

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kemampuan membaca Al-Quran merupakan hal yang fundamental dalam kehidupan umat Muslim. Namun, kenyataannya masih banyak masyarakat Indonesia yang belum mampu membaca Al-Qur'an dengan baik dan benar. Hal ini tidak hanya terbatas pada ketidakmampuan membaca teks Arab, tetapi juga mencakup aspek makharij al-huruf (tempat keluarnya huruf), sifat al-huruf, hukum-hukum bacaan (ahkam al-huruf), serta panjang dan pendeknya bacaan (al-mad wa al-qashr).

Berdasarkan hasil riset yang dipaparkan oleh Tim IIQ Jakarta di Gedung DPR/MPR RI, ditemukan bahwa dari 3.111 responden yang tersebar di seluruh Indonesia, sebanyak 72,25% berada pada level cukup dan kurang dalam kemampuan membaca Al-Qur'an. Penilaian ini dilakukan berdasarkan empat parameter penting dalam ilmu tajwid. Temuan ini menunjukkan bahwa mayoritas masyarakat masih memerlukan bantuan berupa alat bantu evaluasi pelafalan yang mendukung peningkatan kemampuan baca Al-Qur'an [1].

Lebih lanjut, dalam sebuah penelitian oleh Effendi, disebutkan bahwa terdapat lebih dari 80% anak usia sekolah yang belum memiliki kemampuan membaca Al-Quran, meskipun pada usia sekolah dasar mereka seharusnya sudah mampu melakukannya. Angka ini menunjukkan bahwa buta huruf Al-Quran dan hijaiyah masih menjadi persoalan yang signifikan di kalangan generasi muda, terutama pada tahap pendidikan dasar [2].

Melihat fakta tersebut, dibutuhkan suatu pendekatan evaluasi pelafalan yang lebih objektif, konsisten, dan dapat diakses secara mandiri terhadap pelafalan huruf hijaiyah. Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI), khususnya di bidang pengolahan suara (speech processing), membuka peluang besar untuk menghadirkan sistem penilaian pelafalan yang inovatif dan objektif.

Penelitian terbaru yang dilakukan oleh Fawwaz Ijlal Muqsith berjudul "Klasifikasi Pengucapan Huruf Hijaiyah Berbasis Android Menggunakan CNN dengan Fitur Mel-Spectrogram" telah berhasil mengembangkan sistem klasifikasi pelafalan huruf hijaiyah menggunakan *Convolutional neural network* (CNN) dengan ekstraksi fitur Mel-Frequency Cepstral Coefficients (MFCC) pada dataset 8.904 sampel yang mencakup 84 kelas ($28 \text{ huruf} \times 3 \text{ harakat}$). Sistem tersebut diimplementasikan dalam aplikasi Android menggunakan backend Flask API dan menghasilkan rata-rata akurasi keseluruhan sebesar 83,80% [3].

Meskipun hasil tersebut cukup menjanjikan, penelitian Fawwaz masih menunjukkan keterbatasan yang signifikan pada kelas-kelas huruf dengan kemiripan fonetik tinggi, seperti *alif_dommah* (34,88%) dan *ghaiin_dommah* (39,53%). Hal ini mengindikasikan bahwa fitur MFCC, meskipun efektif secara umum, memiliki keterbatasan dalam merepresentasikan pola spektral yang cukup diskriminatif untuk membedakan huruf-huruf tersebut. Selain itu, penggunaan arsitektur CNN yang dibangun dari awal (*from scratch*) berpotensi menghasilkan kemampuan generalisasi yang terbatas pada dataset berukuran sedang. Kondisi ini menunjukkan perlunya eksplorasi metode ekstraksi fitur yang lebih kaya informasi temporal-frekuensi, seperti Mel Frequency Cepstral Coefficient, serta pemanfaatan arsitektur yang telah terlatih melalui *transfer learning*, seperti **MobileNetV2**, untuk meningkatkan akurasi klasifikasi pelafalan huruf hijaiyah.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan yang akan dibahas dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana mengintegrasikan MFCC dan **MobileNetV2** untuk mengklasifikasikan huruf hijaiyah yang diucapkan pengguna?
2. Bagaimana pengaruh penggunaan Log-Mel Spectrogram sebagai alternatif MFCC terhadap performa model **MobileNetV2** dalam klasifikasi huruf hijaiyah?

