

# Pengenalan Aktivitas Manusia Dalam Ruang Dengan Convolutional Neural Networks

Lina<sup>1\*</sup>, Michelle Augustine<sup>2</sup>, Richard Stephen<sup>3</sup>, Lorico Salim<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, Jakarta Barat, DKI Jakarta

Email: <sup>1\*</sup>lina@fti.untar.ac.id, <sup>2</sup>michelle.535200055@stu.untar.ac.id, <sup>3</sup>richard.535200058@stu.untar.ac.id, <sup>4</sup>lorico.535200015@stu.untar.ac.id

(Naskah masuk: 24 Okt 2023, direvisi: 4 Jan 2024, 22 Jan 2024, diterima: 24 Jan 2024)

## Abstrak

Kemajuan teknologi yang telah berkembang sangat pesat dapat membantu pekerjaan manusia dalam berbagai bidang. Salah satu kegiatan yang dapat diotomatisasi adalah aktivitas pengawasan kegiatan manusia dalam sebuah ruangan, misalnya bagi lansia ataupun orang cacat. Umumnya kegiatan pengawasan membutuhkan tenaga manusia untuk memantau kejadian dan aktivitas di lokasi tertentu secara *live* maupun direkam melalui kamera CCTV. Dengan perkembangan teknologi, seluruh kegiatan pengenalan terhadap aktivitas manusia dapat dilakukan secara otomatis. Sistem dapat mengenali aktivitas duduk, berdiri, belajar, mengangkat tangan, serta bertepuk tangan. Sistem dibuat dengan menggunakan metode *Convolutional Neural Network* (CNN). Eksperimen akan dilakukan dengan data uji yang berasal dari internet maupun rekaman video yang dilakukan oleh tim di lapangan. Hasil eksperimen menunjukkan bahwa sistem yang dikembangkan dapat mengenali aktivitas manusia dengan nilai akurasi sebesar 86,65% untuk data uji dari internet dan 96% untuk data uji dari IP camera yang terpasang pada sebuah lokasi. Selain itu, sistem juga dapat menghasilkan sebuah rangkuman catatan aktivitas terhadap seluruh kegiatan yang terjadi di dalam lokasi tersebut pada kisaran waktu yang ditetapkan.

**Kata Kunci:** Pengenalan Aktivitas Manusia, Pengawasan, *You Only Look Once*, *Convolutional Neural Network*.

## *Human Activity Recognition in Indoor Spaces Using Convolutional Neural Networks*

### *Abstract*

*Technological advances that have developed very rapidly can help human work in various fields. One activity that can be automated is the activity of monitoring human activities in a room, e.g. for elderly or disabled people. Generally, surveillance activities require human power to monitor events and activities at certain locations live or recorded via CCTV cameras. With technological developments, all recognition activities for human activities can be carried out automatically. The system can recognize sitting, standing, studying, raising hands and clapping. The system was created using the Convolutional Neural Network (CNN) method. Experiments will be carried out using test data originating from the internet and video recordings carried out by the team in the field. The experimental results show that the system developed can recognize human activity with an accuracy value of 86.65% for test data from the internet and 96% for test data from IP cameras installed at a location. Apart from that, the system can also produce a summary of activity records for all activities that occur in that location within a specified time range.*

**Keywords:** Activity Logbook, Activity Recording Video, Human Activity Recognition, Supervision, *You Only Look Once*

**I. PENDAHULUAN**

Topik pengenalan aktivitas manusia telah menjadi bidang penelitian utama dalam dekade terakhir. Kemajuan dalam teknologi pemrosesan gambar memungkinkan pengenalan aktivitas manusia dengan kamera untuk pemantauan aktivitas sehari-hari dalam berbagai aplikasi seperti bantuan darurat bagi lansia dan orang cacat, bantuan kognitif dan keselamatan [1]. Pengenalan aktivitas mampu mengidentifikasi aktivitas dari beberapa informasi tertentu yang ditangkap dan diperoleh oleh kamera atau sensor [2]. Pemantauan terhadap suatu peristiwa khususnya aktivitas manusia mempunyai tujuan besar untuk mendokumentasikan peristiwa pada waktu dan tempat tertentu. Hal ini menjadi sangat penting karena adanya keterbatasan dalam pemantauan yang harus dilakukan secara manual dan langsung oleh pekerja manusia. Penggunaan tenaga manusia dapat dikategorikan sebagai tugas yang tidak efisien, sehingga pengenalan aktivitas manusia secara otomatis akan menjadi solusi yang berguna dan nyaman [3].

