

# Literature Review: Metode Jaringan Neural Konvolusi (CNN) Untuk Pemrosesan Gambar

<sup>1</sup>Grace Lamudur Arta Sihombing

<sup>1</sup>Institut Agama Kristen Negeri (IAKN) Tarutung  
[graceshbg@gmail.com](mailto:graceshbg@gmail.com)

**Abstract.** Convolutional Neural Networks (CNN) adalah salah satu jenis model arsitektur jaringan saraf tiruan yang efektif dipergunakan dalam pemrosesan gambar, pengenalan pola klasifikasi citra. Sejak metode ini diperkenalkan hingga saat ini mengalami banyak sekali perkembangan dan peningkatan, Namun model ini juga memiliki beberapa kelemahan. Penelitian ini dilakukan dengan tinjauan literatur dari berbagai sumber yaitu: jurnal nasional dan internasional yang bereputasi yang bisa memberikan pemahaman yang mendalam tentang bagaimana model CNN memproses gambar, seperti apa tren dan pencapaian terbaru penggunaan model ini. Dengan mengidentifikasi dan menganalisis area penelitian yang menarik, penelitian ini dapat memberikan saran dan wawasan tentang kemungkinan arah pengembangan CNN di masa depan. Dalam penelitian ini ditemukan bahwa penggunaan CNN dalam pemrosesan gambar sudah mengalami banyak kemajuan yang signifikan. Namun ada beberapa hal yang menarik untuk dikembangkan teknologi lebih lanjut yaitu 1) Peningkatan efisiensi dan efektifitas CNN dalam memproses data gambar dengan resolusi tinggi. 2) Integrasi CNN dengan pembelajaran mendalam (deep learning) atau *Visison Transformer* dalam pemrosesan gambar. 3) Penerapan CNN dalam berbagai domain aplikasi selain medis misalnya keamanan, otomatisasi dll.

Keyword: Convolution Neural Network, Image Processing

## 1. Pendahuluan

Convolutional Neural Network (CNN) merupakan pengembangan dari metode Deep Neural Network - Multilayer Perceptron (MLP)[1] yang didesain untuk mengolah data dua dimensi dan memiliki kedalaman jaringan yang tinggi. CNN adalah jenis arsitektur jaringan saraf tiruan yang paling sering digunakan dalam pemrosesan gambar dan pengenalan pola visual. CNN sudah terbukti sangat efektif dalam tugas-tugas seperti klasifikasi gambar, deteksi objek, segmentasi gambar, dan telah mencapai terobosan besar dalam skema pengenalan gambar [2] [3], dan dapat dipergunakan diberbagai aplikasi[4]. Pada penelitian Mawan [5] menggunakan CNN untuk klasifikasi motif batik dengan akurasi sebesar 65% dengan menggabungkan *Grayscale* dapat meningkatkan akurasi sebesar 5%. Pada penelitian Bariyah at al [6] juga menemukan CNN memiliki akurasi sebesar 91,41%.

Convolutional Neural Networks (CNN) telah berhasil diterapkan dalam berbagai tugas pemrosesan gambar. Nugroho [7] menggunakan CNN untuk mengenali ekspresi wajah manusia dengan akurasi 90% pada pelatihan dan 65% pada validasi. Sabilla [8] mengusulkan arsitektur CNN yang dimodifikasi untuk klasifikasi buah, mencapai akurasi 96% menggunakan *Support Vector Machine (SVM)* pada lapisan yang terhubung penuh. Pratama [9] mengembangkan model CNN untuk menentukan "gagrak" wayang kulit, mencapai akurasi 92,27%. Azmi [10] mengimplementasikan CNN untuk klasifikasi batik tanah liat Sumatera Barat, memperoleh tingkat akurasi 50%. Penelitian-penelitian ini menunjukkan keefektifan CNN dalam tugas-tugas pemrosesan gambar, dengan berbagai tingkat keberhasilan. Namun, waktu komputasi CNN dapat menjadi kelemahan, terutama pada arsitektur dengan jumlah lapisan dan filter yang banyak[3],[8],[11]. Sehingga pada artikel ini, akan dibahas beberapa kelemahan CNN yang menjadi peluang dan untuk penelitian selanjutnya.

## 2. Metodologi

Metode yang digunakan dalam penelitian ini adalah tinjauan literatur review. Tinjauan pustaka adalah proses pencarian dan penelitian literatur dengan membaca dan mengevaluasi berbagai jurnal, buku, dan naskah terbitan lain yang berkaitan dengan topik penelitian guna menyusun artikel dengan topik tertentu[12]. Melakukan literatur review secara mendalam melalui jurnal dan buku dengan kata kunci Convolutional Neural Networks dan pemrosesan gambar (image processing). Teori yang digunakan adalah terbitan terbaru untuk menemukan teori yang sudah ada dan yang akan dikembangkan.

