

IMPLEMENTASI MODEL RECOGNISI SUARA MENGGUNAKAN METODE CONVOLUTIONAL RECURRENT NEURAL NETWORK (CRNN)

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Serjana Komputer



OLEH:

NANDA HARSANA OCTAVYA

09011381722142

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2021**

LEMBAR PENGESAHAN

IMPLEMENTASI MODEL RECOGNISI SUARA MENGGUNAKAN METODE CONVOLUTIONAL RECURRENT NEURAL NETWORK (CRNN)

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

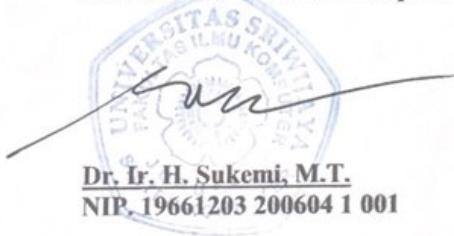
Oleh :

NANDA HARSANA OCTAVYA
09011381722142

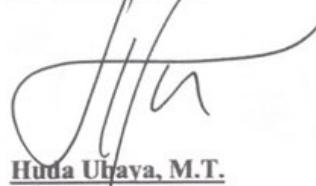
Palembang, 1^VJuli 2021

Mengetahui,

Ketua Jurusan Sistem Komputer


Dr. Ir. H. Sukemi, M.T.
NIP. 19661203 200604 1 001

Pembimbing Tugas Akhir


Huda Ubaya, M.T.
NIP. 198106162012121003

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Selasa

Tanggal : 06 Juli 2021

Tim Penguji:

1. Ketua Sidang : Sarmayanta Sembiring, M.T



2. Sekretaris Sidang : Rendyansyah, M.T



3. Penguji Sidang : Ahmad Zarkasi, M.T

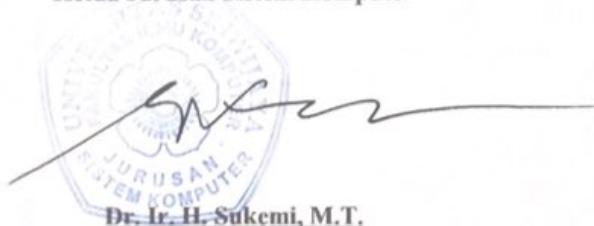


4. Pembimbing : Huda Ubaya, M.T




Mengetahui, 28/7/21

Ketua Jurusan Sistem Komputer



NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Nanda Harsana Octavya
NIM : 09011381722142
Judul : Implementasi Model Rekognisi Suara Menggunakan Metode *Convolutional Recurrent Neural Network (CRNN)*

Hasil Pengecekan Software *iThenticate/Turnitin* : 7%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan.



Palembang, Juli 2021



Nanda Harsana Octavya

09011381722142

HALAMAN PERSEMBAHAN

“Anda Ragu-Ragu Kembali Sekarang Juga”

“Hasbunallah Wanikmal Wakil Ni'mal Maula Wani'man Nasir”

“Alhamdulillah akhirnya bisa melewati 4 tahun yang penuh suka cita ini, semoga pilihan masuk SK tidak salah dan bisa menjembatangi ke jenjang yang lebih suskses 黑洞 (Hwaiting)”

“Kupersembahkan skripsi ini untuk diriku sendiri di masa depan, untuk orang yang selalu mendukungku disaat diriku susah maupun senang, orang yang selalu memperjuangkan ku dengan jerih payah mereka, orang yang selalu ikhlas mendoakanku sehingga aku sampai titik yaitu PAPA dan MAMAKU”

KATA PENGANTAR

Segala puji dan syukur atas kehadiran Allah Subhanahu wa Ta'ala, karena berkat Rahmat dan Karunia-Nya lah sehingga penulis dapat menyelesaikan penyusunan laporan Proposal Tugas Akhir ini yang berjudul "**Implementasi Model Rekognisi Suara Menggunakan Metode *Convolutional Recurrent Neural Network (CRNN)***".

Laporan ini merupakan salah satu syarat untuk memenuhi sebagian kurikulum dan syarat kelulusan Jurusan Sistem Komputer Unggulan serta memperoleh gelar Serjana Komputer.

