

Pengenalan Makhraj Huruf dalam Pelafalan Al-Qur'an Menggunakan Convolutional Neural Network (CNN)

Ismail¹,Diah Puspito Wulandari.¹,Reza Fuad Rachmadi.¹

¹Dept. Teknik Komputer Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

Email: ismail.17072@mhs.its.ac.id,diah@te.its.ac.id,fuad@its.ac.id

Abstrak

Dalam kajian Al-Qur'an, unsur paling mendasar yang perlu diketahui adalah huruf-huruf Al-Qur'an itu sendiri, di mana setiap huruf memiliki tempat keluarnya masing-masing saat diucapkan yang disebut sebagai makhraj huruf. Namun, dalam penerapannya, masyarakat Indonesia masih mengalami kesulitan dalam melafalkan makhraj huruf secara benar akibat perbedaan bahasa yang digunakan dalam kehidupan sehari-hari. Kurangnya fasilitas pendukung yang mampu beradaptasi dengan perkembangan zaman juga menjadi kendala mendasar. Oleh karena itu, pembelajaran makhraj Al-Qur'an masih sangat bergantung pada pembelajaran manual melalui para ahli. Berdasarkan permasalahan tersebut, penelitian ini mengusulkan suatu pendekatan yang dapat membantu proses pengenalan makhraj huruf sebagai sarana pendukung bagi umat Muslim dalam mempelajari Al-Qur'an. Metode *deep learning* yang digunakan dalam penerapan pembelajaran ini adalah Convolutional Neural Network (CNN), yang akan mengenali sinyal audio yang sebelumnya telah dikonversi ke dalam bentuk spektrogram. Melalui penelitian ini, diharapkan umat Muslim dapat lebih mudah mempelajari dan mengenali makhraj huruf berdasarkan hasil pengenalan audio.

Keyword: CNN, Deep Learning, Makhraj Huruf, Pengenalan Ucapan

Diterima Redaksi: 15-02-2024 Selesai Revisi: 30-02-2024 Diterbitkan Online: 15-03-2024
DOI: <https://doi.org/10.59378/jcenim.v2i1.42>

I. PENDAHULUAN

Membaca Al-Qur'an tidak sama dengan membaca sebagian besar buku pada umumnya. Al-Qur'an merupakan salah satu kitab suci yang diturunkan oleh Allah sebagai pedoman hidup bagi umat Islam yang harus dibaca dengan baik dan benar. Dalam membaca Al-Qur'an terdapat aturan-aturan yang harus dipatuhi oleh setiap pembaca. Aturan-aturan tersebut diklasifikasikan ke dalam tiga tingkatan bacaan Al-Qur'an berdasarkan kecepatan membaca. Menurut Imam Ibnu Al-Jazary dalam Mandzumah Thayyibatun Nasyr, tiga tingkatan bacaan tersebut meliputi tahqiq atau tingkat bacaan lambat, yaitu membaca Al-Qur'an dengan jelas dalam pengucapan huruf dan harakatnya, teliti dalam panjang dan pendek harakat, membaca secara rinci, serta memperhatikan tanda-tanda waqaf. Tingkatan kedua adalah tadwir atau tingkat bacaan sedang, yaitu membaca Al-Qur'an dengan memanjangkan hukum mad tetapi tidak penuh. Tingkatan ini biasanya digunakan dalam bacaan salat. Tingkatan ketiga adalah hadr atau tingkat bacaan cepat, yaitu membaca Al-Qur'an dengan cepat dan singkat, namun tetap memperhatikan hukum-hukum tajwid [1].

Dalam mempelajari Al-Qur'an, unsur paling dasar yang harus diketahui adalah huruf-huruf Al-Qur'an itu sendiri. Huruf-huruf Al-Qur'an yang dikenal sebagai huruf hijaiyah berjumlah 29 huruf. Setiap huruf memiliki tempat keluarnya masing-masing ketika dilafalkan yang disebut dengan makhraj huruf. Tempat keluarnya huruf hijaiyah disebut makhraj huruf. Dalam kitab Al-Jazariyah, Syekh Abu Khair Syamsuddin Al-Jazary menyatakan bahwa tempat keluarnya huruf hijaiyah berjumlah 17 tempat. Ketujuh belas tempat tersebut kemudian digeneralisasikan oleh para ulama tajwid menjadi lima tempat utama (mawdli') [1].

Perbedaan makhraj ini menimbulkan kesulitan ketika seseorang membaca Al-Qur'an, khususnya bagi masyarakat Indonesia. Hal ini disebabkan adanya perbedaan antara bahasa Indonesia dan bahasa Arab sebagai bahasa yang digunakan dalam Al-Qur'an. Perbedaan tersebut terletak pada lidah penutur bahasa Indonesia yang tidak terbiasa dengan kosakata bahasa Arab. Selain itu, terdapat beberapa huruf dalam bahasa Arab yang pelafalannya tidak dikenal dalam bahasa Indonesia, sehingga menyebabkan kesulitan

bagi umat Islam di Indonesia dalam melafalkan makhraj huruf secara benar. Permasalahan tersebut menimbulkan banyak ambiguitas dalam pelafalan setiap huruf Arab atau yang disebut huruf hijaiyah. Banyak fonem yang terdengar samar bahkan tampak sama antahuruf, yang dapat mengubah makna atau isi ayat-ayat Al-Qur'an itu sendiri.

Saat ini, faktor tempat dan waktu juga menjadi hal yang sangat krusial. Selain belajar secara langsung dengan seorang pembimbing profesional, diperlukan pula fasilitas atau fitur yang sangat mendukung dan fleksibel sesuai dengan perkembangan zaman. Fitur tersebut diharapkan dapat digunakan sebagai sarana pendukung dalam pembelajaran Al-Qur'an tanpa harus mengkhawatirkan keterbatasan tempat dan waktu.

Penulis juga telah mempelajari beberapa literatur yang membahas topik serupa, baik dari segi objek penelitian maupun metode yang digunakan. Salah satunya adalah penelitian berjudul "Makhraj Recognition for Al-Quran Recitation using MFCC" [2]. Penelitian tersebut dilakukan oleh N. W. Arshad, S. M. Sukri, L. N. Muhammad, H. Ahmad, R. Hamid, F. Naim, dan N. Z. A. Naharuddin pada tahun 2013. Dalam penelitian tersebut, data diambil dari 10 orang yang terdiri dari 5 laki-laki dan 5 perempuan. Setiap orang mengucapkan kombinasi huruf hijaiyah yang sama sebanyak 10 kali sehingga diperoleh lebih dari 2000 sampel, dengan 20 sampel digunakan sebagai data uji untuk keperluan pelatihan. Selain itu, terdapat pula penelitian berjudul "Pengenalan Huruf Hijayyah Berbasis Pengolahan Sinyal Suara Dengan Metode Mel Cepstrum Frequency Cepstrum Coefficient (MFCC)" [3]. Dalam penelitian tersebut, digunakan 600 data uji dan 900 data latih. Data disimpan dalam format .wav dengan durasi setiap huruf selama 1 detik. Proses pengambilan keputusan dilakukan menggunakan metode logika if-then-else pada perangkat lunak Matlab. Nilai cepstrum/channel dari data uji yang berada di luar rentang nilai cepstrum/channel model akan dikategorikan sebagai data pengucapan yang tidak benar pada GUI.

