LAPORAN PENELITIAN ISPO

Pengembangan Sistem Autentikasi Berbasis Wi-Fi Sensing Menggunakan Data CSI dan Algoritma *Machine Learning* dengan Mikrokontroler ESP32

disusun oleh:

Muhammad Fathan Haroki
Oruzgan Abimanyu Adi

Rekayasa Teknologi dan Robotika

Sekolah Menengah Atas Unggulan Rushd

Sragen, Jawa Tengah

2025

ABSTRAK

Tingginya angka kriminalitas menjadikan isu kejahatan dan kriminalitas untuk menjadi perhatian khusus. Sistem keamanan yang efektif, efisien, dan adaptif terhadap kondisi lingkungan diperlukan untuk mengatasi permasalahan tersebut. Beberapa sistem yang telah diusulkan, seperti *Circuit-Closed Television* (CCTV), sistem berbasis NFC (*Near Field Communication*), dan GPS (*Global Positioning System*), masih belum secara maksimal mengatasi permasalahan yang ada. Untuk melampaui keterbatasan ini, dikembangkan sistem autentikasi berbasis *Wi-Fi sensing* menggunakan mikrokontroler ESP32. Data CSI (*Channel State Information*) yang merupakan status informasi sinyal *Wi-Fi* diekstrak menggunakan ESP32 dengan *framework* ESP-IDF. Model *machine learning* SVM dengan pendekatan *two-stage* dikembangkan untuk mengklasifikasi data CSI tersebut. Pada tahap pertama, model mengidentifikasi apakah individu yang terdeteksi merupakan *intruder* atau individu terdaftar. Jika individu terdaftar, tahap kedua dilakukan untuk mengklasifikasi berdasarkan karakteristik sinyal yang dipantulkan. Hasil eksperimen menunjukkan bahwa sistem ini mencapai akurasi 100% hingga 96% dalam mengidentifikasi 2 hingga 10 individu terdaftar. Selain itu, sistem juga mampu mengidentifikasi individu yang tidak dikenal dengan akurasi sebesar 90% pada dataset yang terdiri dari 9 individu terdaftar. Sistem ini memiliki beberapa keunggulan, seperti performa model yang lebih baik, efisiensi yang tinggi, biaya rendah, dan kemudahan dalam implementasi dikarenakan penggunaan mikrokontroler ESP32 sebagai alat dengan portabilitas tinggi.

Keywords: ESP32, *Wi-Fi Sensing*, *User Authentication*, *Machine Learning*, *Channel State Information*.

BAB I
PENDAHULUAN

1.1 LATAR BELAKANG

Dalam beberapa tahun terakhir, teknologi mengalami perkembangan pesat untuk mengatasi masalah-masalah kriminalitas (Anderez et al., 2021). Banyaknya inovasi sensor dalam bidang keamanan untuk mendeteksi individu yang tidak dikenali seperti CCTV (*Circuit-Closed Television*) (Piza, Welsh, Farrington, & Thomas, 2019), NFC (*Near Field Communication*) (Chen et al., 2015), *Audio* (Lawrence, La Vigne, Goff, & Thompson, 2018), dan GPS (*Global Positioning System*) (Griffiths, Johnson, & Chetty, 2017) semakin memudahkan banyak pihak dalam mengatasi dan mengurangi tindakan kriminal. Namun, masih terdapat beberapa kekurangan yang perlu diperhatikan.

Dalam hal sensor berbasis kamera, terdapat beberapa kelemahan seperti ketergantungan pada cahaya dan juga kebutuhan garis pandang yang jelas. Sensor audio juga menimbulkan masalah privasi yang mengganggu kenyamanan pengguna. *Wearable sensors* seperti *Smart Ring* dalam mendeteksi pihak yang dikenali terkadang memberikan ketidaknyamanan terhadap pengguna, dan juga sering kali memerlukan pemakaian dengan cara tertentu untuk pengoperasian yang akurat. Sensor biometrik seperti *fingerprint* juga cukup rentan terhadap serangan *hacking* dan dapat dieksploitasi menggunakan data yang palsu.

Teknologi *wireless* telah menjadi salah satu solusi dalam mengatasi berbagai keterbatasan dan permasalahan yang ada. Salah satu contohnya adalah penggunaan sinyal RF *(Radio Frequency)* sebagai sensor yang dapat digunakan dalam bidang keamanan (Anderez et al., 2021). *Wi-Fi sensing* belakangan ini telah mendapatkan popularitasnya dengan bertambahnya penggunaan *Wi-Fi* dalam beberapa tahun terakhir. Para peneliti mulai memanfaatkan sinyal *Wi-Fi* sebagai sensor alternatif untuk mendeteksi dan memahami pergerakan, posisi dan aktivitas objek atau individu. Setelah sinyal *Wi-Fi* berpropagasi dari sebuah *transmitter*, sinyal tersebut akan mengalami beberapa pantulan (*reflection),* penyerapan (*absorption*)*,* dan tersebar (*scatter*) terhadap objek atau manusia sebelum sampai pada *receiver* (Suroso, Adiyatma, & Cherntanomwong, 2023). Perbedaan fisik pada manusia memengaruhi cara sinyal berpropagasi, yang dapat digunakan sebagai parameter pendeteksi individu dari mengklasifikasi sinyal tersebut dengan bantuan algoritma *Machine Learning* (Wang et al., 2021). Sinyal *Wi-Fi* akan memberikan dua jenis informasi mengenai kualitas sinyal atau kondisi sinyal, yaitu: CSI (*Channel State Information*) dan RSSI (*Received Signal Strength Indicator*).

RSSI merupakan ukuran seberapa kuat sebuah sinyal diterima. RSSI mudah diperoleh di semua perangkat *Wi-Fi* tanpa perangkat keras tambahan. Namun, stabilitas dari RSSI tidak terjamin, walaupun berada di area *indoor* (Parameswaran, Husain, & Upadhyaya, n.d.). Oleh karena itu, RSSI tidak terlalu efektif dalam digunakan sebagai parameter untuk mengenali suatu individu, sedangkan CSI berisi informasi status sinyal *Wi-Fi* yang menjelaskan bagaimana sinyal itu berpropagasi. Data tersebut akan menangkap informasi kompleks mengenai transmisi suatu sinyal. Informasi yang diberikan oleh data CSI ini penting untuk mengoptimalkan sinyal *Wi-Fi* dan memahami karakteristik lingkungan dari tempat terjadinya transmisi sinyal (Mosharaf, Kwak, & Choi, 2024). CSI merupakan matriks yang terdiri dari nilai-nilai kompleks berisi amplitudo dan pergeseran fasa jalur *multipath* pada *Wi-Fi*. Akses langsung ke data *Channel State Information* (CSI) dari perangkat *Wi-Fi* terbatas pada *hardware* dan *software* tertentu, sehingga menjadi tidak praktis untuk penggunaan skala besar yang dimana akan menjadi tantangan bagi orang-orang untuk mengambil data CSI.

Untuk menangani keterbatasan di atas, solusi alternatif yang dapat dilakukan yaitu dengan pengembangan sebuah sistem autentikasi manusia yang dapat mengenali pihak yang dikenal maupun tidak dikenal dengan teknologi *Wi-Fi sensing* yang memanfaatkan data CSI. Penggunaan mikrokontroler ESP32 sebagai alat *receiver* diaplikasikan dikarenakan beberapa kelebihan seperti

low-cost, mudah dalam mendapatkan data CSI dengan ESP-IDF, sistem yang tidak terlalu kompleks, perangkat yang ringan dan mudah dalam pengoperasiannya. Untuk memproses sinyal yang lebih akurat dan minim *noise*, akan diterapkan algoritma DWT (*Discrete Wavelet Transform*) pada data CSI sebelum akhirnya diklasifikasi dengan *machine learning*. algoritma SVM (*Support Vector Machine*) yang merupakan model yang umum digunakan secara luas, akan digunakan dalam penelitian ini untuk mendapatkan hasil klasifikasi yang efektif.

1.2 RUMUSAN MASALAH

Penelitian ini akan menjawab beberapa pertanyaan-pertanyaan masalah berikut.

1. Bagaimana cara mengambil dan mendapatkan data CSI dari sebuah mikrokontroler ESP32?
2. Bagaimana fisik pada manusia mempengaruhi perubahan data CSI pada *Wi-Fi*?
3. Bagaimana cara memproses dan mengolah data CSI agar bisa digunakan sebagai alat untuk mengidentifikasi suatu individu?
4. Apakah penggunaan data CSI efektif sebagai alat untuk mengidentifikasi suatu individu?

1.3 TUJUAN PENELITIAN

Penelitian ini bertujuan untuk membuat sebuah sistem untuk mengidentifikasi individu dari data CSI berdasarkan fitur fisik individu yang dikumpulkan dari ESP32. data CSI tersebut akan diproses dan diolah dengan *machine learning* untuk mengklasifikasi setiap CSI data. Penelitian ini juga bertujuan untuk.

1. Menjelaskan bagaimana cara mengambil dan mendapatkan data CSI dari sebuah mikrokontroler ESP32.
2. Menjelaskan apakah perbedaan fisik pada manusia sangat berpengaruh terhadap perubahan data CSI pada *Wi-Fi*.
3. Menjelaskan bagaimana cara memproses dan mengolah data CSI agar bisa digunakan sebagai alat untuk mengidentifikasi suatu individu.
4. Membuktikan apakah penggunaan data CSI efektif sebagai alat untuk mengidentifikasi suatu individu.

1.4 MANFAAT PENELITIAN

Berdasarkan penelitian yang telah dilakukan, diharapkan dapat memberikan manfaat-manfaat sebagai berikut.

1. Menjadikan ESP32 sebagai alat alternatif untuk menerima data CSI dari sebuah router.
2. Mengusulkan alat autentikasi alternatif guna menyelesaikan masalah-masalah yang alat saat ini.
3. *Wi-Fi Sensing* menggunakan ESP32 sebagai alat alternatif pengidentifikasi manusia.
4. Membuka peluang untuk meneliti lebih dalam mengenai *Wi-Fi* CSI.

BAB II
TINJAUAN PUSTAKA

2.1 Wi-Fi Sensing dalam Human Identification

Wi-Fi memiliki kemampuan untuk mengukur perubahan pada lingkungan sinyal berpropagasi, yang memungkinkan sinyal *Wi-Fi* untuk mendeteksi pergerakan manusia, aktivitas manusia, bahkan identitas manusia dalam sistem yang *device-free* (Lv et al., 2019). Teknologi *Wi-Fi Sensing* memiliki kelebihan seperti sistem *sensing* yang pasif, tidak membutuhkan *LoS*, dan juga dapat beroperasi tanpa kebutuhan cahaya. *Wi-Fi sensing* dapat digambarkan sebagai penggunaan parameter sinyal *Wi-Fi* untuk mendeteksi lingkungan dimana sinyal *Wi-Fi* berpropagasi, seperti dengan mengamati perubahan properti *channel* yang dikenal dengan RSSI (*Received Signal Strength Indicator*) dan juga CSI (*Channel State Information*) (Suroso, Adiyatma, & Cherntanomwong, 2023). *Wi-Fi sensing* memiliki beberapa kelebihan seperti dapat diimplementasikan dengan infrastruktur *Wi-Fi* yang sudah ada, *fine grained localization* (jika menggunakan CSI), dapat menentukan multi-target, tidak intrusif dan juga dapat beradaptasi dengan perubahan lingkungan.

