

# Integrasi Metode Convolutional Neural Networks dengan Arsitektur Model PoseNet untuk Pengembangan Sistem Klasifikasi Gerakan serta Monitoring Repetisi pada Olahraga Bulu Tangkis

Benny Karnadi <sup>1)</sup> Chairisni Lubis <sup>2)</sup> Agus Budi Dharmawan <sup>3)</sup>

<sup>1)2)</sup> Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara  
Jl. Letjen S. Parman No.1, Jakarta  
email : benny.535200044@stu.untar.ac.id<sup>1)</sup>, chairisnil@fti.untar.ac.id <sup>2)</sup> agusd@fti.untar.ac.id <sup>3)</sup>

## ABSTRACT

*This application or a recognition of movement classification system and monitoring repetitions in badminton system is designed for students and coaches to practice movement and stroke techniques in badminton as well as a joint evaluation medium for users who practice using this application, also with the hope that it can help children who want to practice Badminton can be more flexible in terms of time and also efficient for coaches to be able to reach a wider range of students and produce more talented athletes. One of the branches of science used in designing this application is Deep Learning with the Convolutional Neural Network (CNN) method with the MobileNetv2 architecture used in designing this badminton movement classification application, as well as using PoseNet model integration. The training results that can be achieved using the Convolutional Neural Network method with the MobileNetv2 architecture obtain an accuracy score in the range of 90%, and test results can be achieved with an accuracy score of 93%.*

## Key words

*Deep Learning, Convolutional Neural Network, MobileNet, PoseNet, Badminton.*

## 1. Pendahuluan

Olahraga merupakan bagian penting dalam menjaga kesehatan fisik dan mental manusia. Olahraga bulu tangkis merupakan salah satu olahraga yang terkenal di dunia, Indonesia menjadi salah satu negara yang menduduki peringkat atas bersaing dengan banyak negara. Namun dengan banyak nya generasi muda yang ingin berlatih, menyebabkan kurangnya pelatih untuk dapat melatih generasi muda dan ini akan menjadi suatu titik permasalahan yang akan membesar dikemudian hari. Gerakan dalam bermain badminton cukup bervariasi terlepas dari penguasaan raket dan lapangan baik dalam menyerang dan bertahan memiliki pola yang

teratur dan dalam di perhitungan dalam konsep matematika. Teknik gerakan dalam bermain bulu tangkis sendiri menjadi faktor penting bagi individu untuk dapat bermain dengan baik . Ini di mana teknologi kecerdasan buatan berperan dalam membantu kehidupan manusia.

Teknologi dengan cabang ilmu kecerdasan buatan diperlukan sebagai suatu kemajuan teknologi yang dapat memahami dan menganalisis seperti otak manusia menggunakan pelatihan berupa gambar, video, maupun suara sebagai bahan untuk dilatih. Computer Vision sudah berkembang pesat sejak dahulu kala, banyak peneliti yang terus mengembangkannya agar menjadi suatu sistem sempurna. Salah satu metode yang diciptakan yaitu Convolutional Neural Network (CNN) dan akan digunakan dalam perancangan ini di bantu dengan PoseNet sebagai pengenalan struktur tubuh.

Convolutional Neural Network, sebuah metode dalam machine learning, dapat digunakan untuk mengenali gerakan berdasarkan visualisasi atlet saat berolahraga. PoseNet merupakan kebalikannya, memantau dan menganalisis posisi tubuh selama latihan. Integrasi kedua teknologi ini dalam sistem pengajaran dan pelatihan olahraga, seperti bulu tangkis, dapat memberikan umpan balik instan kepada atlet tentang teknik dan jumlah repetisi mereka dalam latihan. memakan waktu apalagi jika sedang sibuk dan tidak memiliki banyak waktu.

