

TEXT RECOGNITION DENGAN KLASIFIKASI NEURAL NETWORK DAN TEXT-TO-SPEECH PADA HURUF ALPHABET

Edy Umar¹, Titin Fatimah²

^{1, 2)} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur
Jl. Raya Ciledug, Petukangan Utara, Pesanggrahan, Jakarta Selatan 12260

Telp. (021) 5853753, Fax. (021) 5866369

¹edyumar@yahoo.com, ² titin.fatimah@budiluhur.ac.id

ABSTRAK

Perkembangan teknologi informasi berdampak kepada membaurnya keilmuan atau keahlian tertentu. Bidang pengolahan citra tidak menjadi satu-satunya metoda dalam pemecahan suatu masalah, tetapi saat ini pengolahan citra dikombinasikan dengan kecerdasan tiruan untuk meneliti ataupun mencari suatu solusi dalam berbagai aplikasi. Media pembelajaran berbasis komputer sekarang ini sudah berkembang sangat pesat, banyak aplikasi yang dibuat dan dikembangkan untuk menunjang proses pembelajaran menjadi interaktif contohnya adalah aplikasi pengolahan citra digital, speech synthesis atau biasa disebut text-to-speech. Serta aplikasi yang menggunakan metoda jaringan syaraf tiruan. Pengolahan citra digital ini sudah meliputi teknik pengenalan karakter seperti karakter alfanumerik, karakter tulisan tangan, karakter huruf kanji, dan lain-lain. Teknik pengenalan karakter ini secara umum dikenal dengan teknologi OCR (Optical Character Recognition). Sedangkan text-to-speech merupakan suatu proses mengkonversi suatu teks tertulis menjadi ucapan. Dalam banyaknya penyedia pengenalan pola gambar, masih kurangnya pengenalan pola model dalam huruf Alphabet. Hal ini bisa dijadikan sebagai salah satu cara alternatif untuk belajar. Pada penelitian ini akan dikembangkan aplikasi pengenalan karakter untuk mengenali karakter huruf Alphabet. Algoritma yang dipresentasikan sebagai contoh nyata dalam program OCR akan diimplementasikan dalam bahasa pemrograman Java dengan tingkat keakurasian tertentu dalam pengoperasiannya, bisa didapatkan hasil (output) karakter yang tepat sesuai dengan learning class yang dibentuk melalui menu "OCR Train". Setelah selesai dikembangkan, aplikasi ini kemudian akan diuji dengan cara membentuk karakter huruf Alphabet yang disimpan sebagai gambar/image. Lalu dengan membuka file saraf buatan sesuai jenis tulisan tersebut diharapkan dapat mengenali karakter teks dalam gambar/image yang diuji.

Kata Kunci : Optical Character Recognition, Backpropagation, Speech.

I. PENDAHULUAN

Teknik pengenalan karakter ini secara umum dikenal dengan teknologi OCR (Optical Character Recognition). Sedangkan text-to-speech merupakan suatu proses mengkonversi suatu teks tertulis menjadi ucapan.

Jaringan saraf tiruan adalah sistem pemroses informasi yang memiliki karakteristik mirip dengan jaringan saraf manusia. Model pembelajaran perlu dilakukan pada suatu jaringan saraf tiruan sebelum digunakan untuk menyelesaikan masalah dengan memeriksa dan memperbaiki setiap kesalahan yang terjadi selama proses pembelajaran. Pada suatu tingkatan tertentu jaringan saraf tiruan dapat memberikan tanggapan yang benar walaupun masukan yang diberikan terdapat derau atau berubah oleh suatu keadaan. Kelebihan dari jaringan saraf tiruan adalah kemampuan mengenali dengan cara belajar dari pola gambar yang diajarkan.

Dalam banyaknya penyedia pengenalan pola gambar, pengenalan pola dalam huruf alphabet masih jarang. Dimana huruf alphabet merupakan salah satu huruf yang digunakan di banyak negara. Dengan menambahkan beberapa learning class tentang karakter huruf alphabet di dalamnya, prototipe text recognition dengan klasifikasi neural network dan text-to-speech diharapkan bisa dijadikan salah satu alternatif pembelajaran dengan media komputer.