2.2 RSSI (Received Signal Strength Indicator)

Sistem Identifikasi Manusia berbasis *Wi-Fi Sensing* terdahulu memanfaatkan *Received Signal Strength Indicator* (RSSI) dari lapisan *Media Access Control* (MAC) karena kemudahan dalam perolehan data. Dalam area *indoor*, sinyal *Wi-Fi* akan berpropagasi menuju *receiver* melalui berbagai jalur (*multipath*). Setiap jalur berkontribusi pada perubahan dan pergerakan sinyal yang berbeda. Dan RSSI menunjukkan hasil penjumlahan energi sinyal *Wi-Fi* dari lintasan *multipath* seperti lintasan *Line of Sight* (LOS) antara pemancar dan penerima, serta beberapa jalur pantulan yang disebabkan oleh dinding, *furniture*, dan orang-orang dalam satuan desibel (dB). Untuk mengukur nilai RSSI, dapat digunakan persamaan berikut (Kurniawati et al., 2023) :

$$RSSI = 10log_2(||V||^2) \quad (1)$$

Dimana V adalah nilai tegangan voltase *baseband* sinyal yang diukur pada *receiver* dalam waktu tertentu. Parameter ini memiliki sifat berfluktuasi tinggi disebabkan lingkungan dan waktu dari transmisi. RSSI dapat diterapkan pada *localization* dalam ruangan untuk teknik *device-free* dan jarak jauh. Karena implementasinya yang cepat dan mudah, RSSI unggul sebagai parameter untuk pengaplikasian *Wi-Fi Sensing* di dalam ruangan.

2.3 CSI (Channel State Information)

Mayoritas sistem komunikasi *wireless modern* (termasuk Wi-Fi) menggunakan teknik yang disebut dengan OFDM (*Orthogonal Frequency Division Multiplexing*) yang merupakan skema modulasi. Dalam teknik OFDM, saluran dibagi menjadi beberapa sub-saluran (*subcarrier*) yang lebih kecil, masing-masing saluran beroperasi pada frekuensi yang berbeda secara paralel. Teknik tersebut dapat menambah efisiensi transmisi data (Burimas, Horanont, Thapa, & Lamichhane, 2024). Masing-masing sinyal yang dikirim melalui *subcarrier* dapat digambarkan dengan persamaan (2) (Wang et al., 2021):

$$Y = H \times X + N \quad (2)$$

Y merupakan vektor sinyal pada *receiver*, X mewakili sinyal vektor saat dikirim, N adalah *noise* yang memengaruhi sinyal selama transmisi, dan H merupakan *matrix* dari CSI (*Channel State Information*) seperti berikut:

$$H = [H_1, H_2, ..., H_k]$$

H_k merupakan *value* dari CSI untuk *subcarrier* ke-k. Dan masing-masing *value* dapat didefinisikan dengan (3) (Wang et al., 2021):

$$H_k = |H_k| e^{j \sin(\angle H_k)} \quad (3)$$

$|H_k|$ mempresentasikan informasi amplitudo dan $\angle H_k$ mempresentasikan informasi fasa. CSI secara terus menerus mendeteksi respons frekuensi *subcarrier* OFDM dan menangkap berbagai perubahan lingkungan, seperti frequency selective fading, shadowing, *multipath*, *destructive*, dan gangguan *constructive*. Berdasarkan perubahan sinyal ini, dapat diidentifikasi identitas individu dengan mengklasifikasi pola gangguan yang terjadi kepada setiap pengguna.

2.4 Pengambilan Data CSI

Saat ini, *Wi-Fi* merupakan teknologi *wireless* yang dapat diakses dari berbagai *device*, namun mayoritas perangkat dengan fitur *Wi-Fi* tidak mampu mengekstrak data CSI, sehingga banyak solusi yang memerlukan modifikasi *firmware* dan membutuhkan perangkat tambahan. Seperti Linux 802.11n (Halperin, Hu, Sheth, & Wetherall, 2011) yang memiliki *Network Interface Card* (NIC) Intel 5300 namun hanya mengumpulkan hingga 30 *subcarrier* dan memerlukan modifikasi *firmware*. Alat yang lain, Atheros CSI (Xie, Li, & Li, 2015) bekerja dengan NIC Atheros 802.11 dan mendapatkan semua 56 *subcarrier* untuk bandwidth 20 MHz tanpa mengutak-atik *firmware*. Meskipun demikian, solusi berbasis NIC yang disebutkan di atas tidak mendukung operasi mandiri dan tetap tidak praktis untuk penggunaan skala besar (Atif, Muralidharan, Ko, & Yoo, 2020). Ekstraktor CSI Nexmon menggunakan *chipset* Broadcom di *smartphone* Android Nexmon 5 untuk mendapatkan data CSI dari semua 56 *subcarrier* di *bandwidth* 20 MHz sebagai solusi untuk dapat dijalankan secara mandiri. Namun, solusi berbasis Nexmon memerlukan modifikasi dan dapat mencabut garansi perangkat.

Dikarenakan harganya yang jauh lebih murah, dan juga mudah dalam mendapatkan data CSI, ESP32 dijadikan sebagai salah satu solusi untuk melampaui masalah ini. Dengan menggunakan *framework* ESP-IDF dengan bahasa pemrograman C, dapat dibuat sebuah *codebase* yang digunakan untuk mengekstrak data CSI. Dengan ini, ESP32 menawarkan solusi alat *Wi-Fi* Sensing yang dapat digunakan dalam skala yang besar.

2.5 Preprocessing

Data CSI yang didapat dari alat-alat pada bagian 2.4 adalah data yang masih mentah, data tersebut harus masih diolah sedemikian rupa agar bisa digunakan sebagai input *machine learning* untuk diklasifikasi, proses ini disebut dengan *preprocessing*. Proses ini pada umumnya dibagi menjadi tiga bagian, pemilihan sinyal (2.5.1), *noise reduction* (2.5.2) dan juga ekstraksi fitur (2.5.3).

2.5.1 Pemilihan Sinyal

Pemilihan tipe sinyal dapat berpengaruh terhadap akurasi dalam sistem autentikasi. CSI data merupakan matriks saluran yang terdiri dari informasi amplitudo dan fasa, pada umumnya para peneliti menggunakan amplitudo atau fasa atau keduanya agar diproses lebih lanjut. Seperti pada HumanFi (Ming, Feng, & Bu, 2019) yang menggunakan amplitudo dan fasa untuk mendeteksi variasi sinyal, dan juga mendapatkan akurasi identifikasi yang tinggi.

2.5.2 Noise Reduction

Data mentah dari CSI memiliki gangguan (*noise*) yang cukup signifikan, dikarenakan perangkat *Wi-Fi* mendeteksi gangguan sinyal dari lingkungan yang ramai dan pantulan dari berbagai objek. Oleh karena itu, gangguan tersebut harus atau setidaknya dikurangi untuk mendapatkan sinyal yang bersih untuk diproses lebih lanjut (Wang et al., 2021). Pada umumnya ada beberapa cara untuk mengurangi atau menghilangkan gangguan tersebut, seperti PCA (*Principal Component Analysis*) dimana teknik tersebut mereduksi dimensi yang kompleks dari sinyal tetapi tetap menjaga fitur-fitur sinyal yang signifikan, *butterworth filter* yang digunakan untuk menghilangkan gangguan frekuensi yang tinggi, DWT (*Discrete Wavelet Transform*) yang dibagi menjadi tiga bagian, yaitu *decomposition*, *thresholding* dan *reconstruction* (Wang et al., 2021).

2.5.3 Ekstraksi Fitur

Ekstraksi sinyal adalah proses untuk mengekstrak target fitur dari pengukuran CSI mentah atau yang telah diproses sebelumnya. Proses ini dilakukan untuk mengambil fitur penting yang berisi informasi ketika sinyal memantulkan karakteristik unik dari manusia dan menyingkirkan bagian-bagian sinyal yang tidak terlalu dibutuhkan. Proses ini dapat meningkatkan efisiensi dari sistem autentikasi. Fitur-fitur berguna yang dimaksud seperti fitur dari domain waktu dan domain frekuensi.

Fitur dari domain frekuensi dapat diambil setelah mengubah series dari CSI menjadi domain frekuensi dengan menggunakan FFT (*Fast Fourier Transform*), fitur-fitur dari domain frekuensi seperti *median*, *root mean square*, dan kurtosis. Fitur dari domain waktu diambil langsung dari bentuk sinyal yang asli.

2.6 Machine Learning dalam Wi-Fi Sensing

Setelah data CSI melewati proses *preprocessing*, akan digunakan fitur yang telah di ekstrak dari data CSI sebagai fitur input dalam *machine learning*. Klasifikasi ini dilakukan untuk mengidentifikasi identitas manusia berdasarkan perbedaan fitur-fitur sinyal. Algoritma yang umum digunakan seperti *Support Vector Machine* (SVM), akan digunakan dan berhasil memberikan akurasi yang tinggi. SVM adalah klasifikasi algoritma ampuh yang telah digunakan dalam *Wi-Fi Sensing* dengan data CSI. SVM berupaya untuk menemukan *hyperplane* yang paling optimal untuk memisahkan objek dari kelas yang berbeda dalam ruang fitur. SVM dilatih untuk mengklasifikasikan objek berdasarkan pola dan variasi dalam data CSI yang sesuai masing-masing individu. Dengan menentukan fungsi kernel yang sesuai, SVM dengan efektif dapat memahami hubungan yang kompleks pada CSI dengan masing-masing individu (Ali et al., 2023).

BAB III
BAHAN DAN METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada bulan November 2024 sampai dengan bulan Februari 2025, dan dilaksanakan di laboratorium fisika SMA Unggulan Rushd.

3.2 Alat dan Bahan

Tabel 3.1, Alat dan Bahan

Tools	*Qty*	Fungsi
ESP32	1	ESP32 berfungsi sebagai penerima sinyal *Wi-Fi* dan mengirimkan data tersebut ke laptop
Laptop	1	Difungsikan untuk mengolah data CSI
Wi-Fi *Router*	1	Berfungsi untuk memancarkan sinyal *Wi-Fi* yang akan diterima oleh ESP32
VSCode	-	*Platform* untuk memprogram ESP32 untuk membaca sinyal CSI; *Platform* pembuatan model *Machine Learning*
SciKit	-	*Framework* yang digunakan untuk membantu dalam pembuatan model *machine learning* SVM
ESP-IDF	-	*Framework* yang dimanfaatkan untuk membantu dalam pembuatan *codebase* pengambilan data CSI dari ESP32

3.3 Rancangan dan Prosedur Penelitian

Model penelitian yang diterapkan merupakan model pengembangan ADDIE yang meliputi tahapan *Analysis, Design, Development, Implementation,* dan *Evaluation* (Thim-Mabrey, 2006). Dalam tahap analisis, akan diungkapkan masalah masalah yang disebabkan oleh perangkat *human identification* pada saat ini. Dalam tahap *design*, dijelaskan cara kerja dari alat yang akan dikembangkan, pada tahap *development, akan* dijelaskan secara rinci tahapan yang mulai dari menyusun alat hingga bagaimana data diproses, pada tahap *implementation* akan dijelaskan bagaimana eksperimen dilakukan dengan alat yang telah dirancang, dan terakhir pada tahap *evaluation*, akan dipaparkan bagaimana sistem ini diuji untuk membuktikan apakah data yang diperoleh akurat dengan kejadian nyata.

3.3.1 Analysis

Berbagai alat sistem keamanan telah diaplikasikan untuk menjaga dan mengurangi angka kriminal saat ini, namun sistem-sistem keamanan yang ada saat ini memiliki beberapa kelemahan yang dapat merugikan berbagai pihak.