II. TINJAUAN PUSTAKA

A. Makhraj Huruf

Dalam bahasa Arab, kata "makhraj" berasal dari fi'il madhi yang berarti "keluar". Kemudian dari bentuk fi'il madhi tersebut berubah menjadi isim makan (kata yang menunjukkan keterangan tempat), sehingga berubah menjadi "makhraj" yang berarti "tempat keluar". Dengan demikian, makhraj huruf adalah tempat keluarnya huruf Al-Qur'an ketika dibaca atau dilafalkan. Huruf-huruf Al-Qur'an atau yang juga disebut huruf hijaiyah berjumlah 29 huruf, di mana setiap huruf memiliki tempat keluarnya masing-masing [4].

Hal ini sangat penting karena apabila pelafalan suatu huruf tidak keluar dari tempat aslinya, maka dapat terjadi perubahan makna ayat yang dibaca serta menimbulkan kerancuan bunyi antahuruf. Pembagian makhraj huruf sendiri terbagi menjadi dua, yaitu secara khusus terdapat 17 makhraj, dan secara umum terdapat 5 makhraj, yaitu Al-Jauf (rongga mulut), Al-Halq (tenggorokan), Al-Lisan (lidah), Ash-Syafatain (kedua bibir), dan yang terakhir adalah Al-Khaisyum (pangkal hidung).

Al-Halq	Al-Jauf	Al-Khaisyum	Asy-syafatain	Al-Lisan
أ	ا	ن	ف	ق
هـ	ي	م	و	ك
ع	و		ب	ل
ح			م	ن
ج				ت
د				ث
ر				ج
				د
				ظ
				ي

Gambar 1: Pembagian Huruf Berdasarkan Kelas Makhraj

B. Spektrogram

Spektrogram digunakan untuk menggambarkan spektrum frekuensi suatu sinyal yang berubah terhadap waktu. Spektrogram juga dikenal sebagai sonograf atau voicegram ketika diterapkan pada sinyal audio. Spektrogram umumnya direpresentasikan dalam bentuk heat-map, yaitu visualisasi intensitas sinyal yang ditunjukkan melalui variasi warna atau tingkat kecerahan tertentu [5]. Apabila amplitudo suatu sinyal rendah, maka ditunjukkan dengan warna spektrum biru yang semakin gelap. Sebaliknya, semakin tinggi amplitudo sinyal, maka warna spektrum akan semakin terang.

C. Convolution Neural Network (CNN)

Arsitektur CNN memiliki kemiripan dengan cara neuron saling terhubung dalam otak manusia, yang terinspirasi dari jaringan saraf biologis, khususnya korteks visual yang terletak di korteks serebral. Setiap neuron merespons rangsangan hanya pada area terbatas dari bidang pandang yang disebut sebagai receptive field. Proses prapemrosesan yang dibutuhkan pada CNN jauh lebih sedikit dibandingkan dengan algoritma klasifikasi lainnya. Pada metode sebelumnya, filter dirancang secara manual dan memerlukan pelatihan yang memadai, sedangkan CNN memiliki kemampuan untuk mempelajari filter atau fitur tersebut secara otomatis [6].

Peran CNN adalah mereduksi citra menjadi bentuk yang lebih mudah diproses tanpa menghilangkan fitur-fitur penting yang dibutuhkan untuk memperoleh prediksi yang baik. Dalam CNN, agar hasil keluaran citra tidak kehilangan terlalu banyak dimensi atau dengan kata lain hasil keluaran tidak jauh berbeda dari masukan, maka diperlukan penambahan padding. Hal ini sangat penting dalam perancangan arsitektur yang tidak hanya mampu mempelajari fitur dengan baik, tetapi juga dapat diskalakan untuk kumpulan data yang sangat besar.

D. Graphical User Interface (GUI)

Graphical User Interface atau yang lebih dikenal sebagai GUI merupakan sistem komponen visual interaktif pada perangkat lunak komputer. GUI menampilkan objek-objek yang menyampaikan informasi serta merepresentasikan tindakan yang dapat dilakukan oleh pengguna. Objek-objek tersebut dapat berubah warna, ukuran, atau tingkat visibilitas ketika pengguna berinteraksi dengannya.

GUI mencakup berbagai objek seperti ikon, kursor, dan tombol. Elemen grafis ini terkadang dilengkapi dengan suara atau efek visual seperti transparansi dan bayangan. Dengan menggunakan objek-objek tersebut, pengguna dapat mengoperasikan komputer tanpa harus mengetahui perintah-perintah secara langsung [7].

III. PERANCANGAN DAN IMPLEMENTASI

Tujuan dari penelitian ini adalah membangun sebuah model pengenalan makhraj huruf pada pembacaan Al-Qur'an menggunakan Convolutional Neural Network (CNN). Model yang dibangun mencakup pengenalan jenis-jenis makhraj huruf serta hasil kedekatan pelafalan setiap huruf dari jenis makhraj yang berbeda sebagaimana ditunjukkan pada Gambar 1. Gambar 2 menyajikan cara kerja sistem pada penelitian ini.

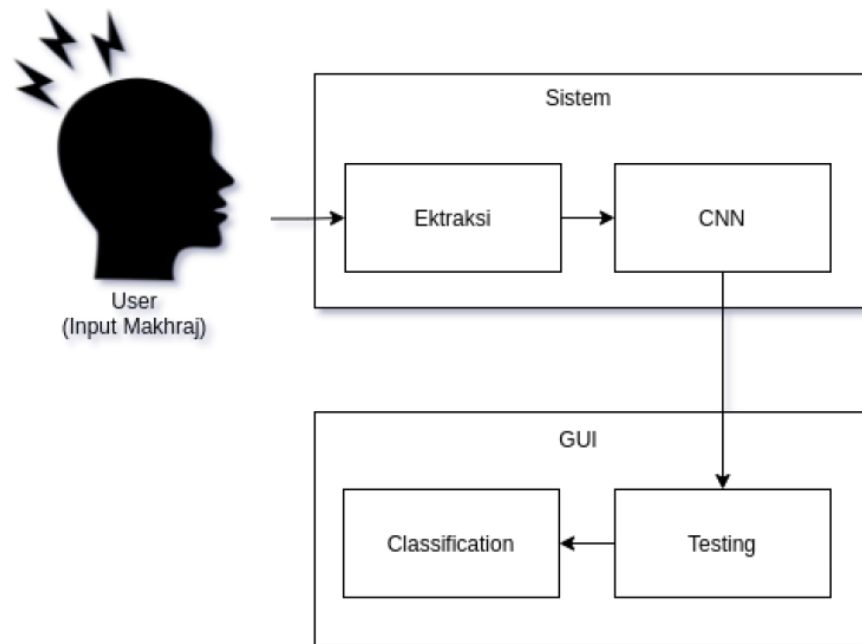
Agar sistem yang dibangun dapat mencapai tujuan tersebut, diperlukan perancangan sistem dalam bentuk alur implementasi. Berdasarkan kebutuhan dan proses yang dilakukan, perancangan ini dapat dibagi ke dalam beberapa subsistem sebagaimana ditunjukkan pada Gambar 3.