3. Bagaimana performa model dalam mengklasifikasikan huruf hijaiyah secara akurat?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk merancang dan mengimplementasikan sistem penilaian pelafalan huruf hijaiyah dengan memanfaatkan teknologi pengenalan suara berbasis kecerdasan buatan. Secara khusus, penelitian ini bertujuan untuk :

1. Mengintegrasikan MFCC dan **MobileNetV2** untuk membangun model klasifikasi pelafalan huruf hijaiyah
2. Membandingkan performa ekstraksi fitur MFCC dan Log-Mel Spectrogram pada arsitektur **MobileNetV2** yang identik untuk menentukan representasi fitur yang lebih optimal.
3. Mengevaluasi performa model klasifikasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*

1.4. Batasan Masalah

Agar penelitian ini terarah dan sesuai dengan tujuan yang ingin dicapai, maka ruang lingkup penelitian dibatasi pada beberapa hal sebagai berikut:

1. Penelitian hanya menilai pelafalan huruf hijaiyah tunggal, tidak mencakup kata, ayat, atau aturan tajwid lanjutan.
2. Penelitian ditujukan untuk anak remaja (SMP/SMA) sebagai target utama pengguna yang belajar mengenal huruf hijaiyah.
3. Ekstraksi fitur suara dilakukan menggunakan *Mel-Frequency Cepstral Coefficients (MFCC)*.
4. Model yang digunakan adalah **MobileNetV2** sebagai arsitektur CNN utama untuk proses penilaian.
5. Output sistem berupa label kelas huruf hijaiyah yang dikenali dari suara pengguna, bukan skor numerik.

6. Aplikasi hanya berfungsi untuk pengambilan suara dan menampilkan hasil penilaian, tanpa fokus pada desain antarmuka yang kompleks.
7. Sistem dapat diakses melalui perangkat web dan *mobile* berbasis Flutter, namun optimasi performa aplikasi *mobile* bukan fokus utama penelitian.

