

BAB II TINJAUAN PUSTAKA

Penelitian ini menjelaskan landasan teori yang berkaitan dengan metode dan dataset yang akan dibuat. Bagian teori membahas mengenai komponen-komponen apa saja yang mendukung semua teori lain yang berhubungan dengan penelitian ini.

2.1 Penelitian Terdahulu

Berikut penelitian yang didasari oleh penelitian sebelumnya dan dijadikan acuan pada tugas akhir ini. Penelitian terdahulu ini meliputi metode penelitian, dataset penelitian dan hasil dari penelitian yang akan digunakan.

Tabel 2.1 Penelitian Terdahulu

No	Penulis	Tahun	Judul	Hasil
1.	Wahyu Andhyka Kusuma, Agus Eko Minarno, dan Mochamad Satriyo Wibowo	2020	Triaxial Accelerometer-based Human Activity Recognition using 1D Convolution Neural Network [4]	Penelitian ini menggunakan dataset WISDM. Hasil akurasi pada penelitian ini adalah 95,9%. Aktivitas dinamis seperti duduk mencapai tingkat akurasi tertinggi yaitu 100%, sedangkan aktivitas turun tangga mencapai tingkat akurasi terendah yaitu 85%.

2.	Hou, Chengli	2020	A Study IMU-Based Human Activity Recognition Using Deep Learning and Traditional Machine Learning [5]	Perbandingan deep learning dan machine learning tradisional menggunakan dataset WISDM menghasilkan akurasi 90% pada deep learning, dan 87% pada machine learning tradisional.
3.	Chew Yong Shan, Pang Ying Han, dan Ooi Shih Yin	2020	Deep Analysis for Smartphone-based Human Activity Recognition [6]	Pada penelitian ini hasil akurasi dataset UCI 90% dan WISDM 87% menggunakan metode CNN dan bi-directional LSTM.
4.	Srivastava, Nitish Hinton, Geoffrey Krizhevsky, Alex Sutskever, Ilya Salakhutdinov, Ruslan	2018	Dropout: A simple way to prevent neural networks from overfitting [7]	Menggunakan dropout disemua layer dan meningkatkan <i>fraction</i> dropout dari 0.0 (tidak ada dropout) menjadi 0.9 dengan <i>step_size</i> 0.1 dan menggunakan masing-masing 20 epoch.

2.2 Human Activity Recognition

Human Activity Recognition ialah pengenalan pola aktivitas manusia yang banyak digunakan dalam aktivitas yang mengacu pada pergerakan seluruh tubuh yang dapat diletakkan di bagian tubuh, lebih tepatnya semacam aktivitas berjalan, berlari, naik tangga, turun tangga, duduk dan berdiri. Pengenalan pola aktivitas manusia banyak digunakan karena aplikasinya di bidang kesehatan, lingkungan, dan lainnya[8].

2.3 Karakteristik Dataset

Penelitian ini menggunakan dataset WISDM atau disebut juga *Wireless Sensor Data Mining* yang memiliki 1.098.207 sampel [9]. Setiap aktivitas dari tubuh manusia diambil sampelnya dengan rate sejumlah 20Hz. Objek eksperimen dari WISDM merupakan 6 aktivitas yang diambil dari 36 subjek. Adapun beberapa aktivitas itu terdiri dari *Walking, Jogging, Upstair, Downstair, Sitting, Standing* yang semuanya termasuk dalam pengumpulan data.

Sensor akselerometer digunakan untuk mengukur pola gaya akselerasi. Gaya akselerasi dibagi menjadi dua bagian yaitu statis dan dinamis. Gaya statis ialah gaya gravitasi yang terus menerus seperti halnya tanpa gerakan, sedangkan gaya dinamis ialah gaya yang merasakan gerakan dan getaran.

Aktivitas ini termasuk aktivitas yang digunakan dalam kehidupan sehari-hari. Untuk setiap aktivitas memiliki tiga sumbu yaitu; sumbu x, y, dan z. setiap sumbu mewakili pergerakan. Sumbu x mewakili gerakan horizontal kaki, sumbu y mewakili gerakan ke atas dan kebawah dan sumbu z mewakili gerakan ke depan.

