

## MAKHRAJ 'AIN PRONUNCIATION ERROR DETECTION USING MEL FREQUENCY CEPSTRAL COEFFICIENT AND MODIFIED VGG-16

Ibnu Kasyful Haq<sup>\*1</sup>, Agi Prasetiadi<sup>2</sup>

<sup>1,2</sup>Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto, Indonesia  
Email: <sup>1</sup>[ibnukasyfulhaq@gmail.com](mailto:ibnukasyfulhaq@gmail.com), <sup>2</sup>[agi@ittelkom-pwt.ac.id](mailto:agi@ittelkom-pwt.ac.id)

(Naskah masuk: 26 Juni 2022, Revisi: 31 Juli 2022, Diterbitkan: 10 Februari 2023)

### Abstract

Based on research conducted by the Institute of Qur'anic Sciences (IIQ) as many as 65% of Muslims in Indonesia are illiterate in the Qur'an. In previous studies, research was conducted on the detection of Arabic word pronunciation errors against non-natives using the Mel Frequency Cepstral Coefficient (MFCC) and Support Vector Machine (SVM) methods with a test result of 54.6%. Due to the low accuracy results in previous studies, this study aims to design and build a system that can correct the accuracy of the pronunciation of makhraj letter 'ain with the method used is a combination of MFCC and Convolutional Neural Network (CNN) with a vgg-16 structure that has been modified. The dataset used is 1,600 voice recordings divided into two categories of the correct pronunciation of the letter 'ain and incorrect pronunciation of the letter 'ain and four variations of pronunciation with different vowels with a total data of 800 records in each category. This study conducted several experiments on variations of the CNN kernel. The results of the training model that produced the best accuracy in all variations were the training model on kernels 16, 32, 64 with a final accuracy rate of 100% for all variations with 96% accuracy validation. In the fathah variation, the validation accuracy is 94%. In the variation of dhommah and the variation of kasrah obtained a validation accuracy of 97%. Therefore, this study succeeded in distinguishing the sound of the pronunciation of the letter 'ain with different vowels and measuring the accuracy of the pronunciation of the letter 'ain. Implementing the modified vgg-16 produces high accuracy and validation values for each speech variation during the model train process.

**Keywords:** CNN, Hijaiyah, Makhraj, MFCC

## DETEKSI KESALAHAN PENGUCAPAN MAKHRAJ HURUF 'AIN MENGGUNAKAN MEL FREQUENCY CEPSTRAL COEFFICIENT DAN VGG-16 TERMODIFIKASI

### Abstrak

Berdasarkan riset yang dilakukan Institut Ilmu Al-Qur'an (IIQ) sebanyak 65% umat Islam di Indonesia mengalami buta aksara Al-Qur'an. Pada penelitian terdahulu dilakukan analisis terhadap pendeteksian kesalahan pengucapan kata bahasa Arab terhadap bukan penduduk asli menggunakan metode *Mel Frequency Cepstral Coefficient* (MFCC) dan *Support Vector Machine* (SVM) dengan hasil uji sebesar 54,6%. Oleh karena hasil akurasi yang rendah pada penelitian sebelumnya, maka penelitian ini bertujuan untuk merancang dan membangun sistem yang dapat mengoreksi tingkat ketepatan pengucapan makhraj huruf 'ain dengan metode yang digunakan adalah gabungan antara MFCC dan *Convolutional Neural Network* (CNN) dengan struktur VGG-16 yang telah dimodifikasi. *Dataset* yang digunakan adalah 1.600 rekaman suara yang terbagi menjadi dua kategori pengucapan huruf 'ain yang benar dan pengucapan huruf 'ain yang salah dan empat variasi pengucapan yang berbeda harakat dengan jumlah data masing-masing kategori berjumlah 800 rekaman. Penelitian ini melakukan beberapa percobaan variasi kernel CNN dan hasil *training* model yang menghasilkan akurasi terbaik pada seluruh variasi adalah *training model* pada kernel 16, 32, 64 dengan tingkat akurasi akhir 100% pada seluruh variasi dengan validasi akurasi 96%. Pada variasi *fathah* didapatkan validasi akurasi sebesar 94%. Pada variasi *dhommah* dan pada variasi *kasrah* didapatkan validasi akurasi sebesar 97%. Oleh sebab itu, penelitian ini berhasil membedakan suara pengucapan huruf 'ain yang berbeda harakat dan mengukur tingkat akurasi kebenaran pengucapan pada rekaman huruf 'ain tersebut. Pengimplementasian vgg-16 yang telah dimodifikasi menghasilkan nilai akurasi dan validasi akurasi yang tinggi pada masing-masing variasi ucapan saat proses *train model*.