A. Masukan Audio

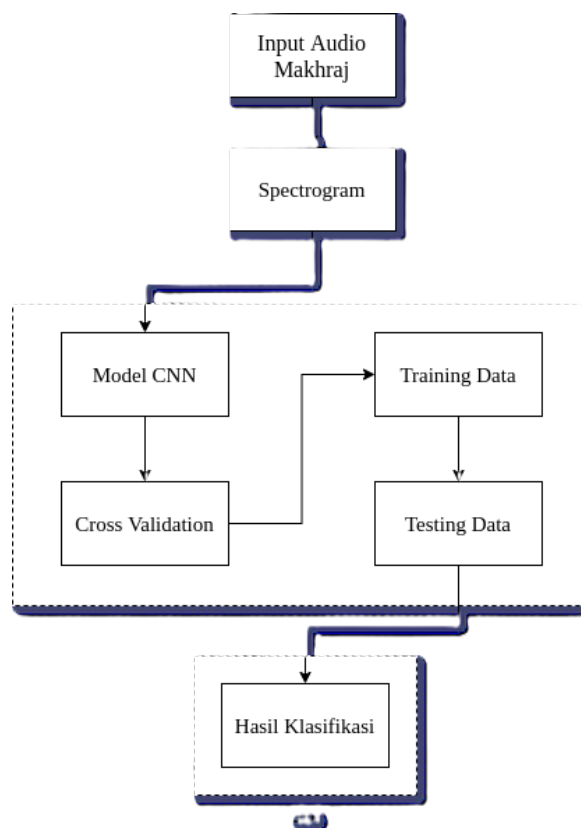
Pada tahap ini, data masukan berupa audio yang diperoleh dari tiga penutur asli bahasa Arab. Data berupa audio pengucapan huruf hijaiyah yang dibagi ke dalam lima (5) kelas sesuai dengan jenis makhraj, di mana setiap kelas terdiri dari 250 data. Audio memiliki format atau ekstensi .wav dengan durasi 1,20 detik untuk setiap data. Data audio yang telah dipotong ini memiliki laju pengambilan sampel sebesar 22.050 kHz. Rincian jumlah dataset dapat dilihat pada Tabel 1.

B. Spectrogram

Agar dapat diproses oleh CNN, data masukan berupa audio diekstraksi menjadi spektrogram, karena CNN bekerja pada citra dua dimensi.



Gambar 2: Desain Sistem

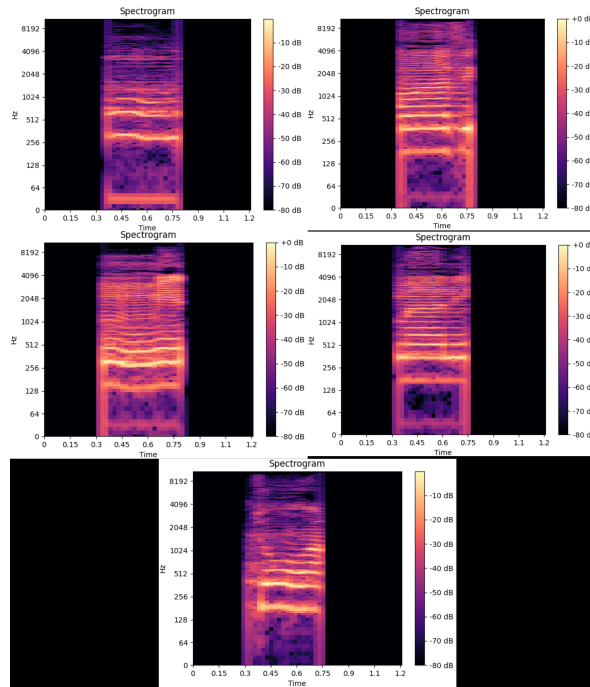


Gambar 3: Blok Diagram penelitian

Audio diekstraksi menggunakan metode Short Time Fourier Transform (STFT). Bentuk citra yang dihasilkan dari ekstraksi STFT berupa spektrogram yang diproses melalui tahapan frame blocking, windowing, dan Fast Fourier Transform. Gambar 4 merupakan visualisasi transformasi sinyal audio menjadi spektrogram.

Tabel 1: Number of Datasets on Each Makhraj Class

No.	Makhraj Class	Number of Data
1	Al-Halq	250
2	Al-Jauf	250
3	Al-Khaisyum	250
4	Asy-Syafatain	250
5	Al-Lisan	250
	Total	1250



Gambar 4: Spectrogram Samples of Each Makhraj Class

C. Convolution Neural Network

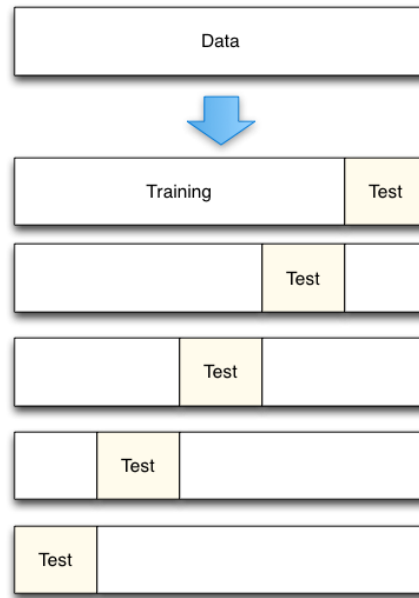
Pada penelitian ini, metode klasifikasi yang digunakan adalah Convolutional Neural Network (CNN). Setelah data spektrogram diperoleh dari proses sebelumnya, data tersebut diproses melalui CNN untuk dilakukan pelatihan dan pengujian. Data uji merupakan data yang akan dikenali berdasarkan data latih [8].