I.1. Masalah

1. Identifikasi Masalah

Berdasarkan pemaparan dari latar belakang di atas, maka masalah yang dapat diidentifikasi adalah sebagai berikut :

- Masih kurangnya model pengenalan huruf alphabet.
- Banyaknya karakter yang dapat dibaca dalam satu image.

2. Batasan Masalah

Aplikasi yang akan dikembangkan dengan bahasa pemrograman Java dan berbasis desktop ini hanya akan mengenali huruf Alphabet saja. Sehingga tidak memungkinkan untuk mengenal jenis tulisan lainnya. Batasan masalah dalam penelitian ini adalah sebagai berikut :

- Melakukan pengenalan huruf alphabet dan membaca huruf alphabet.
- Membuat model jaringan saraf tiruan sebagai learning class serta membuat pembelajaran karakter dalam satu image maksimal satu karakter.
- Teks diutamakan ejaan bahasa Inggris sesuai speech yang ada.
- Menambahkan text-to-speech ke dalam aplikasi dengan teknik pengambilan teks pada speech synthesis.
- Inputan gambar training hanya berupa .bmp dan inputan gambar pada demo hanya berupa .gif.

3. Tujuan Penulisan

Tujuan penulisan ini untuk mempelajari dan mengembangkan sebuah aplikasi perangkat lunak yang dapat membaca karakter berdasarkan jenis tulisan huruf Aphabet dan membuat model pengolahan citra menjadi text dan dikonversi menjadi suara yang diharapkan menjadi bahan referensi untuk pengembangan dalam membuat aplikasi pada tahap selanjutnya.

II. LANDASAN TEORI

2.1 Pengolahan Citra

Pengolahan citra merupakan proses pengolahan dan analisis citra yang banyak melibatkan persepsi visual. Proses ini mempunyai ciri data masukan dan informasi keluaran yang berbentuk citra. Istilah pengolahan citra digital secara umum didefinisikan sebagai pemrosesan citra dua dimensi dengan komputer. Dalam definisi yang lebih luas, pengolahan citra digital juga mencakup semua data dua dimensi [1].

Operasi yang dilakukan untuk mentransformasikan suatu citra menjadi citra lain dapat dikategorikan berdasarkan tujuan transformasi maupun cakupan operasi yang dilakukan terhadap citra [2].

2.2 Morfologi

Dengan pendekatan morfologi, kita memandang suatu citra sebagai himpunan posisi (x,y) yang bernilai 1 atau 0 [3]. Operasi morfologi merupakan operasi yang digunakan untuk meningkatkan bentuk (struktur) sehingga lebih mudah untuk dikenali [4]. Secara umum, pemrosesan citra secara morfologi dilakukan dengan cara *mem-passing* sebuah *structuring element* terhadap sebuah citra dengan cara yang hampir sama dengan konvolusi. *Structuring element* dapat diibaratkan dengan *mask* pada pemrosesan citra biasa (bukan secara morfologi) [5].

2.3 OCR (Optical Character Recognition)

Citra atau gambar atau *image* merupakan suatu yang menggambarkan objek dan biasanya dalam bentuk dua dimensi. Citra merupakan suatu representasi kemiripan dari suatu objek atau benda. Citra digital didefinisikan sebagai representasi diskrit dari data spasial (tata letak) dan intensitas (warna) informasi.

2.4 Character Encoding

Character encoding adalah suatu metode untuk memasang suatu karakter dengan suatu simbol alami. Simbol alami yang dimaksud dapat berupa angka, huruf atau bahkan pulsa elektrik. Tujuan proses *character encoding* adalah untuk memfasilitasi penyimpanan teks dalam komputer ataupun pengiriman teks melalui jaringan telekomunikasi. Contoh dari proses pengkodean semacam ini adalah kode morse, yang memasang latin dengan panjang atau pendek suatu sinyal morse.