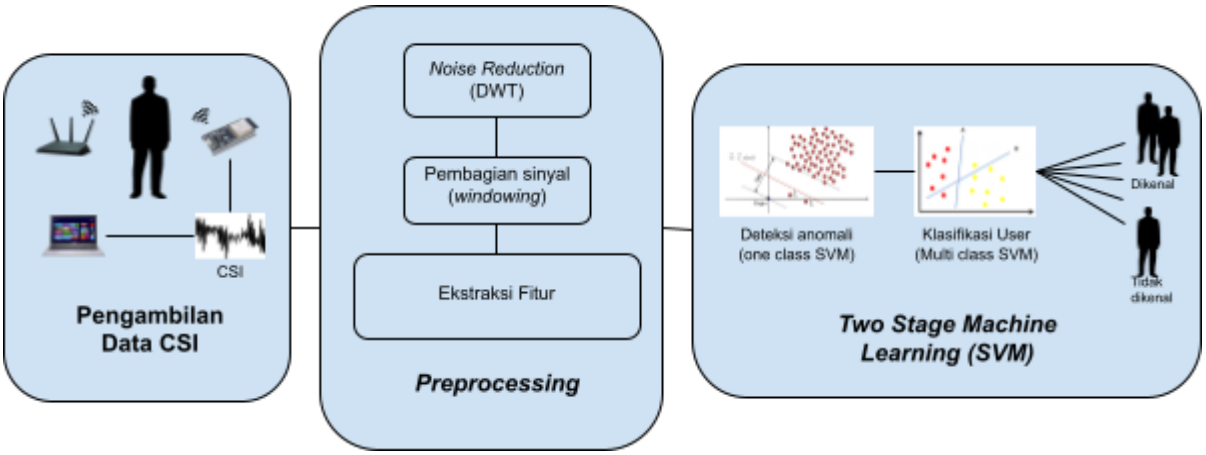
Tabel 3.2
Permasalahan sistem keamanan konvensional

No.	Alat	Permasalahan
1.	CCTV	a. Kekhawatiran privasi b. Visibilitas c. Kondisi cahaya d. Penggunaan listrik yang tinggi
2.	NFC	a. Penggunaan mengharuskan aksi dari pengguna b. Hilangnya alat pengenal
3.	GPS	a. Tidak dapat digunakan dalam area *Indoor*

Dari tabel yang berisi permasalahan yang ditemukan dari alat-alat keamanan saat ini (Anderez et al., 2021), dapat dilihat berbagai keterbatasan. Oleh karena itu sistem keamanan yang inovatif dapat dirancang untuk mengintegrasikan berbagai teknologi, memanfaatkan algoritma pembelajaran mesin dan juga alat yang *low cost* untuk meningkatkan efektivitas, penjagaan privasi, serta mengurangi ketergantungan pada pengguna.

3.3.2 Design

Dalam penelitian ini dikembangkan sistem identifikasi manusia dengan menggunakan sinyal *Wi-Fi*. Sebuah *transmitter* yang berupa router *Wi-Fi* dan ESP32 sebagai *receiver* yang diletakkan secara berseberangan, agar sinyal dapat berpropagasi di antara dua perangkat. Area di antara *transmitter* dan *receiver* akan menjadi area utama sensor dimana sinyal akan terpantul dan terganggu oleh karakteristik fisik manusia. ESP32 akan dihubungkan dengan komputer sehingga data CSI yang merupakan informasi status dari sinyal *Wi-Fi* akan diklasifikasi berdasarkan data individu yang terlatih menggunakan *machine learning*, seperti yang dipaparkan pada Gambar 3.1.



Gambar 3.1
Desain Sistem Human Identification.

3.3.3 Development

Data CSI yang sudah diambil oleh ESP32 akan dikirim melewati kabel USB (*Universal Serial Bus*) menuju laptop untuk diproses lebih lanjut, namun data ini baru akan diproses ketika seluruh data yang dibutuhkan sudah diambil. CSI akan diekstrak dari *receiver* ESP32 dengan memanfaatkan

framework ESP-IDF dan disimpan dalam format CSV. *Timestamp* juga diberikan pada setiap *frame* CSI. Data CSI kemudian diproses yang akan dijelaskan pada bagian 3.4.

3.3.4 Implementation

Dalam upaya mengukur efektivitas sistem yang dikembangkan, dilakukan eksperimen dengan melibatkan 10 orang relawan. Sebuah Tx (*transmitter*) yang berupa TP-Link dan ESP32 sebagai Rx (*receiver*) diletakkan berseberangan, sepuluh orang tersebut akan berdiri di antara Tx dan Rx secara bergantian untuk mengumpulkan data CSI yang dibutuhkan. Data akan melewati fase *preprocessing* dan akan dibagi menjadi 2 dataset seperti yang dijelaskan diatas. Setelah dataset terbentuk model akan dilatih berdasarkan dataset tersebut.

3.3.5 Evaluation

Untuk mengevaluasi sistem yang dibuat, efektifitas dan keakuratan terhadap model yang telah dilatih menggunakan dataset akan diuji. Pengujian dilakukan dengan individu lain yang tidak termasuk dalam dataset sebagai *intruder* untuk melihat apakah model dapat mengklasifikasi dan membedakan mereka dari individu yang telah dikenal. Model juga akan diuji dengan mengklasifikasi masing-masing dari sepuluh orang yang digunakan dalam pelatihan model.

3.4 Pengolahan dan Analisis Data

Setelah data CSI dikumpulkan dalam pengembangan sistem ini, proses *windowing* dilakukan untuk mengorganisir data menjadi bagian-bagian yang dapat dianalisis. Setiap *window* terdiri dari 30 *frame* CSI berurutan, mewakili jarak waktu tertentu dari sinyal yang diterima.

Pada tahap *preprocessing*, dilakukan ekstraksi nilai amplitudo dari setiap nilai kompleks *subcarrier* untuk dijadikan salah satu fitur input bagi model *machine learning*. Untuk mengurangi *noise* pada data CSI sekaligus mempertahankan informasi penting, diterapkan *Discrete Wavelet Transform* (DWT) pada nilai amplitudo. Dalam menentukan input fitur, digunakan seluruh nilai amplitudo yang sudah diterapkan metode DWT.

Setelah fase *preprocessing* selesai dilalui, CSI akan dibagi menjadi 2 *dataset*. 70% data digunakan untuk pelatihan dan 30% data untuk pengujian. Untuk mengklasifikasi data berdasarkan setiap individu, diperlukan model *machine learning* untuk mengklasifikasi masing masing data CSI. Pendekatan *two-stage machine learning* digunakan dalam sistem ini untuk meningkatkan performa dan mendeteksi individu yang tidak dikenal dari pendekatan sistem ini. Pada *stage* pertama, *One Class* SVM digunakan untuk mendeteksi pola data yang tidak normal, *stage* ini bertujuan untuk mendeteksi apakah sebuah individu dikenal ataupun tidak dikenal. *Multi Class* SVM diaplikasikan dalam *stage* kedua dari model ini, algoritma ini akan mengklasifikasi data CSI berdasarkan setiap individu yang dikenal.

Untuk mengevaluasi hasil kinerja model sistem ini, penggunaan metrik seperti *precision*, *F1-score* dan *recall* digunakan berfungsi untuk menilai kinerja dari model. *Precision* mengukur proporsi prediksi positif yang benar, sedangkan *recall* mengukur proporsi prediksi positif yang berhasil diklasifikasi oleh model. *F1-score* juga akan digunakan agar memberikan gambaran yang lebih detail mengenai kinerja model.

BAB IV
HASIL DAN PEMBAHASAN

4.1 Penyiapan Eksperimen

Pengujian dilakukan di ruang laboratorium fisika SMA Unggulan Rushd. Penempatan perangkat yang digunakan, seperti router TP-Link (Tx) dan juga ESP32 (Rx) di laboratorium fisika dapat dilihat pada Gambar 4.1. Subjek berdiri di celah antara dua meja sebagai titik pengambilan data. *Dataset* akan diambil dari 10 subjek, yang terdiri dari 10 laki-laki. Untuk pengambilan data setiap subjek, ESP32 akan mengambil data CSI yang melewati tubuh subjek tersebut dengan kecepatan pengambilan sampel sebesar 100 Hz dalam waktu 30 detik.

Gambar 4.1, Penempatan Tx, Rx, dan juga tempat berdirinya subjek di Laboratorium Fisika.

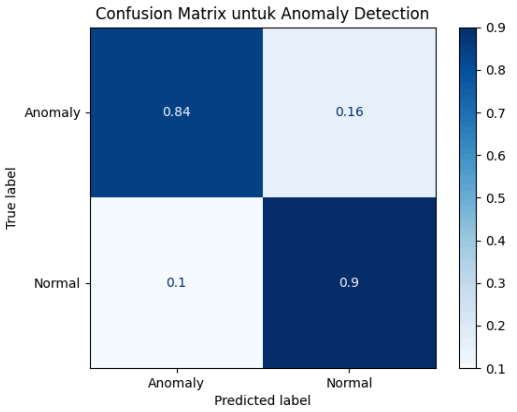
4.2 Evaluasi Eksperimen

Untuk mengukur performa sistem yang dikembangkan, dilakukan evaluasi terhadap kemampuan model dalam mendeteksi *intruder*, mengidentifikasi individu dalam *dataset*, serta menganalisis pengaruh jumlah individu terhadap akurasi klasifikasi. Dilakukan juga perbandingan performa dengan penelitian sebelumnya untuk melihat keunggulan metode yang digunakan dalam penelitian ini. Penilaian performa model dilakukan menggunakan beberapa metrik standar dalam klasifikasi, yaitu F1-*score*, *precision*, *recall*, dan *confusion matrix*.

4.2.1 Evaluasi Identifikasi Individu Tak Dikenal

Identifikasi individu yang tidak dikenal merupakan langkah awal yang krusial dalam sebuah sistem autentikasi keamanan. Sebelum sistem autentikasi dapat mengklasifikasi individu yang terdaftar, sistem tersebut harus terlebih dahulu mampu menentukan apakah individu tersebut terdaftar atau tidak. Setelah sistem berhasil mengidentifikasi bahwa individu tersebut adalah orang yang terdaftar, langkah selanjutnya adalah melakukan klasifikasi terhadap individu tersebut dibandingkan dengan individu terdaftar lainnya.

Untuk mencapai tujuan tersebut, digunakan model *Support Vector Machine* (SVM) dengan pendekatan *One Class* SVM. Model ini memanfaatkan kernel *Radial Basis Function* (RBF) untuk meningkatkan kemampuan klasifikasi dalam ruang fitur yang tidak linier. Parameter *nu* diatur pada nilai 0.1, yang berfungsi untuk mengontrol proporsi data yang diizinkan untuk dianggap sebagai *outlier*. Pemilihan parameter tersebut didasarkan pada proses *tuning* yang dilakukan untuk mendapatkan nilai parameter yang memberikan hasil performa terbaik. Dengan pengaturan ini, model diharapkan dapat secara efektif membedakan antara individu yang terdaftar dan yang tidak terdaftar, serta meningkatkan akurasi sistem autentikasi dalam mengidentifikasi individu yang sah. Subjek yang tak dikenal akan dipilih secara acak dari *dataset* yang tersedia.



Gambar 4.2, Hasil Confusion Matrix pada model.

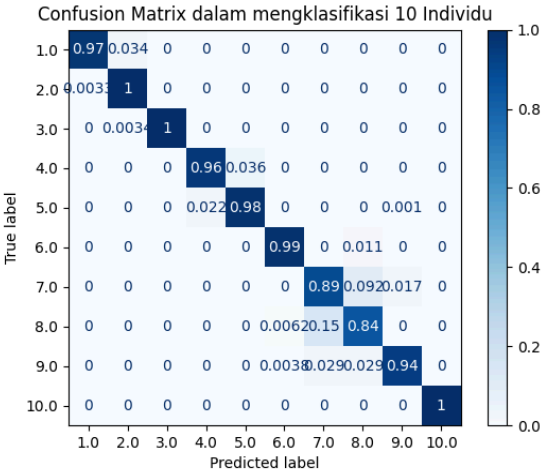
Terlihat pada Gambar 4.2, model One Class SVM yang dikembangkan ini dapat mengidentifikasi *intruder* dengan cukup baik, model mendapatkan 84% untuk TP (*True Positive*) dan 90% untuk TN (*True Negative*). Namun, masih terdapat beberapa kesalahan klasifikasi, seperti FP (*False Positive*) sebesar 10% dan *False Negative* sebesar 16%. Kesalahan klasifikasi diperkirakan dapat terjadi dikarenakan user yang terdaftar memiliki karakteristik yang sangat bervariasi, mengingat bahwa siswa-siswa SMA Unggulan Rushd berasal dari berbagai macam daerah, yang membuat model kesulitan membuat batasan yang jelas antara user dan intruder (Liu et al., n.d.). Namun, dapat disimpulkan bahwa hasil ini menunjukkan bahwa model *One Class* SVM dengan kernel RBF dan parameter *nu* 0.1 cukup efektif untuk menentukan anomali yang terdapat pada *dataset*.

