

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Sistem gerak merupakan perilaku dimana manusia dapat melakukan gerakan yang dikehendaki oleh satu kesatuan tubuh yang terdiri dari tulang, persendian, dan otot [54]. Menurut Direktorat Jendral Kependudukan dan Pencatatan Sipil (Dukcapil) terdapat 30,16 juta jiwa penduduk lansia yang berumur 60 tahun keatas di Indonesia pada tahun 2021 [45] dan berdasarkan data berjalan 2020 dari Biro Pusat Statistik (BPS) menyatakan bahwa jumlah penyandang disabilitas di Indonesia mencapai 22,5 juta jiwa [46]. Masyarakat usia 90 tahun, pada wanita 32% dan laki-laki 17% mengalami patah tulang panggul dan 12-20% meninggal akibat komplikasi. Pada usia 65 tahun, massa tulang menurun 10% dan 20% pada usia 80 tahun. Pada wanita, kehilangan massa tulang lebih tinggi, 15-20% usia 65 tahun dan 30% usia 80 tahun. Pada laki-laki kehilangan massa tulang 1% per tahun sesudah usia 50 tahun, sedangkan wanita pada usia 30 tahun, dengan laju penurunan 2-3% per tahun sesudah mengalami menopause [48]. Rehabilitasi merupakan kegiatan yang dapat membantu para penderita penyandang penyakit ringan, serius atau cacat yang memerlukan pengobatan medis untuk mencapai kesehatan fisik psikologis yang maksimal [49]. Berdasarkan data tersebut, rehabilitas sangat di perlukan bagi orang yang memiliki gangguan pada sistem gerak tubuh dan membutuhkan alat bantu berupa alat medis atau aplikasi dengan sistem deteksi yang berfokus pada pengenalan aktivitas manusia, agar mempermudah proses terapi pada pasien [48]. Sistem *Human Activity Recognition* (HAR) merupakan proses penafsiran gerak manusia yang memungkinkan sistem dapat mendeteksi aktivitas sehari-hari manusia [6] seperti berjalan, duduk, berbaring dan berdiri. Selain itu, pengenalan aktivitas manusia dapat diaplikasikan pada pengenalan sistem perawatan manusia *Activity of Daily Living* (ADL) yang dapat diimplementasikan pada bidang medis sebagai alat bantu proses terapi pada perawatan lanjut usia (lansia). Selain itu, sistem dapat dijadikan alat untuk membantu para dokter dalam menyimpan data pasien, seperti pasien yang mengalami cacat, baik itu penyakit sejak lahir atau akibat kecelakaan. Sehingga membutuhkan pengobatan terapi dalam proses

pengembalian sistem gerak tubuh. Sistem ini memungkinkan dapat membantu penyembuhan dan memudahkan staf medis dalam melakukan pekerjaan [5]. Sehingga, diharapkan dengan penerapan sistem pengenalan aktivitas manusia dapat membantu dan mengurangi tingkat pasien cacat dengan cara pengaplikasian sistem pada proses penyembuhan terapi.

Perkembangan dunia teknologi berbasis *Artificial Intelligence* (AI) kini telah banyak membantu dan memudahkan kehidupan manusia. Salah satu perubahan yang sangat terasa yaitu perkembangan teknologi pada bidang deteksi objek dengan memanfaatkan sistem pada komputer. Deteksi gambar, pengenalan wajah dan bahkan pengenalan aktivitas manusia sehari-hari sedang ramai diperbincangkan. Teknologi yang memanfaatkan komputer dan visi mesin salah satunya adalah HAR. Sistem HAR dikelompokkan menjadi sistem berbasis sensor dan visi video. Dengan pengembangan teknologi *machine learning* dan *release depth sensors* seperti Microsoft Kinect dan NVIDIA's GPUSs memungkinkan sistem untuk mengidentifikasi perilaku manusia menggunakan perangkat video dan sensor gerak dengan cepat [9], namun penggunaan visi video masih terbilang rumit karena objek tidak dapat bergerak bebas dengan keterbatasan kamera dan memerlukan biaya besar. Sistem HAR berbasis sensor *smartphone* telah dilengkapi oleh sensor *accelerometer* dan *gyroscope*, hal tersebut menguntungkan pengguna agar dapat bergerak dengan bebas tanpa khawatir objek tidak terdeteksi, sehingga dapat digunakan sebagai *input* sistem dan membantu proses berjalannya sistem deteksi dengan baik [7]. Namun penggunaan *smartphone* masih memiliki kekurangan dalam hal melakukan komputasi secara *real-time* dengan algoritma yang kompleks karena kapasitas *power supply* lebih sedikit dan kemampuan *processor* yang terbatas. Dunia telah melangkah menuju masyarakat menua yang menghadapi masalah ekonomi dan sosial, sehingga meningkatnya permintaan perawatan kesehatan lansia. Saat ini banyak aplikasi yang membutuhkan fasilitas teknologi berbasis sensor seperti pada aplikasi kesehatan yang di terapkan pada *activity reminder*, deteksi jatuh, instruksi rehabilitas, dan evaluasi kesehatan dalam sistem kehidupan [10]. Alat medis dengan kemampuan sensor dan aktuator dapat menjadi solusi teknologi yang sangat membantu. Namun tidak hanya itu, lansia harus memiliki kemampuan melakukan ADL untuk membantu menjaga

