

Comparative Analysis Of Efficient Image Segmentation Technique For Text Recognition And Human Skin Recognition

Segmentation of
Image Technique
for Recognition

81

Septian Cahyadi, Febri Damatraseta, Lodryck Lodefikus S.

Program Studi Teknologi Informasi, Institut Bisnis dan Informatika Kesatuan

E-Mail: schayadi@ibik.ac.id

Submitted:
JANUARI 2021

Accepted:
JULI 2021

ABSTRACT

Computer Vision and Pattern Recognition is one of the most interesting research subject on computer science, especially in case of reading or recognition of objects in realtime from the camera device. Object detection has wide range of segments, in this study we will try to find where the better methodologies for detecting a text and human skin. This study aims to develop a computer vision technology that will be used to help people with disabilities, especially illiterate (tuna aksara) and deaf (penyandang tuli) to recognize and learn the letters of the alphabet (A-Z). Based on our research, it is found that the best method and technique used for text recognition is Convolutional Neural Network with achievement accuracy reaches 93%, the next best achievement obtained OCR method, which reached 98% on the reading plate number. And also OCR method are 88% with stable image reading and good lighting conditions as well as the standard font type of a book. Meanwhile, best method and technique to detect human skin is by using Skin Color Segmentation: CIELab color space with accuracy of 96.87%. While the algorithm for classification using Convolutional Neural Network (CNN), the accuracy rate of 98%

Keywords: computer vision, segmentation, object recognition, text recognition, skin color detection, motion detection, disability application

ABSTRAK

Computer Vision dan Pattern Recognition adalah salah satu penelitian yang paling menarik subjek ilmu komputer, terutama dalam hal membaca atau pengenalan objek secara realtime dari perangkat kamera. Deteksi objek memiliki jangkauan segmen yang luas, dalam penelitian ini kami akan mencoba menemukan metodologi mana yang lebih baik untuk mendeteksi teks dan kulit manusia. Ini Studi ini bertujuan untuk mengembangkan teknologi visi komputer yang akan digunakan untuk membantu orang dengan disabilitas, terutama yang buta huruf (tuna aksara) dan tuna rungu (penyandang tuli) untuk mengenali dan mempelajari huruf-huruf alfabet (A-Z). Berdasarkan penelitian kami, ditemukan bahwa metode terbaik dan teknik yang digunakan untuk pengenalan teks adalah Convolutional Neural Network dengan prestasi akurasi mencapai 93%, pencapaian terbaik berikutnya diperoleh metode OCR, yaitu mencapai 98% di pelat nomor bacaan. Dan juga metode OCR adalah 88% dengan pembacaan gambar yang stabil dan kondisi pencahayaan yang baik serta jenis font standar sebuah buku. Sementara itu, metode terbaik dan teknik untuk mendeteksi kulit manusia adalah dengan menggunakan Skin Color Segmentation: CIELab color space dengan akurasi 96,87%. Sedangkan algoritma untuk klasifikasinya menggunakan Convolutional Neural Jaringan (CNN), tingkat akurasi 98%

Kata Kunci : computer vision, segmentation, object recognition, text recognition, skin color detection, motion detection, disability application

JIKES

Jurnal Informatika
Kesatuan

Vol. 1 No. 1, 2021

page. 81-90

IBI Kesatuan

E-ISSN 2807-4335

DOI: 10.37641/jikes.v1i1.775

PENDAHULUAN

Object Detection merupakan topik dari Computer Vision dan Pattern Recognition yang sangat populer karena masih terus dikembangkan, terutama pembacaan secara Real-time dari perangkat kamera. Object detection sendiri merupakan langkah awal untuk pengembangan penelitian lanjutan seperti pengenalan teks dan kulit untuk tujuan yang beragam. Sebelum masuk pada tahapan lanjutan tersebut tentunya dibutuhkan tahapan segmentasi terlebih dahulu untuk menyeleksi bagian yang akan diekstrak, yakni dengan mendeteksi objek yang tepat yang akan diproses. Pada hasil laporan konferensi ICDAR (International Conference of Document Analysis and Recognition) tahun 2015, untuk tantangan dengan tema Text on Video, hasil tertinggi dari pengenalan teks pada video belum mencapai 50% atau lebih tepatnya sebesar 45.18% untuk Text Localization dan 41.84% untuk End-to-end Recognition. Dalam perkembangannya sudah ada penerapan metode terbaru untuk menjawab tantangan Text Recognition dalam video hingga mencapai angka presentasi sebesar 60%, meski telah ada peningkatan yang cukup signifikan namun dilihat dari nilai presentase tersebut masih memungkinkan penerapan metode lain yang lebih baik lagi. Berbeda dengan hasil tantangan untuk membaca teks pada still-image baik focused text maupun incidental text, presentase yang diperoleh sudah mencapai angka lebih dari 90%. Hal tersebut dikarenakan scene text sangat sulit dibaca dan dikenali, karena teks yang muncul pada video dipengaruhi oleh pencahayaan yang tidak rata, background gambar yang bertekstur atau kemiringan posisi saat video diambil. Karena hal tersebut, pembacaan teks dalam video merupakan tantangan bagi para peneliti.

Sementara pada deteksi objek untuk mengenali pola pada kulit manusia memiliki tingkat kesulitan yang berbeda, kulit manusia memiliki warna yang khusus dan dapat dijabarkan kedalam sebuah bentuk ruang warna kulit manusia. Secara umum warna kulit manusia dapat dilihat pada area wajah, leher, tangan dan kaki. Beberapa penelitian telah berhasil menganalisa atau membuat ruang warna kulit berbasis pixel baik pada citra maupun video diantaranya adalah ruang warna RGB, HIS/HSV, TSL, YcbCr, dan lain sebagainya. Penelitian yang menggunakan ruang warna kulit diantaranya yaitu, mendeteksi gesture pada wajah pada sebuah video dengan kondisi cahaya yang berbeda dengan latar belakang yang berbeda, yang nantinya akan disimpan kedalam memori[5]. Ruang warna kulit juga dapat digunakan untuk mendeteksi gesture dari tangan dengan cara mengekstrak fitur pada tangan untuk merduce noise pada latar belakang yang kompleks dengan kondisi cahaya yang berbeda-beda, cara yang digunakan dengan memilih vektor dari pixel dalam blok disekitar warna kulit untuk mendapatkan pixel kulit yang baru.

Teknologi Computer Vision juga dapat dimanfaatkan oleh para penyandang disabilitas misalnya dapat membantu para tuna aksara dan penyandang tuli untuk mengenal dan belajar mengenai huruf alfabet. Penerapan teknologi ini dapat sangat membantu para penyandang disabilitas untuk dapat menjalani kehidupan didunia ini agar lebih mudah. Pada penelitian ini kami mencoba melakukan percobaan untuk mencari metode mana yang lebih baik untuk menerapkan teknologi ini. Penggunaan segmentasi pengenalan terhadap sebuah text akan digunakan untuk membantu penelitian terhadap para penyandang tuna aksara, sedangkan untuk para penyandang penyandang tuli akan menggunakan segmentasi pengenalan terhadap warna kulit yang nantinya akan digunakan untuk mendeteksi pergerakan tangan untuk membaca simbol dari huruf alfabet Bahasa Isyarat Indonesia (BISINDO).

Berdasarkan latar belakang dan penjelasan yang menjadi dasar pengembangan, maka penelitian ini bertujuan untuk:

1. Menerapkan algoritma yang efisien untuk sistem tertanam dalam melakukan deteksi pada teks.
2. Menerapkan algoritma yang efisien dalam mengklasifikasi karakter pada sebuah citra pada sistem tertanam.
3. Mengembangkan model pengenalan pola menjadi menjadi huruf agar berjalan optimal pada sebuah sistem tertanam.

4. Melakukan percobaan terhadap sistem image processing untuk mendapatkan sebuah edge dari Skin Color Segmentation CIELab dengan menerapkan teknik operasi Sobel.

METODE PENELITIAN

Penelitian ini merupakan penelitian deskriptif kualitatif. Pada penelitian ini mencoba melakukan perbandingan terhadap segmentasi yang digunakan untuk mengenali sebuah obyek yang diambil dari image ataupun video. Obyek yang akan dikenali berupa pengenalan terhadap bentuk text dan pengenalan terhadap manusia. Berbagai metode akan disajikan dalam penelitian ini mulai dari segmentasi bagaimana mendeteksi sebuah text dan mendeteksi kontur manusia.

Untuk mendeteksi objek pada penilitan yang ditujukan untuk mengenali text ada beberapa metode yang paling populer dilihat dari jurnal yang menjadi referensi kami. Metode yang terpopuler pertama adalah Neural Network, dengan beragam variasinya, seperti halnya Artificial Neural Network, Convolutional Neural Network, maupun Recurrent Neural Network. Metode ini umumnya melakukan pendeteksian awal dengan membaca frame, salah satunya menggunakan MESR based detector untuk mengambil frame yang tepat untuk digunakan dan paling mudah untuk dibaca. Selanjutnya barulah digunakan metode Neural Network untuk mengelompokkan dan mengklasifikasi teks untuk kemudian dilakukan segmentasi. Setelah segmentasi dilakukan barulah dapat digunakan untuk melakukan pengenalan kata.

Untuk mendeteksi kontur pada manusia dapat dilakukan dengan cara melakukan pendeteksian terhadap pixel warna kulit di area kontur manusia yang tersimpan pada gambar atau video. Pendeteksian terhadap pixel warna kulit diambil setelah melakukan konversi citra RGB kedalam model ruang warna. Selain menggunakan segmentasi terhadap warna kulit dapat menggunakan teknik-teknik tertentu seperti menghitung kedalaman sebuah citra dan lain sebagainya. Penggunaan segmentasi warna kulit merupakan sebuah fitur berbasis pixel sederhana namun kemampuannya sangat kuat.

HASIL DAN PEMBAHASAN

Pertama-tama penulis mencari penelitian apa saja yang telah dilakukan berkaitan dengan pengolahan citra untuk text recognition. Pencarian ini mengarahkan penulis pada sebuah judul bertajuk "*Estimating hidden parameters for text localization and recognition*" yang ditulis oleh Neumann L dan Matas J dari universitas Prague, Republik Ceko pada tahun 2011. Sebuah metode baru untuk pembacaan *text line formation* diajukan guna melokalisasi dan mengenali teks dengan lebih baik (Neumann et al., 2011). Neumann meningkatkan kemampuan segmentasi dengan menambahkan *hidden text line parameters*, yakni menambahkan *height ratio*, *centroid angle* dan *text direction* dan teknik ini disebut MSERs (Maximally Stable Extremal Regions) hasilnya didapatkan nilai akurasi sebesar 60% dengan menggunakan dataset ICDAR 2003

Satu tahun setelah penelitan sebelumnya, (Neumann L et al., 2011) kembali melakukan penelitian yang serupa, dengan judul "*Real-Time Scene Text Localization and Recognition*". Pada penelitian kali ini Neumann dan Matas menggunakan pendekatan yang berbeda. Untuk mengurangi distorsi pada gambar yang diambil secara real-time digunakan metode Extremal Region (ERs). Detektor ER melakukan perbaikan pada citra yang blur, iluminasi, warna dan variasi texture dan mampu menangani cahaya dengan kontras rendah. Metode ini diujikan pada dataset ICDAR 2011 dan juga dataset Street View Text dan mendapatkan hasil yang luar biasa pada tahapan *recall* namun pada tahap pengenalan text, pendekatan ini mendapatkan akurasi sebesar 68.7%. Di tahun 2016 (Neumann et al., 2016) juga mengajukan kembali teknik ini dengan perbaikan dengan menambahkan syntethic font yang telah dilatih untuk proses klasifikasi, hasilnya sungguh sebuah penigkatan yang signifikan. Pengujian dilakukan menggunakan dataset ICDAR 2013 meraih nilai akurasi sebesar 76.3% untuk *text localization* dan untuk *case-sensitive end-to-end text recognition* meraih hasil 45.2%. Pada dataset ICDAR 2015 hasil yang diraih

untuk *end-to-end text recognition* sebesar 41.8% dan juga pada SVT diraih hasil 68.1% untuk *end-to-end recognition*.

Selang beberapa waktu kemudian, ditahun yang sama (Neuman et al., 2011) memperbaharui metode yang mereka gunakan pada *text recognition*. Melalui sebuah jurnal dengan judul "*A method for text localization and recognition in real-world images*" Penelitian dilakukan dengan mengembangkan metode MSERs (Maximally Stable Extremal Regions) untuk mendapatkan text line hypothesis formation dan menggunakan SVM untuk klasifikasi. Teknik yang dilakukan diawali dengan mendeteksi karakter pada proyeksi red channel, kemudian melakukan induksi grafik (karakter diwarnai hijau dan yang bukan karakter diwarnai merah), tepi yang panjangnya lebih dari 300px dihilangkan pada gambar agar mudah dibaca). Setelah itu tahapan dilanjutkan dengan melakukan hipotesis terhadap konten (teks). Pada tahapan hipotesis digunakan algoritma SVM dengan Radial Basis Function (RBF) untuk mengklasifikasi hurur. Tiap baris dari bagian-bagian huruf dibaca terpisah kemudian menentukan arah garis tersebut, baru kemudian melakukan klasifikasi. Penelitian ini mendapatkan hasil akurasi yang belum terlalu baik, yakni sebesar 67%. Metode ini juga nampaknya mengalami kendala dari sisi kecepatan proses, dimana tiap karakter membutuhkan waktu pengenalan sebanyak 89 detik. Dataset yang digunakan pada penelitian ini adalah ICDAR 2003. Pada tahun 2015, Neumann dan Matas telah mengembangkan lagi metode ini dengan menambahkan *character refinement* sehingga meraih hasil yang lebih tinggi dari sisi akurasi, yakni sebesar 77.1%.