4.2.2 Evaluasi Klasifikasi Individu dalam Dataset

Setelah data CSI melewati model yang dapat mengidentifikasi apakah data tersebut merupakan data yang terdaftar atau tidak, data CSI yang merupakan data terdaftar akan melewati langkah selanjutnya yaitu data akan diklasifikasi berdasarkan data training. Pada tahap ini, model SVM juga digunakan, namun dengan pendekatan *multi class* untuk menangani klasifikasi individu lebih dari 2. Untuk mengevaluasi kinerja SVM ini, metrik seperti F1 *score*, *recall*, *precision*, dan *confusion matrix* diterapkan untuk setiap subjek.

Tabel 4.1, Hasil metrik klasifikasi pada model.

Nomor Individu	Precision	Recall	f1-score
1	99%	97%	98%
2	98%	100%	99%
3	100%	100%	100%
4	97%	96%	97%
5	97%	98%	97%
6	99%	99%	99%
7	80%	89%	84%
8	87%	84%	86%
9	99%	94%	96%
10	100%	100%	100%



Gambar 4.3, Hasil confusion matrix pada dataset.

Tabel 4.2, Data berat dan tinggi badan volunteer

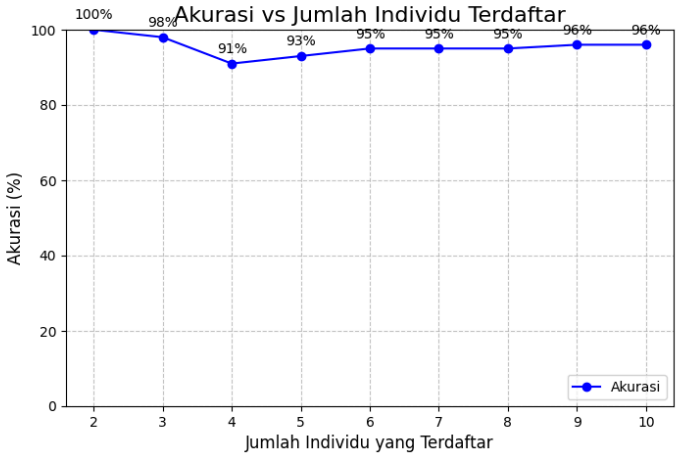
Nomor Volunteer	Berat Badan (kg)	Tinggi Badan (cm)
1	54	163
2	50	161
3	54	177
4	58	164
5	65	172
6	65	167
7	47	163
8	47	164
9	53	173
10	67	164

Hasil performa klasifikasi sepuluh relawan pada eksperimen di atas dapat dilihat dari metrik pada Tabel 4.1 dan juga *confusion matrix* pada Gambar 4.3. Dapat dilihat bahwa hasil akurasi dari klasifikasi data testing mendapatkan akurasi yang bervariasi dari 89% sampai 100% pada confusion matrix. Relawan 2, 3 dan 10 memiliki akurasi tertinggi sebesar 100%. Hal ini dapat terjadi dikarenakan mereka merupakan relawan dengan ciri-ciri fisik yang paling berbeda. Relawan 2 merupakan orang dengan tinggi badan terkecil, Relawan 3 memiliki tinggi badan tertinggi dan Relawan 10 memiliki berat badan terbesar. Dapat dilihat juga bahwa model memiliki kesulitan dalam membedakan Relawan 7 dan 8. Model memperoleh akurasi sebesar 84% dalam mengidentifikasi individu nomor 8. Namun, model juga mengalami kesalahan klasifikasi sebesar 15%, dimana individu nomor 8 secara keliru diidentifikasi sebagai nomor 7. Hasil ini didapat dikarenakan relawan 7 dan 8 memiliki ciri fisik yang hampir mirip. Dapat dilihat pada tabel 4.2, bahwa tinggi badan mereka tak berbeda jauh, serta berat badan mereka juga tak berbeda jauh. Hal ini menunjukkan bahwa tinggi

badan dan juga berat badan yang merupakan ciri fisik memengaruhi transmisi sinyal dengan pola yang berbeda.

4.2.3 Pengaruh Jumlah Individu Terdaftar Terhadap Performa Model

Dalam bagian ini, dibahas bagaimana jumlah individu yang terdaftar dalam *dataset* pada sistem identifikasi memengaruhi performa model klasifikasi. Model akan dilatih menggunakan data CSI dengan variasi jumlah individu, mulai dari 2 individu hingga 10 individu. Pemilihan individu dilakukan secara berurutan, dimulai dari relawan ke-10 dan secara bertahap mencakup relawan dengan nomor lebih kecil hingga relawan ke-1.

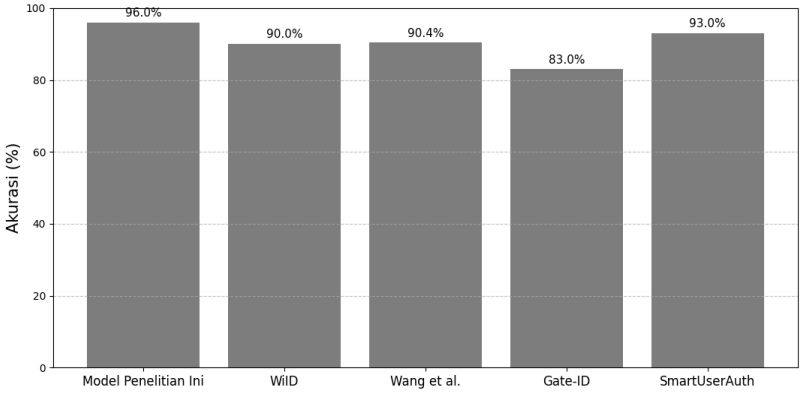


Gambar 4.4, Diagram garis mengenai angka akurasi model dengan berbagai jumlah individu terdaftar.

Berdasarkan Gambar 4.4, ketika jumlah individu pada dataset berjumlah 2 sampai 3 orang, model mendapatkan akurasi sebesar 100% dan 98%. Namun, saat model melakukan klasifikasi terhadap 4 individu, angka akurasi mengalami penurunan yang cukup signifikan. Hal ini disebabkan oleh individu ke-4 yang merupakan relawan ke-7, dan juga individu ke-3 yang merupakan relawan ke-8 memiliki karakteristik fisik yang hampir sama, sehingga model mengalami kesulitan dalam memprediksi individu. Seiring bertambahnya jumlah individu pada *dataset*, angka akurasi kembali meningkat secara perlahan hingga mencapai angka akurasi 96% pada 10 individu.

4.2.3 Perbandingan dengan Penelitian Sebelumnya

Untuk membuktikan performa yang baik dari sistem autentikasi dengan menggunakan mikrokontroler ESP32 ini, hasil performa dari penelitian ini akan dibandingkan dengan hasil performa penelitian yang lain. Beberapa penelitian serupa sebelumnya yang menggunakan 10 subjek sebagai dataset dipilih, seperti WiID (Zheng et al., 2017), penelitian yang dilakukan oleh Wang (Wang et al., 2018), Gate-ID (Zhang et al., 2021), dan SmartUserAuth (Shi et al., 2017) akan dibandingkan dengan performa akurasi model penelitian ini.



Gambar 4.5

Diagram batang yang mewakili hasil performa akurasi dengan berbagai penelitian terdahulu.

WiID (Zheng et al., 2017) menggunakan data CSI dari perangkat *Wi-Fi* komersial dari aktivitas umum manusia seperti gerakan dan interaksi untuk mengidentifikasi identitas individu, sedangkan Gate-ID (Zhang et al., 2021) mengidentifikasi individu berdasarkan arah berjalan masing-masing subjek. Ada juga penelitian yang dilakukan oleh Wang (Wang et al., 2018), di mana mereka menawarkan sistem identifikasi manusia melalui pengaruh unik manusia terhadap sinyal *Wi-Fi*, di sisi lain, SmartUserAuth (Shi et al., 2017) memanfaatkan sinyal *Wi-Fi* untuk menangkap karakteristik fisik dan perilaku manusia, baik saat bergerak maupun diam. Sistem ini menggunakan skema autentikasi berbasis *deep learning*.

Dilihat dari Gambar 4.5, ketika sistem yang diusulkan dalam penelitian ini dibandingkan dengan penelitian-penelitian yang disebutkan sebelumnya, dengan respek, sistem ini memberikan performa yang lebih baik. Terdapat beberapa alasan mengapa hal ini dapat terjadi, seperti tingginya variasi perbedaan fitur fisik siswa-siswi SMA Unggulan Rushd. Pada penelitian Gate-ID, tidak terdapat filtering untuk mengurangi *noise* pada data CSI, penelitian tersebut juga menggunakan perbedaan sinyal yang disebabkan oleh pola jalan untuk mengklasifikasi individu, yang bisa berubah tergantung kondisi fisik, kecepatan berjalan, hambatan lingkungan, dan juga ketika mengalami gangguan saat berjalan, Gate-ID juga menggunakan *Attention-Based Deep Learning* Model dengan *ResNet* dan *Bi-LSTM*, yang berisiko *overfitting* terutama karena dataset mereka terbatas. Pada penelitian WiID, digunakan metode *filtering* dengan menerapkan PCA, yang dapat menyebabkan hilangnya informasi-informasi penting pada suatu data. WiID juga menggunakan pengaruh sinyal *Wi-Fi* terhadap aktivitas manusia sebagai instrumen ukur sensor, sama seperti SmartUserAuth. Penelitian-penelitian yang disebutkan ini juga tidak menggunakan ESP32 sebagai *receiver*, yang menyebabkan sistem mereka memiliki kelemahan karena penggunaan ESP32 dalam penelitian ini memungkinkan sistem untuk lebih portabel, hemat daya, dan lebih mudah diimplementasikan dalam berbagai lingkungan.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dalam penelitian ini, telah dikembangkan sistem identifikasi manusia dengan memanfaatkan sinyal *Wi-Fi* yang diperoleh dari mikrokontroler ESP32. Model *machine learning* SVM digunakan untuk mengolah data CSI yang merepresentasikan status sinyal *Wi-Fi*. Data ini terlebih dahulu melalui tahap *preprocessing* untuk mengurangi *noise-noise* yang mengganggu menggunakan algoritma DWT, kemudian data akan diklasifikasi untuk menentukan apakah data berasal dari data individu yang terdaftar menggunakan pendekatan *One Class* SVM. Selanjutnya, klasifikasi dilakukan berdasarkan pengaruh karakteristik fisik individu terhadap sinyal dengan pendekatan *Multi Class* SVM, sehingga dapat diidentifikasi siapa yang memengaruhi sinyal tersebut. Dalam sebuah eksperimen yang telah dilakukan, sistem ini dapat mengidentifikasi individu dengan akurasi rata-rata 100% hingga 96% dalam kelompok 2 hingga 10 orang. Dalam menentukan apakah data berasal dari data individu yang terdaftar, model mendapatkan akurasi rata-rata sebesar 90% dengan jumlah 10 orang terdaftar. Dibandingkan dengan penelitian-penelitian terkait sebelumnya, sistem ini memiliki beberapa keunggulan, seperti hasil performa model yang lebih baik, efisiensi yang lebih tinggi, biaya yang lebih rendah, kemudahan dalam implementasi, dan merupakan sistem yang portabel dikarenakan penggunaan mikrokontroler ESP32.

