

PENERJEMAHAN HURUF CYRILLIC RUSIA KE HURUF LATIN MENGGUNAKAN ALGORITMA SVM (SUPPORT VECTOR MACHINE)

TRANSLATION RUSSIAN CYRILLIC TO LATIN ALPHABET USING SVM (SUPPORT VECTOR MACHINE)

Dian Faruqi Azid, A.Md.T.¹, Budhi Irawan S.Si., M.T.², Casi Setianingsih S.T., M.T.³

^{1,2,3}Prodi S1 Sistem Komputer, Fakultas Teknik Elektro, Telkom University
Bandung, Indonesia

¹faruqiazid76@gmail.com, ²budhiirawan@telkomuniversity.ac.id, ³setiacasie@telkomuniversity.ac.id

Abstrak

Bahasa Rusia adalah bahasa yang cukup banyak digunakan oleh masyarakat didunia untuk berbagai keperluan karena menempati peringkat ke-6 di dunia sebagai bahasa dengan penutur terbanyak. Hal ini yang menjadi salah satu dasar dari penelitian ini selain tingginya minat penduduk dunia untuk menggunakan dan mempelajari bahasa tersebut. Bahasa Rusia tidak menggunakan aksara yang pada umumnya menggunakan alphabet namun menggunakan aksara Cyrillic, dimana bentuk hurufnya berbeda dari huruf pada umumnya sehingga terdapat beberapa kendala untuk mempelajari, memahami dan melafalkannya.

Capture to translate adalah salah satu media yang dibangun untuk menjadi sebuah solusi dari permasalahan tersebut, dibangun berdasarkan pengolahan *image processing*, proses ekstraksi ciri dengan metode deteksi tepi *findcontours* dan kecerdasan buatan (*artificial intelligence*) menggunakan algoritma klasifikasi *Support Vector Machine (SVM)* dengan antar muka aplikasi *mobile* Android yang memanfaatkan kamera perangkat sebagai inputnya.

Didalam penelitian ini, *Capture to translate* dengan menggunakan algoritma klasifikasi *Support Vector Machine (SVM)* mampu menghasilkan tingkat akurasi klasifikasi kata tertinggi sebesar 93,8 % pada pengujian tiga silabel berdasarkan dari pengujian yang telah dilakukan dengan proses *preprocessing*, ekstraksi ciri dan klasifikasi. Kata Kunci : Rusia, Cyrillic, *Capture to Translate*, *Image Processing*, deteksi tepi, *findcontours* Kecerdasan Buatan, *Support Vector Machine (SVM)*, *Android*, *Preprocessing*..

Abstract

Russian is a language that is widely used by people in the world for various purposes because it is ranked 6th in the world as the language with the most speakers. This is one of the foundations of this study in addition to the high interest of the world population to use and learn the language. Russian language does not use alphabets but uses Cyrillic script, where the font is different from the letters in general so there are some obstacles to learn, understand and pronounce it.

Capture to translate is one of the media that is built to be a solution of the problem, built on image processing, feature extraction process with edges detection findcontours method and artificial intelligence using Support Vector Machine (SVM) classification algorithm With Android mobile application interface that utilizes camera device as its input.

In this study, Capture to translate using the Support Vector Machine (SVM) classification algorithm is able to produce a level of word classification accuracy of 93.8% in three syllable based on test that has been done, which is related to preprocessing, feature extraction and classification.

Keywords : Russia, Cyrillic, Capture to Translate, Image Processing, edges detection, findcontours, Artificial Intelligence, Support Vector Machine (SVM), Android, Preprocessing.

1. Pendahuluan

Bahasa Rusia adalah salah satu bahasa dengan penutur terbanyak yang menempati peringkat ke-6 didunia. Bahasa Rusia merupakan satu dari enam bahasa resmi yang digunakan oleh *United Nation (UN)* atau yang biasa dikenal di Indonesia sebagai Perserikatan Bangsa-Bangsa (PBB). Hal ini menunjukkan bahwa tingginya minat penduduk dunia

untuk menggunakan dan mempelajari bahasa tersebut, tetapi terdapat beberapa kendala untuk mempelajari, memahami dan melafalkan bahasa Rusia dikarenakan bahasa ini menggunakan aksara non-latin *Cyrillic*. Permasalahan yang sering terjadi diantaranya seperti masyarakat asing hendak membaca tulisan bahasa Rusia, mengetikkan sebuah nama tempat atau daerah dalam bahasa tersebut, dan melafalkan dengan aksen yang benar dikarenakan bahasa tersebut tidak menggunakan huruf latin.

