

Klasifikasi Irama Bacaan Al-Qur'an Menggunakan Algoritma CNN

Shoffin Nahwa Utama^{1,*}, Johan Ericka Wahyu Prakasa², Wahyu Hariyanto³

^{1,2}Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia

³Program Studi Perpustakaan dan Sains Informasi, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia

¹shoffin@uin-malang.ac.id; ²johan@uin-malang.ac.id; ³wahyuhariyanto@uin-malang.ac.id

* penulis korespondensi

INFO ARTIKEL

Sejarah Artikel

Diterima: 10 April 2025

Direvisi: 22 April 2025

Diterbitkan: 30 April 2025

Kata Kunci

Al-Qur'an

CNN

Klasifikasi

MFCC

Tilawah

ABSTRAK

Klasifikasi nada bacaan Al-Qur'an sangat penting untuk mendukung pembelajaran tajwid, tartil, serta tilawah yang sesuai dengan aturan. Tantangan utama dalam klasifikasi ini terletak pada keberagaman gaya bacaan qari dan kemiripan akustik antar maqam. Penelitian ini bertujuan untuk mengembangkan model klasifikasi otomatis irama bacaan Al-Qur'an menggunakan pendekatan berbasis CNN dengan 8 kelas maqam bacaan. Model CNN dalam penelitian ini memiliki tiga jalur konvolusi dengan ukuran kernel berbeda. Variasi bentuk masukan berupa data audio yang diubah ke dalam representasi *spektrogram* dan *mel-frequency cepstral coefficients* (MFCC). Evaluasi kinerja model pada dataset bacaan Al-Qur'an yang terdiri dari 8 kelas tilawah yaitu Ajam, Bayat, Hijaz, Kurd, Nahawand, Rast, Saba, dan Seka. Hasil pelatihan menunjukkan bahwa metode yang diusulkan mencapai akurasi 92,6%, sedangkan pada proses pengujian didapatkan akurasi sebesar 82,04%. Hasil *confusion matrix* didapatkan nilai akurasi yang diperoleh dalam proses validasi mencapai 80,88%. Nilai presisi, *recall* dan *F1-score* masing-masing adalah 0,82, 0,80, dan 0,81. Dengan hasil ini, pendekatan CNN yang diusulkan terbukti efektif untuk mendukung otomatisasi dan peningkatan akurasi dalam klasifikasi nada bacaan Al-Qur'an.

PENDAHULUAN

Membaca Al-Qur'an dengan tartil dan indah dikenal sebagai tilawah[1][2][3][4] merupakan bidang kajian yang penting dalam studi Islam, terutama dalam ilmu tajwid dan seni melodi bacaan[1]. Maqamat adalah sistem skala melodi yang digunakan oleh qari (pembaca Al-Qur'an) untuk mengatur nada dan ritme bacaan mereka,[5] yang tidak hanya memperindah bacaan tetapi juga membantu dalam memahami dan menghayati makna ayat-ayat suci [6]. Secara tradisional, maqamat diajarkan secara langsung oleh guru qari kepada murid-muridnya, yang membutuhkan waktu dan dedikasi yang panjang.[7][8]

Dalam beberapa dekade terakhir, kemajuan teknologi informasi telah membuka jalan bagi otomatisasi berbagai aspek kehidupan, termasuk klasifikasi irama dalam bacaan Al-Qur'an[2]. Salah satu teknik yang menjanjikan dalam hal ini adalah penggunaan Convolutional Neural Networks (CNN)[7], yang merupakan bagian dari deep learning[6]. CNN telah terbukti sangat efektif dalam pengenalan pola dan klasifikasi dalam berbagai domain, termasuk pengolahan citra, pengenalan suara, dan analisis sinyal [9].

Klasifikasi jenis tilawah bacaan Al-Qur'an telah menjadi fokus penelitian yang signifikan dalam beberapa dekade terakhir. Studi sebelumnya telah berupaya untuk mengidentifikasi pola-pola jenis nada irama bacaan dan mengklasifikasikannya menggunakan berbagai pendekatan, baik secara tradisional maupun melalui metode modern seperti pengolahan sinyal suara dan machine learning[5].

Pertama, penelitian oleh Sakib Shahriar dan Usman Tariq [2], Dengan menggunakan algoritma deep learning, penelitian ini berkonsentrasi pada pembagian delapan maqāmāt (bentuk jamak dari maqām) yang populer. Untuk pelatihan model, berbagai fitur audio diekstraksi, termasuk rasio frekuensi cepstral mel (MFCC), energi, karakteristik spektral, dan kromatik. CNN, LSTM, dan Deep ANN adalah beberapa arsitektur deep learning yang dilatih untuk mengklasifikasikan contoh audio berdasarkan salah satu dari delapan maqāmāt. Hasil dari penggunaan CNN didapatkan Akurasi pengujian adalah 72,1%. Nilai presisi 0,80, recall 0,76, dan f1-score 0,76. Ini adalah penelitian pertama yang membahas klasifikasi maqām dalam tilawah Al-Qur'an. Dataset "Maqam-478" dapat digunakan untuk pengembangan lebih lanjut dari penelitian ini.

