

# PENGENALAN NAMA MODEL SEPATU MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN MODEL MOBILENET

Renaldi Bong<sup>1)</sup> Chairisni Lubis<sup>2)</sup>

<sup>1)2)</sup> Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara  
Jl. Letjen S. Parman No.1, Jakarta  
email : renaldibong@gmail.com<sup>1)</sup>, chairisnil@fti.untar.ac.id<sup>2)</sup>

## ABSTRAK

Aplikasi atau sistem pengenalan merek beserta nama model sepatu ini berguna untuk memberikan informasi tentang nama model sepatu apa yang ingin dicari. Aplikasi ini diharapkan dapat membantu masyarakat yang kurang paham atau kurang mengikuti trend dan perkembangan sepatu. Sehingga dapat memudahkan untuk mengetahui apa nama model sepatu yang dimiliki dan dari informasi tersebut dapat digunakan untuk keperluan lainnya. Salah satu cabang ilmu atau metode yang dapat digunakan untuk membuat sistem ini adalah Deep Learning. Kemudian pada Deep Learning ini terdapat arsitektur yang disebut Convolutional Neural Network. Nantinya sistem ini akan dilatih menggunakan salah satu model arsitektur yang terdapat pada Convolutional Neural Network ini yaitu MobileNet dan versi MobileNet yang digunakan adalah versi 2 atau MobileNetV2. Dari hasil pelatihan dengan model MobileNetV2 diperoleh akurasi kurang lebih 75,23%. Dan hasil pengujian mendapatkan hasil akurasi sekitar 66,67%.

## Kata Kunci

*Computer Vision, Convolutional Neural Network, Deep Learning, MobileNet, Sepatu.*

## 1. Pendahuluan

Salah satu kebutuhan utama manusia adalah sepatu, yang dapat melindungi kaki dari berbagai kondisi lingkungan. Sepatu juga dapat digunakan untuk meningkatkan penampilan dan gaya hidup sehari-hari. Saat ini, industri sepatu di Indonesia berkembang pesat karena munculnya berbagai merek sepatu baru, yang menghasilkan persaingan yang ketat di antara produsen sepatu. Dengan banyaknya jenis desain sepatu dan nama model sepatu yang dibuat, pastinya tidak semua orang terutama orang yang tidak terlalu mengikuti perkembangan sepatu dapat mengenali dan menghafalkan nama-nama model sepatu mereka.

Salah satu contoh masalah yang dapat ditimbulkan adalah ketika seseorang ingin membeli atau mencari

suatu model sepatu di toko online, tetapi tidak mengetahui nama model sepatu tersebut. Masalah tersebut pastinya cukup merepotkan karena salah satu cara untuk bisa mengetahui nama model sepatu tersebut adalah dengan datang ke toko sepatu secara langsung. Di zaman yang serba online ini, pastinya jika harus pergi ke toko sepatu langsung cukup memakan waktu apalagi jika sedang sibuk dan tidak memiliki banyak waktu.

Sehingga, diperlukan sebuah sistem yang dapat mengenali merek dan nama model sepatu ini agar dapat membantu orang awam mengetahui nama model sepatunya. Salah satu pendekatan yang dapat digunakan adalah dengan menggunakan atau membuat sebuah sistem pengenalan merek sepatu secara otomatis. Sistem ini dapat digunakan untuk mengenali merek sepatu dari gambar sepatu yang diberikan atau ditunjukkan pada suatu kamera. Selain itu juga dapat memberitahu nama model sepatu yang dideteksi.

Salah satu cabang ilmu yang dapat diterapkan dalam perancangan sistem ini adalah Computer Vision. Computer Vision adalah salah satu bidang kecerdasan buatan yang mempelajari cara komputer untuk memahami dan memanipulasi gambar. Computer Vision memiliki banyak manfaat dalam kehidupan sehari-hari, seperti pengenalan wajah, gerakan, dan juga salah satunya objek. Convolutional Neural Network (CNN) adalah salah satu teknik Computer Vision yang paling banyak dikembangkan. CNN adalah algoritma machine learning yang dapat mendeteksi gambar dan video. CNN juga merupakan salah satu algoritma machine learning yang memiliki tingkat akurasi tinggi dalam melakukan pengenalan pada suatu objek.