1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

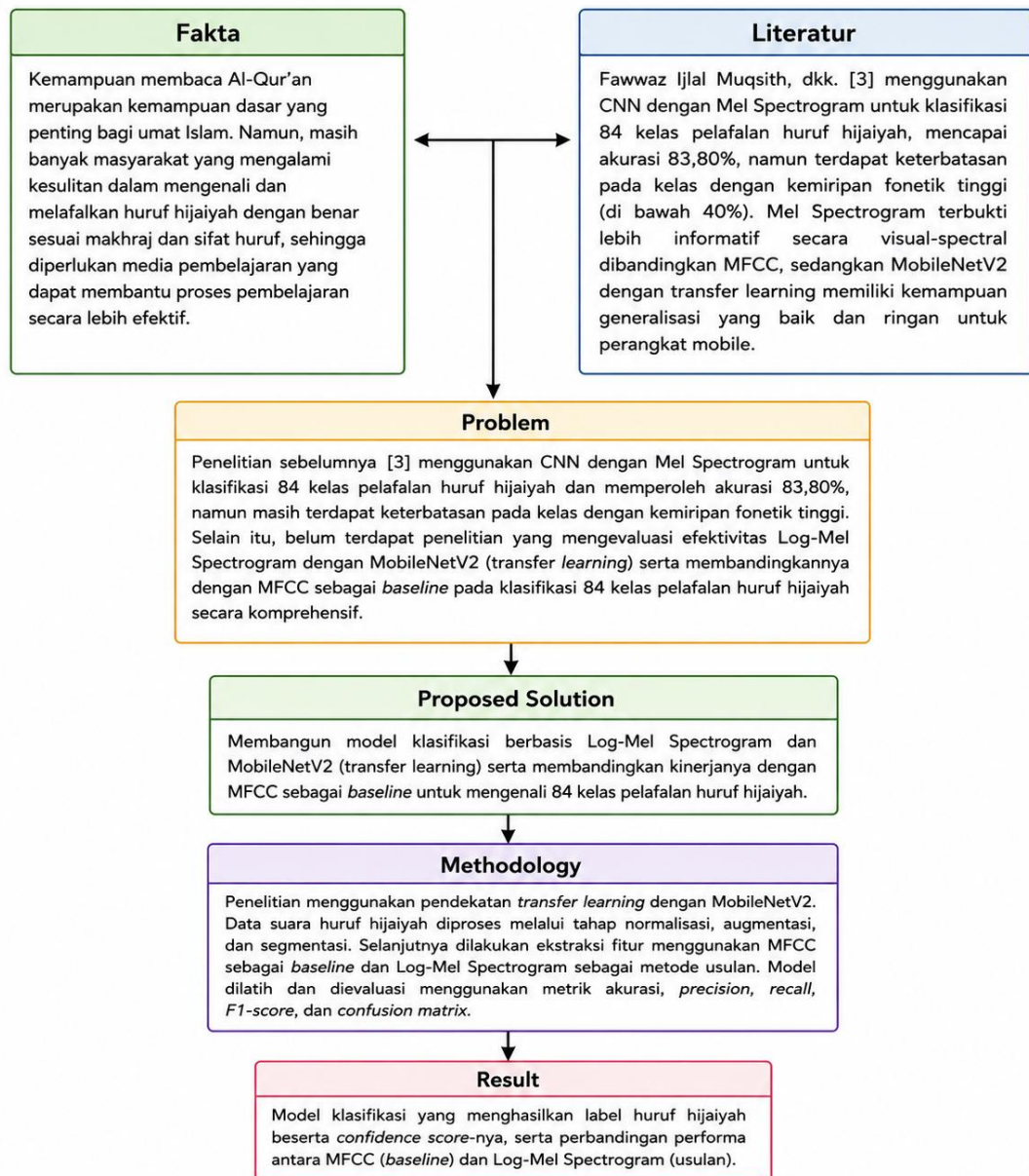
1. Manfaat Teoritis

- Memberikan kontribusi ilmiah dalam penerapan ekstraksi fitur MFCC dan arsitektur **MobileNetV2** pada sistem penilaian pelafalan suara.
- Menjadi referensi bagi penelitian terkait pengenalan suara huruf hijaiyah dan evaluasi pelafalan berbasis deep learning.

2. Manfaat Praktis

- Membantu anak remaja awal dalam belajar mengenal dan melafalkan huruf hijaiyah dengan benar melalui sistem penilaian otomatis.
- Menjadi alat bantu bagi guru SMP/SMA atau pengajar Al-Qur'an dalam memberikan evaluasi pelafalan yang lebih objektif.
- Menjadi contoh implementasi praktis kecerdasan buatan dalam bidang pendidikan keagamaan, khususnya pada pembelajaran makharijul huruf.

1.6. Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini berawal dari fakta bahwa kemampuan membaca Al-Qur'an merupakan keterampilan dasar yang penting bagi umat Islam. Namun, masih banyak peserta didik tingkat Madrasah Tsanawiyah (MTs) dan Madrasah Aliyah (MA) yang mengalami kesulitan dalam mengenali dan melafalkan huruf hijaiyah sesuai dengan makhraj dan sifat huruf yang benar.

Kesulitan tersebut terutama terjadi pada huruf-huruf yang memiliki kemiripan bunyi sehingga sering menimbulkan kesalahan pelafalan. Selain itu, proses evaluasi pelafalan masih bergantung pada penilaian guru atau orang tua sehingga hasil penilaian dapat bersifat subjektif dan tidak selalu konsisten. Oleh karena itu, diperlukan sistem evaluasi pelafalan berbasis teknologi yang mampu memberikan umpan balik secara objektif, konsisten, dan dapat digunakan secara mandiri untuk membantu proses penilaian pelafalan huruf hijaiyah.