Capture to translate adalah salah satu media yang dibangun untuk menjadi sebuah solusi dari permasalahan tersebut, sistem ini dibangun berdasarkan pengolahan *image processing* dan kecerdasan buatan (*Artificial Intelligence*) dengan antar muka aplikasi *mobile* yang memanfaatkan kamera perangkat sebagai inputnya. Pengambilan gambar yang mengandung teks dalam bahasa Rusia akan di proses terlebih dahulu didalam *image processing*. Kemudian sistem ini akan berjalan secara otomatis menjalankan proses transliterasi ke aksara latin. Proses transliterasi ini berjalan diatas *machine learning* yang sebelumnya telah diberikan proses pelatihan kemudian di sinkronisasi menggunakan algoritma klasifikasi *Support Vector Machine (SVM)*.

2. Dasar Teori

2.1 Digital Image Processing

Digital Image Processing adalah kumpulan metode-metode yang mempelajari tentang teknik-teknik mengolah sebuah citra. Citra yang dimaksud disini adalah gambar diam (foto) maupun gambar bergerak (video). Sedangkan digital disini mempunyai maksud bahwa pengolahan citra/gambar dilakukan secara digital menggunakan komputer. Secara matematis, citra merupakan fungsi kontinu (continue) dengan intensitas cahaya pada bidang dua dimensi. Agar dapat diolah dengan komputer digital, maka suatu citra harus dipresentasikan secara numeric dengan nilai-nilai diskrit. Reperesentasi dari fungsi kontinu menjadi nilai-nilai diskrit disebut digitalisasi citra. Sebuah citra digital dapat diwakili oleh sebuah matriks dua dimensi $f(x,y)$ yang terdiri dari M kolom dan N baris, dimana perpotongan antara kolom dan baris disebut piksel (*pixel = picture element*) atau elemen terkecil dari sebuah citra.[1]

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

Gambar 2.1 Citra Digital

Suatu citra $f(x,y)$ dalam fungsi matematis dapat dituliskan sebagai berikut:

$$\begin{aligned} 0 \leq x \leq M-1 \\ 0 \leq y \leq N-1 \dots\dots\dots (2.1) \\ 0 \leq f(x,y) \leq G-1 \end{aligned}$$

Dimana :

M = jumlah piksel baris (row) pada array citra

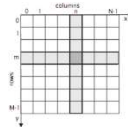
N = jumlah piksel kolom (column) pada array citra

G = nilai skala keabuan (graylevel)

Besarnya nilai M, N dan G pada umumnya merupakan perpangkatan dari dua.

$M = 2^m ; N = 2^n ; G = 2^k$

dimana nilai m, n dan k adalah bilangan bulat positif. Interval (0,G) disebut skala keabuan (grayscale). Besar G tergantung pada proses digitalisasinya. Biasanya keabuan 0 (nol) menyatakan intensitas hitam dan 1 (satu) menyatakan intensitas putih. Untuk citra 8 bit, nilai G sama dengan $28 = 256$ warna (derajat keabuan)[2].



Gambar 2.2 Matriks Citra Digital

Obyek tertentu dapat dideteksi dengan menggunakan pengolahan citra digital ini. Salah satu metode yang digunakan adalah berdasarkan segmentasi warna. Normalisasi RGB adalah salah satu metode segmentasi warna yang memiliki kelebihan yaitu mudah, proses cepat dan efektif pada obyek traffic sign [3].

2.2 Computer Vision

Definisi Kecerdasan Buatan

H. A. Simon [1987] :

“Kecerdasan buatan (*artificial intelligence*) merupakan kawasan penelitian, aplikasi dan instruksi yang terkait dengan pemrograman computer untuk melakukan sesuatu hal yang dalam pandangan manusia adalah cerdas”.

Rich and Knight [1991] :

“Kecerdasan Buatan (*AI*) merupakan sebuah studi tentang bagaimana membuat computer melakukan hal – hal yang pada saat ini dapat dilakukan lebih baik oleh manusia”.