Pada kesempatan ini penulis mengucapkan terima kasih kepada pihak yang telah memberikan bantuan, dorongan, motivasi, semangat dan bimbingan dalam penyusunan laporan ini :

1. Allah SWT, yang telah memberikan rahmat, karunia-Nya serta petunjuk sehingga pelaksanaan dan penulisan Proposal Tugas Akhir ini dapat berjalan dengan lancar.
2. Papa saya Haryanto, Mama saya Asma Helfianah, Adik Saya Aji Herian dan Dinda Harsana beserta keluarga yang selalu mendoakan serta memberikan motivasi, bantuan dan semangat.
3. Bapak Jaidan Jauhari, S.Pd. M.T selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya
4. Bapak Dr.Ir.Sukemi, M.T selaku Ketua Jurusan dan Bapak Muhammad Ali Buchari, M.T selaku Sekertaris Jurusan Sistem Kompuer Fakutas Ilmu Komputer Universitas Sriwijaya

5. Bapak Alm Dr.Reza Firsandaya malik, M.T. selaku Dosen Pembimbing akademik sekaligus Dosen Pembimbing Tugas Akhir Penulis yang telah berjasa dalam kehidupan perkuliahan saya karena selalu membimbing, memotivasi dan mengarahkan saya.
6. Bapak Huda Ubaya, M.T Dosen Pembimbing Tugas Akhir Penulis yang selalu membimbing, memotivasi dan mengarahkan saya.
7. Teman-teman saya di Jurusan Sistem Komputer 2017 Palembang dan di Muara Enim yang telah membantu dan memberikan saya memotivasi selama kuliah.
8. Kakak tingkat Sistem Komputer yang telah membantu dan memberi saya referensi.
9. Jodoh saya yang tertulis di lauhul mahfudz yang secara tidak langsung memberikan saya motivasi menjadi lebih baik lagi.
10. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang selalu memberikan semangat dan bantuan-bantuan yang bermanfaat.

Dalam penyusunan Laporan Tugas Akhir ini penulis menyadari sepenuhnya bahwa laporan ini masih jauh dari sempurna, oleh karena itu penulis mengharapkan saran dan kritik dari semua pihak yang berkenan demi laporan yang lebih baik lagi.

Akhir kata penulis berharap semoga Laporan Tugas Akhir ini dapat bermanfaat serta dapat memberikan pengetahuan dan wawasan bagi semua pihak yang membutuhkannya terutama mahasiswa Sistem Komputer Universitas Sriwijaya.

Palembang, Juli 2021

Penulis,

Nanda Harsana Octavya

NIM. 09011381722142

IMPLEMENTATION OF VOICE RECOGNITION MODEL USING CONVOLUTIONAL RECURRENT NEURAL NETWORK (CRNN) METHOD

NANDA HARSANA OCTAVYA (09011381722142)

*Computer Engineering Department, Computer Science Faculty, Sriwijaya
University
Email : nandaharsana66@gmail.com*

ABSTRACT

Voice recognition is a process that can determine the identity of the user based on the characteristics of the voice through the words conveyed because every human being has a distinctive sound characteristic. Human voices can be recognized by gender. Recognition of gender or gender based on voice can be detected automatically through the characteristics of the voice. The method used in this research is Convolutional Recurrent Neural Network (CRNN). Based on the tuning process that has been carried out, the parameter values that produce the best output are the distribution of 80% training data and 20% testing data, batch_size of 8, learning-rate of 0.0001 and epoch of 300. The model that has been built produces a Training Accuracy value of 99 ,41% while Testing Accuracy is 99.05%. The average training performance value obtained from each class is 99.3% for the value of Specificity, Precision and Recall, while the FI-Score value is 99.4%. The average Testing performance value obtained from each class is 98.9% for the Specificity, Precision and Recall values, while the FI-Score value is 99%.

Keywords: Recognition, Voice, Gender, Convolutional Recurrent Neural Network.

IMPLEMENTASI MODEL RECOGNISI SUARA MENGGUNAKAN METODE CONVOLUTIONAL RECURRENT NEURAL NETWORK (CRNN)

NANDA HARSANA OCTAVYA (09011381722142)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : nandaharsana66@gmail.com