**Kata kunci:** CNN, Hijaiyah, Makhraj, MFCC

## 1. PENDAHULUAN

Al-Qur'an merupakan kitab suci sekaligus pedoman hidup bagi umat Islam yang diturunkan oleh Allah *Subhaanahu Wata'aala* kepada Nabi Muhammad *Shallallahu 'Alaihi wa Sallam* melalui perantara malaikat Jibril *'Alaihissalam*, tercatat dalam *mushaf* yang dikutip kepada kita secara *mutawatir*, membacanya merupakan ibadah, diawali dengan surah *Al-Fatihah* dan diakhiri dengan surah *An-Nas* yang mengandung nilai-nilai ajaran yang relevan untuk kehidupan manusia sepanjang masa [1][2][3][4]. Pembacaan ayat-ayat suci Al-Qur'an haruslah menerapkan kaidah ilmu tajwid yang merupakan ilmu yang mempelajari tempat keluarnya bunyi huruf-huruf hijaiyah atau disebut *makharijul huruf* dan hak dari setiap huruf berupa sifat yang harus dijadikan sebagai acuan, sehingga pembacaan Al-Qur'an akan menjadi benar dan sempurna maknanya sebagaimana yang telah dicontohkan oleh Rasulullah *Shallallahu 'alaihi wa Sallam* [5][6]. Hukum mempelajari ilmu tajwid merupakan *fardhu kifayah* yang artinya tidak wajib dipelajari oleh setiap pembaca Al-Qur'an dan dapat diwakilkan, sedangkan mengamalkan ilmu tajwid ketika membaca ayat-ayat Al-Qur'an merupakan *fardhu 'ain* yang artinya wajib ditunaikan oleh setiap orang yang membacanya dan tidak dapat diwakilkan [7][8][9].

Pembahasan utama yang dipelajari dalam ilmu tajwid adalah *makharijul huruf* dari huruf-huruf hijaiyah yang berjumlah 29 [10][11]. Seseorang yang membaca Al-Qur'an harus mampu melafalkan huruf-huruf hijaiyah sesuai dengan *makharijul huruf*-nya dengan tepat karena kesalahan dalam pengucapan suatu huruf akan berdampak pada makna dari kata yang diucapkan [12]. Apabila seseorang dengan sengaja meninggalkan penerapan ilmu tajwid ketika membaca Al-Qur'an sehingga merubah makna dari ayat yang dibaca, maka hukumnya adalah berdosa karena sesungguhnya Allah *Subhaanahu Wata'aala* menurunkan Al-Qur'an berikut dengan tajwidnya [13][14][15].

Proses pembelajaran Al-Qur'an di era industri 4.0 saat ini memiliki sejumlah problematika yang mencuat disebabkan keterlibatan seluruh faktor pembentuk yang menentukan suatu keberhasilan atau kegagalan proses pembelajaran khususnya terkait pendidik dan peserta didik. Beberapa bentuk manifestasi dari problematika dunia pendidikan yang saat ini terjadi secara nyata di Indonesia adalah metodologi pembelajaran *teacher centered* yang lebih terpusat pada penyampaian materi satu arah dari pendidik, minimnya media pembelajaran dan sarana prasarana yang dimanfaatkan, materi yang disampaikan lebih condong bersifat ekspansif, kurangnya inovasi dan motivasi pendidik dalam manajemen peserta didik, serta model dan strategi yang digunakan bersifat tradisional [16].

Berdasarkan latar belakang diatas, maka muncullah gagasan untuk membuat model

pembelajaran baru dalam dunia pendidikan Al-Qur'an dengan memanfaatkan teknologi *artificial intelligence* sistem deteksi kesalahan pengucapan *makhraj* huruf 'ain menggunakan *Mel Frequency Cepstral Coefficient* (MFCC). Pada penelitian Zulaiha [17] disebutkan bahwa penggabungan dua algoritma machine learning yaitu *Mel Frequency Cepstral Coefficient* dan *Convolutional Neural Network* dapat digunakan untuk mengidentifikasi suara di lingkungan yang bising dengan hasil akurasi mencapai 87,5%. Pada penelitian Ma'mun [18] juga menunjukkan bahwa penggunaan algoritma *Mel Frequency Cepstral Coefficient* dapat mengenali suara bacaan Al-Qur'an melalui ekstraksi ciri rekaman suara dengan rata-rata tingkat akurasi sebesar 51,8% sehingga dapat digunakan sebagai metode pemrosesan pengenalan suara pada penelitian ini. Diharapkan dengan adanya sistem deteksi kesalahan pengucapan *makhraj* huruf 'ain ini, setiap orang dapat mulai mempelajari ilmu tajwid secara praktik dan interaktif secara mandiri dimanapun dan kapanpun, sehingga setiap muslim dapat memenuhi kewajibannya ketika membaca Al-Qur'an dan terhindar dari dosa.