CNN juga memiliki beberapa model yang cukup sering atau banyak digunakan, salah satunya adalah MobileNet. MobileNet adalah model CNN yang dirancang untuk perangkat mobile, namun juga bisa untuk perangkat non-mobile seperti komputer atau laptop. Model MobileNet ini juga memiliki tingkat akurasi yang cukup akurat, sehingga dapat digunakan untuk berbagai aplikasi pengenalan objek, seperti salah satunya pengenalan merek sepatu. Ukuran MobileNet ini juga relatif kecil dan ringan, sehingga memungkinkan penggunaan memori dan daya yang lebih sedikit.

Perancangan sistem ini akan menggunakan CNN dengan model MobileNet untuk membangun sistem pengenalan merek sepatu secara otomatis. Dengan menggunakan model MobileNet ini, diharapkan nantinya sistem ini dapat mengenali merek sebuah sepatu secara real time berdasarkan bentuk dan model sepatunya dari kamera. Dalam pembuatan dan pelatihan sistem ini, data yang akan digunakan berupa gambar-gambar model sepatu dari 9 jenis merek sepatu yang berbeda. Merek tersebut adalah Adidas, Asics, Converse, Crocs, New Balances, Nike, Puma, Reebok, Skechers dimana setiap merek sepatu terdapat 3 model yang dapat dikenali dan dapat dilihat pada Lampiran 1.

Diharapkan bahwa sistem yang dirancang akan membantu orang-orang yang kurang paham tentang model-model sepatu untuk menemukan atau mengetahui nama model dari sepatu yang mereka miliki.

## 2. Dasar Teori

### 2.1 Convolutional Neural Network (CNN)

Salah satu algoritma dalam *Deep Learning* yang dapat digunakan untuk mendeteksi atau mengenali suatu objek adalah *Convolutional Neural Network (CNN)*. CNN juga dapat digunakan untuk tugas-tugas yang berhubungan dengan gambar, seperti klasifikasi gambar ataupun segmentasi gambar. CNN biasanya digunakan untuk mengolah data dua dimensi.

CNN mirip dengan saringan pasir, yang dapat digunakan untuk memisahkan pasir dari batu dan kotoran. Ini bekerja dengan menyaring pasir melalui lubang-lubang kecil. CNN juga bekerja dengan cara yang mirip. CNN akan menyaring gambar melalui filter kecil, yang membantu CNN untuk mengidentifikasi aspek tertentu dari sebuah gambar atau citra.

### 2.2 Input Layer

Ini adalah lapisan di mana data dimasukkan ke dalam model CNN. Sebagai input, CNN biasanya menerima rangkaian foto atau gambar yang belum diproses dengan dimensi tertentu.

### 2.3 Convolutional Layer

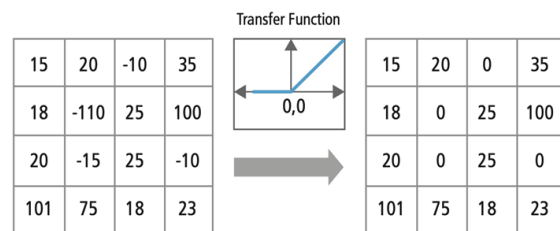
Lapisan ini digunakan untuk mengekstraksi fitur dari kumpulan data input. Matriks yang lebih kecil, biasanya berukuran 2x2, 3x3, atau 5x5 pada gambar masukan dan memiliki filter yang dapat dilatih. Filter yang dihasilkan oleh convolutional layer ini disebut juga Feature Map [1]. Rumus dari *Convolutional Layer* ini adalah sebagai berikut:

$$(I * F)_{m,n} = \sum_{x=0}^{f_1-1} \sum_{y=0}^{f_2-1} I(m+x, n+y) * F(x,y) \quad (1)$$

Dengan I merupakan nilai untuk *input* dan F merupakan nilai untuk filter.