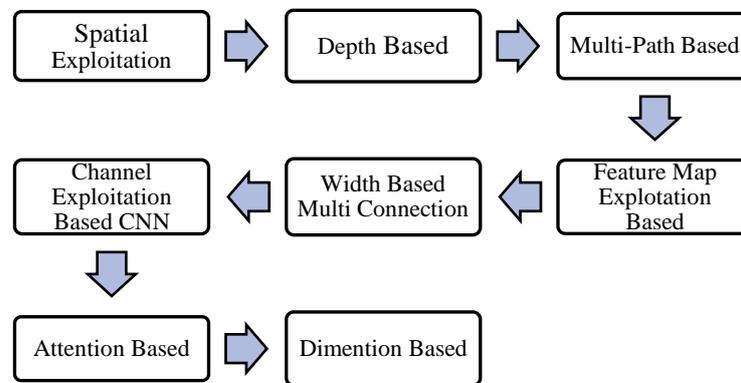
## 3. Pembahasan dan Hasil

### 3.1 Perkembangan CNN dalam Pemrosesan Gambar

Evolusi CNN dalam pemrosesan gambar telah ditandai dengan kemajuan yang signifikan dalam arsitektur dan aplikasi. Bhatt 2021 [13] memberikan gambaran umum yang komprehensif tentang perkembangan CNN, mengkategorikannya ke dalam delapan kelompok dan mendiskusikan kekuatan dan kelemahannya. Potensi teknologi CNN dalam pemrosesan gambar telah diakui di berbagai bidang, termasuk teknologi militer Buzási 2003[14] dan CNN yang dapat mengorganisir diri sendiri Itoh 2005 [15]. Bal 2004 [16] lebih lanjut meningkatkan kemampuan CNN dengan memperkenalkan arsitektur jaringan saraf seluler wavelet dan algoritma pembelajaran, yang telah diuji untuk aplikasi pemrosesan gambar 2D. Evolusi CNN dalam pemrosesan gambar terus berkembang dan lebih baik.

### 3.2 Kemajuan Arsitektur CNN

Berbagai arsitektur varian CNN dijelaskan pada gambar berikut ini, kategori arsitektur varian CNN[13] :



Gambar 1. Kemajuan Arsitektur CNN

#### 3.2.1 CNN Berbasis Eksploitasi Spasial (*Spatial Exploitation-Based CNN*)

Fungsi aktivasi, langkah, ukuran filter, laju pembelajaran, jumlah lapisan, neuron, bias, bobot, dan sebagainya adalah beberapa parameter CNN. Karena operasi konvolusi memperhitungkan lokasi piksel input, berbagai ukuran filter dapat digunakan untuk mempelajari berbagai tingkat korelasi. Granularitas bervariasi di antara ukuran filter yang berbeda; umumnya, filter berukuran besar mengekstrak informasi berbutir kasar, sedangkan filter berukuran kecil mengekstrak informasi berbutir halus. Oleh karena itu, para peneliti mulai menggunakan filter spasial pada awal tahun 2000-an dalam upaya untuk meningkatkan kinerja. Korelasi antara filter spasial dan pembelajaran jaringan ditemukan. Selama waktu ini, sejumlah pengujian menunjukkan bahwa CNN dapat bekerja lebih baik pada detail kasar dan halus dengan memvariasikan filter.

### 3.2.2 *CNN Based on Depth*

Arsitektur Deep CNN memungkinkan jaringan untuk secara efisien mendekati fungsi target dengan menggunakan lebih banyak pemetaan (nonlinier) dan hirarki fitur yang lebih kompleks. [17]. Kedalaman jaringan sekarang menjadi faktor penting dalam pembelajaran yang diawasi. Sistem yang dangkal kurang berhasil dalam merepresentasikan kelas fungsi yang diberikan daripada jaringan yang dalam. Pada tahun 2011, Bengio dan Szegedy menemukan bahwa jaringan yang lebih dalam dapat mempertahankan dampak jaringan dengan biaya yang lebih rendah. [18]. Untuk operasi yang kompleks, jaringan dalam secara komputasi lebih efisien. Kompetisi ILSVRC-2014 melihat kinerja terbaik dari VGG dan Inception, mendukung gagasan bahwa kedalaman adalah faktor penting dalam mengendalikan kapasitas pembelajaran jaringan.

### 3.2.3 *CNN with Multiple Paths*

Salah satu jenis jaringan buatan yang unggul dalam tugas-tugas yang kompleks adalah CNN. Karena gradien yang menghilang, masalah ledakan, atau penurunan kinerja lebih mungkin terjadi ketika kedalaman meningkat daripada overfitting. Kesalahan pengujian dan pelatihan keduanya meningkat sebagai akibat dari masalah gradien yang menghilang [19]. Sebuah teori konektivitas multi-lintasan atau lintas-lapisan disarankan untuk jaringan pembelajaran yang mendalam. Dengan menghilangkan beberapa tahap di antaranya, konektivitas lintas-lapisan, juga dikenal sebagai multi-path, memungkinkan koneksi lintas-lapisan secara analitis dan memungkinkan aliran informasi yang disesuaikan di antara mereka. Koneksi lintas lapisan digunakan untuk mempartisi jaringan ke dalam wilayah yang berbeda. Rute ini memperluas gradien ke lapisan yang lebih rendah, yang memperbaiki masalah gradien yang hilang.

### 3.2.4 *Feature-Map Exploitation Based CNN*

CNN menjadi populer untuk tugas-tugas *machine learning* karena kemampuannya untuk melakukan pembelajaran hierarkis dan ekstraksi fitur otomatis[20]. Kinerja modul klasifikasi, segmentasi, dan deteksi sangat dipengaruhi oleh pemilihan fitur. CNN memilih fitur secara dinamis dengan menyesuaikan bobot yang terkait dengan kernel, yang juga dikenal sebagai mask. Selanjutnya, tahap ekstraksi fitur yang berbeda dilakukan, memungkinkan untuk berbagai jenis fitur (dikenal sebagai peta fitur atau saluran dalam CNN). Namun, beberapa peta fitur memiliki sedikit atau tidak memiliki fungsi dalam diskriminasi objek. Kumpulan fitur yang berlebihan dapat memberikan efek *noise*, yang menyebabkan jaringan menjadi terlalu pas. Hal ini menyiratkan bahwa, selain rekayasa jaringan, pemilihan peta fitur dapat memainkan peran penting dalam meningkatkan generalisasi jaringan.