Berdasarkan kajian literatur, penelitian yang dilakukan oleh Fawwaz Ijlal Muqsith [3] berhasil mengembangkan sistem klasifikasi pelafalan huruf hijaiyah menggunakan kombinasi *Convolutional neural network* (CNN) dan representasi fitur Mel Spectrogram untuk 84 kelas pelafalan huruf hijaiyah. Penelitian tersebut memperoleh akurasi sebesar 83,80%, namun masih menunjukkan keterbatasan pada beberapa kelas huruf yang memiliki kemiripan fonetik tinggi dengan tingkat akurasi di bawah 40%. Keterbatasan serupa juga terlihat pada penelitian Gerhana, dkk. [4] yang menerapkan kombinasi ekstraksi fitur MFCC dan algoritma CNN untuk mengklasifikasikan 28 huruf hijaiyah, namun hanya memperoleh akurasi sebesar 62,45% akibat kemiripan fonetik antarhuruf serta keterbatasan arsitektur CNN yang dibangun dari awal (from scratch) dalam mengekstraksi pola spektral yang kompleks. Di sisi lain, pada domain pengenalan dialek bahasa Arab dengan kondisi data terbatas (low-resource), penelitian Al-Shwayyat dan Gerek [5] menunjukkan bahwa kombinasi MFCC dan CNN mampu mencapai akurasi hingga 91,2%, jauh mengungguli kombinasi Discrete Wavelet Transform dan Recurrent Neural Network yang hanya mencapai 66,5%. Temuan tersebut memperkuat bukti bahwa MFCC merupakan representasi fitur yang cukup kuat ketika dipadukan dengan arsitektur CNN, namun ketiga penelitian di atas juga menunjukkan bahwa performa kombinasi tersebut masih dapat ditingkatkan, terutama pada kelas-kelas dengan kemiripan fonetik tinggi. Selain itu, berbagai penelitian menunjukkan bahwa representasi suara menggunakan Log-Mel Spectrogram mampu mempertahankan informasi spektral yang lebih kaya dibandingkan MFCC karena mampu merepresentasikan pola frekuensi suara secara lebih detail. Selain itu,

arsitektur **MobileNetV2** yang memanfaatkan *transfer learning* terbukti memiliki kemampuan generalisasi yang baik, ukuran model yang ringan, serta efisien untuk diimplementasikan pada perangkat *mobile*.

Meskipun demikian, belum terdapat penelitian yang secara khusus membandingkan performa MFCC sebagai *baseline* dengan Log-Mel Spectrogram sebagai metode usulan menggunakan arsitektur **MobileNetV2** berbasis *transfer learning* pada klasifikasi 84 kelas pelafalan huruf hijaiyah. Baik penelitian Gerhana, dkk. [4] maupun Al-Shwayyat dan Gerek [5] juga belum mengeksplorasi representasi Log-Mel Spectrogram maupun arsitektur berbasis *transfer learning* seperti **MobileNetV2**, sehingga potensi peningkatan performa melalui kedua pendekatan tersebut belum terungkap pada domain pelafalan huruf hijaiyah. Selain itu, efektivitas penggunaan Log-Mel Spectrogram dalam meningkatkan kemampuan model untuk membedakan huruf-huruf yang memiliki kemiripan fonetik tinggi masih perlu dievaluasi lebih lanjut. Kesenjangan penelitian tersebut menunjukkan adanya peluang untuk mengembangkan sistem klasifikasi yang lebih optimal melalui pemanfaatan representasi fitur yang lebih informatif dan model klasifikasi yang memiliki kemampuan generalisasi yang baik.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan pengembangan sistem klasifikasi pelafalan huruf hijaiyah menggunakan Log-Mel Spectrogram sebagai metode usulan dan MFCC sebagai *baseline*. Kedua representasi fitur tersebut digunakan sebagai masukan bagi model **MobileNetV2** berbasis *transfer learning* untuk dibandingkan performanya dalam mengenali 84 kelas pelafalan huruf hijaiyah. Dengan pendekatan ini, penelitian diharapkan dapat memberikan gambaran mengenai representasi fitur yang paling efektif dalam meningkatkan performa klasifikasi pelafalan huruf hijaiyah.