Pengenalan aktivitas manusia juga dapat diintegrasikan menggunakan perangkat *Internet of Things* (IoT). Dikarenakan integrasi ini, sistem memperoleh kemampuan yang lebih baik dan lebih tinggi untuk mengimplementasikan sistem praktis di dunia nyata. Sebagai contoh, sistem akan mengidentifikasi aktivitas manusia dan mengirimkan pesan atau peringatan melalui perangkat IoT [4]. Berbagai metode pengenalan aktivitas manusia telah dikembangkan oleh beberapa peneliti. Pengenalan aktivitas manusia menggunakan DHALR dikemukakan oleh Isah Lawal dan Sophia Bano. Dataset aktivitas diperoleh dengan menempatkan sensor *accelerometer* dan giroskop di beberapa titik pada tubuh manusia [5]. Salah satu metode yang diusulkan Parvis untuk mengidentifikasi aktivitas manusia adalah pendekatan Markov [6]. Kumpulan data UCI HAR, Opportunity, UniMib-SHAR, PAMAP, dan WISDM digunakan untuk melatih model CNN, yang dilatih dengan kerugian lokal (*pred loss* dan *sim loss*) dalam studi pengenalan aktivitas manusia berbasis sensor ini [7].

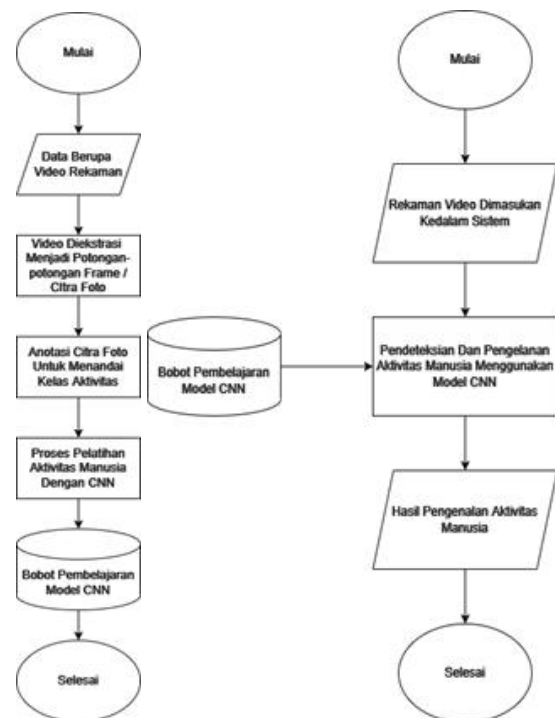
Pengenalan aktivitas manusia menggunakan sensor *smartphone* dan Model *Deep Learning* Berbasis LSTM yang dikemukakan oleh Pornthep Rojanavasu, Anuchit Jitpattanakul dan Sakorn Mekruksavanich [8]. CNN dengan program pengenalan aktivitas manusia berbasis arsitektur Alexnet dikembangkan oleh Kerly Prastika [9]. Selain itu terdapat sistem pengenalan aktivitas manusia dengan tiga kategori dasar menggunakan kategori berbasis tindakan, berbasis interaksi, dan berbasis gerak yang dikemukakan oleh Zawar Hussain dkk. [10]. Juga ada Sistem Pengenalan Aktivitas Manusia menggunakan CNN dengan pengklasifikasi berbasis Fitur HAAR yang dibuat oleh Md. Atikuzzaman dkk [11]. Sistem lain untuk mengenali aktivitas manusia, Ronald Mutegeki dan Dong Seog Han [12] menciptakan sistem berbasis pembelajaran mendalam yang komprehensif yang menggunakan algoritma CNN-LSTM.

Penelitian ini membahas tentang pengenalan aktivitas manusia menggunakan metode *Convolutional Neural Network* (CNN). Sistem yang dikembangkan bertujuan untuk mengenali, mengidentifikasi, dan memperkirakan aktivitas

spesifik dari citra masukan. Perbedaan utama dari penelitian ini dibandingkan penelitian terdahulu adalah dari sisi sistem dapat melakukan klasifikasi aktivitas melalui masukan berupa data video dengan sudut pengambilan gambar yang berbeda-beda. Tujuan pengembangan sistem ini adalah untuk melakukan pengenalan aktivitas manusia pada lokasi *indoor* dimana objek dapat berpindah tempat dengan bebas. Sistem akan menjalani sejumlah tahapan pelatihan agar menghasilkan model yang mengandung fitur penting dari dataset. Selanjutnya sistem akan diuji untuk mengevaluasi efektivitas model dalam memprediksi aktivitas manusia.