## 2. METODE PENELITIAN

### 2.1. Subjek dan Objek Penelitian

Berlandaskan latar belakang yang telah diuraikan pada pendahuluan, maka dibutuhkan subjek penelitian yang memiliki kapabilitas. Pada proses pembelajaran Al-Qur'an dikenal istilah sanad yang merupakan rangkaian bacaan yang dapat dipercaya yang tersambung kepada Rasulullah *shallallahu 'alaihi wa sallam*. Subjek dalam penelitian ini adalah laki-laki yang sudah dan belum memiliki sanad Al-Qur'an. Sedangkan objek penelitian yang diambil berdasarkan batas masalah diatas, yaitu rekaman suara pengucapan huruf 'ain laki-laki yang sudah dan belum memiliki sanad Al-Qur'an.

### 2.2. Fase Pengumpulan Dataset

*Dataset* yang dikumpulkan berupa rekaman suara pengucapan *makhraj* huruf 'ain yang benar dan pengucapan *makhraj* huruf 'ain yang salah yang terdiri dari empat variasi pengucapan, yaitu pengucapan huruf 'ain berharakat *fathah*, huruf 'ain berharakat *dhommah*, huruf 'ain berharakat *kasrah* dan huruf 'ain berharakat *sukun*. Masing-masing informan memberikan 20 pengulangan rekaman pada setiap variasinya. Rekaman suara pengucapan *makhraj* huruf 'ain yang benar didapatkan melalui 10 orang informan yang sudah memiliki sanad Al-Qur'an yang tersambung hingga Rasulullah *Shallallahu 'Alaihi wa Sallam* dengan riwayat Hafs 'an 'Ashim dan pengucapan *makhraj* huruf 'ain yang salah juga didapatkan melalui 10 orang informan yang belum memiliki sanad Al-Qur'an yang tersambung kepada Rasulullah *Shallallahu 'Alaihi wa Sallam* sehingga total informan adalah 20 orang

dan total jumlah rekaman adalah 1.600 rekaman dengan maksimal panjang rekaman adalah 9 detik.

**2.3. Fase Preprocessing**

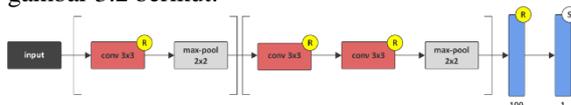
Pada penelitian ini dilakukan lima tahap *preproseccing*. Tahap pertama adalah pengkategorian rekaman suara kedalam empat kategori, yaitu pengucapan huruf 'ain berharakat *fathah*, pengucapan huruf 'ain berharakat *kasrah*, pengucapan huruf 'ain berharakat *dhommah* dan pengucapan huruf 'ain berharakat *sukun*. Masing-masing kategori juga akan dipisah kedalam dua kategori, yaitu pengucapan huruf 'ain yang benar dan pengucapan huruf 'ain yang salah. Tahap kedua adalah pengukuran suara untuk diketahui panjang rekaman terpendek dan terpanjang dari setiap rekaman. Tahap ketiga adalah penentuan nilai tengah agar ketika suara dilakukan *stretching* seluruh rekaman suara memiliki panjang yang seragam. Tahap keempat adalah *stretching* yang dilakukan berdasarkan nilai tengah pada setiap rekaman suara yang bertujuan untuk menyeragamkan panjang data rekaman suara. Tahap terakhir adalah randomisasi untuk pengoptimalisasian kinerja model yang akan dibuat.

**2.4. Fase Konversi Dataset ke MFCC**

Pada proses pengkonversian *dataset* suara ke MFCC terbagai menjadi 3 proses. Proses pertama adalah pengkonversian data suara dari domain waktu kedalam domain frekuensi dua dimensi menggunakan *Fast Fourier Transform* (FFT) untuk kemudian dilakukan proses *Mel Frequency Wrapping* (MFW) untuk mengkonversi domain frekuensi dua dimensi kedalam domain frekuensi tiga dimensi. Output dari proses ini adalah spektogram yang selanjutnya akan di klasifikasikan menggunakan CNN [19].

**2.5. Fase Perancangan Model CNN**

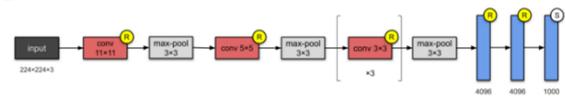
*Convolutional Neural Network* (CNN) adalah salah satu jenis algoritma *Deep Neural Network* yang diimplementasikan pada pengolahan data citra [20]. Data hasil ekstraksi ciri pada MFCC berupa spektogram akan dilakukan proses *training* dan *testing* data menggunakan CNN. Model CNN akan dibuat menggunakan arsitektur VGG-16 termodifikasi yang kemudian diberi nama arsitektur IbNet. Gambar arsitektur IbNet dapat dilihat pada gambar 3.2 berikut:



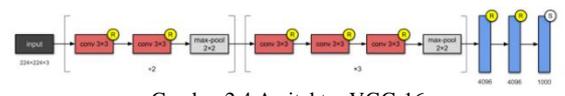
Gambar 2.1 Arsitektur Ibnet

Arsitektur IbNet menggunakan 3 layer konvolusi berukuran 3x3 dengan struktur input, layer konvolusi, max-pooling, layer konvolusi, layer konvolusi, max-pooling kemudian masuk kedalam dense. Struktur IbNet berbeda dengan struktur

AlexNet yang memiliki arsitektur berupa input, layer konvolusi, max-pooling, layer konvolusi, max-pooling, layer konvolusi, max-pooling kemudian masuk kedalam dense. Struktur IbNet juga berbeda dengan asitektur VGG-16 yang berupa input, layer konvolusi, layer konvolusi, max-pooling, layer konvolusi, layer konvolusi, layer konvolusi, max-pooling kemudian masuk kedalam dense. Struktur arsitektur AlexNet dan VGG-16 dapat dilihat pada gambar 3.3 dan 3.4 dibawah ini.



Gambar 2.3 Aritektu AlexNet



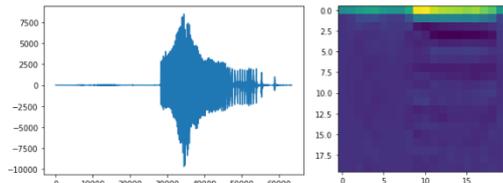
Gambar 2.4 Arsitektur VGG-16

Arsitektur IbNet juga menggunakan 3 lapis *filter* yang diurutkan dari nilai terkecil hingga terbesar. Pada proses pengklasifikasian *dataset* dibagi menjadi tiga dengan perbandingan 80:10:10 dimana 80 merupakan data *training*, 10 adalah data validasi dan 10 merupakan data *testing*. Proses *training* model dilakukan dengan 100 kali *epoch* atau 100 kali putaran menggunakan *optimizer adam*. Model dikatakan bagus apabila *output* membentuk kurva yang *smooth* dengan mengikuti pola *trend* dari data yang di *training*. Hasil riwayat akurasi *training* akan direkam dan diolah menggunakan RegEx dan *google spreadsheet*.

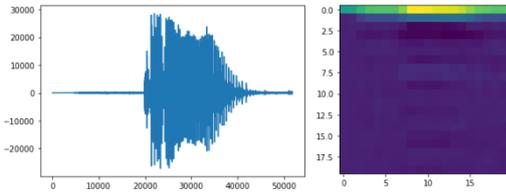
**3. HASIL DAN PEMBAHASAN**

**3.1. Hasil dan Evaluasi Model**

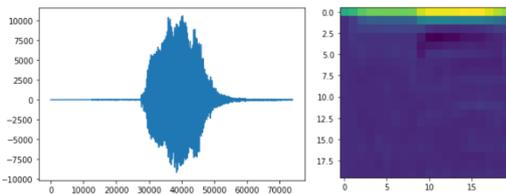
Identifikasi awal perbedaan antara pegucapan huruf 'ain yang benar dan huruf 'ain yang salah ialah dilakukan perepresentasian bentuk gelombang suara dari seluruh kategori dan variasi rekaman huruf 'ain berdasarkan domain waktu dan domain frekuensinya. gambar 3.1 merupakan bentuk representasi rata-rata gelombang suara huruf 'ain *sukun* yang benar, sedangkan gambar 3.2 merepresentasikan bentuk rata-rata gelombang suara huruf 'ain *sukun* yang salah. gambar 3.3 merupakan representasi bentuk rata-rata gelombang suara huruf 'ain *fathah* yang benar, sedangkan gambar 3.4 merepresentasikan bentuk rata-rata gelombang suara huruf 'ain *fathah* yang salah. Gambar 3.5 merupakan bentuk representasi rata-rata gelombang suara huruf 'ain *dhommah* yang benar, sedangkan gambar 3.6 merepresentasikan bentuk rata-rata gelombang suara huruf 'ain *dhommah* yang salah. Gambar 3.7 merupakan rata-rata representasi gelombang suara huruf 'ain *kasrah* yang benar, sedangkan gambar 3.8 merepresentasikan rata-rata gelombang suara huruf 'ain *kasrah* yang salah.