### 3.2.5 *Multi-Connection Depending on the Width*

Selama perkembangan CNN, penekanannya terutama pada peningkatan potensi kedalaman dan efisiensi koneksi dalam regularisasi jaringan selama tahun 2012 hingga 2015. Kawaguchi [21] pada tahun 2019, menemukan bahwa lebar jaringan juga sama pentingnya. Hal ini menyiratkan bahwa, selain kedalaman, lebar merupakan komponen penting dalam mengembangkan filosofi pembelajaran. Hal ini menunjukkan bahwa jaringan saraf dengan fungsi aktivasi ReLU harus cukup lebar untuk mempertahankan properti aproksimasi universal sambil juga meningkatkan kedalaman. Salah satu masalah penting dengan arsitektur jaringan saraf dalam adalah bahwa beberapa lapisan mungkin gagal mempelajari fitur-fitur yang berharga. Meskipun menumpuk banyak lapisan (meningkatkan kedalaman) dapat mempelajari representasi fitur yang bervariasi, hal ini tidak selalu meningkatkan kekuatan pembelajaran *neural network*. Selain itu, setiap jaringan dalam tidak dapat secara sewenang-wenang memperkirakan kelas fungsi kontinu pada himpunan jika lebar maksimum jaringan tidak lebih besar dari dimensi input. Oleh karena itu, fokus penelitian beralih dari desain yang dalam dan sempit ke arsitektur yang lebar dan tipis untuk mengatasi masalah ini.

Karena kapasitasnya untuk ekstraksi fitur otomatis dan pembelajaran hirarkis, CNN telah mendapatkan popularitas untuk aplikasi pembelajaran mesin.[20]. Pemilihan fitur memiliki dampak besar pada seberapa baik modul deteksi, segmentasi, dan klasifikasi bekerja. CNN menggunakan bobot yang terkait dengan kernel, yang sering disebut sebagai mask, untuk memilih fitur secara dinamis. Selanjutnya, banyak fase ekstraksi fitur yang dilakukan, memungkinkan pembuatan berbagai jenis fitur (kadang-kadang disebut sebagai peta fitur atau saluran CNN). Namun, peta fitur tertentu hanya memberikan sedikit atau bahkan tidak ada gunanya dalam diskriminasi objek. Jaringan dapat menjadi terlalu banyak fitur sebagai akibat dari efek noise yang ditimbulkan oleh kumpulan fitur yang ekstensif. Hal ini menunjukkan bahwa pemilihan peta fitur dapat menjadi faktor yang signifikan dalam meningkatkan generalisasi jaringan selain rekayasa jaringan.

### 3.2.6 *Exploitation-Based Feature-Map CNN*

Karena kapasitasnya untuk melakukan pembelajaran hierarkis dan ekstraksi fitur otomatis, CNN telah menjadi solusi dalam masalah visi komputer[22]. Kinerja modul klasifikasi, segmentasi, dan deteksi sangat dipengaruhi oleh pemilihan fitur. CNN memilih fitur secara dinamis dengan menyesuaikan bobot yang dihubungkan dengan kernel, yang juga dikenal sebagai mask. Selain itu, banyak fase ekstraksi fitur yang digunakan dalam CNN untuk menambang berbagai jenis fitur. Namun, beberapa peta fitur memiliki sedikit atau tidak ada signifikansi dalam diskriminasi objek. Kumpulan fitur yang sangat besar dapat memberikan efek *noise*, menyebabkan jaringan menjadi *overfitting*. Hal ini menyatakan bahwa, selain rekayasa jaringan, pemilihan peta fitur menjadi sangat penting dalam meningkatkan generalisasi jaringan.

### 3.2.7 *CNN Based on Attentions*

Tingkat abstraksi yang beragam memainkan peran penting dalam menentukan daya diskriminasi Jaringan Saraf Tiruan. Hirarki abstraksi yang berbeda yang berfokus pada atribut yang relevan dengan lokalisasi dan pengenalan gambar memainkan peran penting dalam pembelajaran. Efek ini dikenal sebagai perhatian dalam sistem visual manusia. Manusia dapat melihat pemandangan apa pun dengan mengintegrasikan pandangan parsial dan berfokus pada aspek yang relevan dengan konteks. Pendekatan ini berfokus pada wilayah tertentu dan memahami berbagai interpretasi item di tempat tertentu, sehingga meningkatkan penangkapan struktur visual. *Recurrent Neural Network (RNN)* [23], [24], [25] dan *Long Short TermMemory (LSTM)* menggabungkan interpretasi yang kurang lebih sebanding. Jaringan RNN dan LSTM menggunakan modul perhatian sebagai fitur progresif[26], dan pencicip baru diberi bobot berdasarkan pengulangannya di putaran sebelumnya. Konsep perhatian dalam jaringan saraf *convolutional* digunakan oleh berbagai ahli untuk meningkatkan representasi dan mengatasi keterbatasan komputasi. Konsep perhatian ini juga berkontribusi pada CNN yang cukup cerdas untuk membedakan item bahkan dalam latar belakang yang sibuk dan skenario yang kompleks.