Kedua, Rizal Hadiyansah dan Rafi Andamira [10] dalam penelitian mereka menggunakan CNN untuk mendeteksi membaca dan menghafal Al-Qur'an. Penggunaan model CNN berhasil membedakan satu ayat dengan ayat lainnya dengan sangat baik. Hasil validasi menunjukkan model mampu mendeteksi dengan benar 57 ayat dari 64 data yang direkam, dengan tingkat akurasi sebesar 89,06%. Ketiga, penelitian [11] Faizan et al, menggunakan MFCC untuk mengkategorikan berbagai ayat Al-Qur'an ke dalam empat emosi yang berbeda berdasarkan kumpulan audio yang telah ditentukan. Kontribusi utama dari makalah ini adalah untuk mencoba memberikan penjelasan ilmiah terhadap sifat menenangkan dari ayat-ayat Al-Qur'an dengan menguraikan fitur-fitur yang mengontrol bagaimana suatu audio dipersepsikan sebagai emosi yang berbeda-beda, oleh manusia. Tujuannya adalah untuk memberikan analisis yang tidak memihak dan menjelaskan alasan di balik emosi yang ditimbulkan oleh ayat-ayat Al-Qur'an kepada manusia dengan menggunakan MFCC dalam model pembelajaran mesin

Terakhir, penelitian M. M. Al Anazi and O. R. Shahin [5] menyarankan penggunaan bahasa pembelajaran mesin untuk mengidentifikasi pembaca Al-Qur'an. Selain itu, suara sepuluh qari terkenal dikumpulkan sebagai data penelitian. MFCC digunakan untuk menganalisis kumpulan data audio. Metode klasifikasi menggunakan jaringan saraf tiruan (ANN) dan KNN. Sistem yang diusulkan divalidasi dengan dua surah terpilih dari Al-Qur'an. Pitch digunakan sebagai fitur untuk melatih KNN dan pengklasifikasi ANN. Hasilnya menunjukkan tingkat akurasi yang sangat baik. Surah ke 7 memiliki akurasi 98,5% dengan bantuan pengklasifikasi ANN, dan Surah 32 memiliki akurasi 97,2% dengan bantuan KNN.

CNN telah terbukti menjadi salah satu algoritma terbaik untuk klasifikasi data berbasis suara dan gambar[10]. Kemampuan CNN dalam menangkap pola-pola kompleks secara otomatis tanpa perlu feature engineering yang mendalam. Dalam konteks klasifikasi irama bacaan Al-Qur'an, penggunaan CNN memungkinkan sistem untuk mengenali dan mengklasifikasikan berbagai maqamat dengan akurasi yang lebih tinggi dibandingkan metode konvensional

Penerapan CNN dalam klasifikasi maqam bacaan Al-Qur'an tidak hanya dapat meningkatkan akurasi dan efisiensi proses, tetapi juga dapat menyediakan alat bantu pengajaran yang berharga bagi para guru dan murid dalam memahami dan mempraktikkan maqamat dengan benar. Studi-studi sebelumnya telah menunjukkan bahwa CNN dapat secara efektif mengenali pola-pola audio yang kompleks, yang sangat sesuai dengan karakteristik maqamat yang melibatkan variasi nada dan ritme yang halus [7]. Penelitian ini memiliki kebaruan pada penerapan model CNN untuk klasifikasi 8 kelas irama tilawah Al-Qur'an. Model yang diusulkan pada penelitian ini diharapkan dapat mengidentifikasi irama tilawah dengan tingkat akurasi yang lebih baik, berdasarkan analisis spektrum daya dari sinyal audio bacaan.

Studi ini tidak hanya memberikan kontribusi signifikan terhadap literatur di bidang pengenalan suara dan pengolahan sinyal, tetapi juga memiliki implikasi praktis yang besar dalam pendidikan dan pelatihan bacaan Al-Qur'an.

METODE

Dataset Eksperiment

Data yang digunakan dalam penelitian ini merupakan kumpulan data audio bacaan Al-Qur'an yang diklasifikasikan berdasarkan irama maqam. Dataset tersebut berasal dari koleksi publik yang dikurasi oleh Faisal Umari[12][13], data yang digunakan berjumlah 3.869 file audio yang masing-masing dikategorikan ke dalam delapan kelas irama tilawah yang berbeda. Setiap file audio menggunakan format file *.wav* dengan durasi 30 detik. Pemilihan durasi yang konsisten ini didasarkan pada kebutuhan algoritma CNN yang digunakan, Agar pembacaan fitur pada data bisa maksimal diperlukan panjang audio yang seragam untuk mempermudah proses pelatihan model dan memastikan hasil yang valid. Dataset yang digunakan adalah dataset publik yang telah dilabeli dengan delapan kelas utama yang umum digunakan di seluruh dunia yang terdiri dari Ajam, Bayat, Hijaz, Kurd, Nahawand, Rast, Saba, dan Seka [12]. Dataset tersebut dipecah menjadi dua bagian, yaitu 80% digunakan untuk data pelatihan dan 20% untuk data pengujian.