Tujuan dari kecerdasan buatan menurut Winston dan Prendergast [1984]:

1. Membuat mesin menjadi lebih pintar (tujuan utama)
2. Memahami apa itu kecerdasan (tujuan ilmiah)
3. Membuat mesin lebih bermanfaat (tujuan entrepreneurial)

AI dapat dipandang dalam berbagai perspektif.

1. Dari perspektif Kecerdasan (*Intelligence*) AI adalah bagaimana membuat mesin yang “cerdas” dan dapat melakukan hal-hal yang sebelumnya dapat dilakukan oleh manusia
2. Dari perspektif bisnis, AI adalah sekelompok alat bantu (*tools*) yang berdaya guna, dan metodologi yang menggunakan tool-tool tersebut guna menyelesaikan masalah-masalah bisnis.
3. Dari perspektif pemrograman (*Programming*), AI termasuk didalamnya adalah studi tentang pemrograman simbolik, pemecahan masalah, proses pencarian (*search*).

2.3 Algoritma Klasifikasi SVM

Klasifikasi adalah sebuah proses untuk membedakan suatu bentuk kelas data yang sudah diketahui maupun yang tidak diketahui dengan tujuan untuk mengelompokkan kelas tersebut berdasarkan hasil prediksi algoritma yang digunakan. Dalam pengklasifikasian data terdapat dua proses yang dilakukan yaitu:

1. Proses training Pada proses training digunakan training set yang telah diketahui label-labelnya untuk membangun model atau fungsi.
2. Proses testing Untuk mengetahui keakuratan model atau fungsi yang akan dibangun pada proses training, maka digunakan data yang disebut dengan testing set untuk memprediksi label-labelnya.

Support Vector Machine (SVM) dikembangkan oleh Boser, Guyon, dan Vapnik, pertama kali diperkenalkan pada tahun 1992 di Annual Workshop on Computational Learning Theory. Konsep dasar metode SVM sebenarnya merupakan gabungan atau kombinasi dari teori-teori komputasi yang telah ada pada tahun sebelumnya, seperti marginhyperplane (Dyda dan Hart, 1973; Cover, 1965; Vapnik, 1964), kernel diperkenalkan oleh Aronszajn tahun 1950, Lagrange Multiplier yang ditemukan oleh Joseph Louis Lagrange pada tahun 1766, dan demikian juga dengan konsep-konsep pendukung lain. Menurut Fachrurrazi (2011) SVM merupakan suatu teknik untuk melakukan prediksi, baik prediksi dalam kasus regresi maupun klasifikasi. Teknik SVM digunakan untuk mendapatkan fungsi pemisah (hyperplane) yang optimal untuk memisahkan observasi yang memiliki nilai variabel target yang berbeda (William, 2011). Hyperplane ini dapat berupa line pada two dimension dan dapat berupa flat plane pada multiple dimension. Support Vector Machine (SVM) adalah sistem pembelajaran yang pengklasifikasiannya menggunakan ruang hipotesis berupa fungsi-fungsi linear dalam sebuah ruang fitur (feature space) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik (Nello Christianini dan John S. Taylor, 2000).

Dalam konsep SVM berusaha menemukan fungsi pemisah (hyperplane) terbaik diantara fungsi yang tidak terbatas jumlahnya. Hyperplane pemisah terbaik antara kedua kelas dapat ditemukan dengan mengukur margin hyperplane tersebut dan mencari titik maksimalnya. Adapun data yang berada pada bidang pembatas disebut support vector. Secara matematika, konsep dasar SVM yaitu :

$$\min \frac{1}{2} |w|^2$$

$$s.t \ y_i(x_i \cdot w + b) - 1 \geq 0 \dots\dots\dots(2.2)$$

Dimana $(x_i \cdot W + b) \geq 1$ untuk kelas 1 dan $(x_i \cdot W + b) \geq -1$ untuk kelas 2. Xi adalah data set, yi adalah output dari data xi, dan w, b adalah parameter yang dicari nilainya. Formulasi optimasi SVM untuk kasus klasifikasi dua kelas dibedakan menjadi klasifikasi linear dan non-linear. Kelebihan Support Vector Machine (SVM) adalah SVM tidak mengalami overfitting karena training perlu dilakukan sekali saja dan mendapatkan solusi optimal. [7]