1) Dataset Splitting: Distribusi dataset dibagi ke dalam dua (2) subset, yaitu data latih dan data uji. Salah satu teknik pembagian data yang digunakan adalah K-Fold Cross Validation. Pada metode k-fold cross validation, dataset dibagi menjadi k-subset, di mana satu bagian digunakan sebagai data uji dan bagian lainnya sebagai data latih. Untuk setiap fold, model akan dilatih sebanyak k kali dengan pembagian data latih dan data uji yang berbeda secara acak. Visualisasi lebih rinci mengenai k-fold cross validation ditunjukkan pada Gambar 5.

Jumlah fold yang digunakan adalah 5 fold ($k = 5$), sehingga dataset dibagi menjadi 5 bagian dengan ukuran masing-masing 20%. Dari pembagian tersebut diperoleh 1000 data latih dan 250 data uji yang akan dibagi secara acak pada setiap fold dan setiap iterasi. Rincian pembagian data dapat dilihat pada Tabel 2.

2) Build a CNN Model: Setelah dataset dibagi, dibuat model Convolutional Neural Network untuk mengklasifikasikan lima (5) jenis makhraj. Model CNN yang digunakan memiliki 13 lapisan dengan ukuran citra masukan sebesar $87 \times 13 \times 1$. Arsitektur CNN yang digunakan ditunjukkan pada Tabel 3.

Pada arsitektur tersebut, data masukan dua dimensi diproses melalui lapisan Feature Learning dan kemudian melalui Flatten. Hasil Flatten digunakan sebagai masukan pada tiga lapisan Fully Connected, sehingga dari proses tersebut dihasilkan persentase klasifikasi lima (5) kelas makhraj.



Gambar 5: Illustration of K-Fold Cross Validation [9]

Tabel 2: Distribution of Dataset

Dataset	Training	Testing
Al-Halq	200	50
Al-Jauf	200	50
Al-Khaisyum	200	50
Asy-Syafatain	200	50
Al-Lisan	200	50
Total	1000	250

Tabel 3: Applied CNN Architecture

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 85, 11, 64)	640
batch normalization	(None, 85, 11, 64)	256
max pooling2d	(None, 43, 6, 64)	0
conv2d 1 (Conv2D)	(None, 41, 4, 32)	18464
batch normalization 1	(None, 41, 4, 32)	128
max pooling2d 1	(None, 21, 2, 32)	0
conv2d 2 (Conv2D)	(None, 20, 1, 32)	4128
batch normalization 2	(None, 20, 1, 32)	128
max pooling2d 2	(None, 10, 1, 32)	0
Flatten	(None, 320)	0
Dense	(None, 64)	20544
dropout (Dropout)	(None, 64)	0
dense1 (Dense)	(None, 5)	325

Total parameter: 44,613. Parameter yang dapat dilatih: 44,357. Parameter yang tidak dapat dilatih: 256.

3) Training: Untuk melatih model jaringan saraf, metrik model diperbarui melalui beberapa perhitungan pada data. Ketika ukuran data besar, proses pelatihan membutuhkan waktu yang lama dan konsumsi daya yang tinggi. Beberapa hyperparameter yang digunakan sebelum pelatihan adalah sebagai berikut:

a) *Batch Size*

Batch size adalah jumlah sampel data pada setiap iterasi pelatihan model. Jika batch size yang digunakan adalah n , maka model menghitung rata-rata kesalahan dan memperbarui bobot setiap n sampel data [10]. Pada penelitian ini digunakan batch size sebesar 64.

b) *Epoch*

Epoch menyatakan berapa kali algoritma memproses seluruh dataset. Setiap kali seluruh sampel dalam dataset telah diproses, maka satu epoch telah selesai [10]. Pada penelitian ini digunakan sebanyak 50 epoch.

c) *Iteration*

Jumlah iterasi yang dibutuhkan oleh batch untuk menyelesaikan satu epoch disebut iteration. Dengan demikian dapat disimpulkan bahwa $\text{batch size} \times \text{jumlah iterasi} = \text{epoch}$ [10].

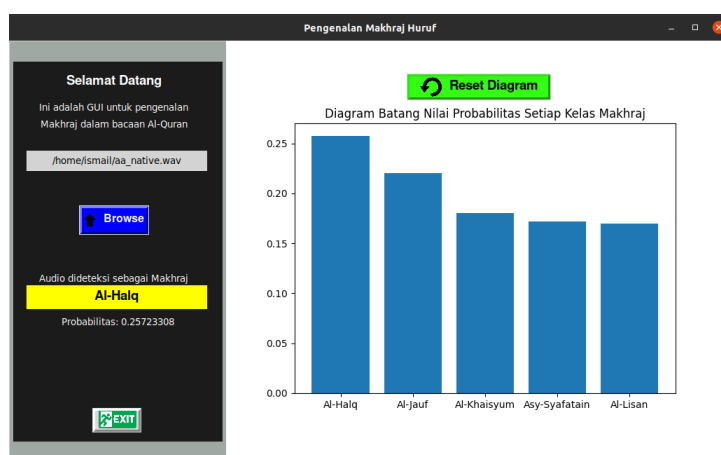
d) *Learning Rate*

Learning rate merupakan salah satu hyperparameter pelatihan pada jaringan saraf untuk menghitung nilai koreksi bobot selama proses pelatihan, dengan rentang nilai antara nol (0) hingga satu (1) [11]. Semakin besar nilai learning rate, bobot yang dihasilkan cenderung kurang optimal, sedangkan semakin kecil learning rate, bobot yang dihasilkan semakin optimal dan akurat, namun membutuhkan waktu pelatihan yang lebih lama.

4) Testing: Untuk mengetahui efektivitas metode CNN yang digunakan, dilakukan tahap pengujian. Pengujian dilakukan menggunakan data uji dari dataset untuk mengetahui kinerja model yang telah dibuat. Selanjutnya, model diuji menggunakan GUI, di mana pengguna dapat langsung memasukkan data berupa potongan audio makhraj untuk mengetahui nilai kedekatan terhadap lima jenis makhraj.

D. Graphic User Interface (GUI)

Keluaran pada tahap ini merupakan hasil keputusan dari data uji terhadap data latih. Keluaran tersebut ditampilkan melalui GUI sebagaimana ditunjukkan pada Gambar 6. Dari keluaran ini diharapkan dapat ditampilkan nilai kedekatan (akurasi) klasifikasi makhraj dari data audio yang dimasukkan oleh pengguna.