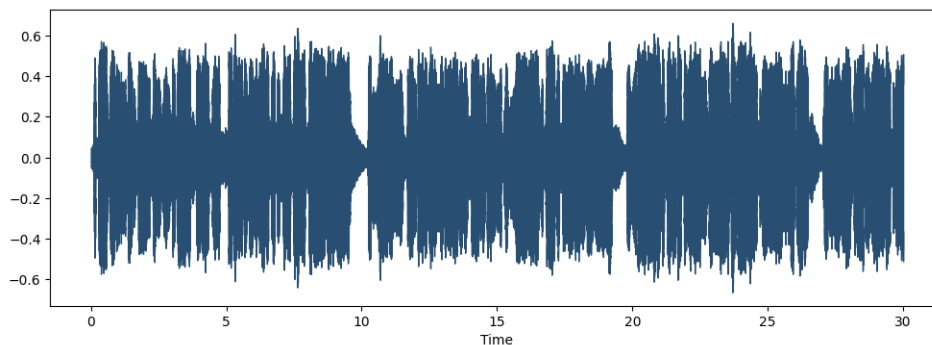
Keunikan dari dataset ini terletak pada tingkat keragaman dan kompleksitasnya. Berbeda dari dataset yang digunakan dalam penelitian-penelitian sebelumnya yang umumnya terdiri dari satu atau dua pembaca dan data yang lebih bersih. Kumpulan data ini lebih representatif terhadap kondisi dunia nyata, namun juga lebih menantang karena mengandung ketidakseimbangan kelas dan noise akustik yang tinggi akibat perbedaan mikrofon, latar belakang suara, dan kualitas rekaman. Jumlah data dalam setiap kelas irama bervariasi. Kelas Hijaz memiliki jumlah data terbanyak dengan 797 file audio, sedangkan kelas Kurd memiliki jumlah data paling sedikit dengan 340 file audio. Perbedaan jumlah data ini dapat memberikan tantangan dalam menjaga keseimbangan pelatihan model. Oleh karena itu, upaya khusus dilakukan untuk memastikan bahwa model tidak bias terhadap maqam-maqam yang memiliki lebih banyak data. Jumlah file audio untuk setiap kelas maqam ditunjukkan pada Tabel 1.

Tabel 1. Jumlah data setiap kelas

No	Kelas irama	Jumlah	Distribusi Kelas
1	Ajam	541	14.0%
2	Bayat	356	9.2%
3	Hijaz	797	20.6%
4	Kurd	340	8.8%
5	Nahawan	350	9.0%
6	Rast	537	13.9%
7	Saba	519	13.4%
8	Seka	429	11.1%

Data pre-processing

Preprocessing data audio dilakukan dengan mengonversi audio tilawah Al-Qur'an menjadi data gambar untuk mendapatkan sinyal audio dalam pengenalan pola dan pembelajaran mesin. Seluruh data audio tilawah dari delapan kelas maqam terlebih dahulu diproses menjadi satu himpunan data yang terpadu sebelum digunakan dalam pelatihan model. Teknik yang digunakan untuk mengonversi audio menjadi representasi gambar adalah dengan menggunakan *Mel-Frequency Cepstral Coefficients* (MFCC).



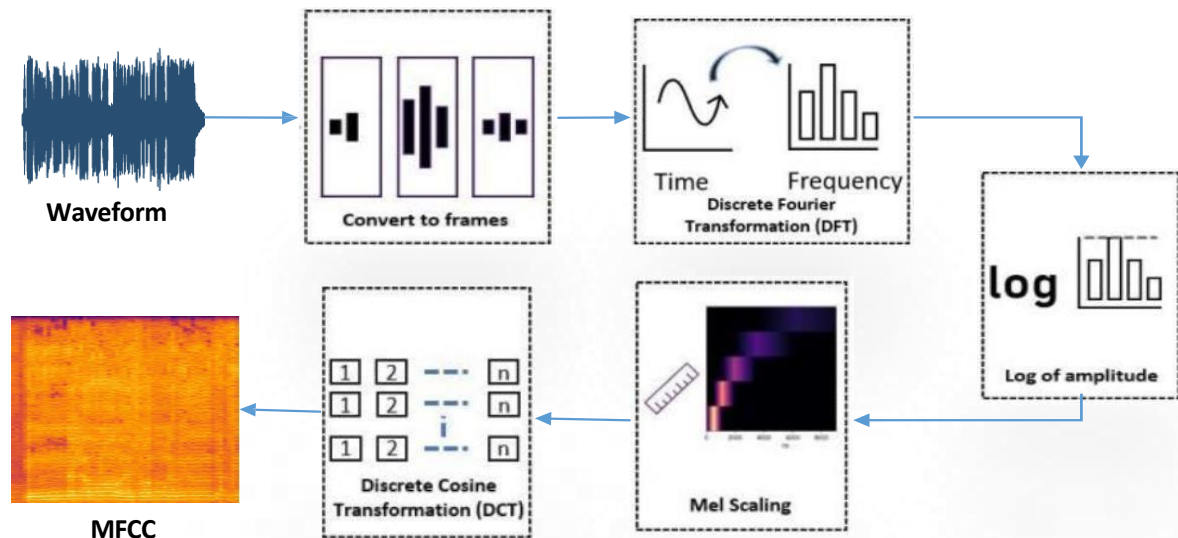
Gambar 1. Sinyal audio irama tilawah Ajam