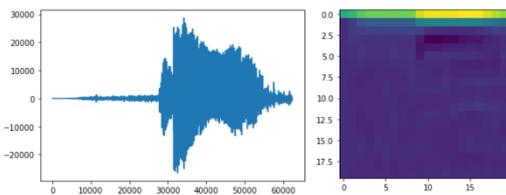
Gambar 3.1. Gelombang suara 'ain sukun' yang benar berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



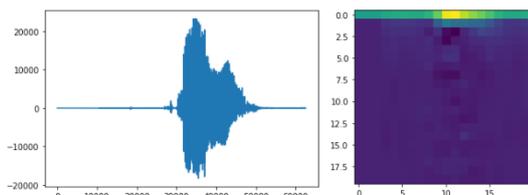
Gambar 3.2. Gelombang suara 'ain sukun' yang salah berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



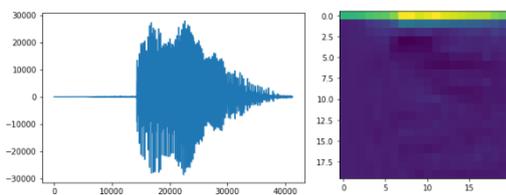
Gambar 3.3. Gelombang suara 'ain fathah' yang benar berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



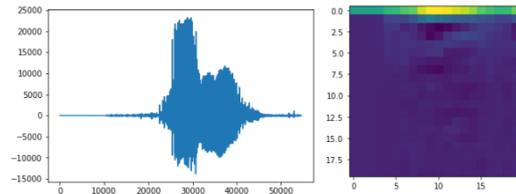
Gambar 3.4. Gelombang suara 'ain fathah' yang salah berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



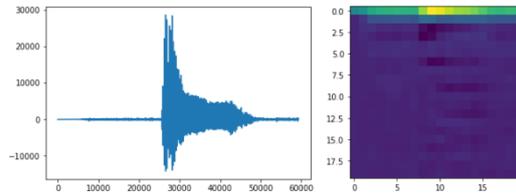
Gambar 3.5. Gelombang suara 'ain dhommah' yang benar berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



Gambar 3.6. Gelombang suara 'ain dhommah' yang salah berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



Gambar 3.7. Gelombang suara 'ain kasrah' yang benar berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)



Gambar 3.8. Gelombang suara 'ain kasrah' yang salah berdasarkan domain waktu (kiri) dan domain frekuensi (kanan)

Berdasarkan gambar 3.1 hingga gambar 3.8 dapat dilihat secara kasat mata bahwa rata-rata pengucapan huruf 'ain yang benar dan pengucapan huruf 'ain yang salah memiliki perbedaan sehingga dapat dilakukan pemodelan *deep learning* untuk mengetahui tingkat akurasi ketepatan pengucapannya.

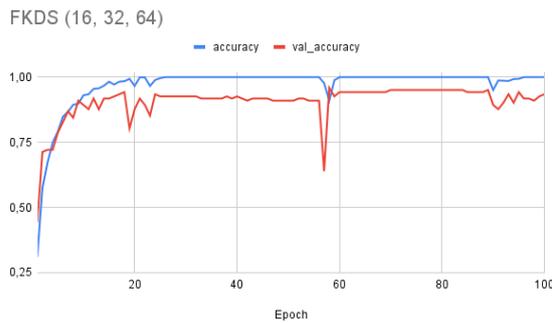
### 3.2 Membedakan Pengucapan Huruf A'in Berbeda Harakat

Penggunaan IbNet pertama kali digunakan untuk membedakan gelombang suara antara pengucapan huruf 'ain berharakat *sukun*, pengucapan huruf 'ain berharakat *fathah*, pengucapan huruf 'ain berharakat *dhommah* dan pengucapan huruf 'ain berharakat *kasrah* yang dapat dilihat pada tabel 3.1 dibawah ini.

Tabel 3. 1 Hasil Uji Coba Kernel Deteksi Perbedaan Harakat

No.	Kernel	Accuracy	Val-Accuracy
1	4, 8, 16	100%	88%
2	8,16,32	100%	89%
3	16,32,64	100%	93%
4	32,64,128	27%	28%

Tabel 3.1 menunjukkan bahwa tingkat akurasi akhir terbaik mencapai 100% dan validasi akurasi mencapai 93% pada kernel 16,32,64. Bentuk kurva dari hasil train kernel 16,32,64 dapat dilihat pada gambar 3.9 dibawah.



Gambar 3.9. Kurva kernel 16, 32, 64 dengan 100 kali epoch

Kurva yang dihasilkan merupakan kurva yang bagus karena tidak mengalami *underfitting* maupun *overfitting*. Kurva juga menyatakan bahwa model berhasil menemukan *pattern* pembelajaran yang tepat untuk membedakan suara pengucapan huruf 'ain berbeda harakat,

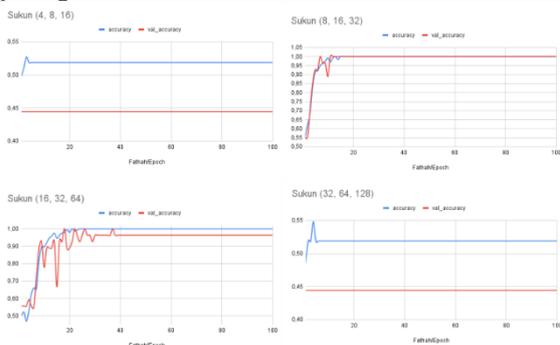
### 3.3. Membedakan Pengucapan Huruf 'ain Benar dan Salah Pada Setiap Variasi

IbNet kemudian digunakan untuk mengklasifikasikan gelombang suara antara pengucapan huruf 'ain yang benar dan pengucapan huruf 'ain yang salah. Hasil perbandingan kernel yang dapat dilihat pada tabel 3.2 hingga 3.5 dan gambar 3.10 hingga gambar 3.13 dibawah ini.