2.4 Deep Learning

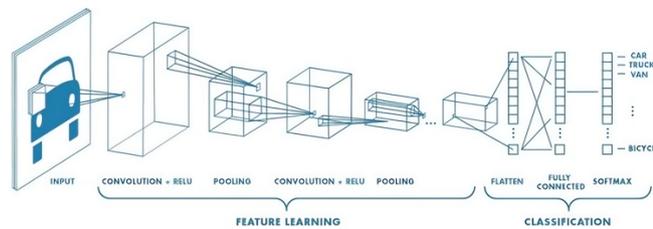
Deep Learning merupakan sebuah teknik pada *machine learning* dengan banyak layer untuk mengolah informasi non-linier serta melakukan ekstraksi fitur, pengenalan pola, dan klasifikasi[10]. *Deep Learning* adalah paradigma yang memungkinkan model komputasi yang terdiri dari beberapa lapisan pemrosesan untuk mempelajari representasi data dengan berbagai tingkat abstraksi. Banyak penelitian yang telah membuktikan penggunaan *deep learning* dapat meningkatkan kinerja banyak orang terutama dalam pengenalan objek ucapan dan visual dan lainnya[11].

2.5 CNN (*Convolutional Neural Network*)

Metode CNN merupakan salah satu metode *deep learning* yang dapat menganalisis data set besar secara akurat[12]. CNN adalah jenis jaringan saraf tiruan yang terdapat dari beberapa lapisan, diantaranya *pooling layer, dense layer, dan fully connected layer*. Oleh karena itu, CNN memiliki variasi yang sangat banyak[13].

2.5.1 Konsep CNN

Cara kerja konsep CNN memiliki persamaan dengan *Multi Layer Perceptron* (MPL), akan tetapi setiap neuron pada CNN ditampilkan dalam bentuk data dua dimensi. *Convolutional Neural Network* juga termasuk bagian jenis dari *Deep Neural Network*. Oleh karena itu sifat prosesnya konvolusi CNN hanya dapat diterapkan pada data yang mempunyai struktur dua dimensi semacam citra dan suara[14]. CNN memiliki istilah lain yaitu, sebuah arsitektur yang bisa dilatih dan terbentuk dari beberapa tahap. Masukan (*input*) dan keluaran (*output*) dari masing-masing tahap adalah terdiri dari beberapa jumlah *array* biasanya disebut juga dengan *feature map*[15]. **Gambar 2.1** menunjukkan sebuah arsitektur sederhana CNN.



Gambar 2.1 Arsitektur Sederhana CNN[16]

2.5.2 Arsitektur Jaringan CNN

CNN terbagi menjadi beberapa layer[17] yaitu:

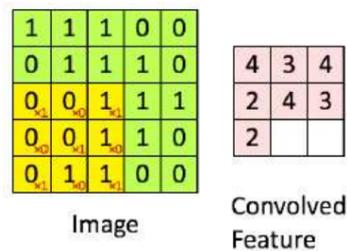
a. *Input Layer*

Input layer merupakan layer pertama yang berfungsi untuk memasukkan data, memuat data, dan melanjutkan layer berikutnya.

b. *Convolution Layer*

Layer dari output sebelumnya melakukan convolution layer dengan operasi konvolusi. Layer konvolusi merupakan dasar dari operasi CNN. Konvolusi memiliki arti matematis yang mengaplikasikan sebuah fungsi pada output dengan fungsi lainnya secara berulang. Dalam pengolahan citra, konvolusi yaitu mengaplikasikan sebuah *kernel* atau yang disebut dengan kotak kuning pada citra diseluruh offset sebagaimana yang dapat ditinjau pada **Gambar 2.2**. Secara keseluruhan, kotak hijau yaitu citra yang akan dikonvolusi. Kernel

akan bergerak ke sudut kiri atas ke kanan bawah. Hasil konvolusi dari citra akan diperlihatkan pada gambar sebelah kanan.



Gambar 2.2 Operasi Konvolusi [14]

c. *Activation Layer*

Aktivasi Layer akan mengenalkan properti non-linier dengan jaringan yang kemudian diikuti oleh lap konvolusi. Tujuan utama dari *activation layer* tersebut untuk mengubah sinyal input yang berasal sebuah node menjadi sinyal output. Beberapa hasil dari output akan diteruskan dengan layer berikutnya. Sejumlah besar dari fungsi aktivasi Rectified Linier Unit (ReLU) diterapkan dalam praktik kehidupan nyata sebab dianggap lebih efisien secara komputasi dan menyatu lebih cepat jika dibandingkan dengan fungsi aktifitas lainnya[18].

Berikut fungsi yang menyatakan bahwa nonlinier ReLU untuk mengambil input x pada persamaan :

$$f(x) = \max(0, x)$$

Fungsi aktivasi softmax tersebut digunakan untuk aktivasi pada layer output.

Layer output termasuk juga bagian dari fullyconnected layer, perbedaan dari kedua layer adalah fungsi dalam penggunaan aktivasi softmax pada layer output dengan aktivasi ReLU pada fully-connected layer.