**II. METODE PENELITIAN**

Metode yang digunakan untuk mengembangkan sistem pengenalan aktivitas manusia pada penelitian ini adalah *Convolution Neural Network* (CNN). Gambar 1 menggambarkan alur pengembangan sistem, mulai dari perolehan dataset hingga hasil pengenalan oleh sistem yang dapat diterapkan.



Gambar 1. Skema Pelatihan dan Pengenalan pada Sistem Pengenalan menggunakan CNN

Sistem yang dirancang pada penelitian ini adalah sistem yang mampu mengklasifikasikan dan mengenali aktivitas manusia dalam sebuah frame video dengan menggunakan metode *Convolutional Neural Network*. *Convolutional Neural Network* (CNN) yang merupakan bagian dari model *Deep Neural Network* merupakan salah satu pengembangan jaringan saraf tiruan paling terkenal yang terinspirasi dari jaringan saraf manusia [13]. Pada tahap awal, sistem menerima masukan berupa data video rekaman yang

diekstraksi menjadi potongan citra dengan pencacah 30 *frame per second*. Selanjutnya dilakukan proses pelatihan aktivitas manusia dengan CNN.

CNN merupakan salah satu jenis dari rangkaian jaringan syaraf buatan yang juga menerapkan jaringan syaraf multi-lapisan. Terdapat banyak lapisan dalam jaringan syaraf multi-layer ini, seperti lapisan konvolusional, ReLU (fungsi aktivasi), dan lapisan *pooling* yang disertakan sebagai pembelajaran fitur dan diikuti dengan klasifikasi, dengan lapisan dan prediksi yang rata dan terhubung sepenuhnya [14][15]. Melalui lapisan tersebut, jaringan syaraf buatan beroperasi untuk memperoleh informasi representatif baru di setiap lapisan.

Lapisan konvolusional berisi kumpulan filter konvolusi (*kernel*). Peta fitur akan dihasilkan sebagai keluaran dari *Convolutional Layer* setelah gambar masukan dinyatakan dalam matriks 2 dimensi [16]. Persamaan 1 di bawah ini untuk menghitung *Convolutional Layer*.

$$x(i, j) = \sum_m \sum_n w_{m,n}^l * o_{i+m, j+n}^{l-1} + b \quad (1)$$

Besarnya keluaran *Convolutional Layer* dapat ditentukan dengan menggunakan Persamaan 2:

$$Output = \frac{w-N}{s} + 1 \quad (2)$$

*Pooling Layer* beroperasi sebagai *sub-sampling* dari peta fitur *Convolutional Layer*. Dengan operasi *down-sampling*, ukuran input dikurangi secara spasial. *Average pooling* dan *max pooling* adalah metode paling terkenal di lapisan ini. Selain itu, ada beberapa jenis metode yang digunakan pada lapisan ini kecuali bagian atas seperti *tree pooling*, *gated pooling*, *min pooling*, *global average pooling* (GAP), *global max pooling* [16]. Perataan mengubah semua matriks 2 dimensi dari peta fitur yang dikumpulkan di *Pooling Layer* menjadi vektor linier [17]. Ini menciptakan keluaran vektor linier berdasarkan nilai baris matriks. Fungsi Aktivasi menentukan apakah neuron perlu diaktifkan menjadi referensi masukan tertentu dan menghasilkan keluaran yang diinginkan. Ini dapat membedakan fitur-fitur yang relevan karena kemampuannya untuk melatih jaringan dari kegagalan propagasi balik [18]. Mayoritas Fungsi Aktivasi hadir dalam tipe *Sigmoid*, *Tanh*, dan *Rectified Linear Units* (ReLU). Persamaan 3 di bawah ini untuk menghitung fungsi *Sigmoid*.

$$f(x)_{sigm} = \frac{1}{1+e^{-x}} \quad (3)$$

Fungsi lainnya, *tanh*, dapat dihitung menggunakan Persamaan 4 dan *Rectified Linear Unit* (ReLU) menggunakan Persamaan 5.