2.4 Bahasa Rusia

Bahasa Rusia adalah salah satu bahasa dengan penutur terbanyak yang menempati peringkat ke-6 didunia. Bahasa Rusia merupakan satu dari enam bahasa resmi yang digunakan oleh United Nation (UN) atau yang biasa dikenal di Indonesia sebagai Perserikatan Bangsa-Bangsa (PBB). Hal ini menunjukkan bahwa tingginya minat penduduk dunia untuk menggunakan dan mempelajari bahasa tersebut akan tetapi terdapat beberapa kendala untuk mempelajari, memahami dan melafalkan Bahasa Rusia dikarenakan bahasa ini menggunakan aksara non-latin Cyrillic.

Аа	Бб	Вв	Гг	Дд	Ее	Ёё	Жж	Зз
а	б	в	г	д	е	ё	ж	з
[a]	[b]	[v]	[g]	[d]	[ye]	[yo]	[z]	[z]
Ии	Йй	Кк	Лл	Мм	Нн	Оо	Пп	Рр
и	й	к	л	м	н	о	п	р
[i]	[j]	[k]	[l]	[m]	[n]	[o]	[p]	[r]
Сс	Тт	Уу	Фф	Хх	Цц	Чч	Шш	Щщ
с	т	у	ф	х	ц	ч	ш	щ
[s]	[t]	[u]	[f]	[x]	[t͡s]	[t͡ʃ]	[ʃ]	[ʃ]
Ъъ	Ыы	Ьь	Ээ	Юю	Яя			
ъ	ы	ь	э	ю	я			
[ɤ]	[ɨ]	[ɨ]	[ɛ]	[ju]	[ja]			

Gambar 2.3 Aksara Cyrillic

2.5 Transliterasi dan Translasi

Transliterasi adalah sebuah proses menerjemahkan sebuah penulisan dari suatu bahasa yang menggunakan aksara non-latin kedalam penulisan latin sehingga mampu dilafalkan oleh masyarakat umum. Sedangkan *translate* diambil dari bahasa inggris yang artinya menerjemahkan suatu bahasa (latin) kedalam bahasa lain yang juga menggunakan aksara penulisan latin.

2.6 Penelitian Sebelumnya

Penelitian berjudul Machine Learning Multi Klasifikasi Citra Digital membahas tentang pembangunan aplikasi *machine learning* multi klasifikasi citra digital sebagai alat bantu bagi para peneliti citra untuk mendapatkan algoritma *learning* yang optimal dalam mengenali multi kelas berbagai objek citra digital. Algoritma yang digunakan dalam penelitian ini adalah algoritma klasifikasi *machine learning* WEKA yaitu *Support Vector Machine*, *K-Nearest Neighbor*, *Naive Bayes*, *C4.5 Decision Tree*, *Logistic Regression* dan *Random Forest*. Aplikasi dibangun menggunakan bahasa pemrograman Java yang didukung oleh *library machine learning* WEKA[8]. Hasil penelitian ini menyimpulkan bahwa algoritma dengan tingkat akurasi terbaik kedua dari beberapa algoritma tersebut adalah *Support Vector Machine*.

Penelitian berjudul Klasifikasi Gambar Berwarna menggunakan *K-Nearest Neighbor* dan *Support Vector machine* juga menjadi acuan dalam penelitian ini. Didalam jurnal penelitian tersebut membahas tentang bagaimana menentukan kategori gambar dalam jumlah yang besar dengan menggunakan komputer secara otomatis berdasarkan klasifikasi warna, tekstur, koefisien *magnitude fourier*, dan koefisien *Discrete Cosine Transform*. Kemudian didalam jurnal tersebut dilakukan pengujian menggunakan *Corel Dataset* dan *feature selection* sehingga menghasilkan tingkat akurasi yang hampir mirip tingkat keakurasiannya antara metode *KNN* dan *SVM* [9].