5.2 Saran

Meskipun sistem yang telah dikembangkan menunjukkan hasil yang baik sebagai sistem autentikasi manusia yang memanfaatkan sinyal *Wi-Fi* dan menggunakan mikrokontroler ESP32, terdapat beberapa aspek yang dapat ditingkatkan untuk penelitian-penelitian selanjutnya, seperti melakukan eksplorasi model yang lebih optimal dalam mendeteksi *outlier* dalam sebuah dataset sehingga akan mendapatkan performa yang lebih baik, mencari *input* fitur yang lebih tepat untuk mencapai akurasi yang lebih tinggi, pengujian eksperimental dimana Rx dan Tx akan ditaruh pada ruangan yang berbeda sehingga sinyal harus melewati dinding, juga pemanfaatan data CSI untuk keperluan lain diluar sistem autentikasi manusia.

UCAPAN TERIMAKASIH

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, serta hidayah-Nya sehingga penulis dapat menyelesaikan penelitian dengan judul "Pengembangan Sistem Autentikasi Berbasis Wi-Fi Sensing Menggunakan Data CSI dan Algoritma Machine Learning dengan Mikrokontroler ESP32" ini dengan baik dan tepat waktu. Dalam penyusunan makalah ini para penulis telah berusaha semaksimal mungkin dan tentunya dengan bantuan berbagai pihak. Untuk itu, penulis tak lupa untuk mengucapkan terima kasih kepada.

1. Orang Tua yang selalu mendukung dalam pengerjaan makalah penelitian ini.

2. Agung Prastyo, S.Pd. sebagai pembimbing dan pembina klub penelitian SMA Unggulan Rushd.

3. Hanun Fithriyah, S.Pd. selaku pembimbing internal makalah penelitian ini.

4. Zahra Dyah Meilani, S.Kom. selaku pembimbing eksternal makalah penelitian ini.

5. Relawan-relawan yang membantu dalam pengerjaan makalah penelitian ini.

6. Pihak-pihak lain yang membantu dalam pengerjaan makalah ini.

Dalam penulisan laporan penelitian ini masih banyak kekurangan dan kesalahan, karena itu segala kritik dan saran yang membangun akan menyempurnakan penulisan laporan penelitian ini serta bermanfaat bagi penulis dan para pembaca.

DAFTAR PUSTAKA

Ali, M., Hendriks, P., Popping, N., Levi, S., & Naveed, A. (2023). *A comparison of machine learning algorithms for Wi-Fi sensing using CSI data*. *Electronics (Switzerland)*, 12(18), 3935. https://doi.org/10.3390/electronics12183935

Anderez, D.O., Kanjo, E., Amnwar, A., Johnson, S., & Lucy, D. (2021). *The rise of technology in crime prevention: Opportunities, challenges and practitioners perspectives*. *arXiv*. http://arxiv.org/abs/2102.04204

Atif, M., Muralidharan, S., Ko, H., & Yoo, B. (2020). Wi-ESP—A tool for CSI-based device-free Wi-Fi sensing (DFWS). Finite Element Analysis and Design, 7(5), 644–656.

Burimas, R., Horanont, T., Thapa, A., & Lamichhane, B. R. (2024). Monitoring the sleep respiratory rate with low-cost Wi-Fi microcontroller in a controlled environment. Applied Sciences, 14(15), 6458. https://doi.org/10.3390/app14156458

Chen, J.J., Jiang, Z.X., Chen, Y.L., Wu, W.T., & Liang, J.M. (2015). Design and realization of an NFC-driven smart home system to support intruder detection and social network integration. Journal of Electronic Science and Technology, 13(2), 163–168.

Griffiths, G., Johnson, S. D., & Chetty, K. (2017). UK-based terrorists' antecedent behavior: A spatial and temporal analysis. Applied Geography, 86, 274–282.

Halperin, D., Hu, W., Sheth, A., & Wetherall, D. (2011). Tool release: Gathering 802.11n traces with channel state information. ACM SIGCOMM Computer Communication Review, 41(1), 53–53.

Kurniawati, N., Nurjihan, SF, Dwitio, D., & Pranoto, A. (2023). Design and construction of a Wi-Fi-based human activity detection system.

Lawrence, D.S., La Vigne, N.G., Goff, M., & Thompson, P.S. (2018). Lessons learned implementing gunshot detection technology: Results of a process evaluation in three major cities. Justice Evaluation Journal, 1(2), 109–129.

Lv, J., Man, D., Yang, W., Gong, L., Du, X., & Yu, M. (2019). Robust device-free intrusion detection using physical layer information of Wi-Fi signals. Applied Sciences, 9(1), 17. https://doi.org/10.3390/app901017

Ming, X., Feng, H., & Bu, Q. (2019). HumanFi: Wi-Fi-based human identification using recurrent neural network. In IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (pp. 640–647). IEEE.

Mosharaf, M., Kwak, J. B., & Choi, W. (2024). Wi-Fi-based human identification with machine learning: A comprehensive survey. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/s24196413

Parameswaran, AT, Husain, MI, & Upadhyaya, S. (n.d.). Is RSSI a reliable parameter in sensor localization algorithms: An experimental study.

Piza, E.L., Welsh, B.C., Farrington, D.P., & Thomas, A.L. (2019). CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis. Criminology & Public Policy, 18(1), 135–159.

Suroso, DJ, Adiyatma, FYM, & Cherntanomwong, P. (2023). Wi-Fi sensing for indoor localization via channel state information: A survey. ELKHA: Journal of Electrical Engineering, 15(2), 152–159.

Thim-Mabrey, C. (2006). Sprachwandel in Übersetzung Bearbeitungen zwischen 1846 und 1999. Neuphilologische Mitteilungen, 107(3), 361–373.

Wang, Z., et al. (2021). A survey of user authentication based on channel state information. Hindawi Limited. https://doi.org/10.1155/2021/6636665

Xie, Y., Li, Z., & Li, M. (2015). Precise power delay profiling with commodity Wi-Fi. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (pp. 53–64). ACM.

Shi, C., Liu, J., Liu, H., & Chen, Y. (2017). Smart User Authentication through Actuation of Daily Activities Leveraging Wi-Fi-enabled IoT. Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing. https://doi.org/10.1145/3084041.3084061

Wang, J., Zhao, Y., Fan, X., Gao, Q., Ma, X., & Wang, H. (2018). Device-Free Identification Using Intrinsic CSI Features. IEEE Transactions on Vehicular Technology, 67(9), 8571–8581. https://doi.org/10.1109/TVT.2018.2853185

Zhang, J., Wei, B., Wu, F., Dong, L., Hu, W., Kanhere, S. S., Luo, C., Yu, S., & Cheng, J. (2021). Gate-ID: Wi-Fi-Based Human Identification Irrespective of Walking Directions in Smart Home. IEEE Internet of Things Journal, 8(9), 7610–7624. https://doi.org/10.1109/JIOT.2020.3040782

Zheng, R., Zhao, Y., & Chen, B. (2017). Device-Free and Robust User Identification in Smart Environment Using Wi-Fi Signal. 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC), 1039–1046. https://doi.org/10.1109/ISPA/IUCC.2017.00158

Liu, J., Yuan, J., & Li, C.-N. (n.d.). *One-class SVM with 0-1 loss*. https://ssrn.com/abstract=5050012

PERNYATAAN PENELITI

Yang bertanda tangan di bawah ini:
Nama : Muhammad Fathan Haroki
Tempat/Tanggal Lahir : Gresik, 06 Februari 2009
NIS : 0095429502
Asal Sekolah : SMA Unggulan Rushd

dengan ini menyatakan sejujurnya bahwa proposal penelitian saya dengan judul: Pengembangan Sistem Autentikasi Berbasis WiFi Sensing Menggunakan Data CSI dan Algoritma Pembelajaran Mesin dengan Mikrokontroler ESP32

bersifat orisinal/bukan hasil tindak plagiarisme/belum pernah dikompetisikan dan/atau tidak sedang diikutkan pada lomba penelitian sejenis/belum pernah mendapatkan penghargaan di tingkat Nasional/Internasional.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, saya bersedia menerima konsekuensi sesuai aturan ISPO.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Dibuat di
Sragen Pada
Tanggal 30 November 2024

Mengetahui,

Hanun Fithriyah, S.Pd
NIP: -

Yang membuat pernyataan

Muhammad Fathan Haroki
NISN: 0095429502

ISPO RESEARCH REPORT

Development of Wi-Fi Sensing-Based Authentication System Using CSI Data and Machine Learning Algorithm with ESP32 Microcontroller

by:

Muhammad Fathan Haroki
Oruzgan Abimanyu Adi

Technology Engineering and Robotics

Unggulan Rushd High School

Sragen, Central Java

2025

ABSTRACT

The high crime rate makes the issue of crime and criminality a major concern. Security systems that are effective, efficient, and adaptive to environmental conditions are needed to overcome these problems. Some systems that have been proposed, such as CCTV (Circuit-Closed Television), NFC (Near Field Communication), and GPS (Global Positioning System) based systems, still do not maximally overcome the existing problems. To overcome these limitations, a Wi-Fi-sensing-based authentication system using an ESP32 microcontroller was developed. CSI (Channel State Information) data which is the status information of a Wi-Fi signal is extracted using ESP32 with ESP-IDF framework. A machine learning SVM model with a two-stage approach was developed to classify the CSI data. In the first stage, the model identifies whether the detected individual is an intruder or a registered individual. If the individual is registered, the second stage is performed to classify based on the characteristics of the reflected signal. Experimental results showed that the system achieved 100% to 96% accuracy in identifying 2 to 10 registered individuals. In addition, the system was also able to identify unknown individuals with an accuracy of 90% on a dataset consisting of 9 registered individuals. The system has several advantages, such as better model performance, high efficiency, low cost, and ease of implementation due to the use of the ESP32 microcontroller as a tool with high portability.

Keywords: ESP32, Wi-Fi Sensing, User Authentication, Machine Learning, Channel State Information.

CHAPTER I
INTRODUCTION

1.1 BACKGROUND

In recent years, technology has developed rapidly to overcome crime problems (Anderez et al., 2021). The many sensor innovations in the security field to detect unrecognized individuals, such as CCTV (Circuit-Closed Television) (Piza, Welsh, Farrington, & Thomas, 2019), NFC (Near Field Communication) (Chen et al., 2015), Audio (Lawrence, La Vigne, Goff, & Thompson, 2018), and GPS (Global Positioning System) (Griffiths, Johnson, & Chetty, 2017) make it easier for many people to overcome and reduce criminal activities. However, there are still some shortcomings that need to be considered.

In the case of camera-based sensors, there are some disadvantages such as dependence on light and also the need for a clear line of sight. Audio sensors also raise privacy concerns that disrupt user comfort. Wearable sensors such as the Smart Ring in detecting the recognized person sometimes cause inconvenience to the user, and also often require wearing in a specific way for accurate operation. Biometric sensors such as fingerprints are also quite vulnerable to hacking attacks and can be exploited using fake data.

Wireless technology has become one of the solutions in overcoming various limitations and problems. One example is the use of RF signals as sensors that can be used in the security field (Anderez et al., 2021). Wi-Fi sensing has recently gained popularity with the increasing use of Wi-Fi in recent years. Researchers have started utilizing Wi-Fi signals as an alternative sensor to detect and understand the movement, position and activity of objects or individuals. After a Wi-Fi signal propagates from a transmitter, it undergoes multiple reflections, absorption, and scatter on objects or people before reaching the receiver (Suroso, Adiyatma, & Cherntanomwong, 2023). The physical differences in humans affect the way the signal propagates, which allow us to detect individuals based on the classification of the signal with the help of Machine Learning algorithms (Wang et al., 2021). Wi-Fi signals will provide two types of information regarding signal quality or signal condition, which are: CSI (Channel State Information) and RSSI (Received Signal Strength Indicator).

RSSI is a measurement of how strong a signal is received. RSSI is easily obtained on all Wi-Fi devices without any additional hardware. However, the stability of RSSI is not ensured, even in indoor areas (Parameswaran, Husain, & Upadhyaya, n.d.). Therefore, RSSI is not very effective in being used as a parameter to recognize an individual. Meanwhile, CSI contains Wi-Fi signal status information that explains how the signal propagates. Wi-Fi signals pass through and include processes such as scattering, attenuation, diffraction, fading, and reflection. These processes capture complex information about the transmission of a signal. The information provided by CSI data is important for optimizing Wi-Fi signals and understanding the environmental characteristics of the place where signal transmission occurs (Mosharaf, Kwak, & Choi, 2024). CSI is a matrix of complex values containing the amplitude and phase shift of multipath paths in Wi-Fi. Direct access to Channel State Information (CSI) data from Wi-Fi devices is limited to specific hardware and software, making it impractical for large-scale use where it would be challenging for people to extract CSI data.