### 2.4 Fungsi Aktivasi

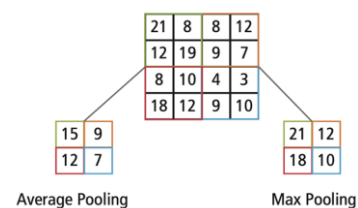
Fungsi Aktivasi ini memiliki fungsi untuk meningkatkan jaringan nonlinier dengan menambahkan suatu fungsi yang disebut fungsi aktivasi ke *output* lapisan sebelumnya. Sebuah contoh fungsi aktivasi yang umum digunakan yakni adalah fungsi aktivasi yang dikenal dengan istilah fungsi Relu. ReLU memiliki rumus  $f(x)=\max(0,x)$  yang mana akan merubah ukuran suatu piksel menjadi 0 apabila nilai piksel tersebut bernilai negatif atau kurang dari 0 [1]. Gambaran fungsi ReLU berdasarkan penjelasan sebelumnya dapat dilihat pada gambar di bawah ini:



Gambar 1. Fungsi ReLU

### 2.5 Pooling Layer

*Pooling Layer* ini bertujuan untuk menurunkan volume atau fitur map yang telah dihasilkan pada lapisan sebelumnya. Sehingga bisa mempercepat komputasi, menghemat memori, dan mencegah terjadinya adanya *overfitting*. *Average Pooling* dan *Max Pooling* adalah dua jenis *Pooling Layer* yang cukup umum digunakan dalam CNN ini. Pada *average pooling*, akan melakukan perhitungan berdasarkan nilai rata-rata dari suatu jendela atau matriks tertentu dari suatu citra. Sedangkan *max pooling* akan mencari nilai terbesar atau maksimum dari suatu jendela atau matriks pada citra [1]. Untuk gambaran yang lebih jelas, bisa dilihat pada gambar berikut ini:



Gambar 2. Average Pooling dan Max Pooling

### 2.6 Fully Connected Layer

Pada lapisan *Fully Connected Layer*, setiap neuron pada lapisan ini akan dihubungkan atau diproses pada setiap neuron yang ada pada lapisan sebelumnya. Kemudian setelah proses telah selesai dilakukan, maka lapisan ini akan menghasilkan sebuah hasil atau *output* [2].

### 2.7 MobileNet

MobileNet merupakan salah satu model dari Convolutional Neural Network (CNN) yang cukup efisien dan ringan terutama untuk perangkat mobile yang memiliki sumber daya terbatas. Namun MobileNet juga dapat digunakan di perangkat non-mobile misalnya seperti laptop. MobileNet dibuat oleh Google pada tahun 2017. Dan itu merupakan versi MobileNet yang pertama atau MobileNetV1. Setahun kemudian, muncul versi MobileNetV1 yang telah ditingkatkan yaitu MobileNetV2. Dan yang terakhir adalah MobileNetV3 yang diluncurkan pada tahun 2019.

MobileNet menggunakan dua buah konvolusi yang dinamakan *depthwise convolution* dan *pointwise convolution* [3]. Kemudian gabungan dari kedua konvolusi tersebut menciptakan sebuah layer yang disebut Depthwise separable convolutions yang merupakan layer utama untuk MobileNet. Arsitektur MobileNet dapat dilihat pada gambar di bawah ini:

Type / Stride	Filter Shape	Input Size	
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3	
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32	
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32	
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64	
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64	
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128	
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128	
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128	
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128	
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256	
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256	
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256	
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256	
5×	Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
	Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512	
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512	
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024	
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024	
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024	
FC / s1	1024 × 1000	1 × 1 × 1024	
Softmax / s1	Classifier	1 × 1 × 1000	

Gambar 3. Arsitektur MobileNet

### 3. Hasil Percobaan

Percobaan atau pengujian dilakukan dengan menggunakan model yang mendapatkan akurasi pelatihan tertinggi. Pelatihan dilakukan dengan mengambil gambar-gambar sepatu yang ada pada Google Images. Total gambar yang didapat sekitar 5177 gambar dengan komposisi 4142 (80%) digunakan untuk pelatihan dan 1035 (20%) untuk validasi dengan jumlah kelas sebanyak 27 kelas.