Penelitian lainnya yang menjadi acuan didalam penelitian ini adalah penelitian yang berjudul Segmentasi Citra Menggunakan *Support Vector Machine (SVM)* Dan *Ellipsoid Region Search Strategy (ERSS)* Arimoto Entropy Berdasarkan Ciri Warna Dan Tekstur. Didalam penelitian tersebut membahas tentang segmentasi citra menjadi beberapa bagian bedasar kriteria kemiripan tertentu dengan tujuan menghasilkan citra *Boundary* yang optimal dengan syarat klasifikasi piksel citra secara linier dan nonlinier. Hasil dari penelitian tersebut menghasilkan tingkat akurasi sebesar 69% yang didapat dari ekstraksi ciri warna dan tekstur yang menggunakan *Gray Level Co-occurrence Matrix (GLCM)* didalam klasifikasi *SVM* yang dikombinasi dengan metode pelabelan Arimoto berbasis *ERSS* [10].

2.7 Deteksi Tepi

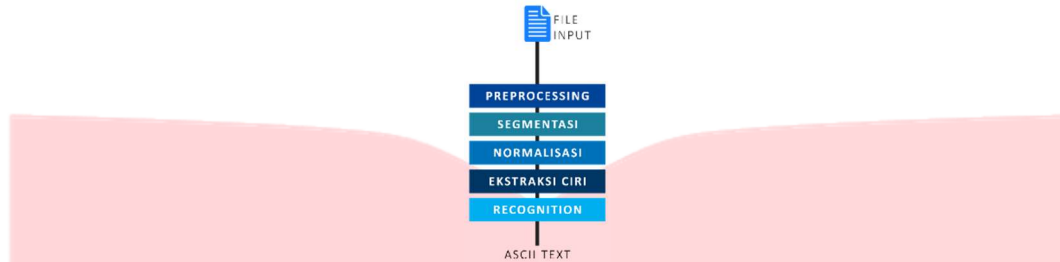
Detektor tepi merupakan proses untuk menemukan perubahan intensitas yang berbeda nyata dalam sebuah citra [11]. Deteksi tepi merupakan pendekatan yang paling umum untuk pendeteksian diskontinuitas nilai intensitas (Prasetyo, [12]).

2.8 OpenCV

OpenCV adalah sebuah library yang berisi fungsi-fungsi pemrograman untuk teknologi Computer Vision secara real time. OpenCV bersifat open source, bebas digunakan untuk hal-hal yang bersifat akademis maupun komersial. Di dalam OpenCV, terdapat interface untuk bahasa pemrograman C, C++, Python, dan nantinya Java yang dapat berjalan pada Windows, Linux, Android dan Mac. Terdapat lebih dari 2500 algoritma dalam OpenCV, digunakan di seluruh dunia, telah lebih dari 2.5 juta kali diunduh, dan digunakan lebih dari 40 ribu orang. Penggunaannya antara lain pada seni interaktif, inspeksi tambang, menampilkan peta di web melalui teknologi robotik. Pada awalnya OpenCV ditulis dengan menggunakan bahasa C namun sekarang secara menyeluruh sudah menggunakan antarmuka bahasa C++ dan seluruh pengembangannya terdapat dalam format bahasa C++. Contoh aplikasi dari OpenCV yaitu interaksi manusia dengan komputer: identifikasi, segmentasi, pengenalan objek, pengenalan wajah, pengenalan gerakan, penelusuran gerakan, gerakan diri, dan pemahaman gerakan, struktur dari gerakan, kalibrasi stereo dan beberapa kamera serta komputasi mendalam, dan robotic[13].

2.9 OCR (Optical Character Recognition)

OCR (Optical Character Recognition) adalah sebuah aplikasi komputer yang digunakan untuk mengidentifikasi citra huruf maupun angka untuk dikonversi ke dalam bentuk file tulisan [14]. Sistem pengenalan huruf ini dapat meningkatkan fleksibilitas atau kemampuan dan kecerdasan sistem komputer. Sistem pengenalan huruf yang cerdas sangat membantu usaha besar – besaran yang saat ini dilakukan banyak pihak yakni usaha digitalisasi informasi dan pengetahuan, misalnya dalam pembuatan koleksi pustaka digital, koleksi sastra kuno digital, dan lain – lain.[15]



Gambar 2.4 Optical Character Recognition

Secara umum proses OCR dapat dilihat pada gambar diatas, dengan penjelasan sebagai berikut [16] :

a. File Input

File input berupa file citra digital dengan format *.bmp atau *.jpg.

b. Preprocessing

Preprocessing merupakan suatu proses untuk menghilangkan bagian – bagian yang tidak diperlukan pada gambar input untuk proses selanjutnya.