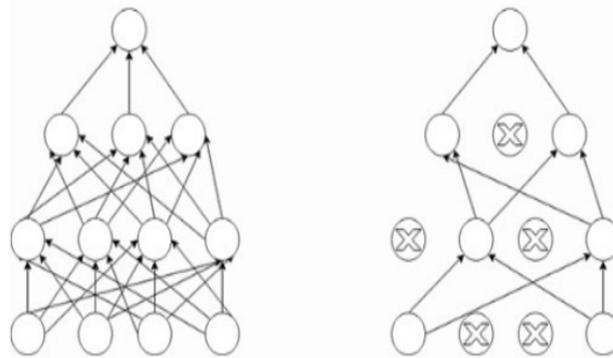
d. *Pooling Layer*

Pooling Layer pada umumnya memperkenalkan dua layer konvolusi secara berurutan dengan mengurangi jumlah parameter dan perhitungan yang dibutuhkan oleh jaringan. Hal tersebut membantu mengatasi overfitting sambil melatih sebuah dataset. Fungsi Max Polling merupakan teknik pooling yang paling sering digunakan sebab

dianggap memberikan hasil yang lebih baik dalam aplikasi kehidupan nyata jika dibandingkan dengan teknik pooling lainnya[18].

e. *Dropout*

Fungsi dropout merupakan teknik yang digunakan dalam melakukan pencegahan terhadap permasalahan overfitting[19]. Layer yang sepenuhnya terhubung menepati sebagian besar ke parameter, karena neuron mengembangkan dan saling ketergantungan satu sama lain. Selama pelatihan yang mengekang kekuatan individu dari setiap neuron menyebabkan pemasangan data pelatihan yang berlebihan [7]. Fungsi dropout dapat dipresentasikan pada fully connected layer di mana cara kerjanya yaitu dengan menonaktifkan beberapa neuron yang tidak diperlukan secara random. **Gambar 2.3** menampilkan ilustrasi dropout.



Gambar 2.3 Ilustrasi Penggunaan Dropout[17]

f. *Fully Connected Layer*

Full connection layer memiliki koneksi penuh terhadap semua layer aktivasi neuron yang terdapat pada layer sebelumnya. Layer terakhir yang terhubung sepenuhnya dapat menghasilkan output bersih pada fungsi aktivasi tersebut yaitu fungsi softmax yang tergantung dari jumlah kelas atau dikategorikan jumlah dari klasifikasi[4][20].

2.6 Evaluasi

Evaluasi deep learning dapat diperoleh dengan cara membandingkan hasil nilai yang sebenarnya dengan nilai hasil prediksi yang didapatkan melalui dataset yang sudah ada[21]. Dengan perbandingan yang dapat ditinjau melalui table confusion matrix[22] pada **Gambar 2.4** dibawah ini.

		True/Actual Class	
		Positive (P)	Negative (N)
Predicted Class	True (T)	True Positive (TP)	False Positive (FP)
	False (F)	False Negative (FN)	True Negative (TN)
		$P=TP+FN$	$N=FP+TN$

s

Gambar 2.4 *Confusion Matrix*[23]

Perbandingan nilai pada **Gambar 4** ini menjelaskan bahwa nilai tersebut dibagi menjadi 4 bagian. Yaitu:

1. *True Positive* (TP) adalah memprediksi nilai dengan nilai yang sebenarnya atau dapat disebut juga sebagai data benar yang sesungguhnya.
2. *True Negative* (TN) adalah nilai yang diprediksi oleh true, sedangkan nilai yang sebenarnya negative atau bisa disebut juga data yang salah.
3. *False Positive* (FP) adalah nilai yang diprediksi false sedangkan nilai yang sebenarnya positive atau dapat disebut sebagai data yang seharusnya diklasifikasikan benar, namun dalam proses pengklasifikasian menjadi salah.
4. *False Negative* (FN) adalah nilai yang diprediksi false sedangkan nilai yang sebenarnya negative atau dapat disebut sebagai data yang seharusnya diklarifikasikan salah, namun dalam hasil klasifikasi menjadi benar.

Confusion matrix dapat juga digunakan dalam menghitung nilai dari accuracy, precision, recall dan fl-score. Hal tersebut akan dijelaskan dalam penjelasan dan persamaan pada berikut ini:

1. *Accuracy* adalah jumlah yang tergolong sebagai kategori bgenar dari prediksi dataset.
2. *Precision* adalah nilai keakuratan yang diperoleh dari rasio prediksi benar dengan semua hasil yang diprediksi benar[21].
3. *Recall* dapat dihitung dari rasio data yang positif benar dengan keseluruhan data yang benar.
4. *F1-Score* adalah hasil dari rata-rata perbandingan antara recall dan precision yang diperoleh[21].