$$f(x)_{tanh} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

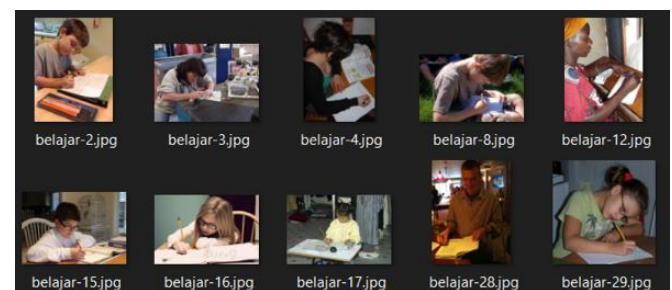
$$f(x)_{ReLU} = \max(0, x) \quad (5)$$

Lapisan yang terhubung sepenuhnya menggunakan teknik dasar Jaringan Syaraf Tiruan (JST) *feed-forward*, yaitu jaringan syaraf *perceptron* multilapis konvensional. Setiap neuron terhubung satu sama lain di lapisan yang terhubung sepenuhnya ini. Vektor dari *Flattening Layer* akan dimasukkan ke dalam lapisan yang terhubung penuh sebagai lapisan masukan, melalui lapisan tersembunyi, dan lapisan yang terhubung sepenuhnya tersebut digabungkan untuk membuat model Jaringan Neural Konvolusional sebagai lapisan keluaran [19]. Luaran dari hasil pelatihan adalah berupa bobot pembelajaran yang merepresentasikan informasi setiap kelas yang ada dalam sistem.

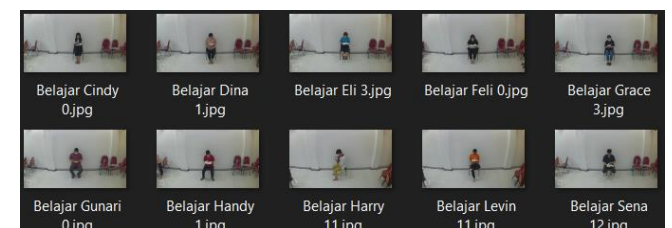
Selanjutnya, pada tahap pengujian, sistem akan menerima masukan berupa rekaman video dari sebuah aktivitas yang dilakukan manusia. Kemudian dilakukan proses pencacahan video menjadi citra dengan 30 fps menggunakan aplikasi *video frame rate converter*. Proses pengenalan dilakukan dengan membandingkan bobot yang diperoleh dengan bobot pembelajaran. Hasil klasifikasi akan mengikuti kelas yang memiliki bobot terdekat dengan kelas pembelajaran.

### III. HASIL DAN PEMBAHASAN

Sistem yang dibangun bertujuan untuk melakukan pengenalan aktivitas manusia dengan *Convolutional Neural Network* (CNN). Sistem ini diharapkan dapat mengenali aktivitas yang terekam dalam video. Dari video rekaman tersebut, dilakukan tahap pencacahan menjadi citra dengan 30 fps. Pada eksperimen yang dilakukan, dataset yang digunakan ada 2 yaitu dataset internet dan dataset *IP Camera*. *Dataset* internet merupakan *dataset* yang didapat dari internet yang berisi kumpulan data gambar-gambar aktivitas manusia. *Dataset IP Camera* merupakan *dataset* yang diambil sendiri dengan menggunakan *IP Camera Xiaomi* dengan tipe *Home Camera*. Contoh dataset citra internet dan *IP Camera* dapat dilihat pada Gambar 2 dan Gambar 3.



Gambar 2. Contoh Citra dari *Dataset* Internet



Gambar 3. Contoh Citra dari *Dataset IP Camera*

Setiap *dataset* terdiri dari kelas aktivitas manusia, seperti duduk, berdiri, belajar, mengangkat tangan, menepuk tangan. Sistem yang dikembangkan dapat mengenali aktivitas manusia secara umum, baik anak-anak maupun dewasa dan tidak dibatasi pada lansia maupun orang cacat saja. *Dataset* ini akan digunakan untuk pelatihan, validasi, dan pengujian. Tabel 1 menunjukkan jumlah citra yang digunakan dalam eksperimen untuk kedua *dataset*.

Tabel 1. Jumlah Citra Pada Dataset Internet dan IP Camera

Jenis Aktivitas	Training		Testing	
	Internet	IP Camera	Internet	IP Camera
Berdiri	190	80	100	40
Duduk	190	80	100	40
Belajar	190	80	100	40
Mengangkat Tangan	190	80	100	40
Menepuk Tangan	190	80	100	40
<b>Total</b>	950	400	500	200

Tabel 2. Arsitektur Model MobileNetV2

Layer	Input
Conv2D 1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	224 x 224 x 3
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	112 x 112 x 32
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	112 x 112 x 16
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	56 x 56 x 24
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	28 x 28 x 32
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	14 x 14 x 64
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	14 x 14 x 96
1 x 1 Conv2D, ReLU6 3 x 3 DwiseConv2D, ReLU6 Linear 1 x 1 Conv2D	7 x 7 x 160
Conv2D 1 x 1	7 x 7 x 320
AvgPool 7 x 7	7 x 7 x 1280
Conv2D 1 x 1	1 x 1 x 1280