Gambar 6: Default Display of GUI

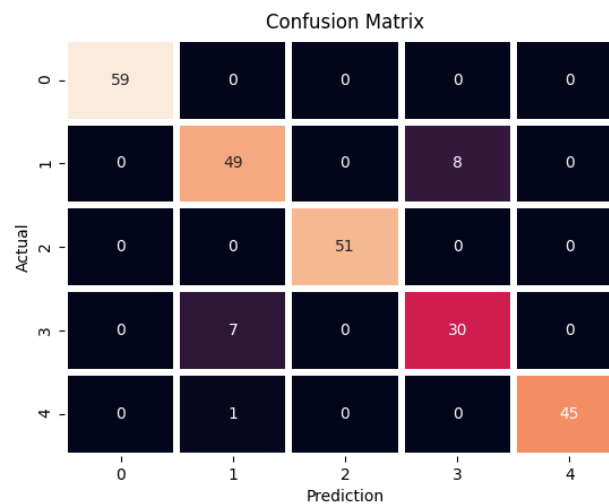
IV. PENGUJIAN DAN HASIL

A. Hasil Pelatihan dan Pengujian Berdasarkan K-Fold Cross Validation

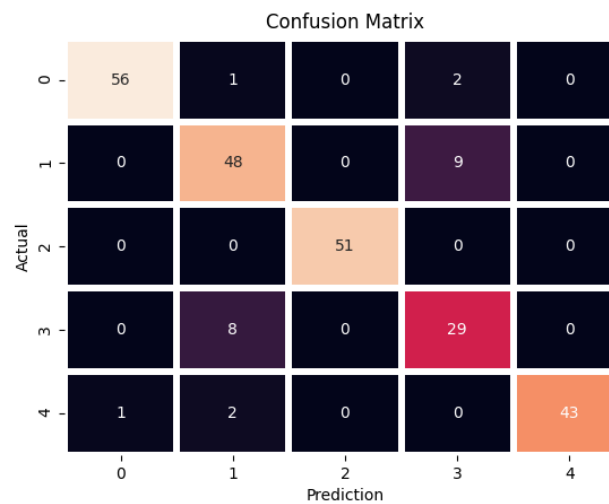
Pada proses ini digunakan sebanyak 5 fold, yang berarti dataset dibagi menjadi 5 bagian dengan ukuran yang sama. Dari pembagian tersebut kemudian dilakukan pengacakan untuk menentukan bagian yang

digunakan sebagai data latih dan data uji. Pada fold pertama, bagian ke- k pertama digunakan sebagai data uji dan bagian lainnya digunakan sebagai data latih. Proses ini berlanjut hingga fold kelima dengan menggunakan epoch sebanyak 50 dan batch size sebesar 64. Dari keseluruhan proses tersebut, model dilatih sebanyak 5 iterasi.

Dari proses tersebut diperoleh hasil pengujian berupa confusion matrix yang ditunjukkan pada Gambar 7, 8, 9, 10, dan 11, serta nilai F-Measure yang disajikan pada Tabel 4.



Gambar 7: Hasil Pengujian pada Fold 1

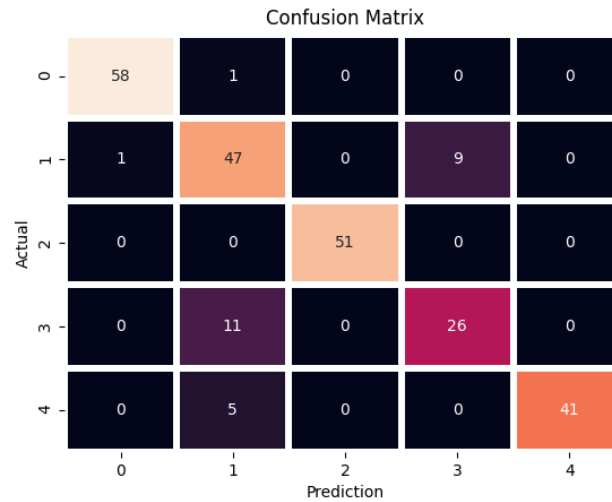


Gambar 8: Hasil Pengujian pada Fold 2

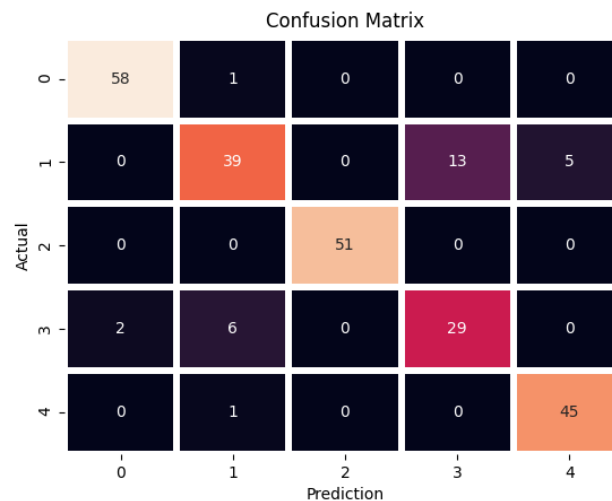
Keterangan:

- 0 : Al-Halq - 1 : Al-Jauf - 2 : Al-Khaisyum - 3 : Asy-Syafatain - 4 : Al-Lisan

Berdasarkan hasil pengujian menggunakan data yang terdapat di dalam dataset, diperoleh hasil pengujian yang relatif tinggi dengan nilai rata-rata sebesar 0.90. Pada hasil pengujian setiap kelas di masing-masing fold, diperoleh variasi data kesalahan. Data kesalahan lebih sering terjadi pada dua kelas makhraj, yaitu Al-Jauf dan Asy-Syafatain. Hal ini disebabkan oleh distribusi huruf pada kedua kelas tersebut yang juga termasuk ke dalam kelas makhraj lainnya. Huruf Alif/Hamzah, misalnya, selain termasuk ke dalam kelas Al-Jauf juga masuk ke dalam kelas Al-Halq, sedangkan huruf Wau pada makhraj Asy-Syafatain juga termasuk ke dalam kelas makhraj Al-Jauf (lihat Gambar 1). Selain itu, terdapat satu kelas yang nilai confusion matrix-nya stabil dari fold pertama hingga fold terakhir, yaitu kelas Al-Khaisyum. Hal ini disebabkan karena makhraj Al-Khaisyum merupakan satu-satunya makhraj



Gambar 9: Hasil Pengujian pada Fold 3



Gambar 10: Hasil Pengujian pada Fold 4

Tabel 4: Rata-Rata Skor K-Fold Cross Validation

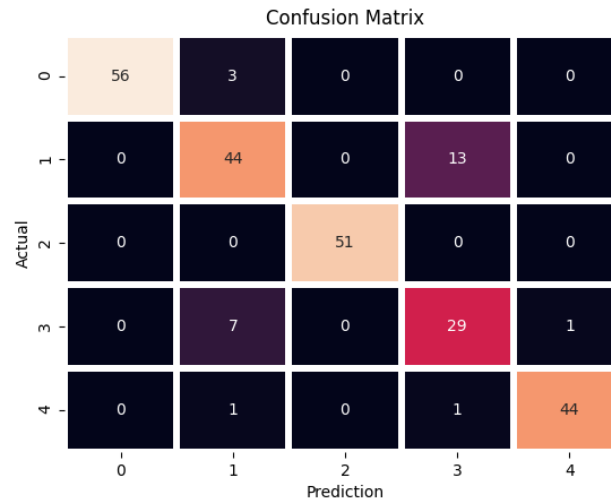
Fold	F-1
1	0.89
2	0.89
3	0.90
4	0.94
5	0.91
Average	0.90

yang bersifat dengung, sehingga bentuk citra yang dihasilkan sangat khas dibandingkan citra dari kelas makhraj lainnya.

