

Resnet

KOMANG JAYA BHASKARA MAHATYA(1103181031)
MUHAMMAD IQBAL THOMAS(1103183187)

The (degradation) problem:

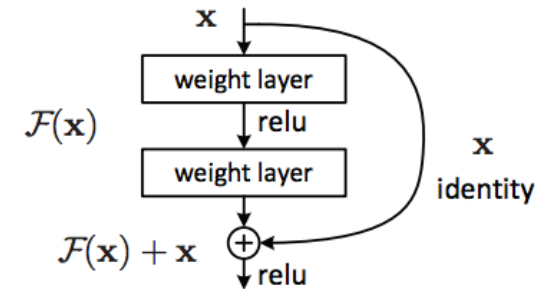
Saat kedalaman jaringan meningkat, akurasi (mungkin tentu saja) jenuh dan kemudian turun dengan cepat. Tanpa diduga, degradasi itu bukan karena overfitting, dan menambahkan lapisan ke model internal yang sesuai menghasilkan kesalahan pelatihan yang lebih tinggi.

The core insight:

Pertimbangkan arsitektur yang lebih datar dan bagian yang lebih dalam yang menambahkan lebih banyak level. Ada solusi model yang lebih dalam menggunakan konstruksi. Lapisan disalin dari model dangkal yang diperiksa dan lapisan yang ditambahkan adalah peta identitas. Kehadiran solusi yang dirancang ini menunjukkan bahwa model yang lebih dalam tidak boleh menyebabkan kesalahan pelatihan yang lebih tinggi daripada model yang lebih dangkal.

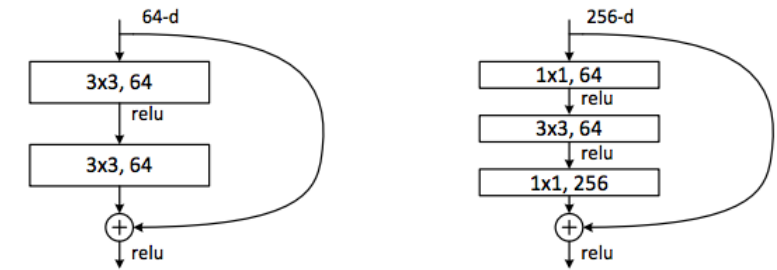
The proposed solution:

Alih-alih mengharapkan setiap tumpukan lapisan untuk mencocokkan pemetaan dasar yang diinginkan secara langsung, buat lapisan itu secara eksplisit cocok dengan sisa pemetaan. Pemetaan asli disusun ulang menjadi $F(x) + x$. Misalkan lebih mudah untuk mengoptimalkan peta residual daripada peta \mathcal{F} yang tidak direferensikan. Dalam kasus ekstrim, ketika peta identitas optimal, lebih mudah untuk mendorong sisanya ke daripada menyesuaikan peta identitas ke tumpukan lapisan nonlinier.



A residual block — the fundamental building block of residual networks. Figure 2:

jaringan dengan mengasumsikan bahwa fungsi optimal yang coba dimodelkan oleh sebuah blok lebih dekat dengan pemetaan identitas daripada peta nol dan akan lebih mudah untuk menemukan pengacakan dengan referensi ke identitas peta daripada kosong peta. Ini menyederhanakan pengoptimalan jaringan hampir tanpa biaya. Dengan demikian, blok berikutnya di jaringan kami bertanggung jawab untuk menyempurnakan output dari blok sebelumnya, daripada harus menghasilkan output yang diinginkan dari awal.



A bottleneck residual block (right) makes residual networks more economical (used hereinafter) but



Wide Residual Networks

Wide (*width refers to the number of channels in a layer*) residual networks upaya untuk mengatasi masalah penggunaan kembali fitur yang semakin berkurang (beberapa blok sisa mempelajari representasi yang berguna atau banyak blok yang berbagi informasi yang sangat sedikit dengan kontribusi kecil untuk tujuan akhir) di jaringan sisa yang sangat dalam (tipis).

ResNeXt

Residual connections sangat membantu untuk menyederhanakan pengoptimalan jaringan, sedangkan transformasi agregat menghasilkan kekuatan representasi yang lebih kuat (seperti yang ditunjukkan oleh fakta bahwa kinerjanya secara konsisten lebih baik daripada rekan-rekan mereka dengan atau tanpa koneksi residual).

Hasilnya adalah arsitektur multi-cabang yang homogen dengan hanya beberapa parameter hiper untuk diatur.

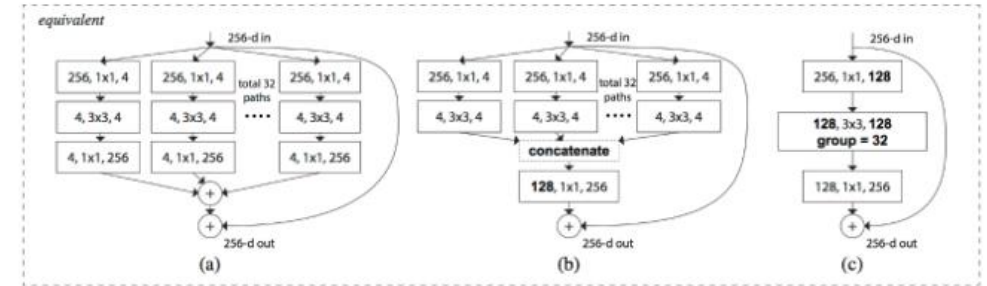


Figure 3. Equivalent building blocks of ResNeXt. (a): Aggregated residual transformations, the same as Fig. 1 right. (b): A block equivalent to (a), implemented as early concatenation. (c): A block equivalent to (a,b), implemented as grouped convolutions [24]. Notations in **bold** text highlight the reformulation changes. A layer is denoted as (# input channels, filter size, # output channels).