Tabel 3. 2 Hasil Uji Coba Kernel *Sukun*

No.	Kernel	Accuracy	Val-Accuracy
1	4, 8, 16	52%	44%
2	8,16,32	100%	100%
3	16,32,64	100%	96%
4	32,64,128	52%	44%

Tabel 3.2 menunjukkan kernel terbaik adalah 16,32,64 dengan tingkat akurasi akhir mencapai 100% dan validasi akurasi mencapai 93%. Kernel 8,16,32 bukan kernel terbaik karena terindikasi mengalami *underfitting* yang berarti model tidak dapat menemukan *pattern* pembelajaran yang tepat sehingga validasi akurasi mencapai 100%. Bentuk kurva dari setiap hasil train tabel 3.2 dapat dilihat pada gambar 3.10 dibawah.



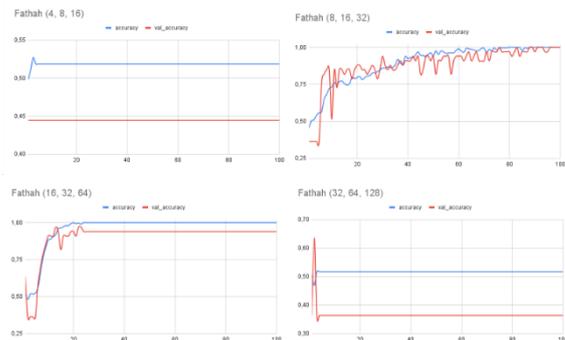
Gambar 3.10. Kurva perbandingan kernel *sukun* 100 kali epoch

Gambar 3.10 merupakan kurva perbandingan dari setiap kernel yang diujikan. Didapatkan tingkat akurasi terbaik pada kernel 16,32,64 dengan tingkat akurasi sebesar 100% dan validasi akurasi 96%. Selanjutnya adalah hasil uji coba kernel *fathah* yang dapat dilihat pada tabel 3.3.

Tabel 3. 3 Hasil Uji Coba Kernel *Fathah*

No.	Kernel	Accuracy	Val-Accuracy
1	4, 8, 16	52%	44%
2	8,16,32	100%	100%
3	16,32,64	100%	94%
4	32,64,128	52%	36%

Tabel 3.3 menunjukkan bahwa kernel terbaik juga terdapat pada kernel 16,32,64 dengan tingkat akurasi akhir mencapai 100% dan validasi akurasi mencapai 94%. Kernel 8,16,32 bukan kernel terbaik juga karena terindikasi mengalami *underfitting* yang berarti model tidak dapat menemukan *pattern* pembelajaran yang tepat sehingga validasi akurasi mencapai 100%. Bentuk kurva dari setiap hasil train tabel 3.3 dapat dilihat pada gambar 3.11.



Gambar 3. 11 Kurva perbandingan kernel *fathah* 100 kali epoch

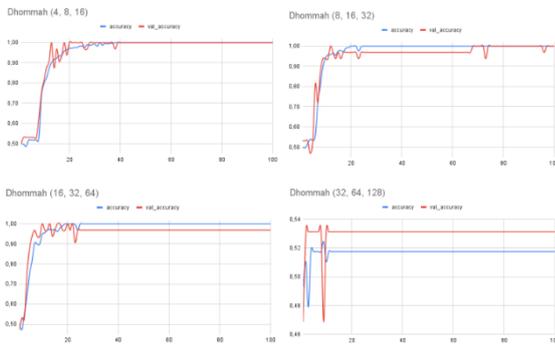
Gambar 3.11 merupakan kurva perbandingan dari setiap kernel yang diujikan. Didapatkan tingkat akurasi terbaik pada kernel 16,32,64 dengan tingkat akurasi sebesar 100% dan validasi akurasi 94%. Selanjutnya adalah hasil uji coba kernel *dhommah* yang dapat dilihat pada tabel 3.4.