B. Hasil Pengujian di Luar Dataset

Pengujian ini dilakukan pada tiga kategori data, yaitu:

1. Data dari penutur asli (*native speaker*)

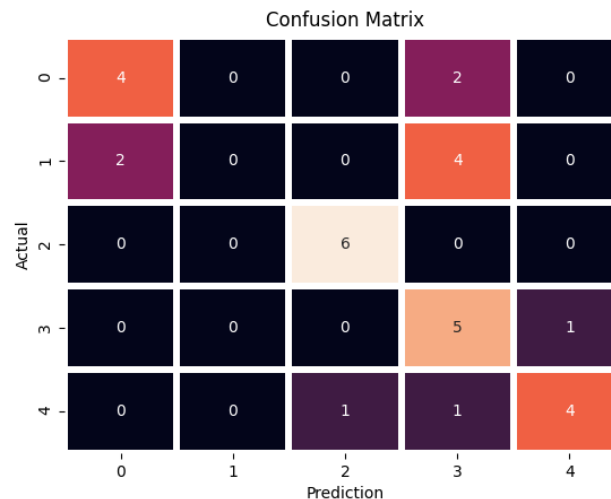


Gambar 11: Hasil Pengujian pada Fold 5

2. Data dari qari *non-native speaker*
3. Data dari non-qari *non-native speaker*

Jumlah data yang digunakan pada masing-masing kategori adalah sebanyak 30 data.

1) Pengujian Data Penutur Asli: Pada tahap ini digunakan data audio yang berasal dari rekaman qari penutur asli bahasa Arab. Pengujian ini bertujuan untuk menguji kinerja model dalam mengenali masukan yang berbeda dari dataset qari penutur asli bahasa Arab yang digunakan pada pelatihan. Hasil pengujian dapat dilihat pada Gambar 12.

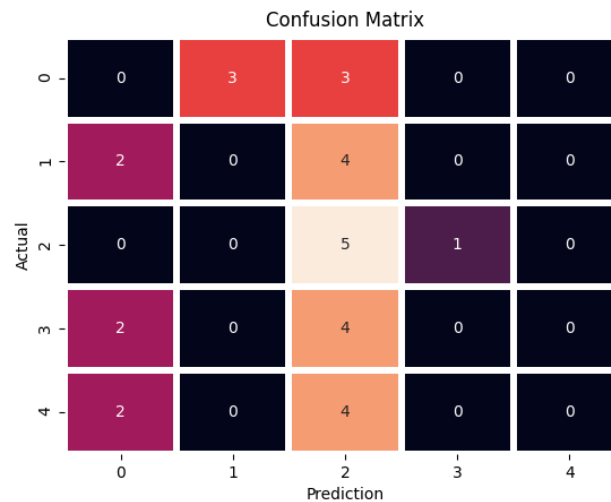


Gambar 12: Confusion Matrix Pengujian Data Penutur Asli

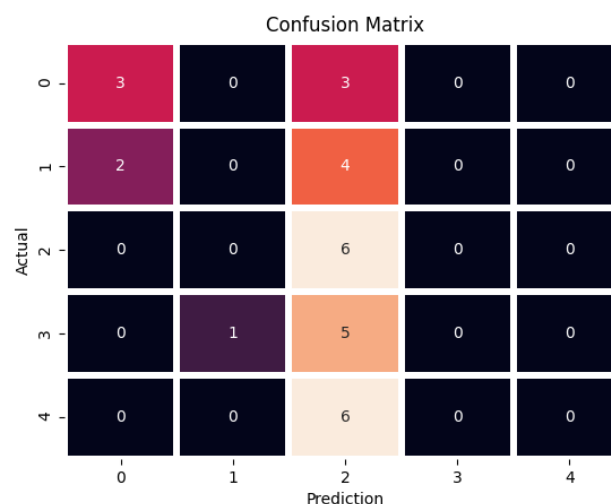
2) Pengujian Data Qari Non-Native Speaker: Pengujian ini bertujuan untuk mengukur kinerja model dalam mengenali masukan yang berbeda dari dataset khusus penutur asli bahasa Arab. Sampel data audio yang digunakan berasal dari penutur asli Indonesia yang merupakan seorang qari (orang yang mahir membaca Al-Qur'an). Hasil pengujian dapat dilihat pada Gambar 13.

3) Pengujian Data Non-Qari Non-Native Speaker: Sampel data audio yang digunakan berasal dari penutur asli Indonesia yang bukan merupakan seorang qari. Hasil pengujian dapat dilihat pada Gambar 14.

4) Skor Pengujian Data di Luar Dataset: Pada hasil pengujian menggunakan data di luar dataset, diperoleh nilai pengujian yang lebih rendah dibandingkan dengan pengujian menggunakan data



Gambar 13: Confusion Matrix Pengujian Data Qari Non-Native Speaker



Gambar 14: Confusion Matrix Pengujian Data Non-Qari Non-Native Speaker

dalam dataset yang memiliki nilai rata-rata sebesar 0.90. Pada hasil pengujian setiap kelas diperoleh nilai kesalahan yang cukup besar. Kesalahan lebih banyak terjadi pada dua kelas terakhir, yaitu Asy-Syafatain dan Al-Lisan. Hal ini disebabkan karena distribusi huruf pada kelas Asy-Syafatain juga termasuk ke dalam kelas makhraj lain, yaitu Al-Jauf (lihat Gambar 1), sehingga menimbulkan ambiguitas dalam proses pengenalan.

Selain itu, terjadinya kesalahan dalam proses pengenalan juga disebabkan oleh faktor distribusi huruf pada setiap kelas makhraj yang tidak merata. Kelas Al-Lisan memiliki jumlah huruf terbanyak, yaitu 18 huruf dari total 29 huruf hijaiyah, sehingga pada saat pengumpulan dataset variasi huruf yang diperoleh relatif lebih sedikit dibandingkan kelas lainnya. Hal ini menyebabkan tingkat pengenalan model terhadap setiap huruf tidak merata.