kehidupan yang sehat aman dan fungsional [11].

Dalam beberapa tahun sebelumnya, algoritma deteksi objek berbasis sensor pada *smartphone* mengalami perkembangan yang signifikan. Pada penelitian Shaohua Wan dkk menggunakan model *Long Short Term Memory* (LSTM) dan UCI HAR yang telah dilatih secara *real-time* sebagai dataset menunjukkan bahwa model tersebut mencapai *accuracy* yang baik, dengan hasil *accuracy* 89,01%, *precision* 89,14%, *recall* 88,99%, dan *f1-score* 88,99% [12]. Namun, masih terdapat kekurangan pada struktur model *artificial neural network* yang digunakan dalam percobaan, sehingga ketepatan sistem pada algoritma pemodelan masih dapat dioptimalkan. LuKun Wang dkk pada penelitian mengenai pengenalan aktivitas manusia berbasis sensor menggunakan model LSTM. Algoritma yang digunakan pada penelitiannya yaitu *decision tree*, *random forest* dan H-LSTM dengan hasil *accuracy* masing - masing 86,33%, 90,84%, dan 91,65%. Namun, kumpulan data saat ini belum dapat mencakup aktivitas manusia secara komprehensif, masih butuh penelitian lebih mengenai H-LSTM untuk mencapai *accuracy* yang optimal [19]. Tahmina zebin dkk, melakukan penelitian terhadap pengenalan aktivitas manusia dari deret waktu inersia menggunakan *batch normalize* pada model LSTM. Pada bagian implementasi LSTM, Zebin dkk menggunakan input data dalam *multi* dimensi yang berisikan tiga sumbu *accelerometer* dan tiga sumbu *gyroscope*. Jumlah kelas yang digunakan adalah enam diantaranya *walking*, *walking upstairs*, *walking downstairs*, *sitting*, *standing*, dan *laying*. Pada dataset UCI HAR, berdasarkan hasil penelitiannya didapatkan akurasi rata rata pengenalan sebesar 92% untuk enam aktivitas menggunakan data input *accelerometer* dan *gyroscope*. Namun, agar sistem terus berkembang dan menghasilkan *output* yang lebih optimal maka akan dilakukan pengujian pada dataset aktivitas fisik yang lebih besar dari UK Biobank [32]. Yu Zhao dkk pada penelitiannya mengenai pengenalan aktivitas manusia menggunakan *wearable sensor* dengan model Bidir-LSTM berhasil menaikkan akurasi menjadi 91,1% pada dataset UCI HAR, pada penelitiannya menyatakan bahwa *window size* merupakan kunci parameter, *window size* kecil tidak menjamin kontinuitas informasi dan *window* yang terlalu besar menyebabkan kesalahan klasifikasi. Maka, nilai-nilai *hyperparameter* harus ditentukan sesuai dengan arsitektur tertentu. Namun,

meskipun pencarian *grid* dapat diterapkan, rentang pencarian harus diubah secara manual dan nilainya selalu berubah. Solusinya, diperlukan cara adaptif agar dapat secara otomatis untuk menyesuaikan proses pencarian dan juga membuat arsitektur *neural network* berkembang [33].