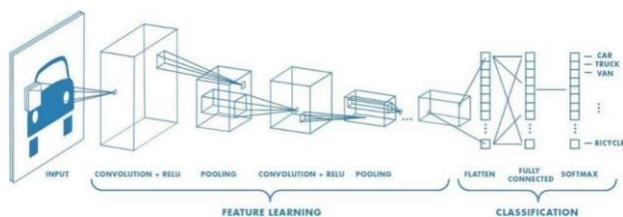
Perancangan sistem nantinya akan berfokus untuk memaksimalkan kemampuan dalam latihan pukulan shuttlecock dengan presisi dan akurat dengan bantuan machine learning model sebagai pengawas gerakan teknik pukulan dasar dalam badminton diantaranya: pukulan servis , pukulan Lob, pukulan Smash, pukulan dropshot, pukulan netting, pukulan drive, dan pukulan backhand. SiDstem akan membaca gerakan apa yang sedang dilakukan dan menilai keakuratan gerakan.

Dengan dirancangnya sistem ini, diharapkan dapat digunakan sebagai media dalam membantu kebutuhan latihan dan menyelesaikan permasalahan yang ada.

## 2. Dasar Teori

### 2.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan jenis khusus dari jaringan saraf tiruan (Artificial Neural Network) yang dikembangkan terutama untuk mengatasi masalah pengolahan citra [1]. CNN pertama kali diperkenalkan oleh Yann LeCun pada tahun 1998. Salah satu komponen kunci dari CNN adalah lapisan konvolusi, yang memungkinkan model untuk secara efisien mengekstrak fitur-fitur penting dari data visual, seperti garis, tepi, tekstur, dan pola [2].



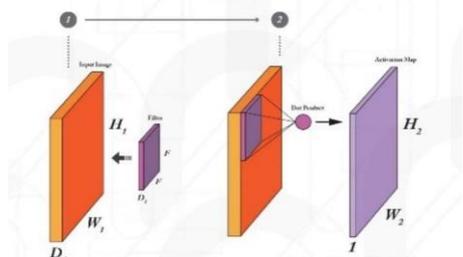
Gambar 1. Arsitektur CNN

### 2.2 Input Layer

Proses CNN dimulai dengan input berupa gambar atau data visual lainnya. Data ini kemudian melewati serangkaian lapisan konvolusi, lapisan pooling, dan lapisan tersembunyi (hidden layers) untuk mengekstrak fitur-fitur dari gambar tersebut.

### 2.3 Convolutional Layer

Convolutional layer adalah komponen kunci dari algoritma CNN. Pada convolutional layer, input gambar dioperasikan oleh filter konvolusi yang mengekstrak fitur dari gambar. Filter konvolusi adalah matriks bobot yang diaplikasikan pada gambar untuk menghasilkan fitur yang relevan untuk tugas yang diberikan. Proses ini dilakukan pada seluruh wilayah gambar untuk menghasilkan output yang berupa map fitur [3].



Gambar 2. Lapisan Konvolusi

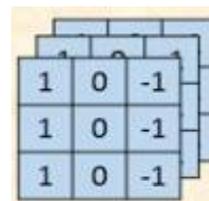
Berikut merupakan rumus matematis dari convolutional layer:

$$W_2 = \frac{W_1 - F}{S} + 1 \tag{1}$$

Dengan  $W_2$  merupakan lebar output,  $W_1$  merupakan lebar input dan  $F$  merupakan nilai untuk filter, serta  $S$  merupakan langkah atau *stride*.

### 2.4 Filter

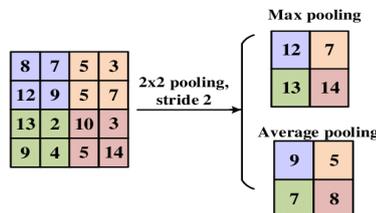
Filter atau kernel merupakan bagian penting dalam CNN, yang berfungsi untuk mengidentifikasi fitur-fitur pada gambar. Filter adalah matriks berukuran kecil yang digeser atau dikali dengan matriks gambar untuk menghasilkan fitur pada gambar. Filter dapat didefinisikan sebagai suatu operasi matematika yang diaplikasikan pada citra untuk menghasilkan citra baru. Ketika filter digeser pada gambar, setiap posisi filter akan menghasilkan suatu nilai yang merepresentasikan fitur pada posisi tersebut. Nilai-nilai ini kemudian digabungkan menjadi suatu citra yang merepresentasikan fitur-fitur pada gambar [4].