Jika dilihat pada Tabel 5, nilai skor pengenalan model terhadap masing-masing kategori menunjukkan pola tertentu. Kategori penutur asli memiliki nilai tertinggi dibandingkan dua kategori lainnya, yaitu sebesar 0.63. Hal ini disebabkan karena jenis data uji yang digunakan berada pada kategori yang sama dengan dataset yang digunakan dalam penelitian ini, yaitu penutur asli bahasa Arab. Posisi kedua adalah kategori qari non-native speaker dengan nilai 0.57, karena pelafalan seorang qari mendekati pelafalan penutur asli. Sementara itu, kategori non-qari non-native speaker memiliki nilai terendah, yaitu 0.30. Hal ini disebabkan karena penutur non-native yang bukan qari memiliki cara pelafalan yang berbeda

dan tidak mendekati penutur asli, sehingga model tidak mampu mengenali jenis audio ini dengan baik.

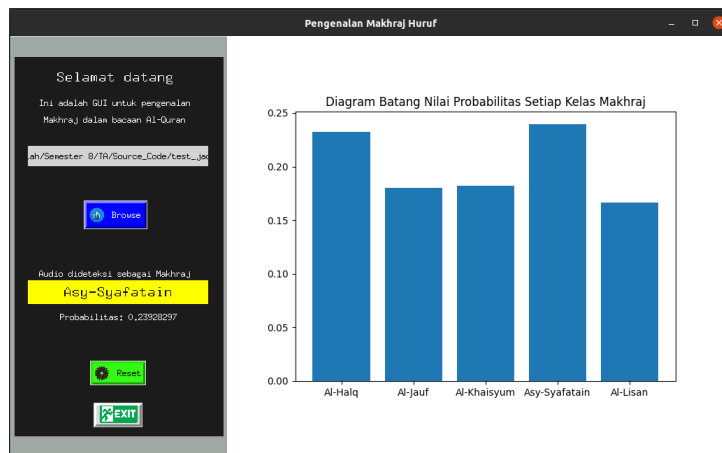
Tabel 5: Skor Pengujian Data di Luar Dataset

Makhraj Class	Native	Qari	Non Qari
Al-Halq	0.67	0.60	0.55
Al-Jauf	0.00	0.00	0.00
Al-Khaisyum	0.92	0.91	0.40
Asy-Syafatain	0.59	0.50	0.00
Al-Lisan	0.73	0.67	0.00
Score	0.63	0.57	0.30

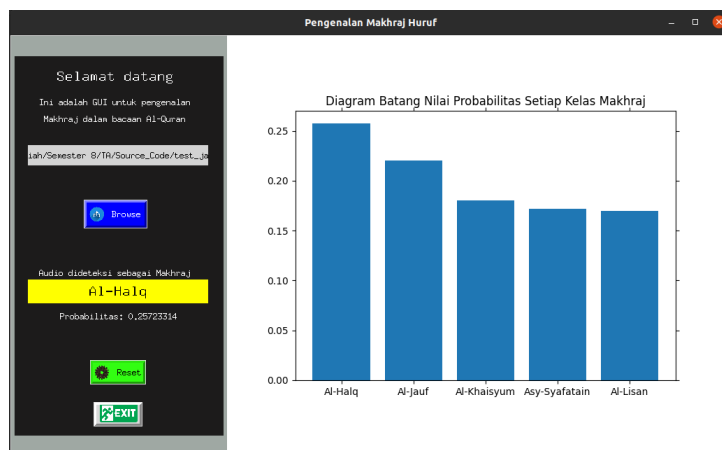
C. Pengujian Graphical User Interface

Pengujian ini dilakukan terhadap GUI yang telah dibangun. Pengujian GUI ini menekankan pada fungsi GUI serta kemudahan pengguna dalam mengetahui makhraj dari data yang dimasukkan dan mengamati nilai kedekatannya terhadap kelas makhraj lainnya.

1) **Pengujian Data Penutur Asli:** Model berhasil mengenali suku kata Faa sebagai makhraj Asy-Syafatain dengan nilai probabilitas sebesar 0.24 (Gambar 15), serta suku kata 'Aa sebagai makhraj Al-Halq dengan nilai probabilitas sebesar 0.26 (Gambar 16).

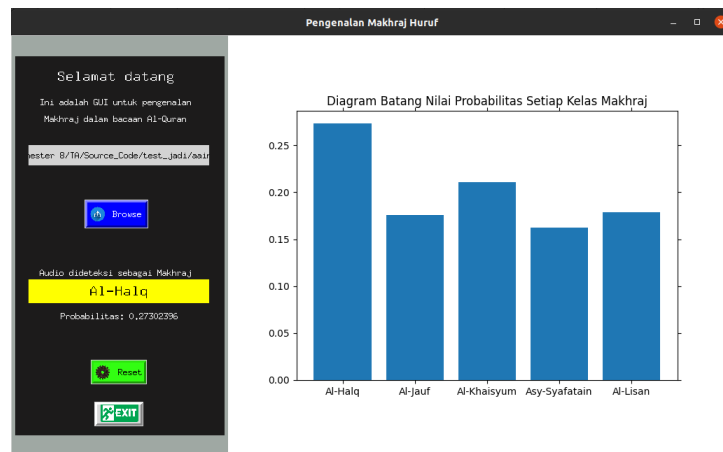


Gambar 15: Tampilan GUI untuk Pengenalan Suku Kata Faa

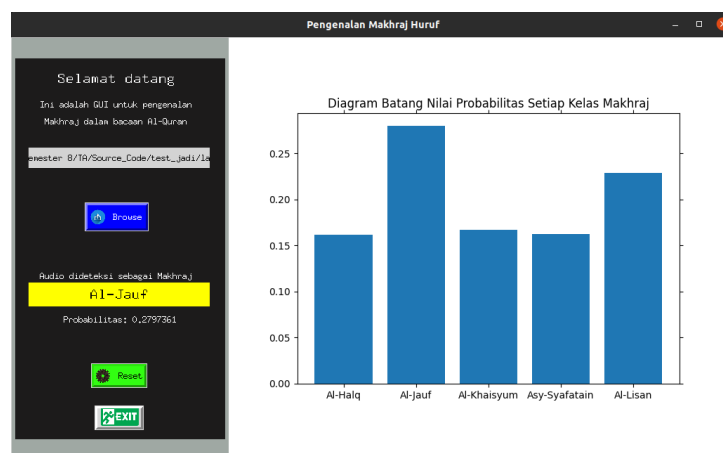


Gambar 16: Tampilan GUI untuk Pengenalan Suku Kata 'Aa

2) Pengujian Data Qari Non-Native Speaker: Model berhasil mengenali suku kata 'Uu sebagai makhraj Al-Halq dengan nilai probabilitas sebesar 0.27 (Gambar 17). Namun, untuk suku kata Laa yang seharusnya termasuk ke dalam kelas makhraj Al-Lisan, model mengenalinya sebagai Al-Jauf dengan nilai probabilitas sebesar 0.28 (Gambar 18). Hal ini disebabkan oleh waktu pemotongan audio yang terlalu singkat sehingga elemen vokal yang diambil hanya sebagian.