2.5 Unicode

Melakukan penalaran *Unicode* adalah suatu standar industri yang dirancang untuk melakukan *encoding* terhadap

karakter universal yang digunakan untuk merepresentasikan teks pada proses komputer. *Unicode* menyediakan suatu cara yang konsisten untuk melakukan *encoding* terhadap berbagai bahasa. Tujuan dari adanya standar ini adalah untuk memudahkan para pengguna komputer ketika berhubungan dengan teks yang berasal dari berbagai bahasa. Selain itu pengguna komputer yang menggunakan berbagai simbol matematis ataupun simbol teknis lainnya, akan sangat dimudahkan dengan penggunaan standar *Unicode*.

2.6 Algoritma Backpropagation

Pengertian *Backpropagation* merupakan sebuah metode sistematis pada jaringan saraf tiruan dengan menggunakan algoritma pembelajaran yang terawasi dan biasanya digunakan oleh *perceptron* dengan banyak layer lapisan untuk mengubah bobot-bobot yang ada pada lapisan tersembunyinya. *Backpropagation* adalah pelatihan jenis terkontrol dimana menggunakan pola penyesuaian bobot untuk mencapai nilai kesalahan yang minimum antara keluaran hasil prediksi dengan keluaran yang nyata.

2.7 Jaringan Saraf Tiruan

Jaringan saraf adalah sebuah prosesor yang terdistribusi *parallel* dan mempunyai kecenderungan untuk menyimpan pengetahuan yang didapatkannya dari pengalaman dan membuatnya tetap tersedia untuk digunakan.[6]

Hal ini menyerupai kerja otak dalam dua hal yaitu:

- Pengetahuan diperoleh oleh jaringan melalui suatu proses belajar.
- Kekuatan hubungan antar sel saraf yang dikenal dengan bobot sinapsis digunakan untuk menyimpan pengetahuan.

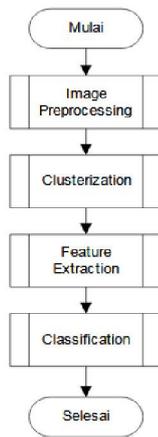
2.8 Speech Synthesis

Speech synthesis adalah transformasi dari teks ke arah suara (*speech*). Transformasi ini mengkonversi teks ke pembeda suara (*speech synthesis*) yang sebisa mungkin dibuat menyerupai suara nyata, disesuaikan dengan aturan-aturan pengucapan bahasa. TTS (*text-to-speech*) dimaksudkan untuk membaca teks elektronik dalam bentuk buku, dan juga untuk menyuarakan teks dengan menggunakan pembeda suara. Sistem ini dapat digunakan sebagai sistem komunikasi.

III. RANCANGAN SISTEM

3.1. Alur Proses Program Keseluruhan

Alur proses berikut ini merupakan alur proses program mulai dari aplikasi ini dijalankan sampai mendapatkan hasil keluaran/output.

