

Penerapan Teknologi *Automatic Speech Recognition* Menggunakan Model *Wav2vec2.0* Sebagai Alat Bantu Untuk Mendeteksi Kesalahan Dalam Membaca Al-Qur'an Berbasis *Mobile*

Muhammad Ihsanudin Syaifullah¹, Ghufron^{2*}

^{1,2}Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Sultan Agung
e-mail: ¹magetan.san154@gmail.com, ²ghufron@unissula.ac.id

**Corresponding Author*

Abstrak

Membaca Al-Qur'an dengan benar adalah suatu kewajiban bagi umat Islam. Namun dalam praktiknya, seringkali terjadi kesalahan dalam pembacaan Al-Qur'an yang dapat mengubah makna dari ayat-ayat suci tersebut dan kebanyakan orang tidak sadar jika ada kesalahan dalam bacaan mereka. Penelitian ini bertujuan untuk mengembangkan aplikasi mobile yang memanfaatkan teknologi Automatic Speech Recognition (ASR) untuk membantu pengguna mengecek bacaan mereka. Model Wav2vec2.0 adalah sebuah model self-supervised learning yang dirancang khusus untuk tugas pemahaman ucapan, dipilih karena kemampuannya yang kuat dalam menangkap representasi audio yang mendalam tanpa memerlukan data berlabel dalam jumlah besar. Model Wav2vec2.0 digunakan untuk menghasilkan transkripsi bacaan Al-Qur'an secara otomatis. Dengan membandingkan transkripsi dengan teks referensi, sistem dapat mengidentifikasi jenis kesalahan bacaan, seperti kesalahan pada harakat dan huruf. Hasil penelitian menunjukkan bahwa model ini mampu mencapai Word Error Rate sebesar 0.367483 dalam mengenali kata-kata dan harakat dalam Al-Qur'an. Hal ini mengindikasikan potensi besar dalam pengembangan alat bantu deteksi kesalahan baca Al-Qur'an yang baik.

Kata Kunci : *Al-Qur'an, Automatic Speech Recognition, Wav2vec2.0, Deteksi Kesalahan Bacaan, Aplikasi Mobile.*

Abstract

Reading the Qur'an correctly is an obligation for Muslims. However, in practice, there are often errors in the recitation of the Qur'an that can change the meaning of the holy verses and most people do not realize if there are errors in their reading. This research aims to develop a mobile application that utilizes Automatic Speech Recognition (ASR) technology to help users check their recitation. The Wav2vec2.0 model is a self-supervised learning model specifically designed for speech understanding tasks, chosen for its strong ability to capture deep audio representations without requiring large amounts of labeled data. The Wav2vec2.0 model is used to automatically generate transcriptions of Qur'anic recitations. By comparing the transcription with a reference text, the system can identify types of reading errors, such as errors in harakat and letters. The results showed that the model was able to achieve a Word Error Rate of 0.367483 in recognizing words and harakat in the Qur'an. This indicates great potential in the development of a good Qur'an reading error detection tool.

Key Word : *Qur'an, Automatic Speech Recognition, Wav2vec2.0, Reading Error Detection, Mobile Application.*

1. PENDAHULUAN

Membaca Al-Qur'an merupakan salah satu ibadah dan hal penting dalam kehidupan umat Islam. Membaca Al-Qur'an dengan benar dan sesuai dengan aturan tajwid adalah suatu kewajiban bagi setiap Muslim. Hal ini dikarenakan, perbedaan kecil dalam suara dalam sebuah kata Arab dapat menghasilkan makna yang berbeda [1]. Namun dalam praktiknya, seringkali terjadi kesalahan dalam pembacaan Al-Qur'an yang dapat mengubah makna dari ayat-ayat suci tersebut. Menurut penelitian yang dilakukan oleh Muhammad Salamun Asngari dan Adis Alena menyebutkan bahwa umat Islam Indonesia sangat sedikit yang mampu dan mahir dalam membaca Al-Qur'an dengan baik dan benar. Persentase umat Islam yang mampu membaca Al-Qur'an dengan baik dan benar yaitu sekitar 20% dari seluruh total jumlah umat islam yang ada [2]. Berdasarkan hasil survei terbaru Ditjen Bimas Islam Kementerian Agama yang dilakukan tahun 2023 menyebutkan, skor Indeks Literasi Al-Qur'an di Indonesia berada di angka 66,038. Survei menunjukkan bahwa sebanyak 61,51% responden mampu mengenali huruf dan *harakat* Al-Qur'an, 59,92% mampu membaca susunan huruf menjadi kata, 48,96% mampu membaca ayat dengan lancar, dan 44,57% dapat membaca Al-Qur'an dengan lancar sesuai tajwid. Sedangkan responden yang belum memiliki literasi baca Al-Qur'an adalah sebesar 38,49% [3].

Adapun beberapa permasalahan yang menyebabkan umat islam Indonesia mengalami kesulitan dalam membaca Al-Quran dengan baik dan benar adalah tidak adanya majelis pembelajaran BTQ (Baca Tulis Al-Qur'an) di tempat tinggal mereka. Jika pun ada, kebanyakan dari mereka tidak pernah mengikuti majelis pembelajaran BTQ di tempat tinggalnya[3]. Kemudian permasalahan selanjutnya adalah media yang digunakan untuk pembelajaran saat ini masih berupa buku materi sehingga kurang variatif dan belum efektif, dan juga masih sulitnya untuk mencari ustadz dan ustadzah yang mengajarkan Al-Qur'an dengan sertifikat sanad [4].

Salah satu cara untuk mengatasi permasalahan tersebut adalah dengan memanfaatkan teknologi, saat ini terdapat beberapa perangkat lunak pembelajaran Al-Qur'an. Namun, terdapat kekurangan dimana pengucapan dan bacaan mereka tidak dapat diperbaiki seperti yang dilakukan oleh seorang guru atau ustadz, karena perangkat lunak atau aplikasi yang ada saat ini tidak dapat mendeteksi kesalahan dalam bacaan mereka [1]. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan suatu sistem yang memanfaatkan teknologi *Automatic Speech Recognition (ASR)* sebagai alat bantu dalam mendeteksi kesalahan bacaan Al-Qur'an melalui perangkat mobile. *Automatic Speech Recognition (ASR)* adalah salah satu pemrosesan ucapan atau suara yang memungkinkan mesin untuk memahami ucapan pengguna dan mengubahnya menjadi serangkaian kata melalui program komputer, sehingga menciptakan semacam komunikasi alami antara manusia dan mesin [5]. Sedangkan *Wav2vec2.0* merupakan model pengenalan ucapan otomatis (*ASR*) yang dapat mengidentifikasi sinyal audio dan mengubahnya menjadi teks dengan akurasi tinggi [6].

Berdasarkan beberapa penelitian sebelumnya yang dilakukan terhadap objek bacaan Al-Qur'an, Penerapan model dengan menggunakan *Wav2vec2.0* untuk mengenali bahasa arab berhasil mencapai akurasi sebesar 96,23% dengan

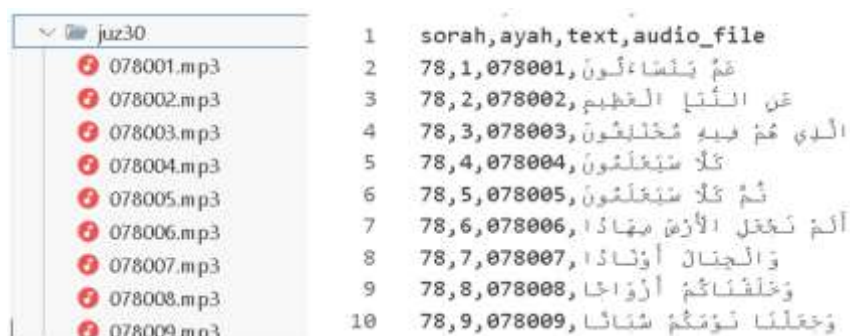
menggunakan augmentasi data. Hasil ini menunjukkan bahwa model ini efektif dalam mengidentifikasi pembicara bahasa arab [7]. Pada penelitian lainnya menyebutkan bahwa Hasil dari prediksi terhadap dataset audio bacaan Al-Qur'an berhasil mencapai tingkat *Word Error Rate (WER)* 1% [8]. Ada juga penelitian tentang penyempurnaan atau *fine-tuning* terhadap model dari XLSR yang disetel dengan *Wav2vec 2.0* pada bahasa Arab menggunakan sejumlah kecil data untuk bahasa Arab, *Word Error Rate (WER)* yang awalnya 54,00%. Setelah disempurnakan parameter model, diperoleh *WER* sebesar 36,69%, hal ini menunjukkan bahwa tingkat akurasi model semakin baik [9]. Model *ASR Wav2vec2.0*, yang telah dilatih khusus untuk mengenali bacaan Al-Qur'an, akan digunakan dalam penelitian ini. Dengan demikian, diharapkan dapat dihasilkan sebuah aplikasi yang mampu memberikan umpan balik secara langsung kepada pengguna mengenai kesalahan bacaan yang dilakukan, sehingga pengguna dapat melakukan koreksi secara mandiri dan meningkatkan kualitas bacaan Al-Qur'an.

2. METODOLOGI PENELITIAN

Dalam penelitian ini, model yang digunakan adalah *Wav2vec2.0*. kemudian dataset yang digunakan berupa kumpulan audio bacaan Al-Qur'an dari pembaca yang berbeda-beda. Dataset ini akan digunakan untuk melatih model *Wav2vec2.0* dan menghasilkan model baru yang dapat mengenali bacaan Al-Qur'an. Kemudian model ini akan diimplementasikan menjadi sebuah sistem *Automatic Speech Recognition (ASR)* berbasis *mobile* untuk mendeteksi kesalahan pembacaan Al-Qur'an.

2.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini didapatkan dari situs *kaggle*. Dataset berupa kumpulan audio bacaan Al-Qur'an dari *qari'* yang berbeda-beda, dan juga *file csv* yang berisi teks transkripsi sebagai referensi dari setiap audio. Data audio terdiri dari berbagai juz dan surat, untuk itu dataset diseleksi secara manual oleh penulis dan hanya diambil surat-surat dari juz 30. Pemilihan ini didasarkan pada pertimbangan efisiensi komputasi, mengingat keterbatasan perangkat yang digunakan. Surat-surat dalam juz 30 umumnya memiliki ukuran *file* yang lebih kecil, sehingga memungkinkan proses pelatihan model dilakukan dalam waktu yang lebih singkat



File Name	Transcription
078001.mp3	1 sorah,ayah,text,audio_file
078002.mp3	2 78,1,078001,عَمُّ يَنْشَأُ لَوْنٌ
078003.mp3	3 78,2,078002,عَنِ النَّبِيِّ الْعَظِيمِ
078004.mp3	4 78,3,078003,الَّذِي هُمْ فِيهِ مُخْلِطُونَ
078005.mp3	5 78,4,078004,كُلًّا مَسْجُودًا
078006.mp3	6 78,5,078005,ثُمَّ كُلًّا مِنِغْلَمُونَ
078007.mp3	7 78,6,078006,أَلَمْ تَخُذْ أَلْأَرْضَ مَهَادًا
078008.mp3	8 78,7,078007,وَالجِنَانَ أَوْثَادًا
078009.mp3	9 78,8,078008,وَخَلَقْنَاكُمْ أَزْوَاجًا
078009.mp3	10 78,9,078009,وَجَعَلْنَا بَيْنَكُمْ سُبُلًا

Gambar 1. File Audio Al-Qur'an dan Teks Transkripsi

Pada Gambar 1 merupakan kumpulan dataset audio bacaan ayat-ayat Al-Qur'an juz 30. Dataset berupa audio dengan ekstensi *.mp3*, Kemudian *file* CSV yang berisi transkripsi dari dataset audio yang akan digunakan sebagai referensi untuk membandingkan hasil pengenalan suara yang dihasilkan oleh model *ASR*.

2. 2. Data Pre-processing

Proses preprocessing data audio Al-Qur'an dilakukan dalam beberapa tahap. Pertama, data dibersihkan dari segala bentuk gangguan suara yang tidak relevan, seperti kebisingan latar belakang. Proses ini bertujuan untuk meningkatkan kualitas audio sehingga model dapat lebih fokus pada fitur-fitur yang relevan. Selanjutnya, audio dibagi menjadi segmen-segmen yang lebih kecil, yakni per ayat. Proses segmentasi ini dilakukan secara manual untuk memastikan akurasi pembagian. Tahap berikutnya adalah normalisasi audio, yang bertujuan untuk menyamaratakan karakteristik audio seperti volume dan *sample rate*. Semua data audio diubah menjadi saluran mono 16-bit dengan laju *sample rate* 16kHz [10]. Normalisasi ini penting agar model dapat belajar pola yang lebih umum dan menghindari bias akibat perbedaan kualitas rekaman. Terakhir, dataset yang telah dinormalisasi dibagi menjadi tiga bagian, yaitu data *training*, *validation*, dan *testing*. Pembagian ini bertujuan untuk melatih model, mengoptimalkan *hyperparameter*, dan menguji kinerja model secara independen sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Rasio Pembagian Dataset

<i>Sub Set</i>	<i>Persentase</i> (%)
<i>Training</i>	70%
<i>Validation</i>	20%
<i>Testing</i>	10%

2. 3. Data Processing

Proses pemrosesan data audio Al-Qur'an yang telah dinormalisasi melibatkan beberapa tahap penting. Pertama, dilakukan ekstraksi fitur dari audio. Proses ini mengubah sinyal audio mentah menjadi representasi numerik yang dapat dipahami oleh model pembelajaran mesin. Model *Wav2vec2.0* yang digunakan dalam penelitian ini memiliki kemampuan unik untuk mengekstrak fitur langsung dari sinyal audio tanpa memerlukan metode ekstraksi fitur tradisional seperti MFCC atau spektrogram [11]. Tahap kedua adalah tokenisasi teks transkripsi. Teks transkripsi dipecah menjadi unit-unit terkecil (token) yang biasanya berupa huruf. Token-token ini kemudian dikonversi menjadi representasi numerik yang unik. Proses tokenisasi memungkinkan model untuk belajar pola bahasa dan memetakan antara representasi audio dan teks. Terakhir, data loader disiapkan untuk mengorganisasikan data menjadi *batch-batch* yang efisien. Setiap batch berisi pasangan data audio dan teks yang sesuai. Data loader ini akan digunakan selama proses pelatihan, validasi, dan pengujian model *Wav2vec2.0*.

2. 4. *Fine-Tuning Model*

Proses *fine-tuning* dalam penelitian ini bertujuan untuk menyesuaikan model pengenalan suara (*ASR*) yang telah dilatih sebelumnya, yaitu "jonatasgrosman/wav2vec2-large-xlsr-53-arabic", agar lebih optimal dalam mengenali suara bacaan Al-Qur'an. Model yang telah dilatih sebelumnya ini memiliki pemahaman yang baik tentang bahasa Arab secara umum. Namun, dengan melakukan *fine-tuning* menggunakan dataset audio Al-Qur'an yang spesifik, model diharapkan dapat meningkatkan akurasi dalam mengenali nuansa dan karakteristik unik dalam bacaan Al-Qur'an. Proses ini melibatkan penyesuaian parameter-parameter dalam model agar dapat lebih baik dalam memetakan antara sinyal audio bacaan Al-Qur'an dengan transkripsi teksnya. Dengan demikian, model yang telah dilakukan *fine-tuning* diharapkan mampu menghasilkan transkripsi teks yang lebih akurat dan relevan dengan konteks bacaan Al-Qur'an.

2. 5. Validasi Model

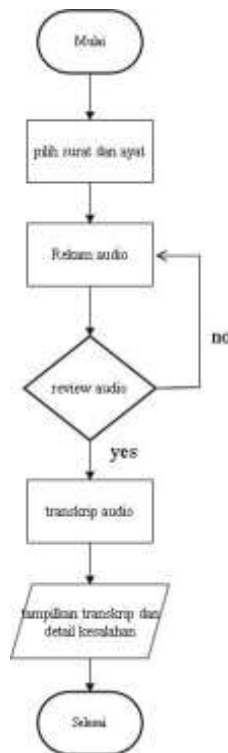
Setelah proses *training* model selesai, model akan divalidasi performanya dalam mengenali bacaan Al-Qur'an. Adapun keluaran dari proses validasi ini adalah berupa matrik kesalahan *WER* (*Word Error Rate*). Persentase kata yang salah, dihapus, atau ditambahkan dalam transkripsi otomatis dibandingkan dengan transkripsi manual. *WER* dihitung dengan menjumlahkan jumlah penyisipan (*insertions*), penghapusan (*deletions*), dan substitusi (*substitutions*) yang diperlukan untuk mengubah teks transkripsi menjadi teks referensi [12].

$$WER = \frac{Insertions + Deletions + Substitutions}{Total\ Kata\ dalam\ Referensi}$$

WER digunakan untuk mengukur tingkat kesalahan dalam mengenali teks yang diucapkan dengan benar. Semakin rendah nilai *WER*, semakin baik kinerja model dalam pengenalan suara dan interpretasi teks bacaan Al-Qur'an.

2. 6. Perancangan Aplikasi

Aplikasi deteksi kesalahan baca Al-Qur'an yang dikembangkan dalam penelitian ini merupakan sebuah aplikasi mobile berbasis *Android* yang dibangun menggunakan kerangka kerja *React Native*. Dengan memanfaatkan fleksibilitas dan efisiensi *React Native*, aplikasi ini berhasil mengintegrasikan teknologi pengenalan suara *Automatic Speech Recognition (ASR)* untuk membantu pengguna dalam memperbaiki bacaan Al-Qur'an. Fitur-fitur utama yang ditawarkan oleh aplikasi ini meliputi pemilihan surat dan ayat, perekaman suara pengguna, peninjauan hasil rekaman, serta penyajian hasil transkripsi dan detail kesalahan bacaan dengan bentuk *list*. Penggunaan *React Native* memungkinkan aplikasi ini untuk berjalan dengan lancar pada berbagai perangkat *Android* [13] sehingga memberikan akses yang mudah bagi pengguna untuk memanfaatkan teknologi *ASR* dalam pembelajaran Al-Qur'an.



Gambar 2. *Flowchart* Aplikasi *Mobile*

Gambar 2 merupakan *flowchart* atau alur sistem pada aplikasi *mobile*, dimulai dengan memilih surat dan ayat kemudian akan masuk ke halaman surat dan ayat yang dipilih, kemudian *user* bisa merekam suara bacaan ayat Al-Qur'an dan meninjau hasil rekaman. Selanjutnya audio akan diproses dan ditranskripsi kemudian akan dideteksi apakah ada kesalahan dalam membaca ayat. Jika terdeteksi ada kesalahan maka sistem akan menampilkan hasil transkripsi dan detail dari kesalahan baca.

2. 7. Validasi Aplikasi

Proses validasi aplikasi *mobile* ini bertujuan untuk mengukur akurasi model dalam membaca Al-Qur'an dan mengidentifikasi jenis-jenis kesalahan yang terjadi. Transkripsi hasil pembacaan Al-Qur'an yang dihasilkan oleh model akan dibandingkan dengan teks Al-Qur'an yang benar sebagai acuan. Perbandingan ini dilakukan menggunakan pustaka *SequenceMatcher* yang mampu mendeteksi perbedaan antara dua buah teks berdasarkan konsep *Edit Distance* dan *Longest Common Subsequence* [14]. Hasil perbandingan kemudian dikategorikan menjadi beberapa jenis kesalahan, yaitu penggantian, penghapusan, dan penambahan. Kesalahan penggantian akan dikategorikan lebih lanjut menjadi kesalahan harakat atau kesalahan huruf, tergantung pada karakter yang diganti. Dengan demikian, melalui proses validasi ini, dapat diperoleh gambaran yang lebih jelas mengenai kinerja model dan jenis kesalahan yang perlu diperbaiki untuk meningkatkan akurasi sistem.

3. HASIL DAN ANALISIS

3.1. Hasil Training Dan Validasi Model

Tabel 2. Hasil Training Model

No Konfigurasi	WER
Konfigurasi 1	0.494553
Konfigurasi 2	0.414254
Konfigurasi 3	0.648107
Konfigurasi 4	0.367483

Berdasarkan Tabel 2, konfigurasi yang menghasilkan nilai *WER* terendah adalah konfigurasi nomor empat dengan *WER* 0.367483. Ini membuktikan bahwa *training* dengan konfigurasi tersebut adalah lebih optimal dan menghasilkan model *Automatic Speech Recognition (ASR)* yang memiliki tingkat kesalahan pengenalan bacaan Al-Qur'an yang lebih rendah dibandingkan dengan konfigurasi lainnya. Berikut adalah hasil transkripsi terhadap beberapa sampel bacaan Al-Qur'an juz 30:

target_text	predicted_text	accuracy
فَقِيلَ الْإِنْسَانُ مَا أَكْفَرَهُ	فَقِيلَ الْإِنْسَانُ مَا أَكْفَرَهُ	0.75
فِي عَمَدٍ مُّمَدَّدَةٍ	فِي عَمَدٍ مُّمَدَّدَةٍ	1
وَإِذَا النُّجُومُ انْكَرَّتْ	وَإِذَا النُّجُومُ انْكَرَّتْ	0.66
فَأَنْتَ لَهُ تَصَدَّقَى	فَأَنْتَرَهُ تَصَدَّقَى	0.33
إِلَّا حَمِيمًا وَضَافًا	إِلَّا حَمِيمًا وَضَافًا	1
يَوْمَئِذٍ تُحَدِّثُ أَخْبَارَهَا	يَوْمَئِذٍ تُحَدِّثُ أَخْبَارَهَا	0.66
وَرَبُّنَا وَنَخْلًا	وَرَبُّنَا وَنَخْلًا	1
وَفَاكِهَةً وَأَبًّا	وَفَاكِهَةً وَأَبًّا	1
وَمَا هُوَ بِقَوْلِ شَيْطَانٍ رَجِيمٍ	وَمَا هُوَ بِقَوْلِ شَيْطَانِ الرَّجِيمِ	0.4
وَاللَّيْلِ إِذَا يَسِرَ	وَاللَّيْلِ إِذَا يَسِرَ	0.66

Gambar 3. Hasil Transkripsi Model

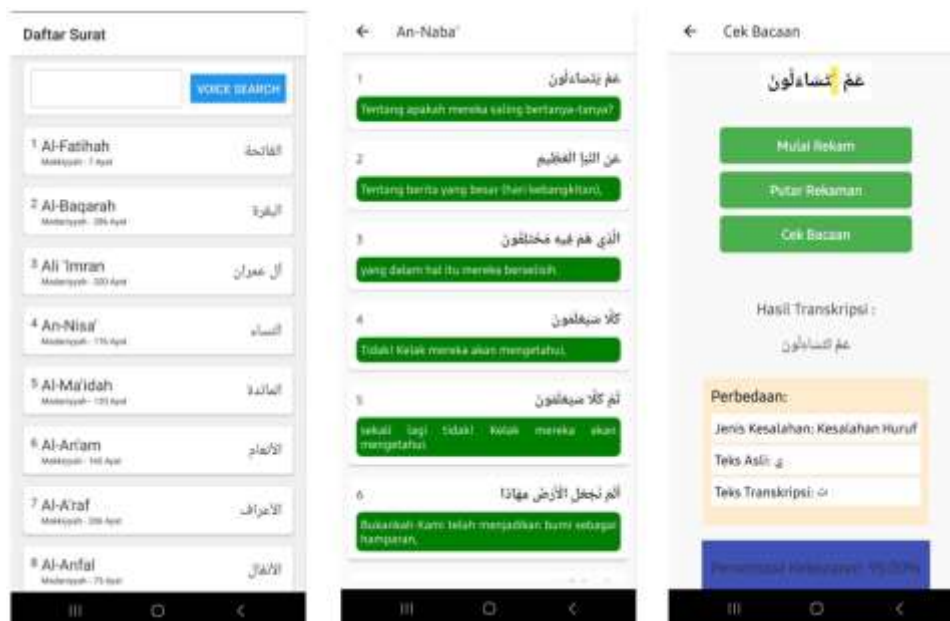
Gambar 3 menampilkan hasil perbandingan komparatif antara teks asli Al-Qur'an dan hasil transkripsi yang dihasilkan oleh model *Automatic Speech Recognition (ASR)*. Setiap baris pada tabel menampilkan segmen teks asli (*target_text*), hasil transkripsi yang diprediksi oleh model (*predicted_text*), serta metrik akurasi yang menunjukkan tingkat kesesuaian antara keduanya (*accuracy*). Nilai akurasi satu mengindikasikan kesempurnaan transkrip, sementara nilai yang lebih rendah menandakan adanya kesalahan atau ketidaktepatan.

Analisis terhadap hasil perbandingan menunjukkan bahwa model *ASR* yang telah dilatih mampu mencapai tingkat akurasi yang cukup baik dalam mentranskripsi teks Al-Qur'an, termasuk dalam hal pengenalan harakat. Hal ini terlihat dari beberapa segmen teks yang berhasil ditranskripsikan dengan sangat akurat. Namun, model masih menghadapi beberapa tantangan, seperti kesalahan dalam mentranskripsi kata "يَقُولُ" menjadi "يَقُومُ" dan kata "أَخْبَارَهَا" menjadi "أَخْبَاعَهَا". Kesalahan-kesalahan ini mengindikasikan bahwa model masih perlu ditingkatkan performanya.

Beberapa faktor yang dapat mempengaruhi akurasi model antara lain kualitas data latih, pemilihan parameter pelatihan, serta kompleksitas bahasa Arab, khususnya dalam konteks bacaan Al-Qur'an. Untuk meningkatkan akurasi model di masa mendatang, dapat dilakukan beberapa upaya seperti memperkaya dataset latih dengan variasi bacaan yang lebih luas, melakukan penyetelan *hyperparameter* model secara lebih cermat, serta mengeksplorasi teknik-teknik pemrosesan bahasa alami yang lebih canggih.

Secara keseluruhan, hasil eksperimen ini memberikan gambaran yang baik mengenai potensi model *ASR* dalam mendukung pembelajaran Al-Qur'an. Meskipun masih terdapat ruang untuk perbaikan, model ini telah menunjukkan kemampuannya dalam memahami dan mentranskripsi teks Al-Qur'an dengan tingkat akurasi yang baik.

3.2. Hasil Implementasi *User Interface*



Gambar 4. Hasil *User Interface*

Gambar 4 merupakan hasil *user interface* (UI) yang telah dikembangkan untuk sistem deteksi kesalahan baca Al-Qur'an dirancang dengan tampilan yang sederhana dan intuitif. Halaman utama menyajikan daftar surat Al-Qur'an secara terstruktur, lengkap dengan informasi terkait. Fitur pencarian yang interaktif memungkinkan pengguna dengan mudah menemukan surat yang diinginkan. Saat memilih sebuah surat, pengguna

akan diarahkan ke halaman yang menampilkan daftar ayat surat tersebut. Di halaman ini, pengguna dapat memilih ayat yang ingin dibaca dan melakukan perekaman suara. Setelah proses perekaman selesai, pengguna dapat memutar ulang rekaman dan melakukan pengecekan. Fitur utama pada halaman cek bacaan adalah tombol Cek Bacaan yang akan memicu proses analisis terhadap rekaman suara. Sistem kemudian akan menampilkan hasil analisis berupa deteksi kesalahan bacaan dan jenis-jenis kesalahan yang terjadi beserta *highlight* atau sorotan pada bagian yang salah dengan memberikan warna yang berbeda. Dengan demikian, Pengguna dapat melakukan evaluasi mandiri terhadap bacaan Al-Qur'an mereka berdasarkan identifikasi kesalahan yang ditampilkan.

3.3. Hasil Pengujian Cek Bacaan

Tabel 3. Hasil Pengujian Aplikasi

Pengujian	Kasus	Hasil	Kesimpulan
Cek bacaan benar	Pengguna melakukan pengecekan terhadap bacaan yang dibaca benar.	Sesuai	Normal
Cek bacaan salah <i>harakat</i>	Pengguna melakukan pengecekan terhadap bacaan yang dibaca dengan salah <i>harakat</i> .	Sesuai	Normal
Cek bacaan salah huruf	Pengguna melakukan pengecekan terhadap bacaan yang dibaca salah huruf.	Sesuai	Normal

Berdasarkan Tabel 3 hasil pengujian sistem aplikasi mulai dari pengujian fungsionalitas setiap halamannya kemudian pengujian terhadap setiap fungsi deteksi kesalahan baca. Semua fungsi berjalan dengan normal dan sesuai dengan fungsionalitasnya, Namun masih ada beberapa ketidak sesuaian pada fungsi deteksi kesalahan dimana terdapat beberapa kasus dimana kesalahan yang dideteksi tidak sesuai dengan kategorinya. Hal ini dikarenakan aturan pengkategorian kesalahan hanya terdapat dua kategori yaitu kesalahan *harakat* dan kesalahan huruf. Sehingga jika terdapat kesalahan diluar kategori tersebut maka sistem akan secara acak memasukkan kesalahan tersebut ke salah satu kategori yang telah ditentukan.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem deteksi kesalahan baca Al-Qur'an yang memanfaatkan teknologi *Automatic Speech Recognition (ASR)* telah berhasil diimplementasikan dalam bentuk aplikasi *mobile*. Model *Wav2vec2.0* yang digunakan menunjukkan kinerja yang cukup baik dalam mengenali bacaan Al-Qur'an, terutama pada juz 30 dengan tingkat *WER* 0.367483. Kemampuan model untuk belajar dari data yang terbatas merupakan keunggulan yang signifikan. Meskipun demikian, masih terdapat ruang untuk peningkatan akurasi,

terutama dalam hal pengenalan harakat dan huruf tertentu. Fitur cek bacaan pada aplikasi ini telah berhasil mendeteksi berbagai jenis kesalahan bacaan, namun perlu dilakukan perbaikan lebih lanjut untuk meningkatkan keandalannya. Secara keseluruhan, penelitian ini menunjukkan potensi besar dari teknologi *Automatic Speech Recognition (ASR)* dalam membantu pengguna melatih dan memperbaiki bacaan Al-Qur'an mereka.

DAFTAR PUSTAKA

- [1] N. O. Balula, M. Rashwan, and S. Abdou, "Automatic Speech Recognition (ASR) Systems for Learning Arabic Language and Al-Quran Recitation: A Review," *International Journal of Computer Science and Mobile Computing*, vol. 10, no. 7, pp. 91–100, 2021, doi: 10.47760/ijcsmc.2021.v10i07.013.
- [2] M. S. Asngari and A. Alena, "Metode Pembelajaran Dalam Meningkatkan Kualitas Bacaan Al-Qur'an," *GUAU: Jurnal Pendidikan Profesi Guru ...*, vol. 2, no. Rida 2017, pp. 305–310, 2022.
- [3] B. Ditjen, "Survei Kemenag, Indeks Literasi Al-Qur'an Kategori Tinggi," kemenag.go.id. Accessed: Jun. 20, 2024. [Online]. Available: <https://kemenag.go.id/nasional/survei-kemenag-indeks-literasi-al-qur-an-kategori-tinggi-w0A7W>
- [4] R. Della and Y. Huda, "Rancang Bangun Aplikasi Media Pembelajaran Tahsin Al-Qur'an Berbasis Android," *Voteteknika (Vocational Teknik Elektronika dan Informatika)*, vol. 7, no. 4, p. 207, 2020, doi: 10.24036/voteteknika.v7i4.106638.
- [5] M. Malik, M. K. Malik, K. Mehmood, and I. Makhdoom, "Automatic speech recognition: a survey," *Multimed Tools Appl*, vol. 80, no. 6, pp. 9411–9457, 2021, doi: 10.1007/s11042-020-10073-7.
- [6] C. Yi, J. Wang, N. Cheng, S. Zhou, and B. Xu, "Applying Wav2vec2.0 to Speech Recognition in Various Low-resource Languages," pp. 1–5, 2020, [Online]. Available: <http://arxiv.org/abs/2012.12121>
- [7] A. Moustafa and S. A. Aly, "Towards an Efficient Voice Identification Using Wav2Vec2.0 and HuBERT Based on the Quran Reciters Dataset," pp. 1–5, 2021, [Online]. Available: <http://arxiv.org/abs/2111.06331>
- [8] Aminudin, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Pengolahan Korpus Dataset Audio Bacaan Al-Qur'an Menggunakan Metode Wav2Vec 2.0," vol. 10, no. 1, pp. 104–109, 2024.
- [9] M. Bakheet, J. M. Peña, and J. Verhook, "Improving Speech Recognition for Arabic language Using Low Amounts of Labeled Data," 2021, [Online]. Available: <https://ep.liu.se/>.
- [10] R. Jain, A. Barcovschi, M. Y. Yiwere, D. Bigioi, P. Corcoran, and H. Cucu, "A WAV2VEC2-Based Experimental Study on Self-Supervised Learning Methods to Improve Child Speech Recognition," *IEEE Access*, vol. 11, no. May, pp. 46938–46948, 2023, doi: 10.1109/ACCESS.2023.3275106.
- [11] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, "wav2vec 2.0: A framework for self-supervised learning of speech representations," *Adv Neural Inf Process Syst*, vol. 2020-Decem, no. Figure 1, pp. 1–19, 2020.

- [12] A. Rahman, M. M. Kabir, M. F. Mridha, M. Alatiyyah, H. F. Alhasson, and S. S. Alharbi, "Arabic Speech Recognition: Advancement and Challenges," *IEEE Access*, vol. 12, pp. 39689–39716, 2024, doi: 10.1109/ACCESS.2024.3376237.
- [13] A. Paul and A. Nalwaya, *React Native for Mobile Development: Harness the Power of React Native to Create Stunning iOS and Android Applications*. Apress, 2019.
- [14] N. V. E. S. Murthy, S. Amarnadh, P. V. G. D. P. Reddy, and N. V. E. S. Murthy, "FreeHand Sketch-based Authenticated Security System using Sequence Matcher," *International Journal of Advanced Science and Technology*, vol. 29, no. 04, pp. 6663–6679, 2020, [Online]. Available: <https://www.researchgate.net/publication/342872000>