Gambar 17: Tampilan GUI untuk Pengenalan Suku Kata 'Uu



Gambar 18: Tampilan GUI untuk Pengenalan Suku Kata Laa

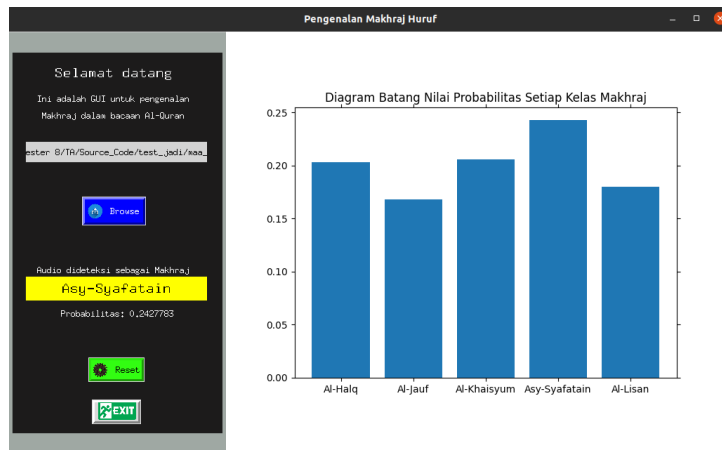
3) Pengujian Data Non-Qari Non-Native Speaker: Model berhasil mengenali suku kata Maa sebagai makhraj Asy-Syafatain dengan nilai probabilitas sebesar 0.24 (Gambar 19), serta suku kata 'Aa sebagai makhraj Al-Halq dengan nilai probabilitas sebesar 0.27 (Gambar 20).

V. KESIMPULAN

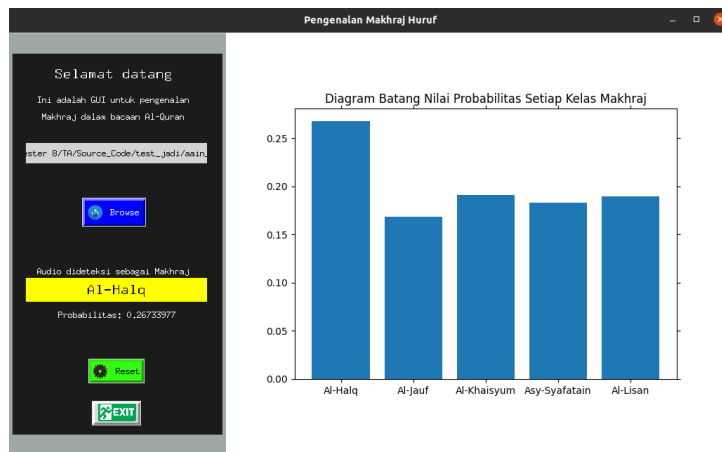
Berdasarkan hasil pengujian yang telah dilakukan, dapat ditarik beberapa kesimpulan dan saran yang diharapkan dapat menjadi bahan pertimbangan dalam pengembangan penelitian ini.

A. Kesimpulan

1. Pada pembagian dataset menggunakan metode cross validation, hasil pengujian tertinggi diperoleh pada fold keempat dengan nilai 0.94. Hal ini menunjukkan bahwa pada fold keempat terdapat distribusi dataset yang paling optimal.
2. Pengujian terhadap data yang terdapat di dalam dataset menghasilkan nilai rata-rata sebesar 0.90, sedangkan pengujian menggunakan data di luar dataset menghasilkan nilai yang jauh lebih rendah.



Gambar 19: Tampilan GUI untuk Pengenalan Suku Kata Maa



Gambar 20: Tampilan GUI untuk Pengenalan Suku Kata 'Aa

3. Data audio yang berasal dari jenis pelafalan yang berbeda memiliki pengaruh terhadap kemampuan model dalam melakukan pengenalan. Penutur asli (native speaker) menghasilkan nilai pengujian yang lebih tinggi dibandingkan penutur non-native.

B. Saran

1. Distribusi dataset sebaiknya dibagi secara merata berdasarkan huruf secara individual agar model memiliki referensi fonetik yang sama kuat dalam melakukan pengenalan setiap huruf.
2. Perlu dilakukan penambahan jumlah dan variasi dataset yang digunakan dalam proses pelatihan dan pengujian, seperti meningkatkan jumlah data dari penutur asli bahasa Arab.

Daftar Pustaka

- [1] M. A. Amir, *ILMU TAJWID PRAKTIS*. Pustaka Baitul Hikmah Harun Ar-Rasyid, 2019.
- [2] N. W. Arshad, S. M. Sukri, L. N. Muhammad, H. Ahmad, R. Hamid, F. Naim, and N. Z. A. Naharuddin, "Makhraj recognition for al-quran recitation using mfcc," *International Journal of Intelligent Information Processing (IJIIP)*, vol. 4, no. 2, 2013.
- [3] H. Bethaningtyas, M. K. A. SP, and S. Suprayogi, "Pengenalan huruf hijayyah berbasis pengolahan sinyal suara dengan metode mel cepstrum frequency cepstrum coefficient (mfcc)," *Majalah Ilmiah Momentum*, vol. 13, no. 2, 2018.

- [4] Meeftha, “Bab tentang makhraj huruf hijaiyah (tempat keluarnya huruf),” <https://tajwid.web.id/bab-tentang-makhraj-huruf-hijaiyah-tempat-keluarnya-huruf>, 2017, accessed on 2020-11-26.
- [5] Wikipedia, “Spectrogram,” <https://en.wikipedia.org/wiki/Spectrogram>, 2020, accessed on 2020-12-09.
- [6] S. Saha, “A comprehensive guide to convolutional neural networks – the eli5 way,” *Towards Data Science*, vol. 15, 2018, accessed on 2021-04-26.
- [7] C. Hope, “What is a gui graphical user interface,” <https://www.computerhope.com/jargon/g/gui.htm>, 2001, accessed on 2021-05-27.
- [8] R. Mulyani, “Pengenalan suara pada sistem notulen rapat menggunakan convolutional neural network (cnn),” Ph.D. dissertation, Universitas Pendidikan Indonesia, 2016.
- [9] R. Ng, “Crossvalidation machine learning deep learning and computer vision,” <https://www.ritchieng.com/machine-learning-cross-validation/>, 2017, accessed on 2017-06-15.
- [10] Androidkt, “What is batch size steps iteration and epoch in the neural network knowledge transfer,” <https://androidkt.com/batch-size-step-iteration-epoch-neural-network/>, 2020, accessed on 2021-05-27.
- [11] R. D. Nurfitra and G. Ariyanto, “Implementasi deep learning berbasis tensorflow untuk pengenalan sidik jari,” *Emitor: Jurnal Teknik Elektro*, vol. 18, no. 1, pp. 22–27, 2018.