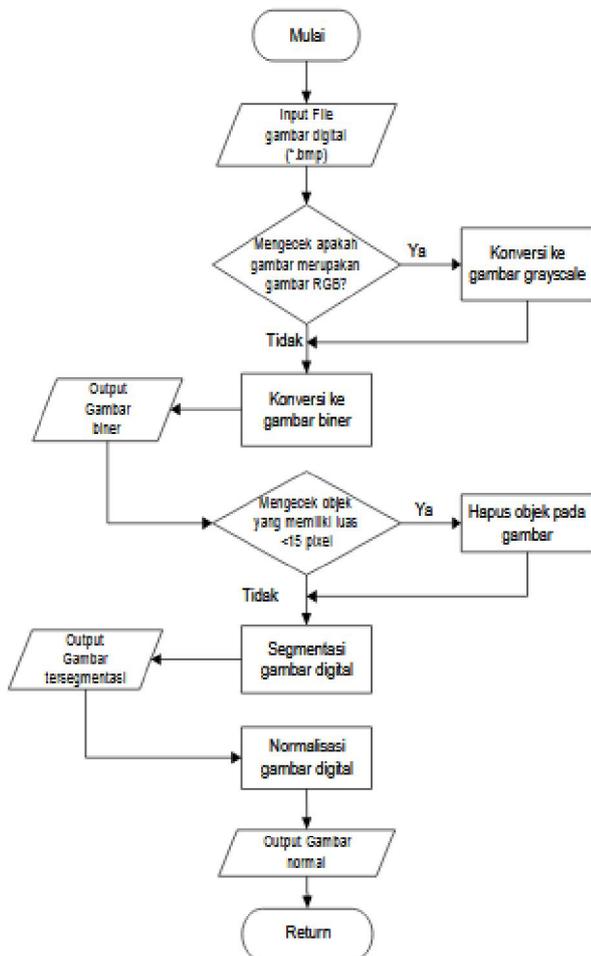


Gambar 1 : Kerangka Alur Proses Program

Dari gambar di atas, proses sistem terdiri dari dua tahapan, yaitu tahapan *preprocessing* yaitu pengolahan citra dan tahapan *postprocessing*.

1. Image Preprocessing

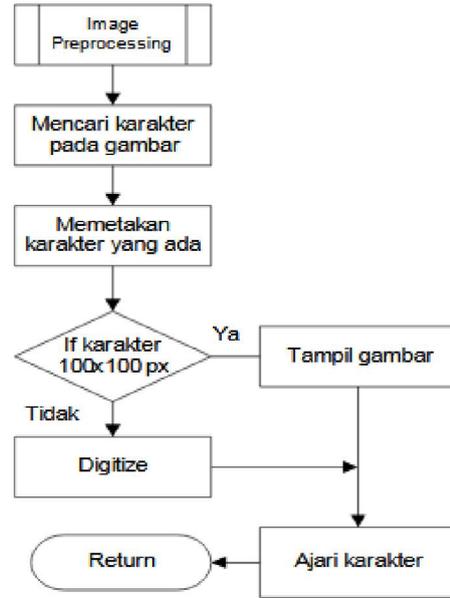
Modul ini merupakan suatu proses untuk menghilangkan bagian-bagian yang tidak diperlukan pada gambar input untuk melakukan proses selanjutnya.



Gambar 2 : Alur Proses Image Preprocessing

2. Clusterization

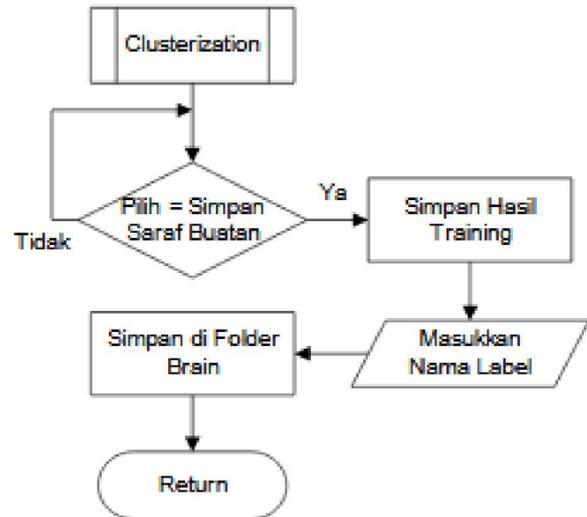
Modul ini bertugas untuk menemukan karakter individu dalam gambar dan menghitung urutan karakter yang benar dalam teks.