To overcome those limitations, an alternative solution is to develop a human authentication system that can recognize both recognized and unrecognized individuals with Wi-Fi sensing technology that utilizes CSI data. The use of ESP32 microcontroller as a receiver is applied due to several advantages such as low-cost, easy to obtain CSI data with ESP-IDF, less complex system, lightweight device and easy to operate. To process the signal more accurately and with less noise, DWT (Discrete Wavelet Transform) is applied to the CSI data before it is classified by machine

learning. Support Vector Machine (SVM) algorithm, which is a widely used model, will be used in this research to get effective classification results.

1.2 PROBLEMS FORMULATION

This research will answer the following problem questions.

1. How to retrieve and obtain CSI data from an ESP32 microcontroller?
2. How does human physique affect the change of CSI signal on Wi-Fi?
3. How to process and process CSI data so that it can be used as a tool to identify an individual?
4. Is the use of CSI data effective as a tool to identify an individual?

1.3 RESEARCH PURPOSES

This research aims to create a system to identify individuals using CSI data based on individual physical features collected from ESP32. The CSI data will be processed and processed with machine learning to classify each CSI data. This research also aims to.

1. Explain how to retrieve and obtain CSI data from an ESP32 microcontroller.
2. Explaining whether physical differences in humans greatly affect changes in the CSI data on Wi-Fi.
3. Explaining how to process CSI data so that it can be used as a tool to identify an individual.
4. Proving whether the use of CSI data is effective as a tool to identify an individual.

1.4 RESEARCH BENEFITS

Based on the research that will be conducted, it can provide the following benefits.

1. Making ESP32 as an alternative tool to receive CSI data from a router.
2. Proposing an alternative authentication tool to solve the problems of the current tools.
3. Wi-Fi Sensing using ESP32 as an alternative human identifier.
4. Opens up opportunities to research more deeply about Wi-Fi CSI.

CHAPTER II
LITERATURE REVIEW

2.1 Wi-Fi Sensing in Human Identification

Wi-Fi has the ability to measure changes in the propagating signal environment, which allows Wi-Fi signals to detect human movement, human activity, and even human identity in a device-free system (Lv et al., 2019). Wi-Fi Sensing technology has advantages such as a passive sensing system, does not require LoS (Line of Sight), and can also operate without the need for light. Wi-Fi sensing can be described as the use of Wi-Fi signal parameters to sense the environment where Wi-Fi signals propagate, such as by observing changes in channel properties known as RSSI (Received Signal Strength Indicator) and also CSI (Channel State Information) (Suroso, Adiyatma, & Cherntanomwong, 2023). Wi-Fi sensing has several advantages such as being able to be implemented with existing Wi-Fi infrastructure, fine-grained localization (if using CSI), being able to determine multi-targets, being non-intrusive and also being able to adapt to environmental changes.

2.2 RSSI (Received Signal Strength Indicator)

Previous Wi-Fi Sensing-based for Human Identification systems utilized the Received Signal Strength Indicator (RSSI) from the Media Access Control (MAC) layer due to the ease of data acquisition. In indoor areas, Wi-Fi signals will propagate to the receiver through multiple paths (multipath). Each path contributes to different signal changes and movements. And RSSI shows the sum of the Wi-Fi signal energy from multipath paths such as the LoS path between the transmitter and receiver, as well as several reflection paths caused by walls, furniture, and people in decibels (dB). To measure the RSSI value, the following equation can be used (Kurniawati et al., 2023):

$$RSSI = 10log_2(||V||^2) \quad (1)$$

Where V is the value of the baseband signal voltage measured at the receiver in a certain time. This parameter has a high fluctuating nature due to the environment and time of transmission. RSSI can be applied to indoor localization for device-free and long-range techniques. Because of its fast and easy implementation, RSSI excels as a parameter for indoor Wi-Fi Sensing applications.

2.3 CSI (Channel State Information)

Most modern wireless communication systems (including Wi-Fi) use a technique called OFDM (Orthogonal Frequency Division Multiplexing) which is a modulation scheme. In OFDM, a channel is divided into several smaller sub-channels (subcarriers), each operating at a different frequency in parallel. This technique can increase the efficiency of data transmission (Burimas, Horanont, Thapa, & Lamichhane, 2024). Each signal sent through a subcarrier can be described by equation (2) (Wang et al., 2021):

$$Y = H \times X + N \quad (2)$$

Where Y is the signal vector at the receiver, X represents the signal vector when transmitted, N is the noise that affects the signal during transmission, and H is the matrix of CSI (Channel State Information) as follows:

$$H = [H_1, H_2, ..., H_k]$$

Which H_k is the value of CSI for the k-th subcarrier. And each value can be defined by (3) (Wang et al., 2021):

$$H_k = |H_k| e^{j \sin(\angle H_k)} \tag{3}$$

$|H_k|$ presenting amplitude information and $\angle H_k$ presenting phase information. CSI continuously detects the frequency response of OFDM subcarriers and captures various environmental changes, such as frequency selective fading, shadowing, multipath, destructive, and constructive interference. Based on these signal changes, individual identities can be identified by classifying the interference patterns that occur to each user.

2.4 CSI Data Collection

Currently, Wi-Fi is a wireless technology that can be accessed from various devices, but the majority of devices with Wi-Fi features are not capable of extracting CSI data, so many solutions require firmware modifications and additional devices. Such as Linux 802.11n (Halperin, Hu, Sheth, & Wetherall, 2011) which has an Intel 5300 Network Interface Card (NIC) only collects up to 30 subcarriers and requires firmware modification. Another tool, the Atheros CSI (Xie, Li, & Li, 2015) works with Atheros 802.11 NICs and gets all 56 subcarriers for 20 MHz bandwidth without tweaking the firmware. However, the mentioned above NIC-based solutions do not support standalone operation and remain impractical for large-scale deployments. (Atif, Muralidharan, Ko, & Yoo, 2020). Nexmon CSI Extractor uses the Broadcom chipset in the Nexmon 5 Android smartphone to obtain CSI data from all 56 subcarriers in 20 MHz bandwidth as a standalone solution. However, the Nexmon-based solution requires modification and may void the device warranty.

Due to its much cheaper price, and also easy to get CSI data, ESP32 is used as one of the solutions to overcome this problem. By using the ESP-IDF framework with the C programming language, a codebase that can extract CSI can be created. With this, ESP32 offers a Wi-Fi Sensing tool solution that can be used on a large scale.

2.5 Preprocessing

The CSI data obtained from the tools in section 2.4 is still raw data, the data must still be processed in such a way that it can be used as machine learning input for classification, this process is called preprocessing. Preprocessing is generally divided into 3 parts, signal selection (2.5.1), noise reduction (2.5.2) and also extracting features (2.5.3)

2.5.1 Signal Segmentation

The selection of signal type can affect the accuracy of the authentication system. CSI data is a channel matrix consisting of amplitude and phase information, in general researchers use amplitude or phase or both for further processing. As in HumanFi (Ming, Feng, & Bu, 2019) which uses amplitude and phase to detect signal variations, and also achieves higher identification accuracy.

2.5.2 Noise Reduction

Raw data from CSI has significant noise, because Wi-Fi devices detect signal interference from crowded environments and reflections from various objects. Therefore, we must eliminate or at least reduce the signal interference in order to obtain a clean signal for further processing (Wang et al., 2021). In general, there are several ways to reduce or eliminate this interference, such as PCA (Principal Component Analysis) where this technique reduces the complex dimensions of the signal but maintains significant signal features, butterworth filter which is used to eliminate high frequency interference, DWT (Discrete Wavelet Transform) which is divided into 3 parts, namely decomposition, thresholding and reconstruction (Wang et al., 2021).

2.5.3 Features Extractions

Feature extraction is the process of extracting important features from raw or pre-processed CSI measurements. This process is done to extract important features that contain information when the signal reflects unique characteristics of humans and remove parts of the signal that are not really needed. This process can improve the efficiency of the authentication system. The useful features in question are features from the time domain and frequency domain.

Features of the frequency domain are extracted after converting the time series of CSI to frequency domain of CSI using FFT (Fast Fourier Transform), the features of the frequency domain such as median, root mean square, and kurtosis. The features of the time domain are taken directly from the original signal shape.

2.6 Machine Learning in Wi-Fi Sensing

After the CSI data goes through the preprocessing, the features that have been extracted from the data will be used as the input features in machine learning. This classification is done to identify human identity based on differences in signal features. Commonly used algorithms such as Support Vector Machine (SVM) will be used and successfully provide high accuracy. SVM is a powerful classification algorithm that has been used in Wi-Fi Sensing with CSI data. SVM attempts to find the most optimal hyperplane to separate objects from different classes in the feature space. SVM is trained to classify objects based on patterns and variations in the CSI data corresponding to each individual. By determining the appropriate kernel function, SVM can effectively understand the complex relationship between CSI and each individual (Ali et al., 2023)

CHAPTER III
RESEARCH MATERIALS AND METHODOLOGY

3.1 Time and Place of Research

This research will be conducted from November 2024 to February 2025, and will be conducted at the physics laboratory of SMA Unggulan Rushd.

3.2 Materials and Tools

Table 3.1, Tools & Materials.

Tools	Qty	Function
ESP32	1	ESP32 functions as a Wi-Fi signal receiver and sends the data to the laptop.
Laptop	1	Where the CSI data is processed
Wi-Fi *Router*	1	Functions to transmit Wi-Fi signals that will be received by the ESP32
VSCode	-	ESP32 programming platform to read CSI signals; Machine Learning model building platform
SciKit	-	Framework used to assist in creating SVM machine learning models
ESP-IDF	-	Framework which is used to help in creating a codebase for retrieving CSI data from the ESP32.

3.3 Research Design and Procedures

This study uses the research and development (R&D) method. The research model applied is the ADDIE development model which includes the stages of Analysis, Design, Development, Implementation, and Evaluation. (Thim-Mabrey, 2006). In the analysis stage, we reveal the problems caused by human identification devices that are often used today. In the design stage, we explain how the tool to be developed works in general, in the development stage we explain in detail the stages starting from compiling the tool to how the data is processed, in the implementation stage we explain how we conduct experiments with the tool that has been designed, and finally in the evaluation stage, we will explain how our system is tested to prove whether the data obtained is accurate with real events.

3.3.1 Analysis

Various security system tools have been applied to maintain and reduce the current crime rate, but the existing security systems still have several weaknesses that can harm various users.

Table 3.2, Problems with current conventional security systems.

No.	Tools	Limitations
1.	CCTV	a. Privacy concerns b. Visibility c. Lighting condition d. High electricity usage
2.	NFC (Near Field Communication)	a. Requires action from the users b. Loss of identification device
3.	GPS	a. Cannot be used in indoor areas

From the table containing the issues identified in current security tools (Anderez et al., 2021), we can observe various limitations. Therefore, an innovative security system can be designed to integrate various technologies, utilize machine learning algorithms, and use low-cost tools to enhance effectiveness, privacy, and reduce reliance on users.

3.3.2 Design

In this study, a human identification system was developed using Wi-Fi signals. A transmitter in the form of a Wi-Fi router and an ESP32 as a receiver are placed opposite each other, so that the signal can propagate between the two devices. The area between the transmitter and receiver will be the main sensor area where the signal will be reflected and disturbed by the physical characteristics of humans. ESP32 will be connected to a computer so that CSI data which is status information from the Wi-Fi signal will be classified based on individual data trained using Machine Learning, as shown in Figure 3.1.

