

ABSTRACT

Winston Renatan (01082170030)

ANDROID APPLICATION TO IDENTIFY AND RECOGNIZE TEXT IMAGES WITH OPTICAL CHARACTER RECOGNITION AND TEXT-TO-SPEECH

(xv + 159 pages; 81 figures; 15 tables; 5 appendices)

Learning style can be distinguished over three categories, namely visual, auditory, and kinesthetic. Auditory learners will utilize hearing as the dominant means of learning. Research shows that Indonesia's literacy rank is quite low compared to 60 other countries. Other research revealed that millennials don't like conventional reading. Based on Indonesia's population proportions, millennials make up 33.75% of the total population. Considering such significant proportion, this research designed an Android application to identify and recognize text images to be converted to spoken text.

The implementation of this application is called Snapspeech which consists of four stages, namely loading asset, reading image input, processing image, and sound synthesis from the recognized text. The loaded asset is a dictionary used in the process of matching words of the recognition result. Valid input is either an image from the gallery or captured through smartphone's camera. The TextRecognizer module is utilized to recognize text and generate digital text output that will be converted into spoken text.

Application testing includes reading ability and accuracy of recognition results. Testing was conducted using smartphone where image is taken with constant distance of $\pm 30\text{cm}$. The sample tested was a text image consisting of 50 words, 6 typefaces from 3 different typography categories, and 5 typeface sizes. The test results showed that Snapspeech have a text recognition accuracy rate of 99.33% and can process unrotated and rotated text image with 90, 180, and 270 degrees. Typefaces testing shows that Cambria and Calibri gave the best results; while Courier New typeface test gave the worst results. The best detection results obtained for typeface size is 14pt. The resulting text of recognition can be read without pause between lines, making it sound more natural. Additional testing was also conducted to see the final Snapspeech application capabilities.

References: 44 (1997—2020).

ABSTRAK

Winston Renatan (01082170030)

APLIKASI ANDROID UNTUK IDENTIFIKASI DAN RECOGNISI GAMBAR TEKS DENGAN *OPTICAL CHARACTER RECOGNITION* DAN *TEXT-TO-SPEECH*

(xv + 159 halaman; 81 gambar; 15 tabel; 5 lampiran)

Gaya belajar (*learning style*) seseorang dapat dibedakan atas tiga kategori, yaitu visual, auditori dan kinestetik. Seseorang dengan gaya belajar auditori akan memanfaatkan pendengaran sebagai sarana yang dominan untuk belajar. Data penelitian menunjukkan bahwa peringkat literasi Indonesia cukup rendah dibandingkan dengan 60 negara lain. Penelitian lain mengungkapkan bahwa generasi milenial tidak menggemari membaca secara konvensional. Berdasarkan proporsi penduduk Indonesia, generasi milenial membentuk 33,75% dari total penduduk. Mempertimbangkan proporsi yang signifikan tersebut, melalui penelitian ini dirancang sebuah aplikasi Android untuk mengidentifikasi dan merekognisi teks pada citra untuk dialihkan menjadi teks yang disuarakan (*spoken text*).

Implementasi aplikasi yang diberi nama Snapspeech terdiri dari empat tahap, yakni pemuatan aset, pembacaan citra masukan, pemrosesan citra, dan sintesis suara dari teks hasil. Aset yang dimuat berupa kamus yang digunakan dalam proses pencocokan kata-kata hasil rekognisi. Masukan yang sahih adalah citra baik yang berada pada galeri maupun tangkapan kamera telepon pintar. Modul TextRecognizer dimanfaatkan untuk merekognisi teks dan menghasilkan luaran berupa teks digital untuk kemudian dikonversi menjadi teks yang diujarkan.

Pengujian aplikasi meliputi kemampuan baca aplikasi dan akurasi hasil rekognisi. Pengujian dilakukan menggunakan telepon pintar di mana citra diambil pada jarak konstan $\pm 30\text{cm}$. Sampel yang diuji berupa citra teks yang terdiri dari 50 kata, 6 *typeface* dari 3 kategori tipografi berbeda, dan 5 ukuran *typeface*. Hasil pengujian menunjukkan bahwa aplikasi Snapspeech memiliki tingkat akurasi rekognisi teks sebesar 99,33% dan dapat memproses citra teks baik yang tidak dirotasi maupun yang dirotasi sebesar 90, 180, dan 270 derajat. Pengujian atas *typeface* Cambria dan Calibri memberi hasil terbaik; sementara hasil uji *typeface* Courier New memberi hasil terburuk. Hasil pendekripsi terbaik diperoleh untuk ukuran *typeface* 14pt. Teks hasil rekognisi dapat diujarkan tanpa jeda antar baris, sehingga hasil pembacaan terdengar lebih natural. Pengujian tambahan juga dilakukan untuk melihat kapabilitas final dari aplikasi Snapspeech.

Referensi: 44 (1997—2020).