Gambar 3. Lapisan Filter

### 2.5 Pooling Layer

Pooling merupakan salah satu komponen utama pada arsitektur CNN yang bertujuan untuk mereduksi dimensi dari feature map hasil dari operasi konvolusi pada layer sebelumnya. Dalam pengolahan citra, pooling sering digunakan setelah operasi konvolusi untuk mengurangi ukuran gambar (downsampling) dan menghilangkan detail-detail yang kurang penting pada citra [5].



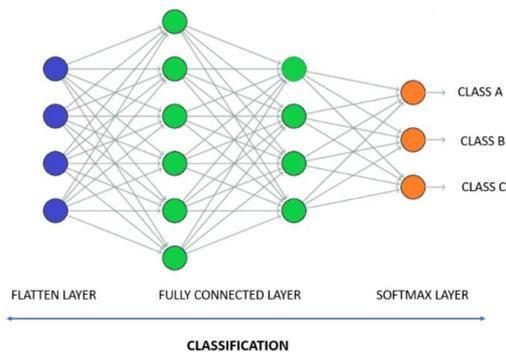
Gambar 4. Lapisan Pooling

Pooling terdiri dari beberapa jenis, di antaranya Max Pooling dan Average Pooling. Max Pooling akan memilih nilai terbesar dalam setiap region dari feature map yang dilewatkan oleh filter, persamaan dari Max Pooling ditunjukkan sebagai berikut:

$$Max\ Pooling(x, i, j, d) = \max(x[i \cdot S : i \cdot S + F, j \cdot S : j \cdot S + F, d]) \tag{2}$$

### 2.6 Fully Connected Layer

Fully Connected Layer (FCL) merupakan layer terakhir dalam sebuah CNN yang bertugas untuk menghubungkan output dari convolutional layer dan pooling layer sebelumnya dengan layer output. FCL juga dikenal sebagai layer dense, karena semua neuron pada layer ini terhubung dengan semua neuron di layer sebelumnya dan layer setelahnya [6].



Gambar 5. Fully Connected Layer

FCL memiliki fungsi untuk melakukan klasifikasi dan prediksi pada data input. Input yang diterima oleh FCL adalah vektor yang dihasilkan dari proses konvolusi dan pooling sebelumnya. Setiap neuron pada FCL melakukan operasi dot product pada vektor input dengan setiap bobot yang ada, kemudian hasilnya dijumlahkan dengan bias. Hasil akhir dari operasi ini kemudian diaktivasi menggunakan fungsi aktivasi seperti ReLU atau softmax. Persamaan dari FCL ditunjukkan sebagai berikut:

$$y = \sigma(W \cdot h + b) \tag{3}$$

Dengan Y merupakan vektor output dari lapisan FCL,  $\sigma$  merupakan fungsi aktivasi, W merupakan matriks bobot, h merupakan vektor input dan b merupakan vektor bias.

### 2.7 Fungsi Aktivasi

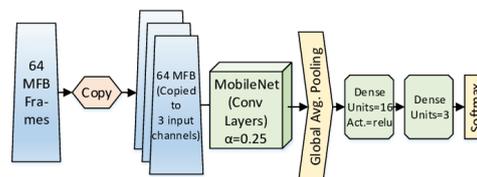
Dalam CNN, fungsi aktivasi digunakan setelah dilakukan konvolusi pada tiap layer. Fungsi aktivasi ini memiliki peran penting dalam menghasilkan non-linearitas pada CNN, sehingga model dapat mempelajari fitur-fitur yang lebih kompleks pada gambar. Beberapa fungsi aktivasi yang sering digunakan pada CNN antara lain ReLU (Rectified Linear Unit), Sigmoid, dan Tanh. ReLU adalah fungsi aktivasi yang paling umum digunakan pada layer convolution dan fully connected pada CNN. Fungsi ini memiliki formula  $f(x) = \max(0,x)$ ,

sehingga nilai input yang kurang dari 0 akan diubah menjadi 0, sedangkan nilai input yang lebih besar dari 0 akan diteruskan tanpa perubahan. Fungsi ReLU ini digunakan karena memiliki keuntungan dalam komputasi yang lebih cepat dan efisien serta mengurangi kemungkinan terjadinya penurunan nilai gradien [7].