Vasundhara Ghate dkk melakukan penelitian terhadap *hybrid deep learning* untuk *smartphone* berbasis sensor, percobaan dilakukan pada WISDM dataset dan UCI HAR dengan hasil yang didapat pada DNN + LSTM dengan 81% pada dataset WISDM dan 87.79% pada UCI dataset. Penelitian ini membuktikan keuntungan menggunakan *hybrid deep learning* untuk pengenalan aktivitas manusia berbasis sensor. Namun, patut dicoba untuk mengubah jumlah nilai pada *hyperparameter* seperti penetapan nilai *batch size*, *learning rate*, *epoch*, dan lainnya [34]. Wen-Hui Chen dkk, menggunakan model LSTM-RNNs dikombinasikan dengan informasi *scene* untuk aktivitas manusia. Dataset yang digunakan yaitu PAMAP2 yang memiliki total kelas 13 yaitu *lying*, *sitting*, *standing*, *walking*, *running*, *watching TV*, *computer work*, *ascending stairs*, *descending stairs*, *vacuum cleaning*, *ironing*, *folding laundr*, dan *house cleaning*. Dengan informasi *scene* yang telah disediakan oleh scene pengenal, maka sistem dapat mengidentifikasi lokasi dan mengesampingkan aktivitas yang tidak mungkin terjadi. Ketepatan sistem dalam mengklasifikasi yaitu sebesar 82,57%, namun hasil masih dapat dioptimalkan lagi agar *output* yang dihasilkan lebih optimal [35].

Fayez Alharbi dkk pada penelitiannya memperkenalkan penggunaan *Wasserstein Generative Adversarial Network* (WGAN) untuk menghasilkan sensor data pada pengenalan aktivitas manusia. Fayez dkk menyelidiki WGAN pada lima kelas berbeda dari aktivitas manusia yang kurang terwakili pada dua kumpulan data publik, kemudian mengevaluasi keragaman dan kualitas *generated synthetic* yang dihasilkan data sensor. Ketika klasifikasi menggunakan CNN 1D dilatih pada *synthetic* dan diuji pada data nyata ditemukan nilai *f1-score* lebih dari 75%, selain itu Fayez dkk melakukan *oversampling set* pelatihan yang tidak seimbang dan menemukan peningkatan kinerja *f1-score* secara keseluruhan antara 7% dan 10%. Pada penelitiannya menggunakan 1D-CNN-classified WGAN-features terhadap fitur yang di produksi *Synthetic Minority Over-sampling*

Technique (SMOTE) terlihat efek keberhasilan dari penggunaan data sensor *synthetic* dengan meningkatnya nilai *f1-score* 1% hingga 2%. Namun, evaluasi serupa menggunakan LSTM tidak menemukan keuntungan langsung dari penelitiannya, karena saat ini belum ada pendekatan atau kerangka kerja yang diakui secara luas untuk evaluasi *synthetic* data sensor [36]. Saurabh Gupta [37] membandingkan kinerja pada sensor *smartphone* dan *smartwatch* menggunakan WISDM dataset. Transformasi data selama *preprocessing* dilakukan menggunakan pendekatan *sliding window*. Tidak ada rekayasa fitur manual dilakukan dalam penelitian ini. Model dasar seperti DeepConvLSTM dan *InceptionTime* dibuat menggunakan AutoML berdasarkan paket McFly *open-source* yang secara signifikan mengurangi upaya dalam membuat model kompleks *deep neural network*. Secara keseluruhan, hasil menunjukkan bahwa *hybrid deep learning model* dapat secara efisien dan otomatis mengekstrak fitur *spatial-temporal* dari *raw* data sensor untuk mengklasifikasikan tindakan manusia yang kompleks. Namun, pada penggunaan model DeepConvLSTM tingkat akurasi lebih rendah yaitu 87,65% pada *smartwatch* dan 75,31% pada *smartphone* [37].

Tanvir Mahmud [38] pada penelitiannya menggunakan arsitektur *Deep Multi-stage LSTM* berdasarkan temporal *Feature Aggregation*. Dataset yang digunakan pada *training* dan evaluasi adalah dataset dari *physionet*. Dataset berisikan beberapa *bio-signal* seperti EKG, PPG dengan beberapa sensor gerak seperti *low noise accelerometer*, *high noise accelerometer*, dan *gyroscope*. Skema memberikan kinerja yang stabil dengan *precision* 84,3%, *recall* 83,4%, *f1-score* 83,9% dan *accuracy* 83,2%. Pada penelitiannya ditemukan bahwa, penggabungan beberapa sensor memberikan peningkatan kinerja yang signifikan dibandingkan dengan sensor tunggal untuk mengintegrasikan beragam representasi dari aktivitas tertentu. Namun, hasil masih dapat dioptimalkan lebih tinggi dan skema ini terbuka untuk mengintegrasikan banyak sensor secara efisien dalam proses pengenalan aktivitas manusia [38]. Rajiv vicent mengidentifikasi aktivitas manusia menggunakan visi video dengan memilih 5 kegiatan dari dataset HMDB51 dan UCF101 dengan membandingkan hasil model Bi-LSTM dan LSTM, kemudian menerapkan pemodelan fitur *inception-v3* sehingga hasil yang didapat saat belatih lebih baik menggunakan model LSTM. Hasil *accuracy* pada

model klasifikasi LSTM dengan inception-v3 80,12%, BiLSTM dengan inception-v3 80,75% dan LSTM dengan Xception 77,02% [9]. Namun, Pada penggunaan visi video masih terdapat keterbatasan pada jarak jangkauan pengguna dari sensor atau objek bergerak bebas dalam adegan dengan derajat yang berbeda sehingga aktivitas tidak dapat dikenali. Hal ini menjadikan sistem HAR *real-time* seperti video tidak praktis [12].