Gambar 3 : Alur Proses Clusterization

3. Feature Extraction

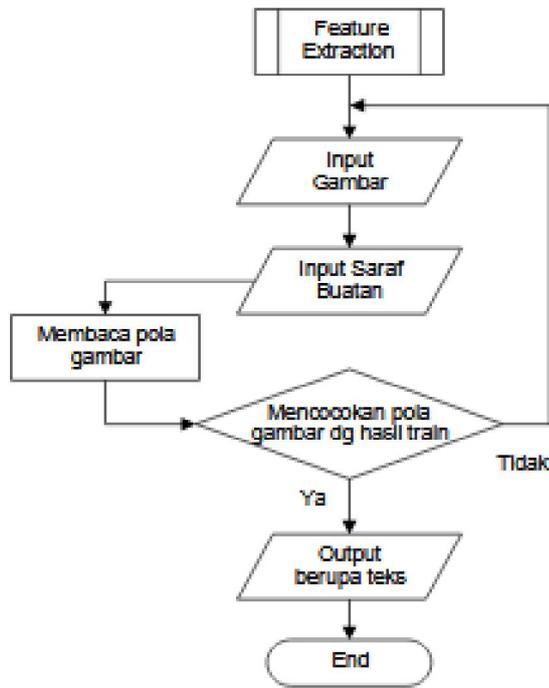
Modul ini bertujuan untuk mengekstrak fitur unik dari karakter individu sehingga dapat diakui oleh modul klasifikasi.



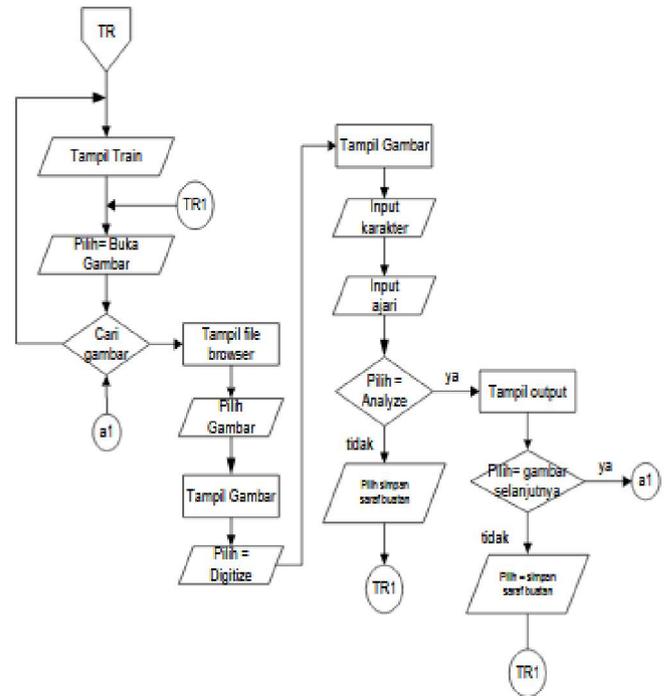
Gambar 4 : Alur Proses Feature Extraction

4. Classification

Ini merupakan modul terakhir pada sistem OCR yang mencoba mengenali karakter menggunakan informasi tentang fitur karakter yang telah diekstraksi pada tahap sebelumnya.



Gambar 5 : Alur Proses Classification

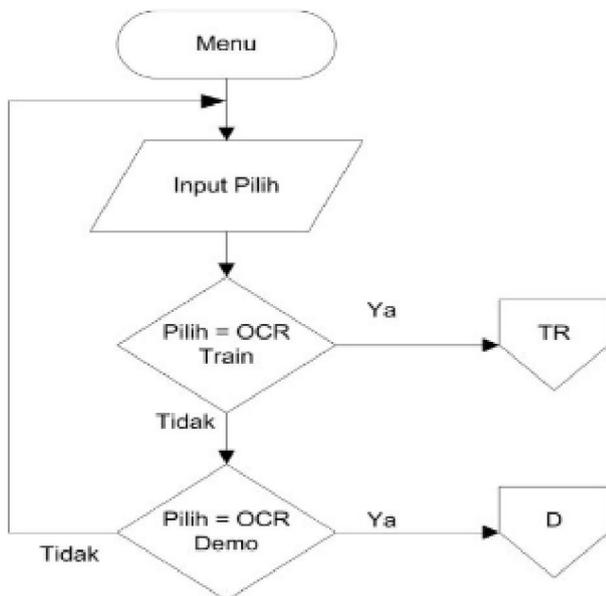


Gambar 7 : Alur Proses Form OCR Train

2.9 Alur Proses Aplikasi/Form

1. Alur Proses Awal Aplikasi

Alur proses awal aplikasi adalah alur saat aplikasi pertama dijalankan.