Tabel 3. 4 Hasil Uji Coba Kernel *Dhommah*

No.	Kernel	Accuracy	Val-Accuracy
1	4, 8, 16	100%	100%
2	8,16,32	100%	100%
3	16,32,64	100%	97%
4	32,64,128	52%	53%

Tabel 3.4 menunjukkan bahwa kernel terbaik juga terdapat pada kernel 16,32,64 dengan tingkat akurasi akhir mencapai 100% dan validasi akurasi mencapai 97%. Kernel 8,16,32 bukan kernel terbaik juga karena terindikasi mengalami *underfitting* yang berarti model tidak dapat menemukan *pattern*

pembelajaran yang tepat sehingga validasi akurasi mencapai 100%. Bentuk kurva dari setiap hasil train tabel 3.3 dapat dilihat pada gambar 3.12.



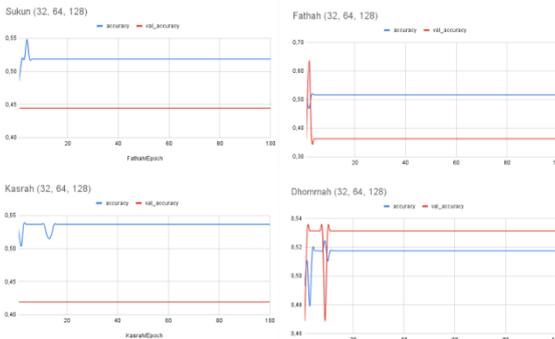
Gambar 3. 12 Kurva perbandingan kernel *dhommah* 100 kali epoch

Gambar 3.12 merupakan kurva perbandingan dari setiap kernel yang diujikan. Didapatkan tingkat akurasi terbaik pada kernel 16,32,64 dengan tingkat akurasi sebesar 100% dan validasi akurasi 97%. Selanjutnya adalah hasil uji coba kernel *kasrah* yang dapat dilihat pada tabel 3.5 dibawah.

Tabel 3. 5 Hasil Uji Coba Kernel *Kasrah*

No.	Kernel	Accuracy	Val-Accuracy
1	4, 8, 16	100%	94%
2	8,16,32	100%	90%
3	16,32,64	100%	97%
4	32,64,128	54%	42%

Tabel 3.5 menunjukkan bahwa kernel terbaik juga terdapat pada kernel 16,32,64 dengan tingkat akurasi akhir mencapai 100% dan validasi akurasi mencapai 97%. Bentuk kurva dari setiap hasil train tabel 3.5 dapat dilihat pada gambar 3.13 dibawah.



Gambar 3. 13 Kurva perbandingan kernel *kasrah* 100 kali epoch

Gambar 3.13 merupakan kurva perbandingan dari setiap kernel yang diujikan. Didapatkan tingkat akurasi terbaik pada kernel 16,32,64 dengan tingkat akurasi sebesar 100% dan validasi akurasi 97%.

#### 4. KESIMPULAN

Kesimpulan yang dapat diambil dari hasil penelitian diatas adalah model pada penelitian ini berhasil mendeteksi kesalahan pelafazan makhrj huruf hijaiyah menggunakan MFCC dan VGG-16 termodifikasi dengan validasi akurasi terbaik adalah model kernel 16, 32, 64 dengan 100 kali epoch yang menghasilkan tingkat validasi akurasi 96% pada variasi *sukun*. 94% pada variasi *fathah*, 97% pada variasi *dhommah* dan *kasrah*.

#### DAFTAR PUSTAKA

- [1] M. A. Prof. Dr. H. Muhammad Chirzin, *Kearifan Al-Qur'an*. Jakarta: PT Gramedia Pustaka Utama, 2020.
- [2] M. Oktapiani, "Tingkat Kecerdasan Spiritual Dan Kemampuan Menghafal Al-Qur'an," *Tahdzib Al-Akhlaq J. Pendidik. Islam*, vol. 3, no. 1, hal. 95–108, 2020, doi: 10.34005/tahdzib.v3i1.861.
- [3] N. D. R. AGUSTIANA, "Implementasi Pembelajaran Menghafal Alqur'an Melalui Metode Tsami' Dan Muroja'ah Di Pondok Pesantren Putri Raudhatul Musthofa Pundensari Rejotangan Tulungagung," IAIN Tulung Agung, 2019.
- [4] F. Aulia, "Hubungan Pemahaman Ilmu Tajwid dengan Kemampuan Membaca Al-Qur'an kelas V di MIN 1 Bandar Lampung.," UIN RADEN INTAN LAMPUNG, 2020. [Daring]. Tersedia pada: <http://dx.doi.org/10.1016/j.ndteint.2014.07.001> <https://doi.org/10.1016/j.ndteint.2017.12.003> <http://dx.doi.org/10.1016/j.matdes.2017.02.024>
- [5] Viska Mutiawani, Maria Ulfa, dan Muslim, "KepoTajwid: Aplikasi Pembelajaran Ilmu Tajwid Berbasis Web Interaktif," *IJAI Indones. J. Appl. Informatics*, 2018.
- [6] Kuswandi, "Korelasi Pemahaman Ilmu Tajwid dengan Kemampuan Membaca Alqur'an Santri Pondok Pesantren 'Ainul Ulum Pulung Ponorogo Tahun Ajaran 2017-2018," IAIN PONOROGO, 2018.
- [7] V. M. - STMIK Nusa Mandiri Jakarta, H. M. N. - AMIK BSI Purwokerto, dan W. R. - AMIK BSI Purwokerto, "Aplikasi Pembelajaran Ilmu Tajwid Berbasis Android," *Evolusi J. Sains dan Manaj.*, vol. 6, no. 1, hal. 91–100, 2018, doi: 10.31294/evolusi.v6i1.3586.
- [8] Suci Shofia, *Tajwid for Children*. Bandung, 2018.
- [9] Khuriyatul Wafiah, "Hubungan Antara Penguasaan Ilmu Tajwid Dengan Kemampuan Membaca Dan Menghafal Al-Qur'an Surah Pendek Santri Putri Kelas Vii Pondok Modern Darul Hikmah Tawangarsi