### 3.2.8 *Dimension-Based CNN*

Lapisan konvolusi klasik mengkodekan informasi saluran dan spasial secara bersamaan, tetapi secara komputasi mahal. Efisiensi konvolusi biasa ditingkatkan dengan diperkenalkannya konvolusi yang dapat dipisahkan (atau dapat dipisahkan berdasarkan kedalaman), yang mengkodekan informasi spasial dan *channel-wise* secara terpisah menggunakan konvolusi *point-wise* dan *depth-wise*. Faktorisasi ini jauh lebih efisien, tetapi menempatkan beban komputasi yang cukup besar pada konvolusi *point-wise*, menjadikannya hambatan komputasi[27].

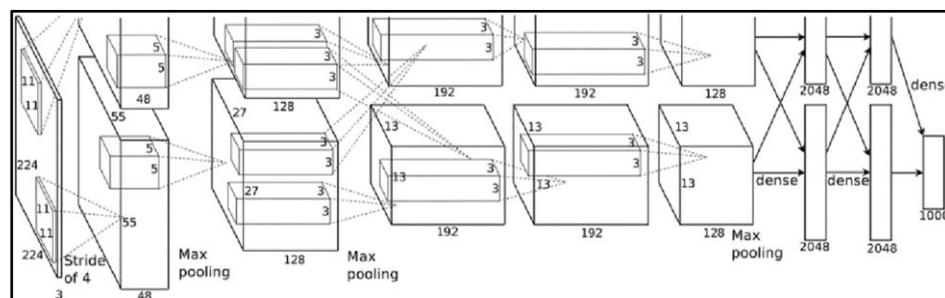
### 3.3. Ringkasan CNN untuk pemrosesan gambar

CNN untuk pemrosesan gambar terus berkembang dari waktu ke waktu, karena banyak penelitian yang terus dikembangkan untuk memperoleh model yang baru yang lebih baik. Berikut ini Tabel 1 Ringkasan CNN untuk pemrosesan gambar [28] :

Jenis Neural Network	Rangking 1	Rangking 5	Jumlah Layer	Proses (G-Ops)
AlexNet	39,7%	18,9%	8	70 M
ZF Net	37,50%	14,8%	8	70 M
VGG Net	25,60%	8,10%	19	155 M
GoogLeNet	29,00%	9,20%	22	10 M
Inception-v3	21,20%	5,60 %	101	35 M
Inception-v4	20,00%	5 %	152	35 M
Inception-ResNet-v2	19,90%	4,90%	467	65 M
ResNet-152	19,38%	4,49 %	152	65 M

#### 3.3.1. AlexNet

Jaringan syaraf pertama yang memenangkan kompetisi klasifikasi gambar ILSVRC adalah AlexNet [29] pada tahun 2012, dengan error ke-5 teratas sebesar 15,31%. Sebagai perbandingan, metode yang tidak menggunakan jaringan syaraf tiruan convolutional mendapatkan kesalahan klasifikasi sebesar 26,1%. AlexNet mengumpulkan teknologi terbaru pada saat itu untuk meningkatkan jaringan. Kemudian pada penelitian Eldem at al [30] tahun 2023 mengembangkan Alexnet. Mengusulkan enam variasi arsitektur AlexNet yang berbeda dengan jumlah lapisan konvolusi, pooling, dan aktivasi ReLU yang berbeda untuk klasifikasi gambar luka. Variasi ini mencakup 3Conv\_Softmax, 3Conv\_SVM, 4Conv\_Softmax, 4Conv\_SVM, 6Conv\_Softmax, dan 6Conv\_SVM. Dengan membandingkan kinerja model yang diusulkan dengan metode *Deep Learning (DL)* tradisional pada dataset yang sama. Hasilnya menunjukkan bahwa model 6Conv\_SVM adalah yang paling sukses, mencapai akurasi 98,85%, sensitivitas 98,86%, dan spesifisitas 99,42%.

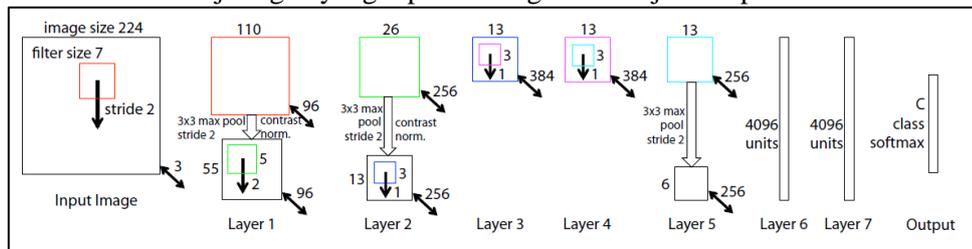


Gambar 2. Arsitektur Alexnet

#### 3.3.2. ZF Net

Jaringan syaraf tiruan convolutional ZF Net [31] adalah pemenang ILSVRC 2013 dengan error ke-5 teratas sebesar 14,8%. Pencapaian utama dari arsitektur ini adalah penciptaan teknik visualisasi filter (jaringan dekonvolusional). Untuk mempelajari perilaku filter pada gambar tertentu menggunakan jaringan saraf terlatih, terlebih dahulu membuat output jaringan, di

lapisan filter yang dipelajari semua bobot = 0, kecuali bobot filter itu sendiri, dan kemudian menerapkan aktivasi yang dihasilkan ke jaringan. Jaringan secara konsisten menggunakan operasi *Unpooling ReLU* dan penyaringan. Operasi *Unpooling* mengembalikan sebagian input dari lapisan sub-sampling yang sesuai dengan mengingat koordinat yang telah dipilih oleh lapisan sub-sampling. Operasi ReLU adalah lapisan biasa yang menggunakan fungsi ReLU. Lapisan pemfilteran melakukan operasi konvolusi dengan bobot dari lapisan konvolusi yang sesuai, tetapi bobot setiap filter "dibalik" secara vertikal dan horizontal. Dengan demikian, aktivasi awal filter bergerak pada arah yang berlawanan sampai ditampilkan dalam ruang gambar asli. Arsitektur jaringan yang dipertimbangkan ditunjukkan pada Gambar berikut ini:



Gambar 3. Arsitektur Alexnet

### 3.3.3. VGG Net

*VGG Net* adalah model jaringan syaraf tiruan convolutional yang memenangkan kompetisi klasifikasi gambar tahun 2014[32]. *VGG Net* yang mengekstrak fitur visual multi-skala dan berbutir halus, yang menunjukkan kinerja kompetitif dalam klasifikasi gambar. Dalam jaringan ini, mereka menolak untuk menggunakan filter yang lebih besar dari 3x3. Karena dalam penelitian membuktikan bahwa lapisan filter 7x7 setara dengan tiga lapisan dengan filter 3x3, dan dalam hal ini 55% lebih sedikit parameter yang digunakan. Demikian pula, lapisan filter 5x5 setara dengan dua lapisan dengan filter 3x3, yang menghemat 22% parameter jaringan. Fitur arsitektur dan organisasi internal jaringan syaraf ini ditunjukkan pada Gambar 4.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

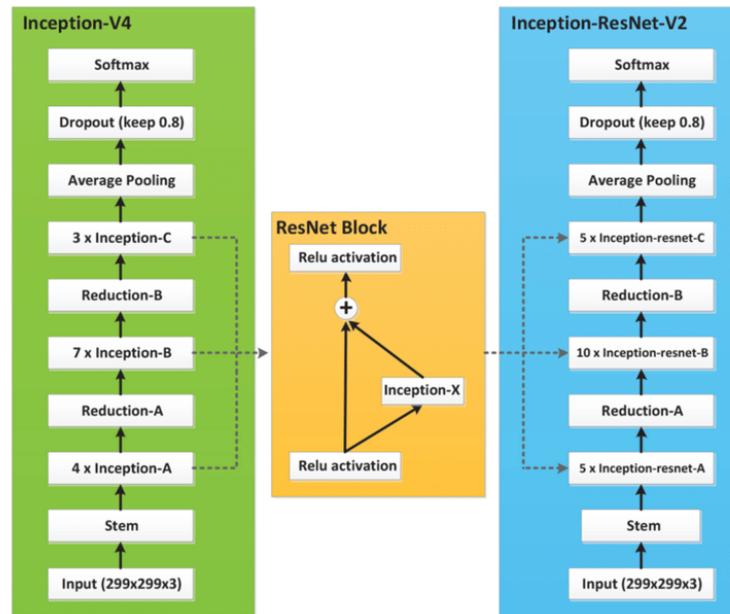
Gambar 4. Variasi yang berbeda dari arsitektur jaringan syaraf tiruan convolutional VGG Net

### 3.3.4. Inception (GoogleNet)



### 3.3.6. Inception-v4 and Inception-ResNet

Setelah keberhasilan penerapan jaringan saraf konvolusional ResNet, versi jaringan Inception berikut diperkenalkan: Inception-v4 dan Inception-ResNet. Dalam kedua kasus tersebut, modul Inception dibagi menjadi modul A, B, dan C. Pada Inception-v4, inovasi utama adalah penggantian *Max Pooling* dengan *Average Pooling* pada *modul Inception* itu sendiri. Untuk Inception-ResNet, melewati koneksi telah ditambahkan ke modul Inception. Dua versi jaringan dirancang *Inception-ResNet-v1*, yang memerlukan lebih sedikit komputasi, dan *Inception-ResNet-v2* masing-masing 35x35, 17x17, dan 8x8. McNeely-White 2019 [38] menemukan bahwa Inception dan ResNet memiliki karakteristik yang sama dalam hal ekstraksi fitur dari gambar. Fitur arsitektur jaringan saraf ini ditunjukkan pada Gambar 7.



Gambar 7. Arsitektur Inception-v4 and Inception-ResNet

### 3.4. Trend CNN Terkini untuk Pemrosesan Gambar

Penggunaan Convolutional Neural Networks (CNN) dalam pemrosesan gambar adalah bidang yang berkembang pesat, dengan fokus pada kemajuan arsitektur dan aplikasi praktis. Berikut ini penjelasan tentang penelitian dengan pengembangan metode CNN dalam pemrosesan gambar yang diambil dari jurnal bereputasi. Hal ini dapat dijadikan sebagai referensi untuk pengembangan penelitian lebih lanjut.