Perbandingan akurasi pada beberapa percobaan *epoch* dan *batch size* dapat dilihat pada tabel berikut ini:

Tabel 1 Perbandingan akurasi pelatihan

Epoch	Batch Size	Akurasi Pelatihan
10	50	36,61 %
50	16	60,80 %

50	50	73,43%
50	64	75,23%

Dengan menggunakan 50 *epoch* dan 64 *batch size*, didapatkan hasil pengujian sebesar 66,67% atau sekitar 18 model sepatu berhasil dikenali dengan benar dan 9 model sepatu salah dikenali.

Model sepatu yang berhasil dikenali ketika diuji adalah Adidas Ultrabounce Women, Adidas EqT Support ADV, Adidas Yeezy 350, Asics Jolt 3, Asics Metarise, Asics Tri Noosa 13, Converse Chuck Taylor All Star Hi Top, Converse Chuck Taylor Low, Converse Run Star Hike, Crocs Classic Clog, Crocs Mega Crush, New Balance 990, Nike Air Force 1 Low, Nike Air Max 97, Puma Interflex Modern, Reebok Royal BB4500, Skechers D'Lites, Skechers Statement.

Dan model sepatu yang salah ketika diuji adalah Crocs Literide, New Balance 420, New Balance FuelCell, Nike Air Jordan, Puma Relax Knit Slip On, Puma Suede, Reebok Royal Complete 3, Reebok Zig Dynamica, Skechers Go Walk.

### 4. Kesimpulan

Kemudian dari hasil pengujian yang dilakukan sebelumnya, didapatkan kesimpulan berupa:

1. Model MobileNetV2 dapat digunakan untuk membuat aplikasi pengenalan objek (merek dan model sepatu) untuk perangkat mobile.
2. Dengan menggunakan model MobileNetV2 ini, akurasi yang didapatkan untuk pengenalan merek dan model sepatu adalah 66,67% dengan jumlah epoch dan batch size yang digunakan untuk pelatihan adalah 50 epoch dan 64 batch size.
3. Pengenalan merek dan model sepatu ini dapat dilakukan secara real-time menggunakan kamera ponsel pada perangkat mobile atau pengenalan juga dapat menggunakan perangkat desktop dengan mengakses situs <https://skripsi-535200045.netlify.app>.
4. Untuk hasil yang lebih maksimal, jumlah sepatu yang dapat dikenali atau ditunjukkan pada kamera hanya 1 buah sepatu.
5. Model sepatu yang memiliki bentuk atau motif yang mirip, dapat menyebabkan kesalahan saat melakukan pengenalan.
6. Warna sepatu juga dapat menyebabkan kesalahan dalam melakukan pengenalan merek dan nama model sepatu.
7. Posisi pengambilan gambar saat melakukan proses pengenalan juga memiliki pengaruh terhadap hasil akhir yang diberikan.

### REFERENSI

- [1] Hibatullah, Alwan., Maliki, Irfan., "Penerapan Metode Convolutional Neural Network pada Pengenalan Pola Citra Sandi Rumput", Universitas Komputer Indonesia, Bandung.

- [2] Sultana, Farhana., Sufian, A., Dutta, Paramartha., Mei 2019, “Advancements in Image Classification using Convolutional Neural Network”, University of Gour Banga, India.
- [3] Ramadhan, Farid Evan., 2020, “Penerapan Image Classification dengan Pre-Trained model MobileNet dengan Client-Side Machine Learning”, Universitas Islam Negeri Syarif Hidayatullah, Jakarta.

**Renaldi Bong**, mahasiswa pada proram studi Fakultas Teknologi Informasi di Universitas Tarumanagara.

**Dra. Chairisni Lubis, M.Kom.**, memperoleh gelar Dra. Dari Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia dan gelar M.Kom dari Universitas Indonesia. Saat ini sebagai Dosen program studi Teknik Informatika, Universitas Tarumanagara.