### 2.8 MobileNet

merupakan salah satu model arsitektur convolutional neural network (CNN) yang sering digunakan sebagai solusi mengatasi kebutuhan akan computing resource berlebih dengan 2 set hyper-parameters dengan tujuan membangun model yang sangat kecil dan latensi rendah, maka dapat lebih ringan dan efisien saat diimplementasikan pada kebutuhan mobile dan embedded applications. Seperti namanya, Mobile, para peneliti dari Google membuat arsitektur CNN yang dapat digunakan untuk ponsel [8].

Perbedaan mendasar antara arsitektur MobileNet dan arsitektur CNN pada umumnya adalah penggunaan lapisan atau layer konvolusi dengan ketebalan filter yang sesuai dengan ketebalan dari input image. MobileNet menggunakan dua buah konvolusi yang dinamakan *depthwise convolution* dan *pointwise convolution* [3]. Gabungan kedua konvolusi diatas kemudian menghasilkan layer yang disebut Depthwise separable convolutions yang merupakan layer utama untuk MobileNet.

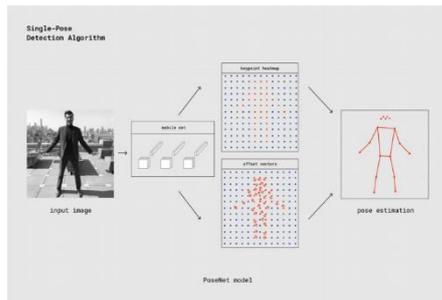


Gambar 6. Arsitektur MobileNet

### 2.9 PoseNet

PoseNet adalah model machine learning yang digunakan untuk mengestimasi posisi tubuh manusia dalam gambar maupun video. PoseNet memanfaatkan arsitektur Convolutional Neural Network (CNN) untuk memprediksi lokasi titik-titik penting tubuh manusia, seperti mata, hidung, telinga, bahu, siku, pergelangan tangan, pinggul, lutut, dan pergelangan kaki. Setiap titik ini dikenal sebagai "keypoint." Secara khusus, PoseNet menghasilkan 17 keypoint untuk setiap gambar atau frame video, yang mencakup sebagian besar area tubuh manusia.

Ke-17 keypoint ini kemudian dapat digunakan untuk mengenali dan memonitor gerakan manusia dalam gambar atau video. PoseNet memiliki berbagai aplikasi, termasuk pelacakan pose dalam waktu nyata, analisis gerakan, dan interaksi manusia dengan komputer. Kemampuannya untuk dengan akurat memperkirakan pose manusia telah menjadikannya alat yang sangat berharga dalam berbagai bidang, termasuk olahraga, kebugaran, dan pengembangan aplikasi interaktif [9].



Gambar 7. Aktualisasi PoseNet

### 3. Hasil Percobaan

Percobaan atau pengujian dilakukan dengan menggunakan model yang diambil secara mandiri oleh perancang. Pelatihan dilakukan dengan merekam setiap gerakan dengan durasi satu menit dilakukan oleh beberapa individu. Total gambar yang telah dibagi menjadi lima frame per detik dari video didapat sekitar 15.240 gambar dengan komposisi pembagian data pelatihan serta pengujian yaitu 80% digunakan sebagai pelatihan dan 20% sebagai validasi dengan jumlah kelas sebanyak 7 kelas.