c. Segmentasi

Segmentasi adalah proses memisahkan area pengamatan (region) pada tiap karakter yang dideteksi.

d. Normalisasi

Normalisasi adalah proses merubah dimensi region tiap karakter dan ketebalan karakter.

e. Ekstraksi Ciri

Ekstraksi ciri adalah proses untuk mengambil ciri – ciri tertentu dari karakter yang diamati.

f. Recognition

Recognition merupakan proses untuk mengenali karakter yang diamati dengan cara membandingkan ciri – ciri karakter yang diperoleh dengan ciri – ciri karakter yang ada pada basis data.

3. Pembahasan

3.1 Gambaran Umum Sistem

Sistem ini dirancang dan dibangun pada sebuah sistem operasi perangkat *mobile* yang berbasis android. Penelitian ini memiliki tujuan membangun sebuah prototipe yang mampu mengimplementasikan *image processing* dan kecerdasan buatan untuk dapat mengenali suatu teks atau aksara non-latin Rusia (Cyrillic) yang berada pada suatu citra sehingga dapat diberikan keterangan meliputi transliterasi dan penerjemahan teks tersebut kedalam bahasa Indonesia. Proses pengenalan teks atau aksara Rusia diawali dengan menentukan citra yang didapatkan langsung oleh pengguna menggunakan media kamera maupun yang terdapat pada media penyimpanan dengan prinsip dasar OCR (*Optical Character Recognition*). Proses akuisisi citra yang sebelumnya telah dilakukan bertujuan untuk menjadikan citra tersebut sebagai teks agar dapat dilanjutkan pada proses transliterasi dan penerjemahan ke dalam bahasa Indonesia.

Berikut adalah gambaran umum dari system yang dirancang dalam penelitian ini :

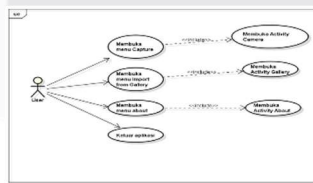


Gambar 3.1 Skema rancangan sistem

3.2 Use Case Diagram

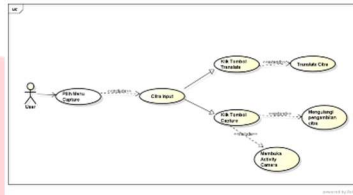
Use Case Diagram merupakan gambaran graphical dari sistem yang melibatkan aktor, use case dan interaksi. Use case diagram digunakan untuk memperkenalkan suatu sistem serta memberi gambaran singkat hubungan didalam sistem tersebut. Penelitian ini menggunakan Use case diagram untuk menjelaskan fungsionalitas sistem, dimana akan dijelaskan melalui use case berikut :

1. Use Case Diagram Sistem



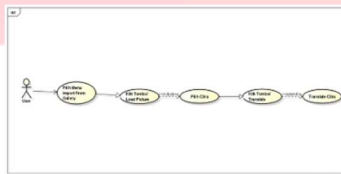
Gambar 3.2 Use Case Diagram

2. Use Case Diagram Menu Capture



Gambar 3.3 Use Case Diagram Menu Capture

3. Use Case Diagram Menu Import from Gallery



Gambar 3.4 Use Case Diagram Menu Import From Gallery

4. Pengujian

4.1 Pengujian Perangkat Lunak

4.1.1 Pengujian Black Box

Pengujian ini bertujuan untuk mengetahui sebuah hasil sistem dan hasil yang diharapkan apabila sistem ini diberikan suatu keadaan masukan tertentu. Tahap pengujian ini menggunakan metoda pengujian *Black Box* yang dimana adalah untuk mengetahui apakah fungsi antarmuka sistem sudah berjalan sesuai yang diharapkan. Pengujian *Black Box* yang dilakukan ini meliputi pengujian fungsi kamera sebagai media input / akuisisi citra, fungsi open gallery sebagai media untuk input / akuisisi citra, fungsi menu dan navigasi, serta fungsi-fungsi lainnya yang berhubungan dengan *user interface*.

1. Skenario Pengujian

Berikut adalah skenario pengujian Black Box pada sitem.