Dalam penelitian ini, preprocessing dilakukan terhadap rekaman audio maqam bacaan Al-Qur'an untuk memastikan bahwa sinyal audio berada dalam kondisi optimal sebelum fitur-fiturnya diekstraksi[14]. Langkah-langkah preprocessing mencakup beberapa tahap penting. Pertama, Konversi Format dilakukan dengan mengubah semua rekaman audio ke dalam format WAV dengan frekuensi sampel 16 kHz. Frekuensi ini dipilih untuk mempertahankan kualitas audio sekaligus mengurangi ukuran data yang diperlukan dalam proses komputasi. Selanjutnya, Normalisasi Amplitudo diterapkan untuk menyamakan tingkat volume di seluruh rekaman. Normalisasi ini berperan penting dalam mengurangi variasi yang mungkin timbul akibat perbedaan volume pada rekaman asli, yang berpotensi memengaruhi kinerja model. Tahap terakhir adalah Pemotongan dan Padding, di mana setiap rekaman audio dipotong atau dipanjangkan menggunakan padding agar memiliki durasi seragam, yaitu 30 detik. Hasil proses sinyal audio dapat dilihat pada gambar 1 yang merupakan sinyal audio irama tilawah Ajam. Langkah ini dilakukan untuk mengatasi perbedaan panjang rekaman yang dapat mempengaruhi konsistensi dan efektivitas pelatihan model. Dengan penerapan langkah-langkah preprocessing ini, diharapkan sinyal audio dapat diproses secara optimal untuk mendukung analisis dan pengenalan pola irama tilawah Al-Qur'an.

Ekstraksi Fitur Menggunakan MFCC

Setelah proses *preprocessing*, langkah berikutnya adalah ekstraksi fitur dari sinyal audio. Fitur-fitur ini adalah representasi digital dari sinyal audio yang berfungsi sebagai input untuk model CNN. Pada penelitian ini, fitur yang diekstraksi adalah MFCC [15]. MFCC merupakan salah satu teknik yang digunakan untuk pengenalan suara dan analisis audio karena kemampuannya dalam menangkap karakteristik frekuensi penting dari sinyal audio [16][14]. Gambar 2 menunjukkan tahapan dalam ekstraksi audio ke MFCC.

Proses ekstraksi fitur audio bacaan Al-Qur'an dimulai dengan *waveform*, yaitu representasi awal sinyal audio dalam domain waktu yang menunjukkan bagaimana amplitudo sinyal berubah seiring waktu [17]. Selanjutnya, sinyal audio yang berkelanjutan dibagi menjadi segmen-segmen pendek yang disebut *frame* dengan durasi setiap *frame* selama 20 ms. Pembagian ini penting karena sinyal audio sering mengalami perubahan dalam waktu singkat, sehingga dengan memecahnya menjadi *frame* memungkinkan analisis yang lebih rinci dan akurat. Setelah pembagian dilakukan, setiap *frame* dikonversi dari domain waktu ke domain frekuensi dengan menggunakan *Discrete Fourier Transform* (DFT). Transformasi ini menghasilkan spektrum frekuensi yang menunjukkan bagaimana energi terdistribusi di berbagai frekuensi dalam setiap *frame*.



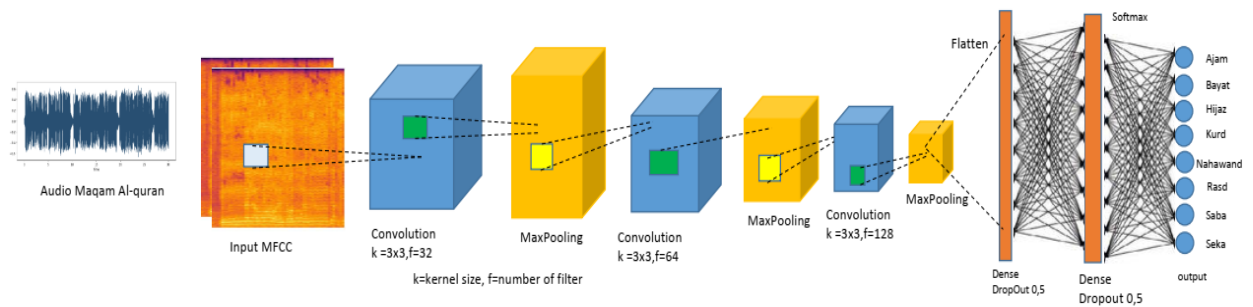
Gambar 2. Proses ekstraksi fitur MFCC

Spektrum frekuensi yang dihasilkan dari DFT kemudian dikonversi ke skala logaritmik. Proses logaritma ini penting karena meniru cara manusia mendengar, di mana sensitivitas terhadap perubahan amplitudo lebih tinggi pada tingkat rendah dibandingkan pada tingkat tinggi. Setelah amplitudo diubah ke skala logaritmik, spektrum tersebut diterapkan ke dalam skala *Mel* melalui filter bank *Mel*. Filter bank ini terdiri dari sejumlah filter segitiga yang berfokus pada rentang frekuensi tertentu untuk menyesuaikan persepsi manusia terhadap frekuensi. Proses ini menghasilkan koefisien *Mel*, yang merupakan representasi energi pada setiap frekuensi *Mel*.