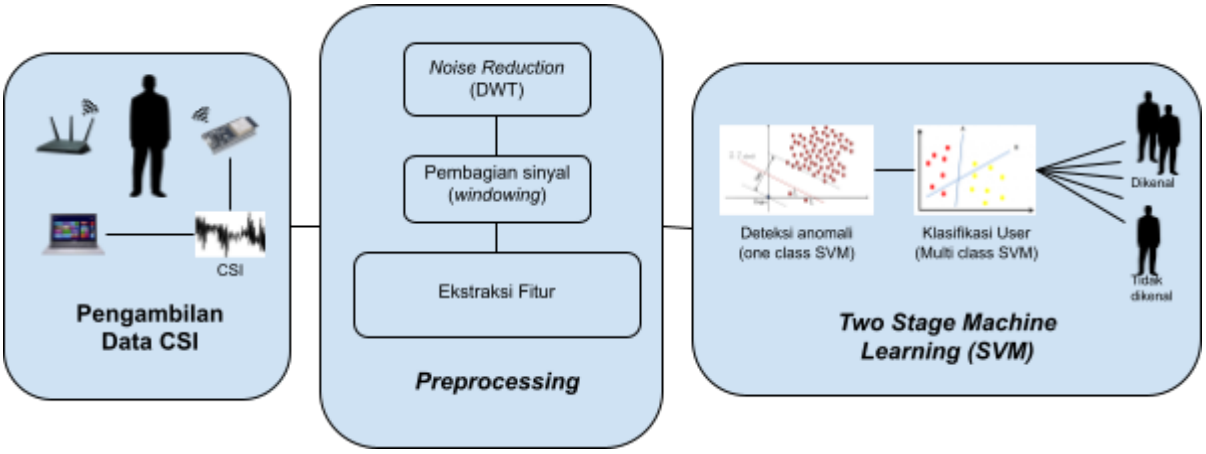


Figure 3.1, Human Identification System Design.

3.3.3 Development

The CSI data that has been taken by the ESP32 will be sent via a USB cable to the laptop for further processing, but this data will only be processed when all the required data has been taken. The CSI will be extracted from the ESP32 receiver using the ESP-IDF framework and stored in CSV format. A timestamp is also given to each CSI frame. The CSI data is then processed which will be explained in section 3.4.

3.3.4 Implementation

In an attempt to measure the effectiveness of the developed system, an experiment was conducted involving 10 volunteers. A Tx (transmitter) in the form of TP-Link and ESP32 as Rx (receiver) were placed opposite each other, the ten people would stand between the Tx and Rx alternately to collect the required CSI data. The data would go through a preprocessing phase and would be divided into 2 datasets as explained above. After the dataset was formed, the model would be trained based on the dataset.

3.3.5 Evaluation

To evaluate the system created, the effectiveness and accuracy of the model that has been trained using the dataset will be tested. Testing is done with other individuals who are not included in the dataset as intruders to see if the model can classify and distinguish them from known individuals. The model will also be tested by classifying each of the ten people used in model training.

3.4 Pengolahan dan Analisis Data

After the CSI data was collected in the development of this system, the windowing process was carried out to organize the data into parts that could be analyzed. Each window consists of 30 consecutive CSI frames, representing a certain time interval of the received signal.

In the preprocessing stage, the amplitude value of each complex number subcarrier was extracted to be used as one of the input features for the machine learning model. To reduce noise in the CSI data while maintaining important information, Discrete Wavelet Transform (DWT) was applied to the amplitude values. In determining the input features, all amplitude values that have been applied to the DWT method are used.

After the preprocessing phase is completed, CSI will be divided into 2 datasets. 70% of the data is used for training and 30% of the data for testing. To classify data based on each individual, a Machine Learning model is needed to classify each CSI data. The two-stage machine learning approach is used in this system to improve performance and detect unknown individuals from this system approach. In the first stage, One Class SVM is used to detect abnormal data patterns, this stage aims to detect whether an individual is known or unknown. Multi Class SVM is applied in the second stage of this model, this algorithm will classify CSI data based on each known individual.

To evaluate the performance results of this system model, the use of metrics such as precision, F1-score and recall are used to assess the performance of the model. Precision measures the proportion of correct positive predictions, while recall measures the proportion of positive predictions that are successfully classified by the model. F1-score will also be used to provide a more detailed picture of the model's performance.

CHAPTER IV
RESULTS AND DISCUSSION

4.1 Experiment Setup

The experiment was conducted in the physics laboratory of SMA Unggulan RUSHD. The placement of the devices used, such as the TP-Link router (Tx) and also the ESP32 (Rx) in the physics laboratory can be seen in Figure 4.1. The subject stands in the gap between two tables as a data collection point. The dataset will be taken from 10 volunteers, consisting of 10 males. For data collection for each subject, the ESP32 will take CSI data that passes through the subject's body with a sampling rate of 100 Hz within 30 seconds.

Figure 4.1, Placement of Tx, Rx, and where the subject stands in the Physics Laboratory.

4.2 Experiment Evaluation

To measure the performance of the developed system, an evaluation was conducted on the model's ability to detect intruders, identify individuals in the dataset, and analyze the effect of the number of individuals on classification accuracy. A performance comparison was also conducted with previous studies to see the advantages of the method used in this study. Model performance assessment was carried out using several standard metrics in classification, namely F1-score, precision, recall, and confusion matrix.

4.2.1 Intruder Detection Performance Evaluation

Identification of unknown individuals is a crucial first step in a secure authentication system. Before an authentication system can classify an enrolled individual, it must first be able to determine whether the individual is enrolled or not. Once the system has successfully identified that the individual is an enrolled person, the next step is to classify the individual in comparison to other enrolled individuals.

To achieve this goal, a Support Vector Machine (SVM) model with a One-Class SVM approach is used. This model utilizes the Radial Basis Function (RBF) kernel to improve classification capabilities in a non-linear feature space. The nu parameter is set to 0.1, which functions to control the proportion of data that is allowed to be considered an outlier. The selection of these parameters is based on the tuning process carried out to obtain parameter values that provide the best performance results. With this setting, the model is expected to be able to effectively distinguish between registered and unregistered individuals, as well as improve the accuracy of the authentication system in identifying legitimate individuals. Unknown subjects will be randomly selected from the available dataset.

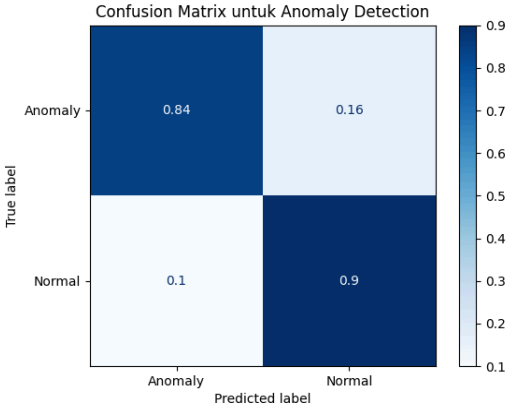


Figure 4.2, Confusion Matrix results on the model.

As seen in Figure 4.2, the developed One Class SVM model can identify intruders quite well, the model gets 84% for TP (True Positive) and 90% for TN (True Negative). However, there are still some classification errors, such as FP (False Positive) of 10% and False Negative of 16%. Classification errors are estimated to occur because registered users have very varied characteristics, considering that Rushd Senior High School students come from various regions, which makes it difficult for the model to make clear boundaries between users and intruders (Liu et al., n.d.). However, it can be concluded that these results indicate that the One Class SVM model with RBF kernel and nu parameter 0.1 is quite effective in determining anomalies in the dataset.

4.2.2 Evaluation of Individual Classification in Dataset

After the CSI data passes through a model that can identify whether the data is registered data or not, the CSI data that is registered data will go through the next step, the data will be classified based on the training data. At this step, the SVM model is also used, but with a multi-class approach to handle the classification of more than 2 individuals. To evaluate the performance of this SVM, metrics such as F1 score, recall, precision, and confusion matrix are applied to each subject.

Table 4.1, Classification metric results table.

Nomor Individu	Precision	Recall	f1-score
1	99%	97%	98%
2	98%	100%	99%
3	100%	100%	100%
4	97%	96%	97%
5	97%	98%	97%
6	99%	99%	99%
7	80%	89%	84%
8	87%	84%	86%
9	99%	94%	96%
10	100%	100%	100%

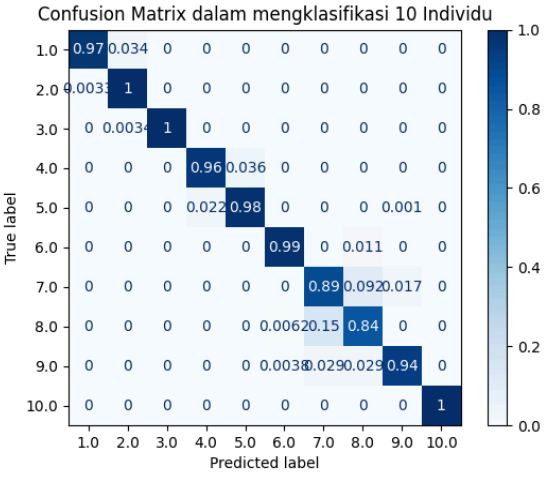


Figure 4.3, Result of confusion matrix on the dataset.

Table 4.2, Weight and height data of volunteers

Volunteer's Number	Weight (kg)	Height (cm)
1	54	163
2	50	161
3	54	177
4	58	164
5	65	172
6	65	167
7	47	163
8	47	164
9	53	173
10	67	164

The results of the classification performance of ten volunteers in the experiment above can be seen from the metrics in Table 4.1 and also the confusion matrix in Figure 4.3. It can be seen that the accuracy results of the testing data classification obtained varying accuracies from 89% to 100% in the confusion matrix. Volunteers 2, 3 and 10 had the highest accuracy of 100%. This can happen because they are volunteers with the most different physical characteristics. Volunteer 2 is the person with the smallest height, Volunteer 3 has the highest height and Volunteer 10 has the largest weight. It can also be seen that the model has difficulty in distinguishing Volunteers 7 and 8. The model obtained an accuracy of 84% in identifying individual number 8. However, the model also experienced a 15% misclassification, where individual number 8 was mistakenly identified as number 7. This result was obtained because Volunteers 7 and 8 have almost similar physical characteristics. Based on Table 4.2, Their heights are not much different, and their weights are also not much different. This shows that height and weight, which are physical characteristics, affect signal transmission with different patterns.

4.2.3 Effect of Number of Registered Individuals on Model Performance

In this section, it is discussed how the number of individuals registered in the dataset in an identification system affects the performance of the classification model. The model will be trained using CSI data with varying numbers of individuals, ranging from 2 individuals to 10 individuals. The selection of individuals is done sequentially, starting from the 10th volunteer and gradually covering volunteers with smaller numbers up to the 1st volunteer.

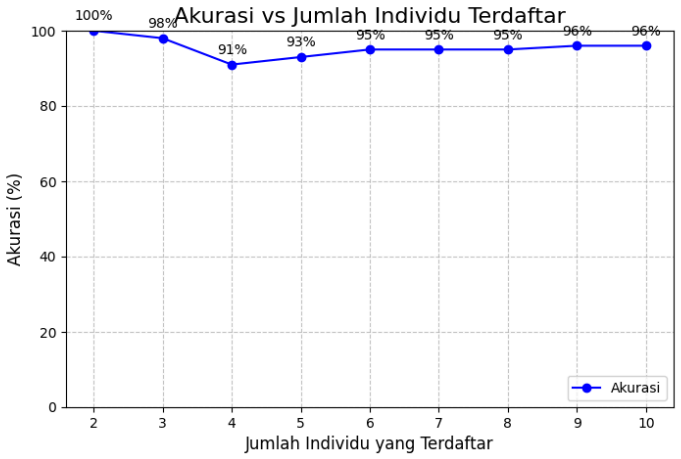


Figure 4.4, Line chart of model accuracy figures with varying numbers of enrolled individuals.