Tabel 3. Hasil Pelatihan Konfigurasi Model Internet

Model	Loss	Accuracy	Validation Loss	Validation Accuracy
1	0,0047	100,00%	0,5937	83,68%
2	0,0108	97,95%	8,5882	35,79%

3	0,0081	99,81%	3,3962	45,79%
4	0,0042	100,00%	5,3317	37,89%
5	0,0574	98,48%	4,2983	47,37%
6	0,0014	100,00%	3,0568	42,11%
7	0,1652	94,08%	3,5617	37,89%
8	0,0052	98,00%	4,0322	58,00%

Tabel 4. Hasil Pelatihan Konfigurasi Model IP Camera

Model	Loss	Accuracy	Validation Loss	Validation Accuracy
1	0,5494	85,35%	1,1553	55,71%
2	0,0006	100,00%	17,2627	21,43%
3	0,0015	100,00%	0,4597	85,71%
4	0,0003	100,00%	0,7083	81,43%
5	0,0207	99,47%	1,5214	71,43%
6	0,0007	100,00%	0,1670	91,43%
7	0,0864	95,76%	0,8401	78,57%
8	0,00004	100,00%	0,3270	88,57%

Pada tahap pelatihan, parameter dan *hyperparameter* model harus dikonfigurasi agar sistem dapat menghasilkan keluaran yang optimal. Pada eksperimen yang dilakukan, berbagai konfigurasi lapisan model diujicobakan terhadap arsitektur *MobileNetV2*. Arsitektur dari *MobileNetV2* tertera pada Tabel 2.

Terdapat berbagai metrik yang dapat digunakan untuk menentukan model terbaik dari sebuah sistem. Pada pengujian ini dilakukan pengukuran dengan metrik *accuracy* dan *loss*. *Accuracy* diukur dalam bentuk prosentase yang diperoleh dengan jumlah dari prediksi benar terhadap keseluruhan jumlah data. Jika dilakukan terhadap data validasi, maka akan diperoleh nilai *validation accuracy*. Formula perhitungan nilai *accuracy* tertera pada Persamaan 6 sebagai berikut:

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{6}$$

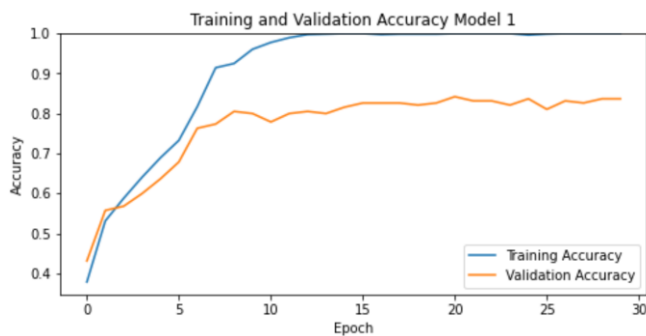
Dimana nilai TP adalah *true positive* yang merupakan jumlah data positif yang terdeteksi benar, TN adalah *true negative* yang merupakan jumlah data negatif yang terdeteksi benar, FP adalah *false positive* yaitu data negatif yang terdeteksi sebagai data positif, dan FN adalah *false negative* yaitu data positif yang terdeteksi sebagai data negatif.

Untuk nilai *loss*, dilakukan perhitungan probabilitas prediksi sejauh mana prediksi berbeda dari nilai kebenarannya. Berbeda dengan akurasi, nilai *loss* bukanlah nilai prosentase, melainkan jumlah *error* yang dilakukan setiap sampel pada *training* atau *validation* set. Metrik *loss* diukur dengan rumus pada Persamaan 7 sebagai berikut:

$$loss = \frac{1}{n} \sum_i^C (y_i - \hat{y}_i)^2 \tag{7}$$

Nilai *loss* menghitung rata-rata kuadrat selisih antara nilai prediksi dengan nilai sebenarnya. Berdasarkan kedua hasil pelatihan *dataset* yang tertera pada Tabel 3 dan 4, dilakukan

analisis nilai *Validation Loss* dan nilai *Validation Accuracy*. Untuk Model Internet, Akurasi Validasi tertinggi sebesar 83,68% pada Model 1 dengan *Validation Loss* 0,5937. Secara detail, hasil grafik Akurasi dan *Loss* dari Model 1 yang merupakan model dengan akurasi tertinggi dan *loss* terendah pada *Dataset* Internet tertera pada Gambar 4 dan Gambar 5. Sedangkan untuk *IP Camera*, Model 6 mendapatkan *Validation Accuracy* tertinggi yaitu 91,43% dengan *Validation Loss* terendah 0,1670. Secara detail hasil grafik Akurasi dan *Loss* dari Model 6 yang memiliki akurasi tertinggi dan *loss* terendah untuk *Dataset* IP Camera tertera pada Gambar 6 dan Gambar 7. Konfigurasi parameter untuk pembentukan Model 1 tertera pada Tabel 5, sedangkan Konfigurasi Parameter untuk pembentukan Model 6 tertera pada Tabel 6.