ABSTRAK

Rekognisi suara adalah suatu proses yang dapat mengetahui identitas pengguna berdasarkan karakteristik suara melalui kata yang disampaikan karena setiap manusia memiliki ciri khas suara yang berbeda-beda. Suara manusia dapat direkognisi berdasarkan *gender*. Rekognisi *gender* atau jenis kelamin berdasarkan suara dapat dideteksi secara otomatis melalui karakteristik suaranya. Metode yang digunakan pada penelitian ini adalah *Convolutional Recurrent Neural Network (CRNN)*. Berdasarkan proses *tuning* yang telah dilakukan nilai parameter yang menghasilkan *output* terbaik adalah pembagian 80% data *training* dan 20% data *testing*, *batch_size* sebesar 8, *learning-rate* sebesar 0,0001 dan *epoch* sebesar 300. Model yang telah dibangun menghasilkan nilai Akurasi *Training* sebesar 99,41 % sedangkan Akurasi *Testing* sebesar 99,05%. Nilai performa *Training* yang didapatkan dari masing-masing class rata-ratanya sebesar 99,3% untuk nilai *Specificity*, *Precision* dan *Recall*, Sedangkan nilai *FI-Score* sebesar 99,4%. Nilai performa *Testing* yang didapatkan dari masing-masing class rata-ratanya sebesar 98,9% untuk nilai *Specificity*, *Precision* dan *Recall*, Sedangkan nilai *FI-Score* sebesar 99%.

Kata Kunci : Rekognisi, Suara, *Gender*, *Convolutional Recurrent Neural Network*.

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
ABSTRACT	viii
ABSTRAK	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan	3
1.5 Metode Penelitian.....	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Suara.....	7
2.2 Rekognisi Suara	8
2.2.1 Perbedaan Suara Laki-laki dan Perempuan.....	8
2.3 <i>Machine Learning</i>	9
2.4 <i>Deep Learning</i>	10

2.5 Convolutional Neural Network (CNN)	12
2.5.1 Convolutional Layer	13
2.5.2 Pooling Layer.....	15
2.5.3 Fully Connected Layer.....	16
2.6 Recurrent Neural Network (RNN).....	16
2.6.1 Long Short Term Memory (LSTM).....	18
2.7 Convolutional Recurrent Neural Network (CRNN)	20
2.8 Validasi Performa.....	22
2.8.1 Akurasi	23
2.8.2 Sensitivitas	23
2.8.3 Specificity	23
2.8.4 Precision.....	23
2.8.5 FI-Score.....	24
BAB III METODOLOGI PENELITIAN	25
3.1 Pendahuluan	25
3.2 Kerangka Kerja	25
3.3 Persiapan Data.....	27
3.4 Data <i>Pre-processing</i>	28
3.5 Pembagian Data <i>Training</i> dan <i>Testing</i>	30
3.6 Normalisasi Data Fitur	30
3.7 Penerapan Metode <i>Convolutional Recurrent Neural Network (CRNN)</i>	31
3.7.1 Proses <i>tuning</i>	33
3.8 <i>Model Predict</i>	35
3.9 Validasi Performa Model	36
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Pendahuluan	37
4.2 Persiapan Data.....	37
4.3 Hasil Klasifikasi	41

4.3.1	Hasil Validasi Berdasarkan Data <i>Training</i> dan Data <i>Testing</i>	41
4.3.1.1	Hasil Pembagian 90% Data <i>Training</i> dan 10% Data <i>Testing</i>	41
4.3.1.2	Hasil Pembagian 80% Data <i>Training</i> dan 20% Data <i>Testing</i>	45
4.3.1.3	Hasil Pembagian 70% Data <i>Training</i> dan 30% Data <i>Testing</i>	48
4.3.2	Hasil Validasi Berdasarkan <i>Batch_size</i>	51
4.3.2.1	Hasil <i>Batch_size</i> 8	51
4.3.2.2	Hasil <i>Batch_size</i> 16	54
4.3.3	Hasil Validasi Berdasarkan <i>Learning_rate</i>	58
4.3.3.1	Hasil <i>Learning_rate</i> 0,00001	58
4.3.3.2	Hasil <i>Learning_rate</i> 0,000001	61
4.3.4	Hasil Validasi Berdasarkan <i>Epochs</i>	65
4.3.4.1	Hasil <i>Epoch</i> 100	65
4.3.4.2	Hasil <i>Epoch</i> 200	68
4.4	Analisa Perbandingan Hasil Klasifikasi.....	72
4.4.1	Analisis Berdasarkan Perbandingan Pembagian Data <i>Training</i> dan Data <i>Testing</i>	72
4.4.2	Analisis Berdasarkan Perbandingan <i>Batch_size</i>	75
4.4.3	Analisis Berdasarkan Perbandingan <i>Learning_rate</i>	78
4.4.4	Analisis Berdasarkan Perbandingan <i>Epochs</i>	81
4.4.5	Analisis Nilai Parameter Terbaik	84
4.5	Hasil dan Analisis Rekognisi Suara Berdasarkan <i>Gender</i>	86
BAB V KESIMPULAN SEMENTARA	90
5.1	Kesimpulan	90
5.2	Saran.....	91
DAFTAR PUSTAKA	92

