakan nilai mean dan std dari dataset tersebut.

V. ANALYSIS AND INTERPRETATION

Secara umum, dapat dilihat bahwa semua model memberikan performa yang cukup baik. Dari hasil tersebut, terlihat bahwa DenseNet memperoleh performa yang paling rendah pada bagian validation accuracy dan test accuracy masing-masing sebesar 90.40% performa yang sedikit lebih baik jika dibandingkan dengan ResNet dimana terjadi peningkatan sebesar 0.32% validation accuracy dan 0.53% kita dapat melihat performa dari masing-masing model ResNet mampu mengungguli model DenseNet dimana terjadi perbedaan hingga lebih dari 1 accuracy dan perbedaan lebih dari 1

Model	Val Accuracy	Test Accuracy	Num Parameters
GoogleNet	90.40%	89.70%	260,650
ResNet	91.84%	91.06%	272,378
ResNetPreAct	91.80%	91.07%	272,250
DenseNet	90.72%	90.23%	239,146

Fig. 4. Simulation Results

VI. CONCLUSIONS AND RECOMMENDATIONS

Kami memberikan perspektif gabungan dari penjumlahan padat untuk memfasilitasi pemahaman tentang perbedaan inti antara ResNet dan DenseNet. Kami menunjukkan bahwa perbedaan inti terletak pada apakah parameter konvolusi dibagikan untuk peta fitur sebelumnya. Kami mengusulkan pintasan ternormalisasi berbobot padat sebagai metode koneksi padat alternatif, yang mengungguli dua teknik koneksi padat yang ada: pintasan identitas di ResNet dan rangkaian padat di DenseNet. Kami menemukan bahwa Penjumlahan padat dari keluaran agregasi memberikan kinerja yang lebih baik

daripada keluaran blok konvolusi. Singkatnya, pintasan padat mengatasi masalah penurunan kapasitas representasional di ResNet sambil menghindari kelemahan membutuhkan lebih banyak sumber daya GPU di DenseNet. DSNet yang diusulkan telah dievaluasi pada beberapa kumpulan data tolok ukur untuk menunjukkan kinerja yang unggul daripada rekanannya ResNet.

REFERENCES

- [1] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In CVPR, 2015.
- [2] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In CVPR, 2016.
- [3] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022, 2016. index=52&id=1557231641&src=Mode=1&sid=3&fmt=3&Inst=PRODVType=PQDRQT=309 Accessed on: Aug. 3, 2014.
- [4] Jiawang Bai, Bin Chen, Yiming Li, Dongxian Wu, Weiwei Guo, Shu-tao Xia, and En-hui Yang. Targeted attack for deep hashing based retrieval. In Proceedings of the European Conference on Computer Vision (ECCV), 202.
- [5] Philipp Benz, Chaoning Zhang, Tooba Imtiaz, and In So Kweon. Double targeted universal adversarial perturbations. arXiv preprint arXiv:2010.03288, 2020.