No	Tahun & Jurnal	Penulis	Metode	Dataset	Deskripsi
1. [39]	2021, Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing	Qimin Cheng, Yuzhuo Zhou , Peng Fu , Yuan Xu, and Liang Zhang	Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), dan Semantic Alignment Modul (SAM)	UCMerced-LandUse-Captions, Sydney-Captions, RSICD, dan NWPU-RESISC45-Captions.	Metode ini berhasil mengatasi kompleksitas semantik dan meningkatkan korespondensi antara gambar pemantauan jarak jauh dan teks dengan memperbarui fitur visual dan tekstual.
2. [11]	2021, IEEE Transactions	Weiwei, Shutao Li	Deep Hashing convolutional	UCMD, WHU-RS, dan AID	Metode ini mampu mengatasi beberapa tantangan dalam

	On Geoscience And Remote Sensing	and Jón Atli Benediktsson	neural network (DHCNN)		pengambilan gambar jarak jauh yang visual dan semantik, menghasilkan hasil pengambilan gambar yang lebih baik dengan resolusi objek yang sulit dan jarak jauh. Selain itu, DHCNN juga mampu melakukan klasifikasi label semantik dari gambar yang serupa secara akurat.
3. [40]	2021, IEEE Acces	Surajit Saikia, Laura Fernández-Robles, Eduardo Fidalgo Fernández, Enrique Alegre	CNN Features and Colour Models	COIL-100, INSTRE-M, Paris 6K, dan Revisiting Paris 6K.	Metode ini berhasil menghasilkan deskriptor neural warna yang efektif untuk instance retrieval. Mengatasi tantangan dalam mengambil gambar yang identik, terutama dalam kasus di mana objek memiliki penampilan yang mirip tetapi warna, sudut pandang, atau kondisi pencahayaan yang berbeda.
4. [41]	2023, IEEE Acces Digital Object Identifier	Shutao Tan, Liang Dong, Min Zhang, Ye Zhang	Content-based image retrieval (CBIR) dan Deep Learning CNN. Model pre-trained Mask R-CNN, Konvolusi ResNet atau VGG	Textile images displayed at the International Textile Fabric Expo in China Textile City.	Metode ini berhasil meningkatkan akurasi pengambilan gambar tekstil, terutama dalam mengatasi gangguan latar belakang dan fitur lokal yang berulang dalam gambar tekstil.
5. [42]	2019, IEEE Acces Digital Object Identifier	Famao Ye, Meng Dong, Wei Luo, Xiaoyong Chen, Weidong Min	Convolutional Neural Network (CNN, Ekstraksi fitur manual : Zernike Moments, Haralick, dan Global-Local Pyramid Pattern (GLPP)	University of California, Merced dataset (UCMD) dan PatternNet	Pengembangan metode re-ranking berbasis jaringan saraf konvolusional (CNN) untuk pengambilan gambar berbasis konten untuk citra penginderaan jauh. Metode ini menggunakan dua jarak gambar-ke-kelas: jarak gambar-ke-kelas pelatihan dan jarak gambar-ke-kelas kueri
6. [43]	2022, IEEE Geoscience and Remote Sensing Letters	Famao Ye, Wei Luo, Meng Dong, Dajun Li,	CNN with Content-based remote sensing image retrieval (RSIR)	University of California, Merced data set (UCMD)	Pengembangan metode content-based remote sensing image retrieval (RSIR) menggunakan aturan fuzzy dan jarak

		and Weidong Min	menggunakan aturan fuzzy dan jarak fuzzy		fuzzy. Metode ini bertujuan untuk mengurangi waktu pencarian secara keseluruhan dengan mempertahankan kinerja pengambilan gambar yang tidak terpengaruh.
7. [44]	2022, IEEE Transactions on Computers	Yingying Li, Jianfeng Ma , Yinbin Miao , Huizhong Li, Qiang Yan, YueWang , Ximeng Liu, and Kim-Kwang Raymond Choo	pre-trained Convolutional Neural Network (CNN) with an encrypted index based on the K-means clustering algorithm to improve retrieval efficiency	Caltech256 database	Pengembangan skema Dynamic Verifiable Retrieval over Encrypted Images (DVREI) yang bertujuan untuk meningkatkan akurasi, efisiensi, dan verifikasi hasil pencarian gambar terenkripsi. Temuan utama meliputi penggunaan pre-trained Convolutional Neural Network (CNN) untuk mengekstrak fitur gambar, konstruksi indeks terenkripsi berdasarkan algoritma pengelompokan K-means untuk meningkatkan efisiensi pencarian.

#### 4. Simpulan dan Penutup

Dalam studi literatur ini, beberapa penelitian membahas bahwa CNN telah terbukti efektif mengatasi kompleksitas semantik dan meningkatkan korespondensi antara gambar dengan resolusi objek yang sulit bahkan gambar dengan bentuk yang mirip tetapi warna, sudut pandang, atau kondisi pencahayaan yang berbeda. CNN juga berhasil mengatasi gangguan latar belakang gambar dan fitur lokal yang berulang. Namun CNN juga memiliki kelemahan dalam mengekstrak gambar dengan tingkat resolusi yang rendah yang mengakibatkan rentan terhadap overfitting karena jumlah parameter yang tinggi digunakan dalam model. CNN juga membutuhkan jumlah data yang besar untuk melatih model yang efektif sehingga memerlukan komputasi yang besar, terutama ketika digunakan pada gambar dengan resolusi tinggi atau dalam arsitektur yang lebih dalam. Hal Ini membuat proses pelatihan dan inferensi menjadi mahal dari segi komputasi, membutuhkan sumber daya perangkat keras yang lebih kuat. Sebagai rekomendasi untuk pengembangan model jaringan saraf konvolusional yang paling efektif untuk ekstraksi gambar, klasifikasi gambar, dan analisis fitur arsitekturalnya maka disarankan melakukan lebih banyak lagi percobaan untuk mengeneralisasi dan menganalisis hasil efisiensi penggunaan jaringan saraf dengan metode atau algoritma lainnya (Tabel 1 dan Tabel 2) untuk menemukan metode terbaik dalam mengatasi kelemahan CNN dalam pemrosesan gambar.