DAFTAR GAMBAR

Gambar 2.1	Arsitektur Convolutional Neural Network [18].....	12
Gambar 2.2	<i>Feature Map</i> [17].....	14
Gambar 2.3	Contoh <i>Max Pooling</i> [20]	15
Gambar 2.4	Arsitektur Neural Network [23]	16
Gambar 2.5	<i>Layer</i> pada <i>Recurrent Neural Network</i> [24]	17
Gambar 2.6	Arsitektur <i>Long Short Term Memory</i> [23].....	18
Gambar 2.7	Modul <i>Recurrent Neural Network</i> [23]	19
Gambar 2.8	Arsitektur <i>Convolutional Recurrent Neural Network (CRNN)</i> [9].....	21
Gambar 3.1	Diagram Alur Penelitian	24
Gambar 3.2	Diagram Alur Data <i>Pre-Processsing</i>	29
Gambar 3.3	Diagram Alur <i>Tuning</i>	34
Gambar 3.4	Diagram Alur Model <i>Predict</i>	35
Gambar 4.1	Visualisasi Dataset Berdasarkan Label.....	38
Gambar 4.2	Gambar Histogram Data Fitur	39
Gambar 4.3	Persebaran Data Label Berdasarkan Fitur <i>Meanfun</i> dan <i>IQR</i>	39
Gambar 4.4	Persebaran Data Label Berdasarkan Fitur <i>Meanfun</i> dan <i>Q25</i>	40
Gambar 4.5	Persebaran Data Label Berdasarkan Fitur <i>IQR</i> dan <i>Sd</i>	40
Gambar 4.6	Persebaran Data Label Berdasarkan Fitur <i>Sp.ent</i> dan <i>Q25</i>	40
Gambar 4.7	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 90% Data <i>Training</i> dan 10% Data <i>Testing</i>	44

Gambar 4.8	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 90% Data <i>Training</i> dan 10% Data <i>Testing</i>	44
Gambar 4.9	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 80% Data <i>Training</i> dan 20% Data <i>Testing</i>	47
Gambar 4.10	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 80% Data <i>Training</i> dan 20% Data <i>Testing</i>	47
Gambar 4.11	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 70% Data <i>Training</i> dan 30% Data <i>Testing</i>	50
Gambar 4.12	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Pembagian 70% Data <i>Training</i> dan 30% Data <i>Testing</i>	50
Gambar 4.13	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Batch_size</i> 8	53
Gambar 4.14	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Batch_size</i> 8	54
Gambar 4.15	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Batch_size</i> 16	57
Gambar 4.16	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Batch_size</i> 16	57
Gambar 4.17	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Learning_rate</i> 0,00001	60
Gambar 4.18	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Learning_rate</i> 0,00001	61
Gambar 4.19	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Learning_rate</i> 0,000001	64
Gambar 4.20	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan <i>Learning_rate</i> 0,000001	64
Gambar 4.21	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Epoch 100	67
Gambar 4.22	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Epoch 100	68

Gambar 4.23	Grafik Akurasi Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Epoch 200.....	71
Gambar 4.24	Grafik <i>Loss</i> Pada Proses <i>Training</i> dan <i>Testing</i> Berdasarkan Epoch 200	71
Gambar 4.25	Rangkuman Nilai Akurasi Dari Proses <i>Tuning</i>	85

DAFTAR TABEL

Tabel 1.	Tabel Tipe <i>Convolutional Layer</i>	15
Tabel 2.	Tabel Matrik Konfusi	22
Tabel 3.	Tabel Parameter Dataset.....	27
Tabel 4.	Tabel Jumlah Data <i>Training</i> dan Data <i>Testing</i>	30
Tabel 5.	Tabel Model CRNN	32
Tabel 6.	Tabel Parameter Umum Model CRNN	33
Tabel 7.	Tabel Konfusi Matrik Berdasarkan Pembagian 90% Data <i>Training</i> dan 10% Data <i>Testing</i>	42
Tabel 8.	Tabel Nilai Performa Berdasarkan Pembagian 90% Data <i>Training</i> dan 10% Data <i>Testing</i>	42
Tabel 9.	Tabel Konfusi Matrik Berdasarkan Pembagian 80% Data <i>Training</i> dan 20% Data <i>Testing</i>	45
Tabel 10.	Tabel Nilai Performa Berdasarkan Pembagian 80% Data <i>Training</i> dan 20% Data <i>Testing</i>	46
Tabel 11.	Tabel Konfusi Matrik Berdasarkan Pembagian 70% Data <i>Training</i> dan 30% Data <i>Testing</i>	48
Tabel 12.	Tabel Nilai Performa Berdasarkan Pembagian 70% Data <i>Training</i> dan 30% Data <i>Testing</i>	49
Tabel 13.	Tabel Konfusi Matrik Berdasarkan <i>Batch_size</i> 8.....	51
Tabel 14.	Tabel Nilai Performa Berdasarkan <i>Batch_size</i> 8.....	52
Tabel 15.	Tabel Konfusi Matrik Berdasarkan <i>Batch_size</i> 16.....	55
Tabel 16.	Tabel Nilai Performa Berdasarkan <i>Batch_size</i> 16.....	55