Koefisien *Mel* kemudian diproses menggunakan *Discrete Cosine Transform* (DCT) untuk mengurangi korelasi antar-koefisien dan menghasilkan fitur yang lebih kompak. Hasil dari DCT ini adalah *Mel-Frequency Cepstral Coefficients* (MFCCs), yang berisi 12 koefisien untuk setiap frame. MFCCs menjadi representasi fitur penting yang sering digunakan dalam berbagai aplikasi seperti pengenalan suara dan klasifikasi audio. MFCCs ini juga dapat divisualisasikan sebagai *spectrogram*, yaitu grafik yang menggambarkan perubahan koefisien seiring waktu. Terakhir MFCCs disusun menjadi vektor fitur yang berfungsi sebagai input bagi model CNN untuk klasifikasi irama tilawah. Dengan menggunakan MFCCs, model dapat secara efektif menangkap karakteristik frekuensi yang penting dan berbeda dari setiap kelas tilawah, sehingga sangat membantu dalam proses klasifikasi irama tilawah bacaan Al-Qur'an.

Convolutional Neural Network

Arsitektur CNN untuk klasifikasi irama bacaan Al-Qur'an dalam penelitian ini dirancang untuk memproses data audio yang telah diubah menjadi fitur MFCC. Arsitektur CNN ditunjukkan pada gambar 3. terdiri dari beberapa lapisan konvolusi yang bertujuan mengekstraksi pola dan fitur penting dari representasi *spektral audio* [10]. Setiap lapisan konvolusi dilengkapi dengan sejumlah filter untuk menangkap fitur lokal di dalam data audio, diikuti dengan fungsi aktivasi ReLU (Rectified Linear Unit) untuk menambah non-linearitas dan mengatasi masalah *vanishing gradient* [18]. Lapisan *pooling* diterapkan untuk mengurangi dimensi data, sehingga mengurangi kompleksitas komputasi dan mencegah *overfitting*[19]. Dengan menumpuk beberapa lapisan konvolusi dan *pooling*, model dapat menangkap pola *temporal* dan frekuensi yang relevan dari irama bacaan Al-Qur'an[20].



Gambar 3. Model Arsitektur CNN

Setelah lapisan konvolusi dan *pooling*, data diproses oleh lapisan *Dense* yang berfungsi mengintegrasikan fitur-fitur yang telah diekstraksi menjadi representasi yang lebih abstrak. Lapisan *Dense* pertama memiliki 256 neuron, diikuti dengan lapisan yang lebih kecil dengan 128 dan 64 neuron, yang secara bertahap mereduksi dimensi data untuk meningkatkan efisiensi dalam pembelajaran. Pada lapisan output, model menggunakan fungsi aktivasi *Softmax* dengan 8 neuron, yang sesuai dengan jumlah kelas maqam bacaan Al-Qur'an yang diklasifikasikan. Model ini dilatih menggunakan data audio dengan optimasi berdasarkan *categorical crossentropy* dan *optimizers* seperti *Adam*, untuk mencapai akurasi yang optimal dalam klasifikasi multi-kelas. Tabel 2 menunjukkan parameter-parameter yang digunakan dalam arsitektur CNN.

Tabel 2. Parameter arsitektur CNN

Layer Type	Parameter	Value
Input Layer	Input Shape	(height, width, channels)
Layer 1: Conv2D	Jumlah Filter	32
	Ukuran Kernel	(3, 3)
	Aktivasi	ReLU
Layer 1: MaxPooling2D	Ukuran Pooling	(2, 2)
Conv2D Layer 2	Jumlah Filter	64
	Ukuran Kernel	(3, 3)
	Aktivasi	ReLU
MaxPooling2D Layer 2	Ukuran Pooling	(2, 2)
Conv2D Layer 3	Jumlah Filter	128
	Ukuran Kernel	(3, 3)
	Aktivasi	ReLU
MaxPooling2D Layer 3	Ukuran Pooling	(2, 2)
Flatten Layer	-	-
Dense Layer 1	Jumlah Neuron	256
	Aktivasi	ReLU
	Dropout	0.5
Dense Layer 2	Jumlah Neuron	128
	Aktivasi	ReLU
	Dropout	0.5
Output Layer	Jumlah Neuron	8
	Aktivasi	Softmax
Optimizer	Tipe	Adam
Loss Function	Tipe	Crossentropy
Metrics	-	Accuracy