#### References

- [1] F. Akbar, A. Hidayatno, dan A. Triwiyatno, "Perancangan Program Pengenalan Isyarat Tangan Dengan Metode Convolutional Neural Network (CNN)," *Transient: Jurnal Ilmiah Teknik Elektro*, vol. 9, no. 1, hlm. 26–36, Mar 2020, doi: 10.14710/transient.v9i1.26-36.

- [2] N. Liu, "A Novel Intelligent Image Recognition Scheme based on Fully Convolutional Neural Network," dalam *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Jun 2023, hlm. 945–949. doi: 10.1109/ICCES57224.2023.10192732.
- [3] V. N. Satya dan S. Chimakurthi, "Application of Convolution Neural Network for Digital Image Processing," *Engineering International*, vol. 8, no. 2, 2020.
- [4] S. Coleman, D. Kerr, dan Y. Zhang, "Image Sensing and Processing with Convolutional Neural Networks," *Sensors*, vol. 22, no. 10. MDPI, 1 Mei 2022. doi: 10.3390/s22103612.
- [5] R. Mawan, "Klasifikasi motif batik menggunakan Convolutional Neural Network," *JNANALOKA*, hlm. 45–50, Apr 2020, doi: 10.36802/jnanaloka.2020.v1-no1-45-50.
- [6] T. Bariyah, M. A. Rasyidi, dan N. Ngatini, "Convolutional Neural Network untuk Metode Klasifikasi Multi-Label pada Motif Batik," *Techno.Com*, vol. 20, no. 1, hlm. 155–165, Feb 2021, doi: 10.33633/tc.v20i1.4224.
- [7] P. Nugroho, I. Fenriana, dan R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia," 2020. [Daring]. Tersedia pada: <https://api.semanticscholar.org/CorpusID:228841859>
- [8] I. A. Sabilla, "Arsitektur Convolutional Neural Network (CNN) untuk Klasifikasi Jenis dan Kesegaran Buah pada Neraca Buah," 2020. [Daring]. Tersedia pada: <https://api.semanticscholar.org/CorpusID:229184929>
- [9] A. S. S. Pratama, A. Prasetya Wibawa, dan A. Nur Handayani, "Convolutional Neural Network (CNN) Untuk Menentukan Gagrak Wayang Kulit," *Jurnal Mnemonic*, vol. 5, no. 2, hlm. 98–102, Agu 2022, doi: 10.36040/mnemonic.v5i2.4671.
- [10] K. Azmi, S. Defit, dan S. Sumijan, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," *JURNAL UNITEK*, vol. 16, no. 1, hlm. 28–40, Jun 2023, doi: 10.52072/unitek.v16i1.504.
- [11] W. Song, S. Li, dan J. A. Benediktsson, "Deep Hashing Learning for Visual and Semantic Retrieval of Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, hlm. 9661–9672, Nov 2021, doi: 10.1109/TGRS.2020.3035676.
- [12] A. Marzali, "Menulis Kajian Literatur," *Jurnal Etnografi Indonesia*, 2016.
- [13] D. Bhatt *dkk.*, "CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope," *Electronics (Basel)*, vol. 10, no. 20, hlm. 2470, Okt 2021, doi: 10.3390/electronics10202470.
- [14] T. Buzási, "Image processing in the military technology," 2003. [Daring]. Tersedia pada: <https://api.semanticscholar.org/CorpusID:110195378>
- [15] M. ITOH dan L. O. CHUA, "IMAGE PROCESSING AND SELF-ORGANIZING CNN," *International Journal of Bifurcation and Chaos*, vol. 15, no. 09, hlm. 2939–2958, Sep 2005, doi: 10.1142/S0218127405013794.
- [16] A. Bal, O. N. Ucan, H. Pastaci, dan M. S. Alam, "Wavelet-cellular neural network architecture and learning algorithm," H. H. Szu, M. V. Wickerhauser, B. A. Pearlmutter, dan W. Sweldens, Ed., Apr 2004, hlm. 159. doi: 10.1117/12.542353.
- [17] Y. Bengio, "Deep Learning of Representations: Looking Forward," Mei 2013.
- [18] C. Szegedy, S. Ioffe, V. Vanhoucke, dan A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning."