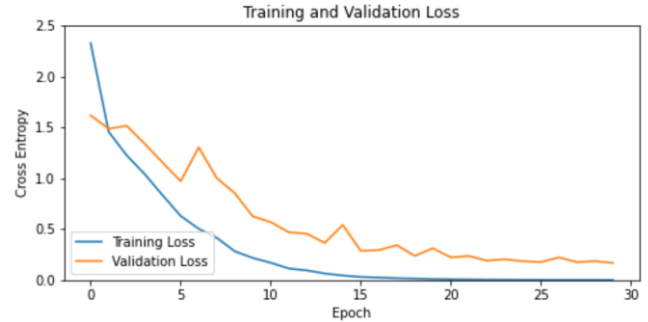
Gambar 4. Grafik Akurasi Untuk Model 1 Pada *Dataset* Internet



Gambar 5. Grafik *Loss* Untuk Model 1 Pada *Dataset* Internet



Gambar 6. Grafik Akurasi Untuk Model 6 Pada *Dataset* IP Camera



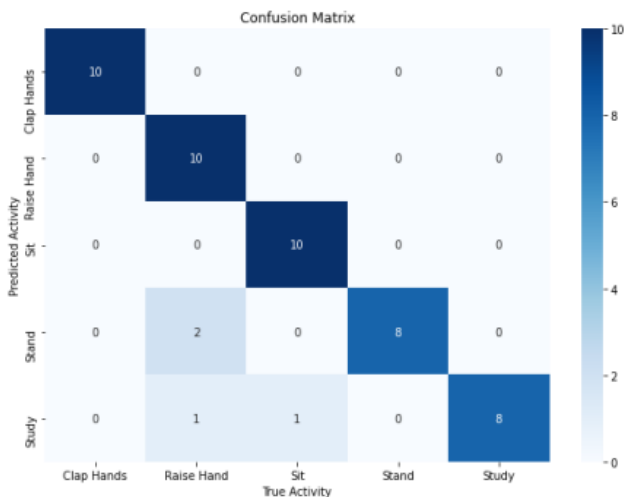
Gambar 7. Grafik *Loss* Untuk Model 6 Pada *Dataset* IP Camera

Tabel 5. Konfigurasi Parameter Untuk Model 1 Pada *Dataset* Internet

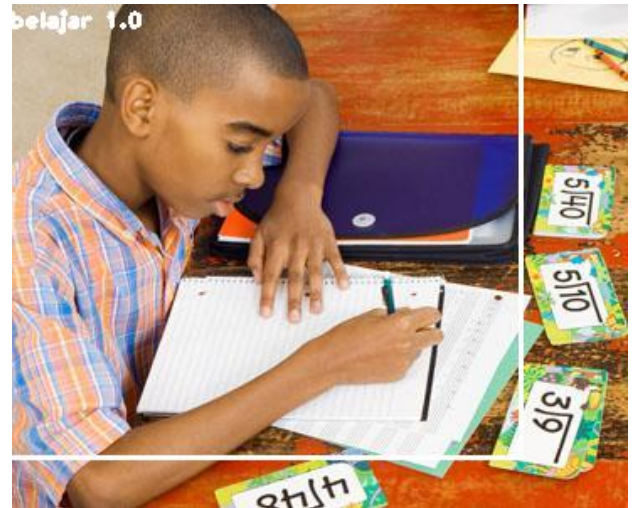
Layer Model	Parameter	Output Shape
MobileNetV2	Tabel 2	(None, 7, 7, 1280)
Conv2D	Filter = 32 Kernel Size = 3x3 Activation = ReLU	(None, 5, 5, 32)
Dropout	0,2	(None, 5, 5, 32)
GlobalAvgPool2D	-	(None, 32)
Dense	Units = 5 Activation = Softmax	(None, 5)