Tabel 17.	Tabel Konfusi Matrik Berdasarkan <i>Learning_rate</i> 0,00001	58
Tabel 18.	Tabel Nilai Performa Berdasarkan <i>Learning_rate</i> 0,00001.....	59
Tabel 19.	Tabel Konfusi Matrik Berdasarkan <i>Learning_rate</i> 0,000001.....	62
Tabel 20.	Tabel Nilai Performa Berdasarkan <i>Learning_rate</i> 0,000001.....	62
Tabel 21.	Tabel Konfusi Matrik Berdasarkan <i>Epoch</i> 100.....	65
Tabel 22.	Tabel Nilai Performa Berdasarkan <i>Epoch</i> 100.....	66
Tabel 23.	Tabel Konfusi Matrik Berdasarkan <i>Epoch</i> 200.....	69
Tabel 24.	Tabel Nilai Performa Berdasarkan <i>Epoch</i> 200.....	69
Tabel 25.	Tabel Perbandingan Konfusi Matrik Berdasarkan Pembagian Data <i>Training</i> dan Data <i>Testing</i>	73
Tabel 26.	Tabel Perbandingan Nilai Performa Berdasarkan Pembagian Data <i>Training</i> dan Data <i>Testing</i>	73
Tabel 27.	Tabel Perbandingan Grafik Akurasi dan <i>Loss</i> Berdasarkan Pembagian Data <i>Training</i> dan Data <i>Testing</i>	75
Tabel 28.	Tabel Perbandingan Konfusi Matrik Berdasarkan <i>Batch_size</i>	76
Tabel 29.	Tabel Perbandingan Nilai Performa Berdasarkan <i>Batch_size</i>	76
Tabel 30.	Tabel Perbandingan Grafik Akurasi dan <i>Loss</i> Berdasarkan <i>Batch_size</i>	78
Tabel 31.	Tabel Perbandingan Konfusi Matrik Berdasarkan <i>Learning_rate</i>	79
Tabel 32.	Tabel Perbandingan Nilai Performa Berdasarkan <i>Learning_rate</i>	79
Tabel 33.	Tabel Perbandingan Grafik Akurasi dan <i>Loss</i> Berdasarkan <i>Learning_rate</i>	81
Tabel 34.	Tabel Perbandingan Konfusi Matrik Berdasarkan <i>Epoch</i>	82
Tabel 35.	Tabel Perbandingan Nilai Performa Berdasarkan <i>Epoch</i>	82
Tabel 36.	Tabel Perbandingan Grafik Akurasi dan <i>Loss</i> Berdasarkan <i>Epoch</i>	84
Tabel 37.	Tabel Hasil Rekognisi Suara Tidak <i>Noise</i> Berdasarkan <i>Gender</i>	87
Tabel 38.	Tabel Hasil Rekognisi Suara <i>Noise</i> Berdasarkan <i>Gender</i>	88

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam dunia informatika dan teknologi informasi suatu sistem dapat diakses dengan cepat menggunakan teknologi *intelligent voice assistants*. *Intelligent voice assistants* dapat mempermudah pengguna dalam mengoprasikan suatu perangkat dengan menggunakan suara, contoh perangkat dari *intelligent voice assistant* yaitu *Siri*, *Google Now*, *Cortona*, *Amazon Echo*, *Alexa* dan lain-lain[1]. Metode yang digunakan dalam mengoprasikan perangkat tersebut adalah dengan mengeksekusi perintah menggunakan suara pengguna[2]. Selain dapat digunakan pada perangkat lunak *intelligent voice assistants* juga dapat digunakan di perangkat keras misalnya pada mesin[3].