Tabel 4.1 Black Box testing

No.	Pengujian	Penjelasan uji
1	Akuisisi citra menggunakan android smartphone berdasarkan media kamera dan galeri	Activity camera Activity gallery Preprocessing dan ekstraksi ciri Hasil deteksi
3	About	Penjelasan informasi developer dan aplikasi

2. Hasil pengujian dan analisis

Berikut ini adalah penjelasan dan analisis hasil pengujian

Tabel 4.2 Pengujian dan Analisis Black Box

No.		Hasil Sistem	Hasil yang diharapkan	Kesimpulan
1	Klik tombol floating action button "capture"	Menampilkan activity camera untuk akuisisi citra	Dapat Menampilkan activity camera untuk akuisisi citra	✓ Diterima Ditolak
	Hasil akuisisi citra dari kamera	Menampilkan hasil citra	Proses akuisisi dapat berjalan	✓ Diterima Ditolak

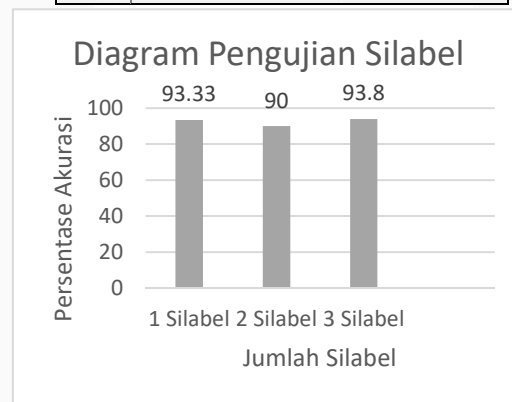
			serta ditampilkan oleh sistem	
2	Klik tombol "Load Picture"	Menampilkan activity gallery untuk akuisisi citra	Dapat menampilkan activity gallery dan dapat mengakuisisi citra	✓ Diterima Ditolak
	Hasil akuisisi citra dari gallery	Menampilkan hasil citra	Proses akuisisi dapat berjalan serta ditampilkan oleh sistem	✓ Diterima Ditolak
3	Melakukan swipe navigation pane	Membuka navigation drawer	Dapat membuka navigation drawer	✓ Diterima Ditolak
4	Klik menu "about"	Menampilkan informasi yang tersedia pada layout about	Dapat menampilkan informasi yang tersedia pada layout about	✓ Diterima Ditolak

4.2 Pengujian Performa Sistem

4.2.1 Pengujian Huruf Dan Silabel

Tabel 4.3 Pengujian Silabel

NO.	SILABEL	AKURASI
1	1 Silabel	93.33 %
2	2 Silabel	90%
3	3 Silabel	93.8%



Gambar 2.4 Diagram Hasil Pengujian Silabel

Berdasarkan hasil pengujian huruf, dapat disimpulkan bahwa tiap-tiap huruf yang dijadikan data uji, mampu dikenali atau ditransliterasikan ke dalam tulisan latin dengan baik. Hal ini dikarenakan data uji tersebut diakuisisi satu persatu dan dalam keadaan yang sedekat mungkin saat penentuan data latih. Sedangkan hasil pengujian silabel, memiliki tingkat akurasi diatas 90% dari masing-masing jumlah huruf yang ditransliterasi benar pada pembagian tiap tiap pengujian silabel. Oleh karena itu pada pengujian tiap silabel ini bobotnya akan berbeda, semakin banyak huruf yang diuji maka akan semakin banyak huruf yang dianggap benar maupun semakin banyak pula resiko huruf yang dianggap salah, tergantung dari seberapa ideal dataset (data training), jarak, sudut, pencahayaan, ukuran resolusi dan lain-lain yang sesuai.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis dari penelitian ini dapat diambil beberapa kesimpulan yaitu :

Capture to Translate adalah sebuah aplikasi yang mampu melakukan translasi dari bahasa Cyrillic Rusia ke huruf latin Indonesia dengan mengimplementasikan *image processing* yang melibatkan proses deteksi tepi *findcontour* untuk ekstraksi ciri dan algoritma klasifikasi SVM. Penelitian ini menghasilkan tingkat akurasi terbaik mencapai 93,8% diraih pada pengujian 3 silabel. Akurasi tersebut didapatkan berdasarkan rata-rata akurasi perhuruf dari pengujian yang telah dilakukan dengan menggunakan resolusi kamera 6MP, jarak 30CM, dan sudut kemiringan objek berkisar diantara 0 derajat serta pencahayaan sekitar 13 lux.