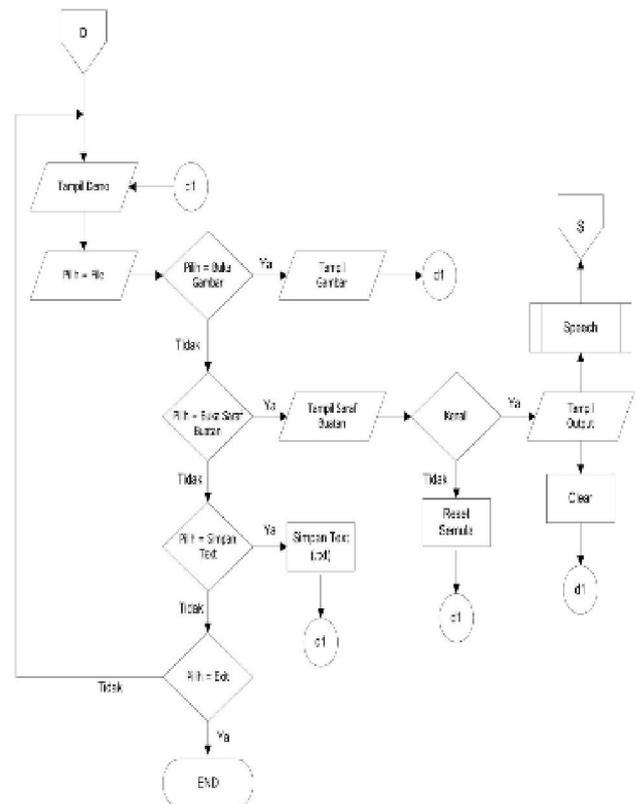
Gambar 6 : Alur Proses Awal Aplikasi

2. Alur Proses Form OCR Train

Pada menu ini diberikan pelatihan terhadap karakter data yang nantinya menjadi suatu input pada form OCR Demo.

3. Alur Proses Form OCR Demo

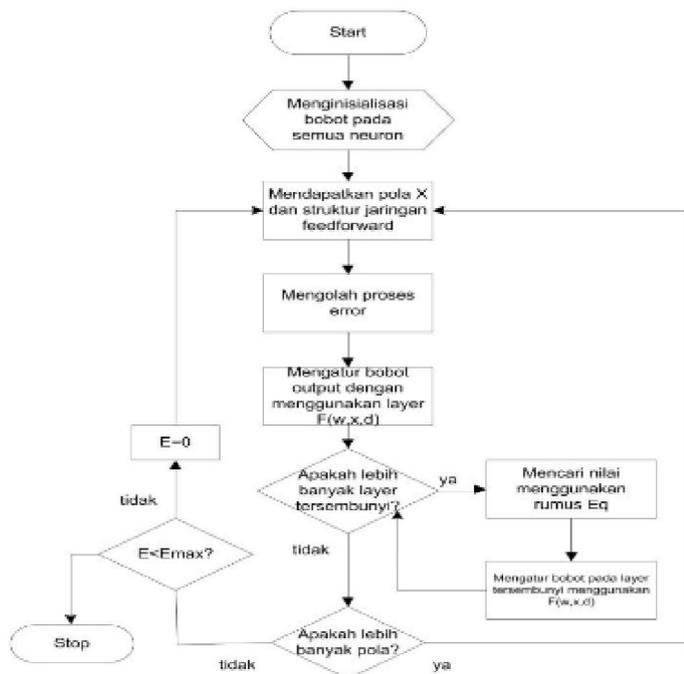
Pada menu ini dijalankan eksekusi terhadap gambar berupa .gif dan membuka saraf buatan yang sudah dibuat pada OCR Train.