Evaluasi

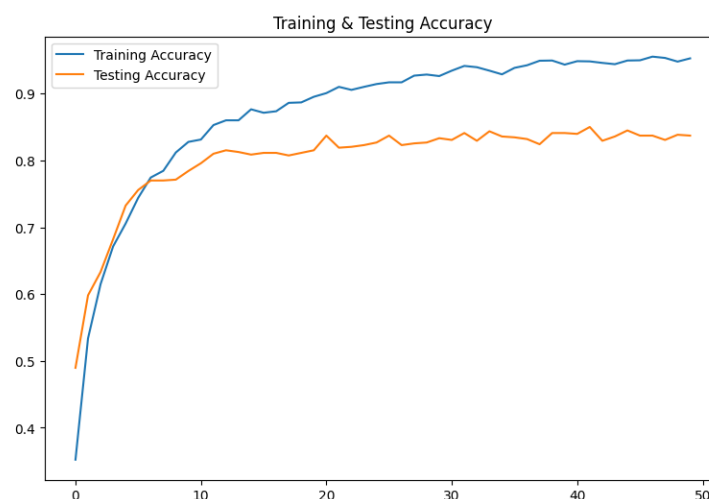
Evaluasi hasil pengujian dilakukan untuk menilai kinerja sistem klasifikasi pada irama Tilawah Al-Qur'an. Dalam penelitian ini, evaluasi sistem dilakukan menggunakan *confusion matrix*. *Confusion matrix* adalah tabel matriks yang menunjukkan performa model klasifikasi pada data uji yang nilai aktualnya sudah diketahui. Tabel ini terdiri dari empat kombinasi antara prediksi dan nilai aktual, yaitu: TP (*true positive*), yang menandakan data positif yang diprediksi dengan benar; TN (*true negative*), untuk data negatif yang diprediksi dengan tepat; FP (*false positive*), yang mengindikasikan data negatif namun diprediksi sebagai positif; dan FN (*false negative*), di mana data positif diprediksi sebagai negatif. Berdasarkan *confusion matrix* ini, dapat dihitung beberapa metrik kinerja, seperti akurasi, presisi, *recall* dan *f1-score*.

HASIL DAN PEMBAHASAN

Dalam penelitian ini, model CNN digunakan untuk melakukan klasifikasi irama tilawah Al-Qur'an. Hasil dari proses pelatihan dan pengujian model dapat dianalisis berdasarkan metrik akurasi dan loss, baik untuk data pelatihan (*training*) maupun validasi (*testing*).

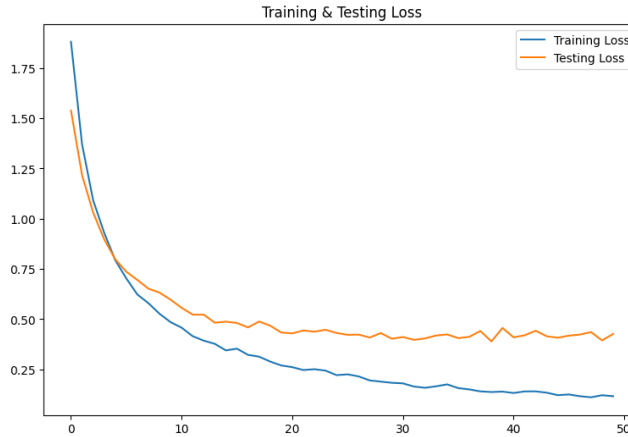
Hasil Pelatihan dan Pengujian Model

Hasil Akurasi pada proses pelatihan data dengan arsitektur yang dibuat mencapai nilai 0.9262, yang berarti bahwa model mampu mengklasifikasikan 92.62% dari data pelatihan dengan benar. Hasil ini menunjukkan model telah mempelajari berbagai pola irama tilawah yang relevan dari data latih yang diberikan dan mampu melakukan klasifikasi dengan tingkat kesalahan yang rendah. Hasil akurasi yang tinggi pada proses pelatihan menandakan bahwa model telah melakukan *fitting* yang baik terhadap data, tanpa mengalami *underfitting* yang biasanya ditandai dengan akurasi yang rendah. Hasil proses pelatihan dapat dilihat pada gambar 4. Akurasi pada data pengujian adalah 0.8204, yang berarti bahwa model mampu mengklasifikasikan 82.04% dari data pengujian validasi dengan benar. Perbedaan akurasi antara data pelatihan 92.62% dan data validasi 82.04% menunjukkan masih ada kesenjangan performa antara pelatihan dan pengujian.



Gambar 4. Hasil Akurasi Proses Training dan Testing

Gambar 5 menunjukkan hasil nilai loss pada proses pelatihan mendapatkan nilai sebesar 0.1940 menunjukkan bahwa kesalahan rata-rata yang dibuat oleh model dalam memprediksi kelas irama tilawah pada data pelatihan cukup rendah. Loss yang rendah pada data pelatihan ini sejalan dengan akurasi yang tinggi, mengindikasikan bahwa model mampu mengurangi kesalahan secara efektif selama proses pelatihan.