Penelitian ini menggunakan metodologi CRISP-DM yang terdiri dari tahapan *Business Understanding*, *Data understanding*, *Data preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Pada tahap *Data preparation* dilakukan proses normalisasi, segmentasi, augmentasi data, serta ekstraksi fitur menggunakan

MFCC sebagai *baseline* dan Log-Mel Spectrogram sebagai metode usulan. Selanjutnya, kedua representasi fitur digunakan sebagai masukan bagi model **MobileNetV2** yang dilatih menggunakan pendekatan *transfer learning*. Kinerja model kemudian dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, serta *confusion matrix* untuk mengetahui kemampuan model dalam mengklasifikasikan setiap kelas pelafalan huruf hijaiyah. Model terbaik yang dihasilkan selanjutnya diintegrasikan ke dalam aplikasi *mobile* berbasis Flutter sebagai media implementasi sistem.

Hasil yang diharapkan dari penelitian ini adalah diperolehnya evaluasi yang komprehensif mengenai performa MFCC dan Log-Mel Spectrogram pada klasifikasi 84 kelas pelafalan huruf hijaiyah menggunakan **MobileNetV2** berbasis *transfer learning*. Selain itu, penelitian ini diharapkan menghasilkan model klasifikasi dengan performa yang baik dalam mengenali pelafalan huruf hijaiyah serta menghasilkan prototipe aplikasi *mobile* yang mampu memberikan prediksi huruf hijaiyah beserta nilai confidence score secara otomatis. Dengan demikian, sistem yang dikembangkan dapat dimanfaatkan sebagai alat bantu evaluasi dan latihan mandiri untuk menilai pelafalan huruf hijaiyah secara lebih objektif dan konsisten.

1.7. Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun untuk memberikan gambaran mengenai isi dan alur pembahasan penelitian secara terstruktur. Adapun sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab I berisi gambaran umum penelitian yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, kerangka pemikiran, serta sistematika penulisan yang digunakan dalam penyusunan skripsi.

BAB II KAJIAN LITERATUR

Bab II membahas landasan teori dan kajian pustaka yang mendukung penelitian. Pada bab ini dijelaskan konsep-konsep yang berkaitan dengan *Convolutional neural network* (CNN), **MobileNetV2**, Mel-Frequency Cepstral Coefficients (MFCC), *transfer learning*, teknik regularisasi, *data augmentation*, huruf hijaiyah, *confusion matrix*, serta metodologi CRISP-DM yang digunakan dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab III menjelaskan metodologi penelitian yang digunakan, meliputi metode penelitian, kebutuhan perangkat keras dan perangkat lunak, pengumpulan dan pemahaman data, tahapan pengolahan data, ekstraksi fitur menggunakan MFCC, perancangan dan pelatihan model **MobileNetV2**, evaluasi model, serta integrasi model ke dalam aplikasi *mobile* berbasis Flutter sesuai tahapan CRISP-DM.

BAB IV HASIL DAN PEMBAHASAN

Bab IV memaparkan hasil implementasi dan pengujian sistem penilaian pelafalan huruf hijaiyah yang telah dikembangkan. Pada bab ini dibahas hasil pelatihan model, evaluasi performa menggunakan metrik yang telah ditentukan, analisis hasil pengujian, serta implementasi model pada aplikasi *mobile*.

BAB V PENUTUP

Bab V berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.

DAFTAR PUSTAKA

Bagian ini memuat seluruh sumber referensi yang digunakan sebagai landasan teori dan pendukung dalam penyusunan penelitian.

LAMPIRAN

Bagian lampiran berisi dokumen pendukung penelitian seperti dataset, hasil pengujian, konfigurasi model, source code, dokumentasi implementasi sistem, serta dokumen lain yang mendukung penelitian.