Pada penelitian Jian Sun dkk [51] melakukan penelitian pada pengenalan aktivitas manusia menggunakan *deep* arsitektur berdasarkan kombinasi *convolutional layer*. Hasil *experiment* pada penggunaan model *baseline* CNN dengan *f1-score* 78,3% dan 90,6% pada model CNN-LSTM-ELM dengan menggunakan dataset Opportunity. Namun, masih perlu dikembangkan dengan meningkatkan kemampuan adaptif pembelajaran sekuensial. Dengan demikian *machine learning* mampu untuk terus *learning* sambil melakukan verifikasi [51]. Pada penelitian Ignatov Andrey melakukan penelitian menggunakan model CNN + stat. Features dengan jumlah 196 filters (C1) menggunakan WISDM dataset dengan hasil eksperimen yang memuaskan yaitu 90.42% pada akurasi. Namun, masih dapat ditingkatkan dengan menekankan eksperimen lintas dataset pada arsitektur *platform-independen* tidak hanya pada konteks pengguna, tetapi juga pada perangkat dengan kalibrasi *accelerometer* yang berbeda [52]. Terlihat penggunaan model CNN pada pengenalan aktivitas manusia memiliki akurasi dibawah dari penelitian LuKun Wang dkk [19] dengan akurasi 91.65% dan penelitian Tahmina Zebin dkk memiliki akurasi sebesar 92%, karena model LSTM memiliki arsitektur yang memuaskan dalam mengatasi klasifikasi deret waktu [18], LSTM memiliki kemampuan untuk menghapus atau menambahkan informasi kedalam sel melalui struktur yang dirancang dengan baik [12] sehingga LSTM dapat mengatasi kekurangan pada model CNN dan menghasilkan output yang lebih optimal.

Pada penelitian Fauzan Zharfan dkk [50] melakukan penelitian alat bantu senam rehabilitasi mandiri untuk lansia menggunakan metode *Euclidean Distance*. Pengujian menunjukkan bahwa pembacaan gestur tangan menggunakan Kinect dapat digunakan untuk membaca gerakan senam rehabilitasi *Range of Motion* (ROM), pembacaan dilakukan pada 20 titik tubuh manusia dan hasil

terdeteksi 16 titik tubuh yang akurat. Namun, penelitian ini masih harus dikembangkan agar dapat mendeteksi titik-titik tubuh lebih akurat dengan memperbaharui *device* Kinect ke versi terbaru dan pengembangan pemrograman agar dataset gerakan rehabilitasi bisa bervariasi bukan hanya gerakan ROM [50]. Pada konteks HAR terdapat lima metode yang memenuhi, yaitu *Deeply connected network* (DFN), *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), *Stacked Autoencoder* (SAE), dan *Restricted Boltzmann Machine* (RBM) [13]. Jika hanya untuk menganalisis dari perspektif ekstraksi fitur, semua metode tersebut serupa, dengan perbedaan jumlah *layers* dan cara *layers* terhubung satu sama lain. Berdasarkan hasil evaluasi pada penelitian Rajiv Vincent dkk, menggunakan dataset berupa video pada metode LSTM dengan fitur inception-v3 mendapatkan akurasi sebesar 80,12% [9]. Sedangkan pada penelitian LuKun Wang dkk, deteksi objek menggunakan sensor *accelerometer* dan *gyroscope* yang tertanam pada *smartphone* menggunakan metode LSTM dapat menunjukkan akurasi yang lebih tinggi 91,65% [19] dan penelitian Yu Zhao dkk mengusulkan untuk pemilihan nilai parameter *window size*. Oleh karena itu, berdasarkan metode dan analisis pada evaluasi sensor tersebut, pada Tugas Akhir ini penulis menggunakan model LSTM berbasis sensor *smartphone* dan *window size* 128 pada UCI HAR sebagai dataset dengan modifikasi pada *hyperparameter* dalam meningkatkan akurasi deteksi pada sistem.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka rumusan masalah dari Tugas Akhir ini sebagai berikut:

1. Menentukan *hyperparameter* yang mampu memberikan kinerja sistem deteksi objek yang baik pada pengenalan aktivitas manusia.
2. Menganalisis pengaruh jumlah nilai fitur *hidden layer* dan iterasi pada model *train* dan validasi sistem.
3. Perbandingan konfigurasi *hyperparameter* terbaik berdasarkan parameter performansi *accuracy*, *precision*, *recall*, dan *f1-score*.