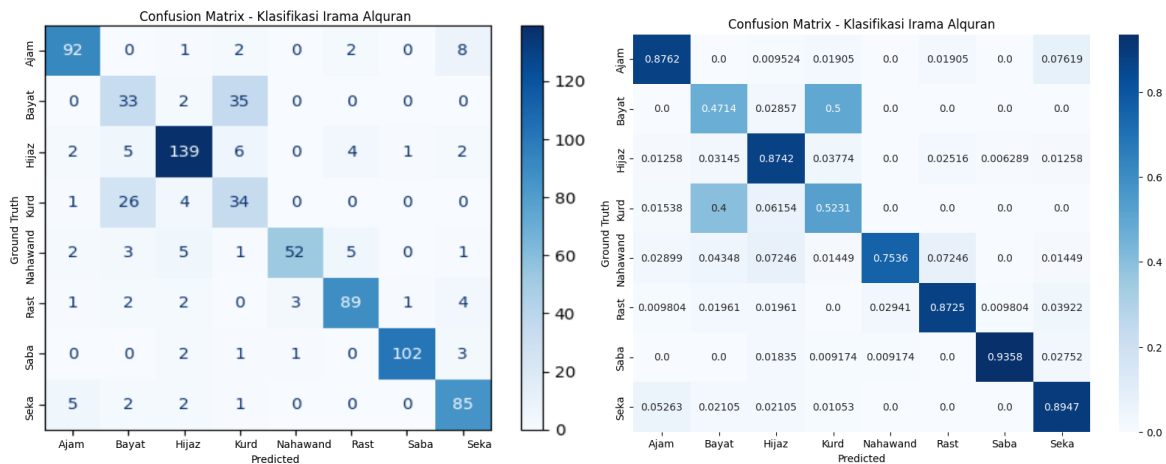


Gambar 5. Hasil nilai loss proses *training* dan *testing*

Untuk nilai *loss* pada data Testing didapatkan hasil sebesar 0.4636 lebih tinggi dibandingkan dengan loss pada data pelatihan (0.1940). Nilai *loss* pada data testing atau pengujian mengalami peningkatan loss, hal ini menguatkan indikasi model mengalami sedikit *overfitting*, di mana model melakukan kesalahan yang lebih besar saat berhadapan dengan data yang pernah dikenali selama pelatihan. *Loss* yang lebih tinggi pada data validasi dibandingkan dengan data pelatihan adalah indikasi bahwa model mungkin telah belajar terlalu banyak dari detail atau *noise* dalam data pelatihan, yang tidak berlaku untuk data baru.

Pembahasan Evaluasi Kinerja Model

Model dievaluasi melalui validasi pengujian hasil pelatihan. Proses validasi dilakukan dengan memprediksi 774 set data yang belum pernah diakses sebelumnya. Hasil prediksi dalam tahap validasi ini dapat dilihat menggunakan matriks *convusion*.



Gambar 6. Hasil Confusion Matrik

Gambar 6 memperlihatkan bahwa hampir semua prediksi sesuai dengan label yang benar. Model berhasil memprediksi 626 data sesuai kelas dari total 774 data yang dicatat dengan akurasi yang baik. Hasil *confusion matrix* didapatkan nilai akurasi yang diperoleh dalam proses validasi mencapai 80,88%. Nilai presisi sebesar 0,82, nilai *recall* 0,80 dan *F1-score* sebesar 0,81. Dari *confusion matrix* ini, terlihat bahwa banyak nilai diagonal yang relatif besar dibandingkan dengan nilai di luar diagonal, yang menunjukkan bahwa model memiliki

performa cukup baik. Akan tetapi, terdapat beberapa kelas yang lebih sering keliru diprediksi, seperti Bayat dan Kurd yang terlihat cukup sering diklasifikasikan sebagai kelas lain. Kelas Hijaz memiliki precision yang cukup tinggi, dengan banyak prediksi yang benar (139) dan sedikit prediksi keliru. Sebaliknya, kelas Kurd memiliki precision rendah karena terdapat cukup banyak data dari kelas ini yang diklasifikasikan sebagai kelas lain khususnya ke kelas Bayat (26). *F1-score* penting untuk kelas dengan distribusi yang tidak seimbang. Dari analisis matriks, terlihat bahwa kelas Hijaz dan Saba memiliki F1-Score yang tinggi karena keduanya menunjukkan *precision* dan *recall* yang tinggi. Sementara itu, kelas Bayat dan Kurd mungkin memiliki *F1-score* yang lebih rendah karena nilai *precision* dan *recall* yang tidak optimal akibat banyaknya prediksi yang salah.

Secara keseluruhan, model menunjukkan performa yang baik pada beberapa kelas tertentu, seperti Hijaz dan Saba, namun perlu perbaikan pada kelas lain seperti Bayat dan Kurd. Ini menunjukkan bahwa model masih memiliki kesulitan dalam membedakan beberapa maqam yang mungkin memiliki karakteristik audio yang mirip. Strategi tambahan seperti peningkatan jumlah data untuk kelas tertentu atau penyesuaian arsitektur model dapat membantu meningkatkan performa di masa mendatang

KESIMPULAN

Berdasarkan hasil tersebut, dapat disimpulkan bahwa model CNN yang digunakan menunjukkan performa yang baik selama proses pelatihan, dengan akurasi yang tinggi dan loss yang rendah. Namun, penurunan akurasi dan peningkatan *loss* pada data validasi menunjukkan bahwa model sedikit mengalami *overfitting*. Ini adalah masalah umum dalam pembelajaran mendalam (*deep learning*), terutama ketika model menjadi terlalu kompleks dan mulai belajar pola-pola spesifik dari data pelatihan yang tidak berlaku secara umum. Hasil evaluasi model dengan matriks *confusion* menunjukkan keberhasilan model dalam mengklasifikasikan 80,88% data yang terbagi dalam 8 kelas.