Based on Figure 4.4, when the number of individuals in the dataset is 2 to 3 people, the model gets an accuracy of 100% and 98%. However, when the model classifies 4 individuals, the accuracy rate decreases quite significantly. This is because the 4th individual who is the 7th volunteer, and also the 3rd individual who is the 8th volunteer have almost the same physical characteristics, so the model has difficulty in predicting individuals. As the number of individuals in the dataset increases, the accuracy rate increases slowly again until it reaches an accuracy rate of 96% at 10 individuals.

4.2.3 Performance Comparison with Previous Research

To prove the good performance of the authentication system using the ESP32 microcontroller, the performance results of this study will be compared with the performance results of other studies. Several previous similar studies using 10 subjects as the dataset were selected, such as WiID (Zheng et al., 2017), research conducted by Wang (Wang et al., 2018), Gate-ID (Zhang et al., 2021), and SmartUserAuth (Shi et al., 2017) will be compared with the accuracy performance of this research model.

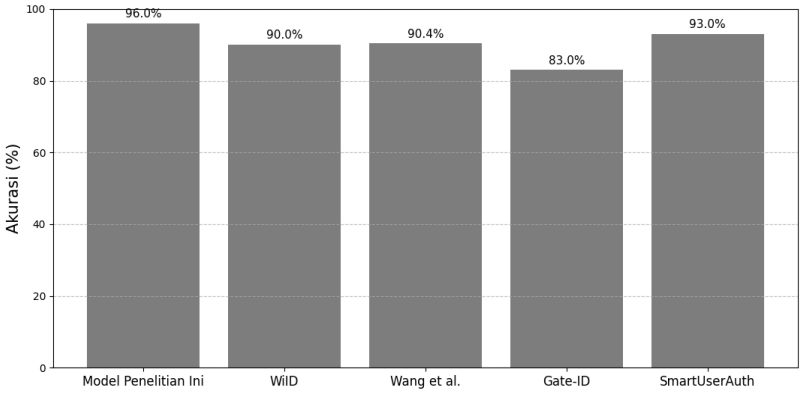


Figure 4.5, Bar chart representing accuracy performance results with various previous studies.

WiID (Zheng et al., 2017) uses CSI data from commercial Wi-Fi devices from common human activities such as movement and interaction to identify individual identities, while Gate-ID (Zhang et al., 2021) identifies individuals based on the walking direction of each subject. There is also a study

conducted by Wang (Wang et al., 2018), where they offer a human identification system through the unique influence of humans on Wi-Fi signals, on the other hand, SmartUserAuth (Shi et al., 2017) utilizes Wi-Fi signals to capture physical characteristics and human behavior, both when moving and stationary. This system uses a deep learning-based authentication scheme.

As seen from Figure 4.5, when the proposed system in this study is compared with the previously mentioned studies, with respect, this system provides better performance. There are several reasons why this can happen, such as the high variation in physical features of students at Unggulan Rushd High School. In the Gate-ID study, there is no filtering to reduce noise in the CSI data, the study also uses signal differences caused by walking patterns to classify individuals, which can change depending on physical conditions, walking speed, environmental obstacles, and also when experiencing disturbances while walking, Gate-ID also uses Attention-Based Deep Learning Model with ResNet and Bi-LSTM, which are at risk of overfitting especially because their dataset is limited. In the WiID study, a filtering method is used by applying PCA, which can cause the loss of important information in the data. WiID also uses the influence of Wi-Fi signals on human activity as a sensor measurement instrument, just like SmartUserAuth. The studies mentioned also did not use ESP32 as a receiver, which causes their systems to have weaknesses because the use of ESP32 in these studies allows the system to be more portable, power efficient, and easier to implement in various environments.

CHAPTER V
CONCLUSION AND SUGGESTION

5.1 Conclusion

In this study, a human authentication system has been developed by utilizing Wi-Fi signals obtained from the ESP32 microcontroller. The SVM machine learning model is used to process CSI data that represents the Wi-Fi signal status. This data first goes through a preprocessing stage to reduce disturbing noise using the DWT algorithm, then the data will be classified to determine whether the data comes from registered individual data using the OneClassSVM approach. Furthermore, classification is carried out based on the influence of individual physical characteristics on the signal with the MultiClassSVM approach, so that it can be identified who influences the signal. In an experiment that has been conducted, this system can identify individuals with an average accuracy of 100% to 96% in groups of 2 to 10 people. In determining whether the data comes from registered individual data, the model gets an average accuracy of 90% with 10 registered people. Compared to previous related studies, this system has several advantages, such as better model performance results, higher efficiency, lower cost, ease of implementation, and is a portable system due to the use of the ESP32 microcontroller.

5.2 Suggestion

Although the developed system shows good results as a human authentication system that utilizes Wi-Fi signals and uses an ESP32 microcontroller, there are several aspects that can be improved for further research, such as exploring a more optimal model in detecting outliers in a dataset so that it will get better performance, finding more appropriate feature inputs to achieve higher accuracy, experimental testing where Rx and Tx will be placed in different rooms so that the signal must pass through the wall, as well as utilizing CSI data for other purposes outside the human authentication system.

ACKNOWLEDGEMENT

Praise be to Allah SWT who has bestowed His grace and guidance so that the author can complete the research entitled "Development of Wi-Fi Sensing-Based Authentication System Using CSI Data and Machine Learning Algorithm with ESP32 Microcontroller" well and on time. In compiling this paper, the authors have tried their best and of course with the help of various parties. For that, the author would like to express his gratitude to.

1. Parents who always support in doing this research paper.
2. Agung Prastyo, S.Pd. as a mentor of the Unggulan Rushd Senior High School research club.
3. Hanun Fithriyah, S.Pd. as the internal supervisor of this research paper.
4. Zahra Dyah Meilani, S.Kom. as the external supervisor of this research paper.
5. Volunteers who helped in the completion of this research paper.
6. Other parties who helped in the preparation of this paper.

In writing this research report, there are still many shortcomings and errors, therefore all constructive criticism and suggestions will improve the writing of this research report and be useful for the author and readers.

REFERENCES

Ali, M., Hendriks, P., Popping, N., Levi, S., & Naveed, A. (2023). *A comparison of machine learning algorithms for Wi-Fi sensing using CSI data*. *Electronics (Switzerland)*, 12(18), 3935. https://doi.org/10.3390/electronics12183935

Anderez, D.O., Kanjo, E., Amnwar, A., Johnson, S., & Lucy, D. (2021). *The rise of technology in crime prevention: Opportunities, challenges and practitioners perspectives*. *arXiv*. http://arxiv.org/abs/2102.04204

Atif, M., Muralidharan, S., Ko, H., & Yoo, B. (2020). Wi-ESP—A tool for CSI-based device-free Wi-Fi sensing (DFWS). Finite Element Analysis and Design, 7(5), 644–656.

Burimas, R., Horanont, T., Thapa, A., & Lamichhane, B. R. (2024). Monitoring the sleep respiratory rate with low-cost Wi-Fi microcontroller in a controlled environment. Applied Sciences, 14(15), 6458. https://doi.org/10.3390/app14156458

Chen, J.J., Jiang, Z.X., Chen, Y.L., Wu, W.T., & Liang, J.M. (2015). Design and realization of an NFC-driven smart home system to support intruder detection and social network integration. Journal of Electronic Science and Technology, 13(2), 163–168.

Griffiths, G., Johnson, S. D., & Chetty, K. (2017). UK-based terrorists' antecedent behavior: A spatial and temporal analysis. Applied Geography, 86, 274–282.

Halperin, D., Hu, W., Sheth, A., & Wetherall, D. (2011). Tool release: Gathering 802.11n traces with channel state information. ACM SIGCOMM Computer Communication Review, 41(1), 53–53.

Kurniawati, N., Nurjihan, SF, Dwitio, D., & Pranoto, A. (2023). Design and construction of a Wi-Fi-based human activity detection system.

Lawrence, D.S., La Vigne, N.G., Goff, M., & Thompson, P.S. (2018). Lessons learned implementing gunshot detection technology: Results of a process evaluation in three major cities. Justice Evaluation Journal, 1(2), 109–129.

Lv, J., Man, D., Yang, W., Gong, L., Du, X., & Yu, M. (2019). Robust device-free intrusion detection using physical layer information of Wi-Fi signals. Applied Sciences, 9(1), 17. https://doi.org/10.3390/app901017

Ming, X., Feng, H., & Bu, Q. (2019). HumanFi: Wi-Fi-based human identification using recurrent neural network. In IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (pp. 640–647). IEEE.

Mosharaf, M., Kwak, J. B., & Choi, W. (2024). Wi-Fi-based human identification with machine learning: A comprehensive survey. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/s24196413

Parameswaran, AT, Husain, MI, & Upadhyaya, S. (n.d.). Is RSSI a reliable parameter in sensor localization algorithms: An experimental study.

Piza, E.L., Welsh, B.C., Farrington, D.P., & Thomas, A.L. (2019). CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis. Criminology & Public Policy, 18(1), 135–159.

Suroso, DJ, Adiyatma, FYM, & Cherntanomwong, P. (2023). Wi-Fi sensing for indoor localization via channel state information: A survey. ELKHA: Journal of Electrical Engineering, 15(2), 152–159.

Thim-Mabrey, C. (2006). Sprachwandel in übersetzung bearbeitungen zwischen 1846 und 1999. Neuphilologische Mitteilungen, 107(3), 361–373.

Wang, Z., et al. (2021). A survey of user authentication based on channel state information. Hindawi Limited. https://doi.org/10.1155/2021/6636665

Xie, Y., Li, Z., & Li, M. (2015). Precise power delay profiling with commodity Wi-Fi. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (pp. 53–64). ACM.

Shi, C., Liu, J., Liu, H., & Chen, Y. (2017). Smart User Authentication through Actuation of Daily Activities Leveraging Wi-Fi-enabled IoT. Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing. https://doi.org/10.1145/3084041.3084061

Wang, J., Zhao, Y., Fan, X., Gao, Q., Ma, X., & Wang, H. (2018). Device-Free Identification Using Intrinsic CSI Features. IEEE Transactions on Vehicular Technology, 67(9), 8571–8581. https://doi.org/10.1109/TVT.2018.2853185

Zhang, J., Wei, B., Wu, F., Dong, L., Hu, W., Kanhere, S. S., Luo, C., Yu, S., & Cheng, J. (2021). Gate-ID: Wi-Fi-Based Human Identification Irrespective of Walking Directions in Smart Home. IEEE Internet of Things Journal, 8(9), 7610–7624. https://doi.org/10.1109/JIOT.2020.3040782

Zheng, R., Zhao, Y., & Chen, B. (2017). Device-Free and Robust User Identification in Smart Environment Using Wi-Fi Signal. 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC), 1039–1046. https://doi.org/10.1109/ISPA/IUCC.2017.00158

Liu, J., Yuan, J., & Li, C.-N. (n.d.). *One-class SVM with 0-1 loss*. https://ssrn.com/abstract=5050012

PERNYATAAN PENELITI

Yang bertanda tangan di bawah ini:
Nama : Muhammad Fathan Haroki
Tempat/Tanggal Lahir : Gresik, 06 Februari 2009
NIS : 0095429502
Asal Sekolah : SMA Unggulan Rushd

dengan ini menyatakan sejujurnya bahwa proposal penelitian saya dengan judul: Pengembangan Sistem Autentikasi Berbasis WiFi Sensing Menggunakan Data CSI dan Algoritma Pembelajaran Mesin dengan Mikrokontroler ESP32

bersifat orisinal/bukan hasil tindak plagiarisme/belum pernah dikompetisikan dan/atau tidak sedang diikutkan pada lomba penelitian sejenis/belum pernah mendapatkan penghargaan di tingkat Nasional/Internasional.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, saya bersedia menerima konsekuensi sesuai aturan ISPO.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Dibuat di
Sragen Pada
Tanggal 30 November 2024

Mengetahui,

Hanun Fithriyah, S.Pd
NIP: -

Yang membuat pernyataan

Muhammad Fathan Haroki
NISN: 0095429502