Pada kasus tertentu suatu perangkat harus mempunyai sistem keamanan yang baik agar tidak sembarang orang bisa mengoprasikan perangkat tersebut sehingga orang-orang yang tidak mempunyai wewenang tidak dapat menggunakanannya misalnya dengan mengenali suara pengguna. Rekognisi suara adalah suatu proses yang dapat mengetahui identitas pengguna berdasarkan karakteristik suara melalui kata yang disampaikan[4] karena setiap manusia memiliki ciri khas suara yang berbeda-beda[5]. Proses pengenalan suara terdiri dari 2 macam yaitu Rekognisi pembicara, mendeteksi suara berdasarkan suara orang yang berbicara dan Rekognisi ucapan, mendeteksi suara berdasarkan kata atau kalimat yang diucap[3].

Deteksi *gender* atau jenis kelamin berdasarkan suara dapat dideteksi secara otomatis melalui karakteristik suaranya[6]. Oleh karena itu diperlukan suatu metode pada mesin yang dapat membangun sistem model rekognisi suara agar bisa

mendeteksi suara dengan cepat. Penelitian sebelumnya mengenai Implementasi Model Rekognisi suara berdasarkan *gender* telah dilakukan yaitu dengan menggunakan *Gaussian Mixture Models* (GMM) dengan menghasilkan akuarasi sebesar 81,18% [7] Namun kesalahan dalam klasifikasi suara nilainya masih besar dan penelitian tersebut hanya sebatas klasifikasi saja tidak sampai merekognisi suara dengan menguji sampel suara menggunakan data suara diluar dataset.

Kemudian penelitian menggunakan *Vector Quantization*[8] yang menghasilkan pengujian *offline* laki-laki mencapai 90% dan wanita mencapai 80%. Sedangkan pada pengujian *online* laki-laki mencapai 67% dan wanita mencapai 61% namun pada penelitian tersebut data diuji diambil dari data *training* atau data didalam dataset bukan suara yang diluar dataset sehingga pada penelitian tersebut data suara yang sebelumnya tidak dijadikan data *training* menghasilkan akurasi yang lebih kecil dibandingkan data suara yang sebelumnya dijadikan data *training* maka dapat dikatakan bahwa sistem tersebut belum terlalu baik untuk merekognisi suara *random* atau diluar data *training*.

Convolutional Recurrent Neural Network (CRNN) merupakan metode yang memiliki keunggulan yaitu pada *convolutional layer* dapat mengekstrak *feature* secara efisien, pada *pooling layer* dapat mengurangi komputasi dan kontrol *overfitting*, Sedangkan pada *recurrent layer* dapat mengekstrak informasi kontekstual dari *feature sequence* yang dihasilkan oleh *convolutional layer* sebelumnya dan dapat menangani *input* yang mempunyai variabel panjang[9] serta dapat memprediksi suara diluar dataset yang digunakan sehingga suara siapa saja dapat direkognisi, dengan demikian metode CRNN perlu digunakan agar meminimalisir kekurangan dan tidak terjadi kendala seperti penelitian sebelumnya.

1.2 Rumusan Masalah

Berdasarkan penelitian sebelumnya, Rekognisi suara dapat diimplementasikan dengan berbagai metode namun terdapat kendala yang menyebabkan akurasi kurang maksimal sehingga diperlukan metode baru yang dapat memberikan akurasi yang baik. Maka didapatkan perumusan masalah yaitu Bagaimana cara mengimplementasikan Model Rekognisi Suara menggunakan metode *Convolutional Recurrent Neural Network (CRNN)*.

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah

1. Metode yang digunakan dalam penelitian ini adalah *Convolutional Recurrent Neural Network (CRNN)*.
2. Model Rokognisi suara yang dibuat akan mengolah suara pembicara sehingga sistem akan mengenali suara pembicara berupa *gender* atau jenis kelamin saja.
3. Menggunakan dataset yang berasal dari *kaggle* berformat csv yang sudah diekstraksi fiturnya.
4. Dataset yang digunakan sudah berlabel *male* (laki-laki) dan *female* (perempuan).

1.4 Tujuan

Adapun Tujuan pada penelitian ini memiliki tujuan sebagai berikut:

1. Merancang sistem Model Rekognisi Suara untuk mengenali identitas suara pembicara berupa *gender*.
2. Memvalidasi kinerja Model Rekognisi Suara menggunakan metode *Convolutional Recurrent Neural Network (CRNN)*.