5.2 Saran

Berikut adalah saran untuk pengembangan sistem ini selanjutnya berdasarkan dari kesimpulan dan hasil yang sebelumnya telah diperoleh.

1. Tingkatkan presentase keakuratan deteksi dan penerjemahan dengan algoritma SVM dan ekstraksi ciri *findcontour* untuk studi kasus klasifikasi bahasa Rusia maupun lainnya.
2. Mengembangkan sistem ini agar dapat melakukan klasifikasi selain dari hasil kata print-out, seperti tulisan tangan.
3. Gunakan kernel lain dalam klasifikasi SVM seperti *RBF*, *sigmoid*, *polynomial*, dll.
4. Gunakan metode lainnya untuk proses segmentasi, ekstraksi ciri, dan klasifikasi untuk translasi dan transliterasi tulisan bahasa Rusia.
5. Kembangkan pada platform smartphone lain.
6. Perhatikan cara menanggulangi noise yang terakuisisi sistem.

DAFTAR PUSTAKA

- [1] T. Sutoyo, E. Mulyanto, V. Suhartono, O.D. Nurhayati, Wijanarto, Teori Pengolahan Citra Digital, Penerbit Andi, Yogyakarta, 2009, p.256.
- [2] Munir, Rinaldi, 2011. "Pengolahan Citra Digital". Bandung: Penerbit Informatika.
- [3] Aryuanto, Somawirata Komang, Limpraptono. F. Yudi, "A New Color Segmentation Method Based on Normalized RGB Chromaticity Diagram", ISSN 2085 – 973, Seminar on Intelligent Technology and Its Applications, 2009.
- [4] H. A. Simon. 1987. Diambil dari buku Kusrini yang berjudul Sistem Pakar Teori dan Aplikasi. Andi Yogyakarta : Yogyakarta.
- [5] Elaine Rich. Kevin Knight. Shivashankar B Nair, 2009, Artificial Intelligence Third Edition, McGraw-Hill Publishing Company Limited, New Delhi.
- [6] V. N. Vapnik, "The Nature of Support Vector Machine". Berlin Heidelberg. 1999.
- [7] Yulli Soelistyani, 2013, Model Peramalan Harga Saham dengan Pendekatan Neural Network Algoritma Multilayer Perceptron (MLP) dan Support Vector Regression (SVR).
- [8] Lukman, Andi, 2014, "Machine Learning Multi Klasifikasi Citra Digital", Konferensi Nasional Ilmu Komputer (KONIK) 2014 Volume: 1, Makassar
- [9] Farsiah, Laina (dkk), 2013, "Klasifikasi Gambar Berwarna menggunakan K-Nearest Neighbor dan Support Vector Machine", Universitas Syiah Kuala, Banda Aceh.
- [10] Ratnasari, Kamilah (dkk), 2016, "Segmentasi Citra Menggunakan Support Vector Machine (SVM) Dan Ellipsoid Region Search Strategy (ERSS) Arimoto Entropy Berdasarkan Ciri Warna Dan Tekstur", Universitas Dr. Soetomo, Surabaya.
- [11] Sutoyo, T, dkk. 2009, Teori Pengolahan Citra Digital, Penerbit Andi: Yogyakarta.
- [12] Prasetyo, E., (2011). Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab, Yogyakarta: Penerbit Andi.
- [13] Kurniawan, Wisnu Rizky. 2015 Purwarupa Sistem Klasifikasi dan Penghitung Jumlah Kendaraan Bermotor Menggunakan Kamera Webcam Berbasis Citra Digital. Yogyakarta. Universitas Gajah Mada.
- [14] Cheriet M., Kharna N., Liy C., Suen C.Y., 2006, "Character Recognition System A Guide for Student and Practioners", John Willey & Sons. Inc.
- [15] Suryo Hartanto, Aris Sugiharto (dkk), 2014, "Optical Character Recognition Menggunakan Algoritma Template Matching Correlation", Jurnal Masyarakat Informatika, Volume 5, Nomor 9, Semarang.
- [16] Sofani, Rachmah, 2009, "Sistem OCR", Institut Teknologi Telkom, Bandung.