Tabel 6. Konfigurasi Parameter untuk Model 6 Pada *Dataset* IP Camera

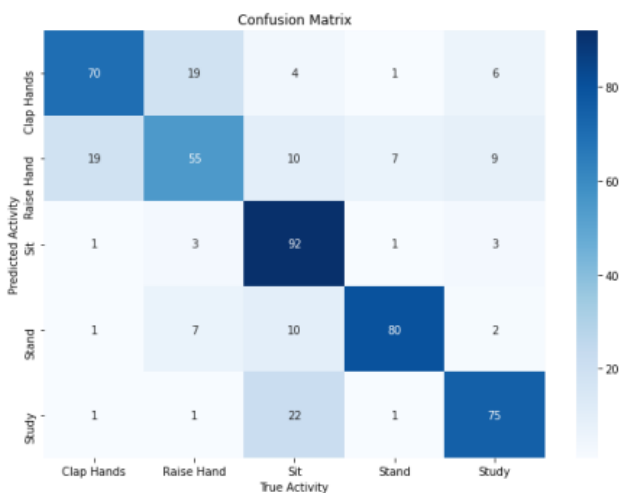
Layer Model	Parameter	Output Shape
Conv2D	Filter = 8 Kernel Size = 7x7 Padding = 0 Activation = ReLU Input Shape = 224x224	(None, 224, 224, 8)
MaxPool2D	Pool Size = 2x2	(None, 112, 112, 8)
Conv2D	Filter = 16 Kernel Size = 5x5 Padding = 0 Activation = ReLU	(None, 112, 112, 16)
MaxPool2D	Pool Size = 2x2	(None, 56, 56, 16)
Flatten	-	(None, 50176)
Dropout	0,2	(None, 50176)
Dense	Units = 5 Activation = Softmax	(None, 5)



Gambar 8. Confusion Matrix Model Dataset IP Camera



Gambar 10. Contoh Hasil Pendeteksian Dataset Internet



Gambar 9. Confusion Matrix Model Dataset Internet



Gambar 11. Contoh Hasil Pendeteksian Dataset IP Camera

Untuk mendapatkan model yang lebih akurat akan dilakukan redefinisi *hyperparameter*. Untuk dataset pertama (IP Camera) menggunakan redefinisi *epoch* = 150 dan ukuran *batch* = 32, sedangkan untuk dataset kedua (Dataset Internet) menggunakan redefinisi *epoch* = 30 dan ukuran *batch* = 128. Kemudian sistem akan diuji menggunakan sisa dataset, yang terdiri dari 50 data yang belum dipelajari dari 5 kelas aktivitas berbeda, dengan 10 kondisi di setiap kelas. Hasil pengujian masing-masing model disajikan dalam bentuk *Confusion Matrix* yang tertera pada Gambar 8 dan Gambar 9. Angka-angka ini memberikan gambaran umum tentang performa model pada data pengujian. Berdasarkan hasil pengujian didapatkan nilai akurasi uji untuk Model IP Camera sebesar 92%, sedangkan nilai akurasi uji untuk Model Internet sebesar 74,4%.

Pengujian dilakukan dengan cara pengenalan aktivitas manusia. Model yang telah dilatih akan dilakukan pengujian untuk melakukan evaluasi hasil dari proses pelatihan. Contoh pengenalan aktivitas manusia dapat dilihat pada Gambar 10 dan Gambar 11.

Hasil Pengujian menunjukkan konfigurasi model yang berbeda maka mendapatkan hasil akurasi yang berbeda juga. Dapat dikatakan bahwa model yang telah dilatih cukup baik dalam mengenali aktivitas manusia.

#### IV. KESIMPULAN DAN SARAN

Sistem yang diusulkan menunjukkan kemampuan dalam mengenali aktivitas manusia secara otomatis dalam video dari kamera dan internet, mencapai tingkat akurasi validasi yang tinggi sebesar 94% pada model 6 untuk dataset Kamera IP dan 83,68% pada model 1 untuk dataset Internet. Kemudian setelah proses pengujian dengan menggunakan *hyperparameter* baru didapatkan nilai akurasi uji sebesar 92% untuk model IP Camera dan 74,4% pada model Internet.

Disarankan untuk menggabungkan kumpulan data yang lebih besar untuk peningkatan sistem, yang dapat membantu mengurangi *overfitting* dan mengaktifkan model

menggunakan sensor yang berbeda. Selain itu, mengintegrasikan berbagai sensor akan meningkatkan fungsionalitas sistem, memungkinkan sistem menangkap informasi yang lebih komprehensif dan memperluas penerapannya pada berbagai skenario dan lingkungan. Dengan menerapkan hal ini, sistem dapat mencapai akurasi yang lebih tinggi dan menjadi lebih kuat dalam mengenali aktivitas manusia.