1.5 Metode Penelitian

Berikut ini adalah metedelogi penelitian yang digunakan dalam penelitian ini:

a. Tahap Pertama (Studi Pustaka/Litelatur)

Pada tahap ini dilakukan untuk mencari dan membacara referensi tentang penelitian yang dilakukan agar dapat menunjang penulisan laporan Tugas Akhir.

b. Tahap Kedua (Persiapan Data)

Pada tahap ini dilakukan dengan cara menganalisis data yang akan digunakan agar data dapat digunakan menggunakan metode CRNN.

c. Tahan Ketiga (Pra Pengolahan Data)

Pada tahap ini data di *pre-processing* terlebih dahulu sebelum masuk ke proses *learning*. Tahapan ini dilakukan dengan cara membagi parameter fitur dan label, mengubah label dari text menjadi numberik dan normalisasi.

d. Tahap Keempat (*Learning*)

Pada tahap ini membahas mengenai proses membangun model *Convolutional Recurrent Neural Network (CRNN)* dalam merekognisi suara.

e. Tahap Kelima (Analisa dan Kesimpulan)

Pada tahap ini menganalisa hasil performa model yang telah dibangun dan menarik kesimpulan berdasarkan penelitian yang telah dilakukan.

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini akan melewati beberapa tahapan sebagai berikut:

BAB I PENDAHULUAN

Pada bab pertama berisikan penjelasan sistematik topik yang diambil meliputi latar belakang, perumusan masalah, tujuan, metedelogi penelitian dan sistematika penulisan

BAB II TINJAUAN PUSTAKA

Pada bab kedua akan menjelaskan dasar teori yang bertujuan untuk menunjang pembahasan dari penelitian ini meliputi dasar teori dari suara, rekognisi suara, *machine learning*, *deep learning*, *Convolutional Neural Network (CNN)*, *Recurrent Neural Network (RNN)*, *Convolutional Recurrent Neural Network* dan validasi performa.

BAB III METEDELOGI PENELITIAN

Pada bab ketiga menjelaskan bagaimana penelitian ini berjalan mulai dari persiapan data, data *pre-processing*, pembagian data *training* dan *testing*, Normalisasi, penerapan metode *Convolutional Recurrent Neural Network (CRNN)*, *model predict* serta validasi performa model.

BAB IV HASIL DAN ANALIS

Pada bab keempat menjelaskan hasil dan analisis penelitian yang telah dilakukan.

BAB V KESIMPULAN

Pada bab kelima memberikan kesimpulan berdasarkan hasil dan analisis penelitian yang telah dilakukan serta saran untuk penulis selanjutnya yang akan dijadikan bahan referensi.

DAFTAR PUSTAKA

- [1] N. Goksel Canbek and M. E. Mutlu, “On the track of Artificial Intelligence: Learning with Intelligent Personal Assistants,” *Int. J. Hum. Sci.*, vol. 13, no. 1, p. 592, 2016, doi: 10.14687/ijhs.v13i1.3549.
- [2] I. Supiandi, “Analisis Digital Assistant versi Cortana , Siri dan Google Now,” pp. 19–23, 1861.
- [3] N. R. Pradipta and H. Kusuma, “Perancangan Perangkat Antarmuka Berbasis Pengenalan Suara pada Purwarupa Mesin Cetak Huruf Braille ITS untuk Operator Tunanetra,” vol. 7, no. 2, 2018.
- [4] E. Noviyantono and J. L. Buliali, “Pengenalan Suara Dengan Ketergantungan Teks Menggunakan Metode Principal Component Analysis Pada Terapan Komputasi Awan PENGENALAN SUARA DENGAN KETERGANTUNGAN TEKS MENGGUNAKAN METODE PRINCIPAL COMPONENT,” no. October, 2015.
- [5] F. Teknik and U. Darma, “FAKULTAS TEKNIK,” 2011.
- [6] Suyanto, *Deep Learning Modernisasi Machine Learning Untuk Bigdata*. Bandung, 2019.
- [7] F. D. Adhinata, D. P. Rakhmadani, and A. J. T. Segara, “Pengenalan Jenis Kelamin Manusia Berbasis Suara Menggunakan MFCC dan GMM,” *Jorunal data Sci. IoT, Mach. Learn. Informatics*, vol. 1, no. 1, pp. 11–12, 2021.
- [8] Y. I. N, A. Ana, and D. Permatasari, “Pengenalan Pembicara untuk Menentukan Gender Menggunakan Metode MFCC dan VQ,” *MIND J.*, vol. 2,