- Tulungagung,” IAIN Tulung Agung, 2019.
- [10] A. Febrianto, J. T. Informatika, dan A. Reality, “View of Penggunaan Teknologi Augmented Reality Dalam Mempelajari Ilmu Tajwid Use Of Augmented Reality Technology In Studying Tajwid Science Using Augmented Reality Technology In Studying Tajwid Science,” vol. 7, no. 1, hal. 83–90, 2018.
- [11] R. J. Rosyanafi, “Pengaruh Media Jigsaw Puzzle Terhadap Minat Belajar Huruf Hijaiyah Anak Usia Dini,” *Ijaz Arab. J. Arab. Learn.*, vol. 1, no. 1, 2018, doi: 10.18860/ijazarabi.v1i1.5016.
- [12] S. Muthmainnah dan I. Nuramaliah, “Analisis Kesalahan Pelafalan Bunyi Huruf 1 Hijaiyyah berdasarkan Makhrijul huruf dalam Membaca Teks Dialog Bahasa Arab Siswa Kelas X SMA Muhammadiyah Limbung,” hal. 1–10, 2019.
- [13] E. Musdzalifah, “Materi Ilmu Tajwid Dalam Kitab Terjemah Matan Jazariyah Karya Syekh Muhammad Bin Muhammad Ibn Al Jazari Dan Implementasinya Dalam Mata Pelajaran Qur’an Hadits Kelas Vii Madrasah Tsanawiyah,” no. November, 2020.
- [14] SAEFUL ANAS, “Kajian Ilmu Tajwid pada Kitab Matan Al-Muqaddimah Al-Jazariyah karya Ibnu Jazari dan Relevansinya dengan Bahan Ajar Al-Qur’an Hadits Kelas VIII Madrasah Tsanawiyah,” IAIN Kudus, 2020.
- [15] A. A. Al Halim dan Wida Nurul 'Azizah, “Upaya Peningkatan Kemampuan Membaca Al-Qur’an Melalui Pengenalan Huruf Hijaiyah Menggunakan Metode Qo’idah Baghdadiyah Ma’a Juz ‘Amma (Turutan) Di Kelas 1a Mi Ma’arif Nu 01 Tritihkulon Tahun Pelajaran 2015/2016,” *Tawadhu*, vol. 2, no. 1, hal. 490, 2018.
- [16] D. Ratnasri, “Problematics of the Qur ’ an Learning in the Industry Era,” vol. 6, no. 1, hal. 72–92, 2020.
- [17] A. Ashar, M. S. Bhatti, dan U. Mushtaq, “Speaker Identification Using a Hybrid CNN-MFCC Approach,” *2020 Int. Conf. Emerg. Trends Smart Technol. ICETST 2020*, 2020, doi: 10.1109/ICETST49965.2020.9080730.
- [18] R. T. Handayanto dan H. Herlawati, “Prediksi Kelas Jamak dengan Deep Learning Berbasis Graphics Processing Units,” *J. Kaji. Ilm.*, vol. 20, no. 1, hal. 67–76, 2020, doi: 10.31599/jki.v20i1.71.
- [19] H. Heriyanto, S. Hartati, dan A. E. Putra, “Ekstraksi Ciri Mel Frequency Cepstral Coefficient (Mfcc) Dan Rerata Coefficient Untuk Pengecekan Bacaan Al-Qur’an,” *Telematika*, vol. 15, no. 2, hal. 99, 2018, doi: 10.31315/telematika.v15i2.3123.
- [20] R. Magdalena *et al.*, “Convolutional Neural Network for Anemia Detection Based on Conjunctiva Palpebral Images,” *J. Tek. Inform.*, vol. 3, no. 2, hal. 349–354, 2022, [Daring]. Tersedia pada: <http://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/197>