- [19] T. Tong, G. Li, X. Liu, dan Q. Gao, "Image Super-Resolution Using Dense Skip Connections," dalam *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, hlm. 4809–4817. doi: 10.1109/ICCV.2017.514.
- [20] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, dan E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J Big Data*, vol. 2, no. 1, Des 2015, doi: 10.1186/s40537-014-0007-7.
- [21] K. Kawaguchi, J. Huang, dan L. P. Kaelbling, "Effect of Depth and Width on Local Minima in Deep Learning," *Neural Comput*, vol. 31, no. 7, hlm. 1462–1498, Jul 2019, doi: 10.1162/neco\_a\_01195.
- [22] Y. LeCun, K. Kavukcuoglu, dan C. Farabet, "Convolutional networks and applications in vision," dalam *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010, hlm. 253–256. doi: 10.1109/ISCAS.2010.5537907.
- [23] F. Ye, H. Xiao, X. Zhao, M. Dong, W. Luo, dan W. Min, "Remote sensing image retrieval using convolutional neural network features and weighted distance," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 10, hlm. 1535–1539, Okt 2018, doi: 10.1109/LGRS.2018.2847303.
- [24] W. Zhou, H. Guan, Z. Li, Z. Shao, dan M. R. Delavar, "Remote Sensing Image Retrieval in the Past Decade: Achievements, Challenges, and Future Directions," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 16, hlm. 1447–1473, 2023, doi: 10.1109/JSTARS.2023.3236662.
- [25] Q. Qin, L. Huang, Z. Wei, K. Xie, dan W. Zhang, "Unsupervised Deep Multi-Similarity Hashing with Semantic Structure for Image Retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 7, hlm. 2852–2865, Jul 2021, doi: 10.1109/TCSVT.2020.3032402.
- [26] S. Iqbal, A. N. Qureshi, M. Alhussein, I. A. Choudhry, K. Aurangzeb, dan T. M. Khan, "Fusion of Textural and Visual Information for Medical Image Modality Retrieval Using Deep Learning-Based Feature Engineering," *IEEE Access*, vol. 11, hlm. 93238–93253, 2023, doi: 10.1109/ACCESS.2023.3310245.
- [27] SAGAR SHARMA, "Activation Functions in Neural Networks," towardsdatascience.COM.
- [28] A. Arsenov, I. Ruban, K. Smelyakov, dan A. Chupryna, "Evolution of Convolutional Neural Network Architecture in Image Classification Problems."
- [29] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." [Daring]. Tersedia pada: <http://code.google.com/p/cuda-convnet/>
- [30] H. Eldem, E. Ülker, dan O. Y. Işıklı, "Alexnet architecture variations with transfer learning for classification of wound images," *Engineering Science and Technology, an International Journal*, vol. 45, Sep 2023, doi: 10.1016/j.jestch.2023.101490.
- [31] M. D. Zeiler dan R. Fergus, "Visualizing and Understanding Convolutional Networks," Nov 2013, [Daring]. Tersedia pada: <http://arxiv.org/abs/1311.2901>
- [32] K. Simonyan dan A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep 2014, [Daring]. Tersedia pada: <http://arxiv.org/abs/1409.1556>
- [33] C. Szegedy dkk., "Going deeper with convolutions," dalam *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun 2015, hlm. 1–9. doi: 10.1109/CVPR.2015.7298594.

- [34] K. Gao, Q. Zhang, dan H. Wang, “A Lightweight Residual-Inception Convolutional Neural Network,” *J Phys Conf Ser*, vol. 1237, no. 3, hlm. 032058, Jun 2019, doi: 10.1088/1742-6596/1237/3/032058.
- [35] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, dan V. K. Asari, “Inception recurrent convolutional neural network for object recognition,” *Mach Vis Appl*, vol. 32, no. 1, hlm. 28, Jan 2021, doi: 10.1007/s00138-020-01157-3.
- [36] Z. Shen dan Y. Liu, “A novel connectivity of deep convolutional neural networks,” dalam *2017 Chinese Automation Congress (CAC)*, IEEE, Okt 2017, hlm. 7779–7783. doi: 10.1109/CAC.2017.8244187.
- [37] S. Bose dan A. Dey, “ResCNN: An alternative implementation of Convolutional Neural Networks,” dalam *2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, IEEE, Nov 2021, hlm. 1–5. doi: 10.1109/UPCON52273.2021.9667654.
- [38] D. G. McNeely-White, J. R. Beveridge, dan B. A. Draper, “Inception and ResNet: Same Training, Same Features,” 2020, hlm. 352–357. doi: 10.1007/978-3-030-25719-4\_45.
- [39] Q. Cheng, Y. Zhou, P. Fu, Y. Xu, dan L. Zhang, “A Deep Semantic Alignment Network for the Cross-Modal Image-Text Retrieval in Remote Sensing,” *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, hlm. 4284–4297, 2021, doi: 10.1109/JSTARS.2021.3070872.
- [40] S. Saikia, L. Fernandez-Robles, E. F. Fernandez, dan E. Alegre, “Colour Neural Descriptors for Instance Retrieval Using CNN Features and Colour Models,” *IEEE Access*, vol. 9, hlm. 23218–23234, 2021, doi: 10.1109/ACCESS.2021.3056330.
- [41] S. Tan, L. Dong, M. Zhang, dan Y. Zhang, “Fine-Grained Retrieval Method of Textile Image,” *IEEE Access*, vol. 11, hlm. 70525–70533, 2023, doi: 10.1109/ACCESS.2023.3287630.
- [42] F. Ye, M. Dong, W. Luo, X. Chen, dan W. Min, “A New Re-Ranking Method Based on Convolutional Neural Network and Two Image-to-Class Distances for Remote Sensing Image Retrieval,” *IEEE Access*, vol. 7, hlm. 141498–141507, 2019, doi: 10.1109/ACCESS.2019.2944253.
- [43] F. Ye, W. Luo, M. Dong, D. Li, dan W. Min, “Content-Based Remote Sensing Image Retrieval Based on Fuzzy Rules and a Fuzzy Distance,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, 2022, doi: 10.1109/LGRS.2020.3030858.
- [44] Y. Li dkk., “DVREI: Dynamic Verifiable Retrieval Over Encrypted Images,” *IEEE Transactions on Computers*, vol. 71, no. 8, hlm. 1755–1769, Agu 2022, doi: 10.1109/TC.2021.3106482.