- no. 1, pp. 34–47, 2018, doi: 10.26760/mindjournal.v2i1.34-47.
- [9] H. Wu and S. Prasad, “Convolutional recurrent neural networks for hyperspectral data classification,” *Remote Sens.*, vol. 9, no. 3, 2017, doi: 10.3390/rs9030298.
- [10] Z. Zhang, “Mechanics of human voice production and control,” *J. Acoust. Soc. Am.*, vol. 140, no. 4, pp. 2614–2635, 2016, doi: 10.1121/1.4964509.
- [11] S. Jeska and K. S. Pascha, “About the authors – Acknowledgements,” *Emergent Timber Technol.*, 2014, doi: 10.1515/9783038216162.175.
- [12] R. B. Dubey, “Automatic Speaker Recognition System,” no. 4, 2012.
- [13] P. S. Ke, A. Perintah, E. Vr, E. Vr, F. Teknik, and U. M. Kudus, “Uji validasi suara berbasis pengenalan suara (,” pp. 801–806, 2017.
- [14] Shiva Shankar, J. Raghaveni, P. Rudraraju, and Y. Vineela Sravya, “Classification of gender by voice recognition using machine learning algorithms,” *J. Crit. Rev.*, vol. 7, no. 9, pp. 1217–1229, 2020, doi: 10.31838/jcr.07.09.222.
- [15] T. Wahyono, *Fundamental Of PYTHON FOR MACHINE LEARNING*. Yogyakarta: Penerbit Gava Media, 2018.
- [16] P. Harrington, *Machine Learning in action*. new york: Manning Publication, 2012.
- [17] A. Çelik *et al.*, “No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title,” *J. Mater. Process. Technol.*, vol. 1, no. 1, pp. 1–8, 2018, [Online]. Available: <http://dx.doi.org/10.1016/j.cirp.2016.06.001> [Ahttps://doi.org/10.1016/j.jmatlet.2019.04.024](http://dx.doi.org/10.1016/j.jmatlet.2019.04.024) [Ahttps://doi.org/10.1016/j.mateclett.2016.12.055](https://doi.org/10.1016/j.mateclett.2016.12.055) [Ahttps://doi.org/10.1016/j.ijfatigue.2019.02.006](https://doi.org/10.1016/j.ijfatigue.2019.02.006) [Ahttps://doi.org/10.1016/j.matlet.2019.04.024](https://doi.org/10.1016/j.matlet.2019.04.024)

- 19.127252%0Ahttp://dx.doi.o.
- [18] M. A. Ismail, D. Ma, M. R. Khairullah, F. Budiman, and D. Purwanto, “Sistem Pengenalan Suara untuk Perintah pada Robot Sepak Bola Beroda,” pp. 2–7.
 - [19] E. N. Arrofiqoh and H. Harintaka, “Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi,” *Geomatika*, vol. 24, no. 2, p. 61, 2018, doi: 10.24895/jig.2018.24-2.810.
 - [20] A. Hibatullah and I. Maliki, “Penerapan Metode Convolutional Neural Network Pada Pengenalan Pola Citra Sandi Rumput,” 2019.
 - [21] D. Yogatama, C. Dyer, W. Ling, and P. Blunsom, “Generative and discriminative text classification with recurrent neural networks,” *arXiv*, 2017.
 - [22] M. R. Firmansyah, R. Ilyas, and F. Kasyidi, “Klasifikasi Kalimat Ilmiah Menggunakan Recurrent Neural Network,” *Pros. 11th Ind. Res. Work. Natl. Semin.*, vol. 11, no. 1, pp. 488–495, 2020.
 - [23] V. K. Ayyadevara and V. K. Ayyadevara, “Recurrent Neural Network,” *Pro Mach. Learn. Algorithms*, pp. 217–257, 2018, doi: 10.1007/978-1-4842-3564-5_10.
 - [24] D. Tarkus, S. R. U. A. Sompie, and A. Jacobus, “Implementasi Metode Recurrent Neural Network pada Pengklasifikasian Kualitas Telur Puyuh,” *J. Tek. Inform.*, vol. 15, no. 2, pp. 137–144, 2020, [Online]. Available: <https://ejournal.unsrat.ac.id/index.php/informatika/article/view/29552>.
 - [25] J. B. Heaton, N. G. Polson, and J. H. Witte, “Deep Learning in Finance,” no. February, pp. 1–20, 2016, [Online]. Available: <http://arxiv.org/abs/1602.06561>.
 - [26] N. Networks, “The 2010 International Joint Conference,” no. Ijcn, 2010.
 - [27] A. Le Calvez and D. Cliff, “Deep Learning can Replicate Adaptive Traders in