Gambar 8 : Alur Proses Form OCR Demo

4. Alur Proses Algoritma Backpropagation

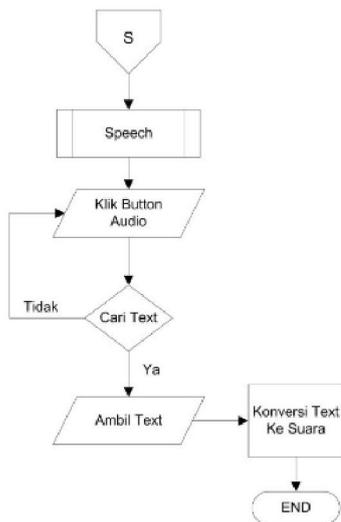
Aplikasi *text recognition* ini menggunakan algoritma *backpropagation*.



Gambar 9 : Alur Proses Algoritma Backpropagation

5. Alur Proses Text-To-Speech

Alur proses ini berjalan pada saat *button* Audio pada OCR Demo diklik.



Gambar 10 : Alur Proses Text-To-Speech

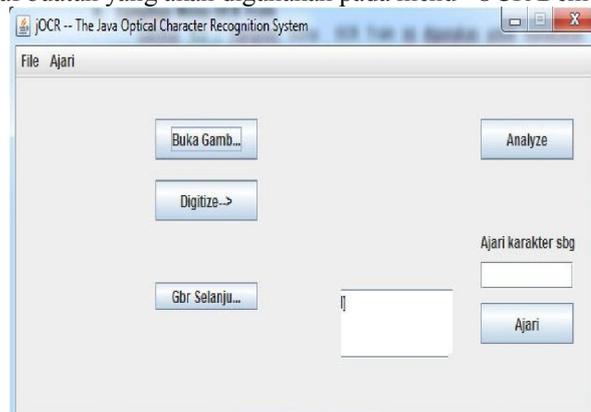
IV. HASIL DAN PEMBAHASAN

4.1. Tampilan Layar Aplikasi

1. Tampilan Layar OCR Train

Tampilan menu OCR Train ini digunakan untuk melakukan *training*/pelatihan teks yang berasal dari gambar

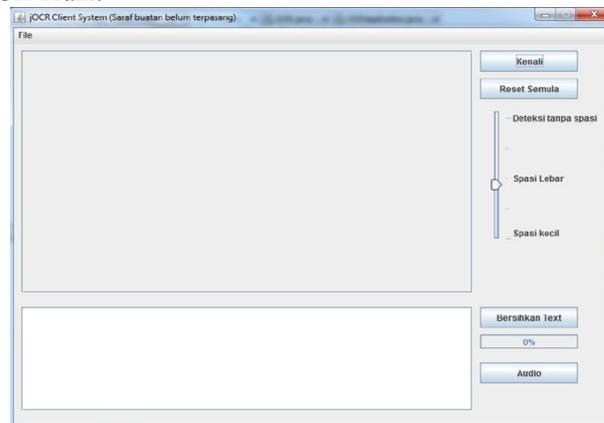
yang berekstensi .bmp. Hasil dari menu “OCR Train” ini adalah saraf buatan yang akan digunakan pada menu “OCR Demo”.



Gambar 11 : Tampilan Layar OCR Train

2. Tampilan Layar OCR Demo

Tampilan menu OCR Demo ini digunakan untuk melakukan pembacaan teks dari gambar yang berekstensi .gif. Karakter pada gambar akan dikenali dengan memanggil/membuka saraf buatan yang telah dibuat pada menu OCR Train.



Gambar 12 : Tampilan Layar OCR Demo

4.2. Hasil Uji Coba Aplikasi

Parameter keberhasilan pada sistem adalah ketepatan pengenalan sistem terhadap beberapa masukan citra karakter. Pengujian dilakukan secara berkelompok. Di mana terdapat 3 kelompok yaitu pengujian pada huruf Georgia dengan saraf buatan Comic Sans dan huruf kapital dengan spasi.