1.3 Tujuan dan Manfaat

Tujuan dan penelitian Tugas Akhir ini adalah sebagai berikut:

1. Mendapatkan parameter performansi yang baik untuk menghasilkan sistem objek deteksi dengan keluaran yang optimal.
2. Menganalisis pengaruh jumlah nilai fitur *hidden layer* dan *iterasi* pada sistem deteksi objek.
3. Menganalisis data perbandingan konfigurasi *hyperparameter* berdasarkan parameter performansi *accuracy*, *precision*, *recall*, dan *f1-score*.

1.4 Batasan Masalah

Untuk membatasi cakupan pengerjaan dan memfokuskan area kerja dalam penelitian ini dibentuk batasan masalah sebagai berikut:

1. Sistem deteksi objek berbasis sensor *accelerometer* dan *gyroscope*.
2. Algoritma deteksi objek yang digunakan adalah LSTM.
3. Berfokus pada analisis jumlah nilai fitur *hidden layer* dan *iterasi* pada sistem pengenalan aktivitas manusia.
4. Parameter performansi yang digunakan *accuracy*, *precision*, *recall*, dan *f1-score*.
5. Dataset yang digunakan berisi 6 aktivitas untuk mengklasifikasi gerakan yaitu *walking*, *walking upstairs*, *walking downstairs*, *sitting*, *standing*, dan *laying*.
6. Dataset digunakan sebagai data *train* 70% dan data *validasi* 30%.
7. Perancangan sistem menggunakan Google Colab dengan bahasa pemrograman python dan *framework* TensorFlow.

1.5 Metode Penelitian

Metode penelitian yang digunakan pada penelitian Tugas Akhir ini adalah sebagai berikut:

1. Studi literatur

Pada tahap ini penulis mengumpulkan berbagai referensi seperti *paper*, yang terkait dengan topik pengenalan aktivitas manusia pada sensor *accelerometer* dan *gyroscope* dan metode LSTM.

2. Pengumpulan data

Pada tahap ini penulis menggunakan dataset public yang diambil dari internet. Dataset yang digunakan berisi 6 kelas yang telah ditentukan yaitu *walking*, *walking upstairs*, *walking downstairs*, *sitting*, *standing*, dan *laying*.

3. Perancangan sistem

Pada tahap ini model LSTM digunakan untuk melatih sistem deteksi objek dengan *input* sinyal *accelerometer* dan *gyroscope* menggunakan bahasa pemrograman python dengan *framework* TensorFlow menggunakan Google Colab untuk proses *training* dan *evaluasi*.

4. Simulasi dan analisis

Pada tahap ini penulis melakukan simulasi sebanyak lima belas kali pada model dengan mengganti *hyperparameter* fitur *hidden layer* dan iterasi. Analisis dilakukan dengan membandingkan hasil konfigurasi menggunakan *accuracy*, *precision*, *recall*, dan *f1-score*.

5. Kesimpulan

Pada tahap ini penulis membuat kesimpulan dari penelitian yang telah dilakukan dan penetapan konfigurasi terbaik untuk diimplementasikan pada sistem pengenalan aktivitas manusia.

1.6 Sistematika Penulisan

Penulisan sistematika Tugas Akhir ditulis sebagai berikut:

• **BAB II DASAR TEORI**

Bab ini berisi landasan teori yang digunakan pada penelitian seperti LSTM, *accelerometer*, *gyroscope*, fitur *hidden layer* dan iterasi.

• **BAB III PERANCANGAN SISTEM**

Bab ini berisi alur perancangan sistem berupa diagram alir, parameter yang menjadi referensi penelitian, dan desain rancangan setiap skenario.

• **BAB IV ANALISIS SIMULASI SISTEM**

Bab ini menampilkan tabel dan grafik untuk mempermudah proses analisis dan pembahasan hasil evaluasi dari performansi model LSTM yang telah dilatih menggunakan UCI HAR dataset.

• **BAB V KESIMPULAN DAN SARAN**

Bab ini memaparkan kesimpulan hasil analisis dan saran untuk pengembangan selanjutnya.