

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dengan pesatnya perkembangan teknologi informasi dan komunikasi, kebutuhan untuk mengakses sumber-sumber pengetahuan secara lebih cepat dan mudah semakin meningkat. Bagi umat Islam, Alquran merupakan salah satu sumber pengetahuan yang sangat penting, meskipun akses terhadap Alquran, baik dalam bentuk cetak maupun digital semakin luas, masih ada sejumlah pengguna yang menghadapi kesulitan, khususnya mereka yang memiliki keterbatasan visual atau literasi, dalam mengakses dan memahami ayat-ayat Alquran. Di sinilah peran teknologi pengenalan suara menjadi relevan, karena dapat menyediakan solusi yang lebih praktis melalui otomatisasi deteksi dan pengenalan ayat Alquran dari rekaman suara. Membaca Al-Qur'an adalah ibadah yang dilakukan oleh setiap muslim, sehingga sudah menjadi kebiasaan bagi umat Islam untuk mempelajarinya. Al-Qur'an memegang peran yang sangat penting dalam kehidupan umat Islam sehari-hari. Oleh sebab itu, sebagaimana yang disampaikan oleh Rochanah dalam kutipan (Meiyanti, 2023), pendidikan Al-Qur'an perlu ditanamkan sejak dini melalui metode hafalan, pembelajaran, dan pengamalan isi Al-Qur'an. Namun, masih banyak anak-anak, remaja, bahkan orang tua muslim yang belum mampu membaca Al-Qur'an, apalagi menghafalnya.

Teknologi pengenalan suara berbasis pembelajaran mesin (machine learning) telah berkembang pesat dalam berbagai bidang, seperti asisten virtual dan penerjemah bahasa. Namun, penerapannya dalam konteks pengenalan ayat-ayat Alquran masih relatif jarang, terutama untuk mendeteksi ayat secara otomatis dari suara. Teknologi ini memiliki potensi besar untuk memudahkan pengguna dalam menemukan, mendengarkan, dan memahami ayat Alquran tanpa harus mencari atau membaca teks secara manual (Harere & Jallad, 2023).

Pendekatan *Region-based Convolutional Neural Networks* (R-CNN) pada dasarnya dikembangkan untuk tugas deteksi objek dalam gambar. R-CNN adalah arsitektur jaringan saraf konvolusi yang sangat efektif dalam mengekstraksi fitur-

fitur dari citra dan mengidentifikasi objek-objek yang ada di dalamnya dengan akurasi yang sangat tinggi. Model ini pertama kali diperkenalkan untuk mendeteksi dan mengklasifikasikan berbagai objek dalam citra berdasarkan wilayah tertentu (*region*), yang diidentifikasi melalui algoritma pencarian wilayah atau *selective search*, dalam konteks pengolahan citra, R-CNN bekerja dengan cara membagi gambar menjadi beberapa wilayah (*regions of interest*) yang mungkin berisi objek, kemudian menerapkan jaringan saraf konvolusi (CNN) pada setiap wilayah ini untuk mengenali dan mengklasifikasikan objek di dalamnya, namun, inovasi dalam ilmu pembelajaran mesin telah memungkinkan adaptasi arsitektur yang sama (R-CNN) untuk diterapkan pada data yang berbeda, termasuk data audio dan tugas pengenalan suara, data audio juga dapat dilihat sebagai data yang memiliki pola dan struktur yang dapat dikenali oleh model seperti halnya gambar (Hadiyansah & Andamira, 2023).

Mengonversi rekaman suara menjadi representasi visual seperti *spektrogram* atau *Mel Frequency Cepstral Coefficients* (MFCC), model ini dapat secara otomatis mendeteksi dan mengenali ayat-ayat Alquran. Meski begitu, adaptasi ini menimbulkan beberapa tantangan, seperti variabilitas kualitas data audio, serta kemampuan model untuk mengatasi perbedaan dalam pengucapan, intonasi, dan kualitas rekaman (Al Harere & Al Jallad, 2023).

Selain itu, dengan beragamnya pengucapan di berbagai wilayah, penting bagi model pengenalan suara untuk dapat mengenali ayat Alquran dengan akurasi tinggi dalam berbagai kondisi nyata. Oleh karena itu, penelitian ini bertujuan untuk tidak hanya mengembangkan model deteksi ayat Alquran berbasis suara, tetapi juga untuk mengatasi tantangan yang berkaitan dengan adaptasi algoritma R-CNN dan perbaikan kualitas serta keragaman data pelatihan.

Penelitian ini diharapkan dapat berkontribusi besar dalam meningkatkan aksesibilitas Alquran bagi berbagai kalangan, termasuk mereka yang memiliki keterbatasan visual, literasi, atau yang ingin mendengarkan ayat-ayat Alquran secara lebih mudah dan interaktif. Berdasarkan dari latar belakang yang telah diuraikan, maka peneliti pengangkat penelitian yang berjudul “**Aplikasi Deteksi**

Ayat Alquran Dengan Menggunakan Model Pengenalan Suara *Machine Learning* R-CNN (Studi Kasus Surah An-Naba)”

1.2 Rumusan Masalah

Berdasarkan dari latar belakang diatas, maka rumusan masalah yang dapat diuraikan adalah :

1. Bagaimana cara mengimplementasikan model R-CNN untuk mendeteksi dan mengenali ayat-ayat Alquran dari rekaman suara?
2. Bagaimana hasil dari implementasi model R-CNN untuk mendeteksi dan mengenali ayat-ayat Alquran dari rekaman suara?

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Penelitian ini dapat memberikan solusi yang lebih inklusif bagi individu yang memiliki keterbatasan visual atau literasi dalam mengakses Alquran, dengan menyediakan sistem pengenalan suara yang otomatis untuk mendeteksi dan memutar ayat-ayat Alquran tanpa memerlukan teks cetak.
2. Penelitian ini menggunakan teknologi pengenalan suara, pengguna dapat lebih mudah menemukan ayat-ayat Alquran dari rekaman suara tanpa harus mencari manual dalam teks. Hal ini bermanfaat bagi mereka yang ingin mendengarkan dan memahami ayat-ayat secara lebih cepat dan efisien.

Penelitian ini memberikan kontribusi dalam memajukan penerapan teknologi pembelajaran mesin, khususnya R-CNN, dalam konteks pengenalan suara untuk tujuan keagamaan. Ini dapat membuka jalan bagi inovasi lebih lanjut dalam teknologi berbasis AI untuk memfasilitasi aksesibilitas dan pembelajaran Alquran di masa depan.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk:

1. Mengimplementasikan model R-CNN untuk mendeteksi dan mengenali ayat-ayat Alquran dari rekaman suara

2. Model pengenalan suara dapat beradaptasi terhadap variasi kualitas rekaman, intonasi, dan pengucapan ayat-ayat Alquran dari berbagai wilayah

1.5 Ruang Lingkup Dan Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Jenis Data: Penelitian ini dibatasi pada penggunaan data audio berupa rekaman suara yang mengandung ayat-ayat Alquran. Variasi dalam bentuk teks atau video tidak menjadi fokus utama.
2. Penelitian ini hanya menggunakan model *Region-based Convolutional Neural Networks* (R-CNN) untuk mendeteksi dan mengenali ayat-ayat Alquran.
3. Penelitian ini difokuskan pada pengenalan suara dari rekaman ayat Alquran yang diucapkan dalam bahasa Arab.
4. Data audio diproses untuk membersihkan noise, menormalisasi volume, dan mengubah rekaman menjadi format yang sesuai untuk analisis lebih lanjut
5. Penelitian ini hanya akan mengonversi data audio menjadi bentuk representasi visual Spektrogram atau *Mel Frequency Cepstral Coefficients* (MFCC).
6. Fokus penelitian lebih kepada peningkatan aksesibilitas bagi pengguna dengan keterbatasan visual atau literasi, sehingga aspek interaktifitas atau fitur tambahan bagi pengguna tanpa keterbatasan ini tidak dibahas secara mendetail.