Tabel 1 : Tabel Pengujian Data

No.	Citra Masukan	Hasil	Jumlah Karakter yang berhasil dibaca
1.	HANDPHONE	HANDPHON E	9
2.	S A L E	SAZE	4
3.	ogep ces	osep ces	7

Pada pengujian setiap kata terlihat bahwa kesalahan sering terjadi pada huruf yang agak mirip. Dengan adanya template sistem masih bisa membedakan antara huruf besar dengan huruf kecil. Tetapi untuk huruf seperti g dan p yang memiliki ekor di bawah sering terjadi kesalahan. Karena untuk beberapa jenis font, bentuknya juga mirip. Untuk tanda baca yang memiliki bentuk amat kecil seperti titik dan koma, masih sering terdapat kesalahan, karena pada proses *thinning* keduanya dapat memiliki bentuk yang sama, di mana sebuah titik yang tebal adakalanya pada proses *thinning* menjadi sebuah garis kecil seperti pada tanda baca koma. Dan masih terdapat kesalahan juga apabila huruf dalam gambar terlalu rapat spasinya.

4.3. Kelebihan dan Kekurangan Program

1. Kelebihan Program

Kelebihan yang dimiliki oleh aplikasi *text recognition* ini, antara lain:

- a. Kemudahan untuk memakai aplikasi ini karena dapat dioperasikan di berbagai *platform* OS.
- b. Aplikasi *text recognition* ini memiliki antar muka (*interface*) yang mudah dikenali dan dikerjakan.
- c. Untuk bisa membaca jenis karakter huruf Alphabet yang lain, user hanya perlu melakukan pelatihan/*training* pengenalan karakter pada form OCR Train.
- d. Aplikasi ini bisa menjadi referensi untuk pengembangan dan penelitian program yang lebih baik lagi.
- e. Dapat mengeluarkan *speech* sesuai *text* yang berhasil.

2. Kekurangan Program

Kekurangan yang dimiliki oleh aplikasi *text recognition* ini, antara lain:

- a. *File* gambar yang di-*training* pada OCR Train hanya bisa dibaca dengan resolusi 100x100 *pixel* dan dengan ekstensi (*.*bmp*).
- b. *File* gambar karakter pada OCR Demo dapat dibuka dengan gambar yang memiliki ekstensi (*.*gif*) saja.
- c. *Speech* masih berupa ejaan dalam bahasa Inggris.

V. KESIMPULAN

Sesuai dengan pembahasan mengenai *prototype text recognition* dengan klasifikasi *neural network* dan *text-to-speech* pada huruf Alphabet, maka kesimpulan yang dapat diambil adalah sebagai berikut:

1. Hasil training dari aplikasi ini belum sempurna, karena tidak dapat mengenali berbagai macam karakter huruf Alphabet secara baik.
2. Maksimal karakter yang mampu dibaca oleh aplikasi adalah 128 karakter (karakter latin), berdasarkan code pages ASCII.
3. Inputan karakter untuk ajari di form train hanya dapat mengajari 1 karakter saja.
4. Aplikasi ini dapat digunakan pada karakter huruf yang lain agar lebih mudah dalam proses belajar terutama yang merupakan 1 karakter.

DAFTAR PUSTAKA

- [1] Riyanto, Citra-Bab 8, http://lecturer.eepisits.edu/riyanto/citra_bab8.pdf, diakses pada tanggal 2 April 2016.
- [2] Ginting, E.D., Deteksi Tepi Menggunakan Metode Canny dengan Netbeans untuk Membedakan Uang Asli dan Uang Palsu, Universitas Gunadarma, 2010.
- [3] Arwariningsih, S.H., Perhitungan Luas dan Keliling Bangun Geometri Menggunakan Pendekatan Morfologi, Seminar Nasional Aplikasi Teknologi Informasi, 2009.
- [4] Usman, A., Pengolahan Citra Digital dan Teknik Pemrogramannya, Penerbit Graha Ilmu, 2005
- [5] Arymurthy, A.M., Diktat Kuliah CITRA, Jakarta: Universitas Inggris. RSI Team. 2004.
- [6] Haykin, S., Neural Networks: A Comprehensive Foundation, NY, Macmillan, 2004