#### REFERENSI

- [1] F. Ullah *et al.*, "An image-based human physical activities recognition in an indoor environment," *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, 2020, pp. 588-593, doi:10.1109/ictc49870.2020.9289314.
- [2] H. C. Nguyen *et al.*, "Deep learning for human activity recognition on 3D human skeleton: Survey and Comparative Study," *Sensors*, 2023, vol. 23, no. 11, p. 5121, doi:10.3390/s23115121.
- [3] H. Najeh, C. Lohr and B. Leduc, "Dynamic segmentation of sensor events for real-time human activity recognition in a smart home context," *Sensors*, 2022, vol. 22, no. 14, p. 5458, doi:10.3390/s22145458.
- [4] N. Ahmed, J. I. Rafiq and M. R. Islam, "Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model," *Sensors*, 2020, vol. 20, no. 1, p. 317, doi:10.3390/s20010317.
- [5] I. A. Lawal and S. Bano, "Deep human activity recognition with localisation of wearable sensors," *IOP IEEE Access*, 2020, vol. 8, p. 155060-155070, doi:10.1109/access.2020.3017681.
- [6] P. Asghari, E. Soleimani and E. Nazerfard, "nline human activity recognition employing hierarchical hidden Markov models," *Journal of Ambient Intelligence and Humanized Computing*, 2020, vol. 11, no. 3, pp. 1141-1152, doi:10.1007/s12652-019-01380-5.
- [7] Q. Teng *et al.*, "The Layer-Wise Training Convolutional Neural Networks Using Local Loss for Sensor-Based Human Activity Recognition," *IEEE Sensors Journal*, 2020, vol. 20, no. 13, pp. 7265-7274, doi:10.1109/jsen.2020.2978772.
- [8] P. Rojanavasu, A. Jitpattanukul and S. Mekruksavanich, "Comparative analysis of LSTM-based deep learning models for HAR using smartphone sensor," *2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering*, 2021, pp. 269-272, doi:10.1109/ectidamtncon51128.2021.9425733.
- [9] K. Prastika and Lina, "Application of individual activity recognition in the room using CNN alexnet method," *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 1007, no. 1, p. 012162, doi:10.1088/1757-899x/1007/1/012162.
- [10] Z. Hussain, Q. Z. Sheng and W. E. Zhang, "A review and categorization of techniques on device-free human activity recognition," *Journal of Network and Computer Applications*, 2020, vol. 102738, p. 167, doi:10.1016/j.jnca.2020.102738.
- [11] M. Atikuzzaman *et al.*, "Human Activity Recognition System from Different Poses with CNN," *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 2020, pp. 1-5, doi:10.1109/sti50764.2020.9350508.
- [12] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition," *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 2020, pp. 362-366, doi:10.1109/icaaic48513.2020.9065078.
- [13] R. D. Nurfitra and G. Ariyanto, "Implementasi deep learning Berbasis Tensorflow Untuk Pengenalan Sidik Jari," *Emitor: Jurnal Teknik Elektro*, 2018, vol. 18, no. 1, pp. 22-27, doi:10.23917/emitor.v18i01.6236.
- [14] R. Yamashita *et al.*, "Convolutional Neural Networks: An overview and application in Radiology," *Insights into Imaging*, 2021, vol. 9, no. 4, pp. 611-629, doi:10.1007/s13244-018-0639-9.
- [15] T. Kattenborn *et al.*, "Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, vol. 173, pp. 24-49, doi:10.1016/j.isprsjprs.2020.12.010.
- [16] L. Alzubaidi *et al.*, "Review of Deep Learning: Concepts, CNN Architectures, challenges, applications, Future Directions," *Journal of Big Data*, 2021, vol. 8, no. 7, pp. 1-74, doi:10.1186/s40537-021-00444-8.
- [17] E. S. Wahyuni and M. Hendri, "Smoke and fire detection base on Convolutional Neural Network," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 2019, vol. 7, no. 3, p. 455, doi:10.26760/elkomika.v7i3.455.
- [18] A. Basavaraju *et al.*, "A machine learning approach to road surface anomaly assessment using smartphone sensors," *IEEE Sensors Journal*, 2020, vol. 20, no. 5, pp. 2635-2647, doi:10.1109/jsen.2019.2952857.
- [19] N. Pathak, "Bridge Health Monitoring using CNN," *2020 International Conference on Convergence to Digital World - Quo Vadis (ICCDW)*, 2020, pp. 1-4, doi:10.1109/iccdw45521.2020.9318674.