Perbandingan akurasi pada beberapa percobaan *epoch* dan *batch size* dengan learning rate 0.001 dapat dilihat pada tabel berikut ini:

Tabel 1 Perbandingan akurasi pelatihan

Epoch	Batch Size	Akurasi Pelatihan
10	32	96,00%
50	16	97,77 %
50	32	96,46%
50	64	96,00%

Dengan menggunakan 50 *epoch* dan 16 *batch size*, didapatkan hasil pengujian sebesar 97.77%

Teradapat 7 gerakan bulu tangkis yang dapat diklasifikasi diantaranya : gerakan *backhand*, gerakan *forehand*, gerakan *smash*, gerakan *drive*, gerakan *dropshot*, gerakan *netting*, dan gerakan servis.

### 4. Kesimpulan

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

1. Bahwa evaluasi dan implementasi aplikasi merupakan tahap penting dalam pengembangan aplikasi berbasis model
2. Dengan implementasi model MobileNetV<sub>2</sub> dengan integrasi PoseNet ini, maka hasil pengujian dan pelatihan menghasilkan akurasi yang cukup baik dengan kisaran diatas 95% yang sudah diuji dengan menggunakan pembandingan beberapa epoch
3. Pengenalan klasifikasi gerakan olahraga bulu tangkis ini dapat dilakukan secara langsung pada saat pelatihan bulu tangkis dengan menggunakan kamera ponsel pada perangkat mobile sehingga akan sangat praktis bagi pengguna serta dapat menjadi evaluasi pemain dikarenakan fitur yang dapat menyimpan hasil klasifikasi pengenalan gerakan.
4. Untuk jumlah pemain yang dapat dikenali saat ini hanya berjumlah 1 orang dalam frame kamera hal ini dilakukan untuk mendapatkan hasil optimal bagi sistem dalam melakukan klasifikasi.
5. Posisi kamera ponsel dalam mengambil video secara langsung juga mempengaruhi tingkat akurasi dalam melakukan klasifikasi.

### REFERENSI

- [1] J. A. Y. D. & L. C. P. Dunmon, "Assessment of Convolutional Neural Networks for Automated Classification of Chest Radiographs," *Radiology*, 2019.
- [2] X. H. Y. & W. Zhang, "A Hybrid convolutional neural network for sketch recognition. Pattern Recognition Letters," 2019.
- [3] A. A. R. F. & M. A. Reshi, "An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification," *Hindawi*, 2021.
- [4] N. K. Uba, "Land Use and Land Cover Classification Using Deep Learning Techniques," *Arizona State University*, 2016.
- [5] E. B. N. & M. Z. Julpan, "Analisis Fungsi Aktivasi Sigmoid Biner dan Sigmoid Bipolar dalam Algoritma Backpropagation pada Prediksi Kemampuan Siswa," *Jurnal Teknovasi*, vol. 02, no. 1, p. 103 – 116, 2015.
- [6] A. M. & W. Z. W. M. Alhassan, "Brain tumor classification in magnetic resonance image using hard swish-based RELU activation function-convolutional neural network," *Neural Computing and Applications*, 2021.
- [7] P. A. S. A. S. R. Ahmad Roihan, "Pemanfaatan Machine Learning dalam Berbagai Bidang;," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 1, pp. 75-82, 2020.

- [8] S. S. Widi Hastomo, "Convolutional Neural Network Arsitektur MobileNet-V2 untuk Mendeteksi Tumor Otak," Seminar Nasional Teknologi Informasi dan Komunikasi STI&K (SeNTIK), Jakarta , 2021.
- [9] D. Oved, "Real-time Human Pose Estimation in the Browser with TensorFlow.js," 07 May 2018. [Online]. Available: <https://blog.tensorflow.org/2018/05/real-time-human-pose-estimation-in.html>. [Diakses 10 September 2023].

**Benny Karnadi**, 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.

**Agus Budi Dharmawan S.kom, M.T., M.sc.** memperoleh gelar M.T dari ITS Surabaya. Kemudian memperoleh gelar M.Sc dari Elektronik Engenering FH Darmstad. Saat ini sebagai Dosen Program studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara.