

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bahasa merupakan alat komunikasi yang penting bagi manusia, karena dengan bahasa kita sebagai masyarakat yang hidup bersosial ingin mengetahui informasi yang kita butuhkan dalam kehidupan sehari-hari. Menurut (Hanny Rindiani : 2015) Bahasa adalah suatu sistem lambang berupa bunyi yang bersifat arbitrer, dan alat komunikasi yang penting bagi manusia, digunakan masyarakat untuk mendapatkan informasi yang dibutuhkan, bekerja sama, berkomunikasi dan mengidentifikasi diri. Pada saat ini masyarakat sudah mulai tertarik untuk mempelajari bahasa-bahasa asing di luar bahasa mereka sendiri, mulai dari bahasa Arab, Belanda, Jepang, Inggris, Perancis, Jerman, Mandarin dan lain-lain. Akan tetapi sekarang ini penelitian untuk bahasa latin sudah mulai sedikit berkurang, saat ini yang sedang berkembang, maju dan banyak di sukai orang untuk di pelajar adalah bahasa oriental seperti, bahasa Arab, Jepang, Mandarin dan Korea, yang sangat banyak peminatnya.

Bahasa Jepang menjadi salah satu bahasa yang menarik untuk dipelajari. Menurut survei yang dilakukan oleh CBI Education & Skills Survey 2012, bahasa Jepang menjadi salah satu dari 10 bahasa yang patut untuk dipelajari karena kebudayaan Jepang banyak menginspirasi masyarakat baik dalam hal yang kecil. Selain itu, bahasa Jepang dianggap penting karena pilar perdagangan dan penelitian di Asia dipegang oleh negara Jepang sehingga untuk dapat berkomunikasi dengan masyarakatnya tentu harus mampu menguasai bahasa oriental dari negara tersebut. Bahasa Jepang merupakan salah satu bahasa yang sulit untuk dipelajari karena penulisan dan pengucapannya yang relatif rumit serta memiliki tiga jenis penulisan yaitu Hiragana, Katakana dan Kanji.

Pembahasan mengenai bahasa sebagai alat komunikasi dan sebagai alata untuk mengetahui informasi yang kita butuhkan baik dalam bentuk percakapan maupun tertulis, tidak lepas dari teknologi informasi dan komunikasi yang hingga saat ini berkembang dengan sangat cepat. Salah satunya adalah perkembangan dalam bidang teknologi *Text To Speech* yang merupakan cabang ilmu dari pemrosesan bahasa alami. *Text To Speech* Sendiri adalah suatu sistem pengkonversi dari teks menjadi suara ucapa ataupun suara yang dihasilkan dari

sistem *cloud Text to Speech API* yang salah satunya bisa kita dapatkan melalui Google Developer untuk membuat pengembangan aplikasi yang akan kita buat.

Dalam proses perancangan, pembuatan aplikasi *Text To Speech* ini perlu ada pembelajaran untuk mengenali karakter huruf Jepang yang cukup sulit untuk dimengerti. Hal ini dikarenakan proses pengenalan memerlukan masukan berupa text dengan karakter huruf Jepang. Untuk mengatasi masalah masukan berupa huruf teks bahasa Jepang, maka teknik *Optical Character Recognition (OCR)* dapat digunakan untuk mengenali karakter huruf Jepang (Hiragana, Katakana dan Kanji). *OCR* adalah teknik untuk mengubah teks non digital menjadi text digital atau dapat diartikan secara harfiah sebagai pengenalan karakter optik, seperti mengenali gambar karakter huruf yang diambil melalui hasil foto atau scan gambar.

Teknologi perangkat mobile *Android* merupakan teknologi yang sedang berkembang pesat di Indonesia dengan peningkatan yang sangat cepat sejak dirilis secara resmi tahun 2007 hingga sekarang. *Android* merupakan sistem operasi berbasis linux yang dirancang untuk perangkat bergerak seperti smart phone dan banyak lagi lainnya. *Android* awalnya dikembangkan oleh *Android Inc*, dengan dukungan finansial dari *Google*, yang kemudian membelinya pada tahun 2005. Sistem operasi ini dirilis secara resmi pada tahun 2007. Secara garis besar *android* adalah sistem operasi open source yang berbasis linux yang menggunakan bahasa pemrograman *Java*.

Pada proses pengenalan karakter menurut penelitian (Rindiani, H. 2015, "Aplikasi *Android* untuk Pengenalan Citra karakter Jepang dengan library *Tesseract*", Ilmu Komputer IPB, Bogor) Pada penelitian ini sistem pengenalan karakter Jepang berbasis *Android* dirancang untuk mengenali citra karakter Jepang dan menerjemahkannya menjadi bahasa Indonesia menggunakan library *Tesseract OCR (Optical Character Recognition)*. Pada penelitian ini terdapat 5 kebutuhan fungsional pengguna yaitu mengambil gambar dengan kamera, mengambil gambar dari galeri, melakukan konversi dari citra ke teks, mengedit teks Jepang hasil *OCR*, dan menampilkan hasil terjemahan teks Jepang. Berdasarkan 10 sampel karakter Jepang yang telah diuji, nilai akurasi yang di dapat dari pengujian gambar dari kamera sekitar 80% dan gambar dari galeri 94%. Selain itu pengujian aplikasi menggunakan kuesioner mencapai 91%.

Berdasarkan pembahasan di atas maka pada penelitian ini akan merancang dan membuat *Aplikasi Text to Speech* Pengucapan kata dalam bahasa

Jepang berbasis Android dengan menggunakan citra karakter huruf Jepang (Hiragana, Katakana dan Kanji) sebagai masukan yang bisa diambil dari citra yang mengandung karakter huruf Jepang melalui kamera. dengan menggunakan *Tesseract OCR Engine* untuk mengenali citra karakter huruf Jepang ke text latin serta menggunakan *Cloud Text to Speech API* untuk mengkonversi dari teks bahasa Jepang ke suara pengucapan kata dalam bahasa Jepang yang benar. dan akan ditambahkan juga *Google cloud Translation API* yang berfungsi untuk menerjemahkan kata bahasa Jepang Ke bahasa Indonesia yang akan lebih mempermudah untuk mempelajari kata-kata yang berhubungan dengan bahasa Jepang. Mesin OCR dan API (*Application Programming Interface*) yang berbasis bahasa pemrograman C/C++ ini akan diterapkan pada platform Android dengan menggunakan bahasa pemrograman Java. Dalam penelitian ini juga mempertimbangkan sudut pengambilan citra katarker huruf Jepang dan tingkat kejelasan dari gambar yang diambil serta melakukan pengujian terhadap *aplikasi Text to Speech* dalam pengucapan kata bahasa Jepang dengan metode Black Box (pengujian yang dilakukan hanya mengamati hasil eksekusi melalui data uji dan memeriksa fungsional dari perangkat lunak) dan Waterfall (proses pengembangan perangkat lunak secara berurutan atau seperti air terjun yang mengalir dari atas ke bawah, seperti analisis kebutuhan, desain sistem, penerapan atau pembuatan (kontruksi), pengujian dan evaluasi).

1.2 Perumusan Masalah

Berdasarkan indikator permasalahan yang dijabarkan dalam latar belakang tersebut, maka dalam penelitian ini masalah yang akan dibahas dalam penelitian adalah sebagai berikut :

1. Bagaimana merancang *Aplikasi Text to Speech* dalam pengucapan kata dalam bahasa Jepang berbasis Android ?
2. Bagaimana *Tesseract OCR Engine* dapat diterapkan dalam pengenalan citra karakter huruf Jepang dan *Cloud Text to Speech API* mengkonversikan text menjadi suara pengucapan bahas Jepang untuk platform Android ?
3. Seberapa akurat dan efektif program Aplikasi *Text to Speech* dalam pengucapan kata bahasa Jepang berbasis Android ini dalam mengenali katrakter dan mengubah teks ke suara pengucapan bahasa Jepang dengan baik ?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini yang akan dilakukan adalah sebagai berikut :

1. Merancang dan membuat program aplikasi *Text to speech* dalam pengucapan kata bahasa Jepang berbasis Android yang menghasilkan suatu program aplikasi mobile yang dapat digunakan untuk mengenali citra karakter huruf Jepang ke teks dan mengubahnya menjadi pengucapan suara dalam bahasa Jepang.
2. Dapat Menerapkan *Tesseract OCR Engine* sebagai pengenalan karakter huruf Jepang dan *Cloud Text To Speech API* sebagai pengkonversi teks ke suara pengucapan bahasa Jepang kedalam platform Android.
3. Mengetahui tingkat keakuratan dan efektifitas metode pengenalan *Tesseract OCR Engine* dan *cloud Text to Speech API* untuk mengenali teks yang akan dirubah ke suara pengucapan dalam bahasa Jepang.

1.4 Manfaat penelitian

Manfaat yang dapat diperoleh dari hasil penelitian ini adalah dapat membantu pengguna yang ingin belajar Bahasa Jepang dengan mengenali karakter-karakter huruf bahasa Jepang (Hiragana, Katakana dan Kanji) dengan menggunakan gambar atau citra karakter yang dapat diambil melalui kamera handphone dan menghasilkan pengucapan suara yang benar serta mengetahui arti dari kata bahasa Jepang yang di kenali, yang mudah untuk dibawa kemana saja.

1.5 Batasan Penelitian

Lingkup dari penelitian ini adalah :

1. Aplikasi ini hanya menggunakan karakter huruf Jepang (Hiragana, Katakana dan Kanji) yang diambil menggunakan kamera, namun bukan menggunakan tulisan tangan.
2. Masukan berupa citra karakter huruf Jepang harus berupa ekstensi gambar yang didukung oleh Android, *Tesseract OCR Engine* dan *Cloud Text to Speech API*, yaitu berupa .jpeg .gif .png dan .bmp.
3. Untuk dapat menjalankan aplikasi ini perlu menggunakan koneksi internet atau online untuk dapat digunakan dengan lebih akurat dan efektif.

4. Batas karakter yang ingin dikenali kesistem kurang lebih mencapai 100 karakter huruf Jepang
5. Batas minimum resolusi gambar yang digunakan dalam pengenalan karakter huruf Jepang adalah 1024 x 768 resolution.
6. Batas jarak maksimu dalam pengambilan gambar melalui kamera mencapai 20 cm dari gambar karakter yang ingin dikenali.

Halaman ini sengaja dikosongkan