REFERENSI

- [1] V. Y. Mafula, A. C. Fauzan, and T. R. Fernando, "Identifikasi Irama Tilawah al-Quran dengan Gaya Mujawwad Menggunakan Naive Bayes Classifier," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 4, no. 2, pp. 242–251, 2022, doi: 10.28926/ilkomnika.v4i2.464.
- [2] S. Shahriar and U. Tariq, "Classifying Maqams of Qur'anic Recitations Using Deep Learning," *IEEE Access*, vol. 9, pp. 117271–117281, 2021, doi: 10.1109/ACCESS.2021.3098415.
- [3] H. A. Annuri, *Panduan Tahsin Tilawah Al-qur'an & Ilmu Tajwid*. Pustaka Al-Kautsar, 2020.
- [4] A. Chairunisa, F. Ramadhan, M. Zulkifli, and others, "Konsep Qira'ah dan Tilawah Menurut Al-Qur'an," *Al-Dirosah J. Pendidik. Agama Islam*, vol. 1, no. 01, 2024.
- [5] M. M. Al Anazi and O. R. Shahin, "A Machine Learning Model for the Identification of the Holy Quran Reciter Utilizing K-Nearest Neighbor and Artificial Neural Networks," *Inf. Sci. Lett.*, vol. 11, no. 4, pp. 1093–1102, 2022, doi: 10.18576/isl/110410.
- [6] D. Omran, A. Kandil, A. ElBialy, S. Samy, and S. fawzy, "CNN for speech recognition case study: Recitation Rules of the holy Quran," *MSA Eng. J.*, vol. 0, no. 0, pp. 0–0, 2023, doi: 10.21608/msaeng.2023.225120.1335.
- [7] S. S. Alrumiah and A. A. Al-Shargabi, "Intelligent Quran Recitation Recognition and Verification: Research Trends and Open Issues," *Arab. J. Sci. Eng.*, vol. 48, no. 8, pp. 9859–9885, 2023, doi: 10.1007/s13369-022-07273-8.
- [8] E. A. S. El-Dahshan, H. M. Mohsen, K. Revett, and A. B. M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5526–5545, 2014, doi: 10.1016/j.eswa.2014.01.021.
- [9] A. Al Harere and K. Al Jallad, "Quran Recitation Recognition using End-to-End Deep Learning," pp. 1–22, 2023, [Online]. Available: <http://arxiv.org/abs/2305.07034>.
- [10] R. Hadiyansah and R. Andamira, "Convolutional Neural Network (CNN) for Detecting Al-Qur'an Reciting and Memorizing," *Khazanah J. Relig. Technol.*, vol. 1, no. 2, pp. 44–48, 2023, doi: 10.15575/kjrt.v1i2.235.

- [11] M. Faizan, M. S. Arif, J. N. Chattha, and F. A. Butt, "Mel-Frequency-based Feature Analysis of Audio Signals in the Context of Holy Quran Recitation," *Arab. J. Sci. Eng.*, vol. 49, no. 5, pp. 6971–6979, 2024, doi: 10.1007/s13369-023-08555-5.
- [12] F. Omari, "3869Maqam_dataset.zip.," *figshare*. 2023, doi: <https://doi.org/10.6084/m9.figshare.23837865.v1>.
- [13] F. Omari, M. Ghantous, and N. Peleg, "Maqam Classification of Quranic Recitations using Deep Learning," 2023, doi: <https://doi.org/10.6084/m9.figshare.24131781.v1>.
- [14] A. M. Alagrami and M. M. Eljazzar, "SMARTAJWEED Automatic Recognition of Arabic Quranic Recitation Rules," no. March, pp. 145–152, 2020, doi: 10.5121/csit.2020.101812.
- [15] G. Samara, E. Al-Daoud, N. Swerki, D. Alzu'bi, and Z. Hu, "The Recognition of Holy Qur'an Reciters Using the MFCCs' Technique and Deep Learning," *Adv. Multimed.*, vol. 2023, Jan. 2023, doi: 10.1155/2023/2642558.
- [16] J. H. Alkhateeb, "A machine learning approach for recognizing the Holy Quran reciter," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 268–271, 2020, doi: 10.14569/IJACSA.2020.0110735.
- [17] R. Sharan, H. Xiong, and S. Berkovsky, "Benchmarking Audio Signal Representation Techniques for Classification with Convolutional Neural Networks," *Sensors*, vol. 21, p. 3434, 2021, doi: 10.3390/s21103434.
- [18] H. H. Tan and K. H. Lim, "Vanishing Gradient Mitigation with Deep Learning Neural Network Optimization," in *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, 2019, pp. 1–4, doi: 10.1109/ICSCC.2019.8843652.
- [19] X. Ying, "An Overview of Overfitting and its Solutions," *J. Phys. Conf. Ser.*, vol. 1168, no. 2, 2019, doi: 10.1088/1742-6596/1168/2/022022.
- [20] M. Faizan, M. S. Arif, J. N. Chattha, and F. A. Butt, "Mel-Frequency-based Feature Analysis of Audio Signals in the Context of Holy Quran Recitation," *Arab. J. Sci. Eng.*, vol. 49, no. 5, pp. 6971–6979, 2024, doi: 10.1007/s13369-023-08555-5.