Satu tahun setelah penelitian sebelumnya, (Neumann L et al., 2011) kembali melakukan penelitian yang serupa, dengan judul "*Real-Time Scene Text Localization and Recognition*". Pada penelitian kali ini Neumann dan Matas menggunakan pendekatan yang berbeda. Untuk mengurangi distorsi pada gambar yang diambil secara real-time digunakan metode Extremal Region (ERs). Detektor ER melakukan perbaikan pada citra yang blur, iluminasi, warna dan variasi texture dan mampu menangani cahaya dengan kontras rendah. Metode ini diujikan pada dataset ICDAR 2011 dan juga dataset Street View Text dan mendapatkan hasil yang luar biasa pada tahapan *recall* namun pada tahap pengenalan text, pendekatan ini mendapatkan akurasi sebesar 68.7%. Di tahun 2016 (Neumann et al., 2016) juga mengajukan kembali teknik ini dengan perbaikan dengan menambahkan syntethic font yang telah dilatih untuk proses klasifikasi, hasilnya sungguh sebuah peningkatan yang signifikan. Pengujian dilakukan menggunakan dataset ICDAR 2013 meraih nilai akurasi sebesar 76.3% untuk *text localization* dan untuk *case-sensitive end-to-end text recognition* meraih hasil 45.2%. Pada dataset ICDAR 2015 hasil yang diraih untuk *end-to-end text recognition* sebesar 41.8% dan juga pada SVT diraih hasil 68.1% untuk *end-to-end recognition*.

Dibayan Chakraborty et al (2013) membahas mengenai teknik pengambilan karakter atau proses scanning sebuah buku dengan lebih cepat menggunakan kamera digital video. Cara ini diasumsikan dapat lebih cepat prosesnya dibanding melakukan proses scanning sebuah buku dengan menggunakan mesin scanner. Terdapat dua metodologi yang digunakan dalam paper ini, terbagi dalam dua tahapan. Tahapan yang pertama adalah menentukan region area yang akan di scan, membaca proses membalikkan buku dan menentukan posisi yang pas untuk dilakukan proses pembacaan karakter. Pada tahapan ini digunakan metode Region of Interest (ROI). Dalam lingkungan yang tidak dapat dikontrol, dimana *background* dari buku mungkin banyak terdapat noise dan juga bertekstur, ROI harus mampu menangkapnya. ROI menggunakan Algoritma Hough Transformation pada sudut gambar untuk mencari garis lurus. Pada Hough Transformation, gambar ditransformasi menjadi sebuah parameter jarak dari koordinat polar dan garis dapat dideteksi menggunakan skema voting. Selanjutnya garis bingkai grayscale dari gambar dideteksi dengan ROI. Setelah region ditentukan, pengenalan karakter dilakukan dengan metode SVM. Pada tahapan ini dilakukan proses ekstraksi dan klasifikasi. Setelah garis bingkai lokalisasi dilakukan, proses dilanjutkan dengan menentukan titik tengah dari buku dan membaca kemiringan dari sudut buku dengan

geometric 2 dimensi. Dengan demikian isi buku dapat di diekstrak untuk kemudian diklasifikasi dengan menggunakan SVM.

Sumber referensi berikutnya, dengan judul jurnal "*Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning*" yang ditulis oleh (Coates et al, 2013) menawarkan metode *feature learning* dalam pengenalan teks secara end-to-end yang diimplementasikan pada dataset ICDAR 2003. Proses awal dari metode ini hampir sama dengan convolution neural network, tetapi metode training ini dapat digunakan untuk melakukan konstruksi dengan cepat pada fitur set berskala besar dengan tuning yang minimal. Tahapan pada metode ini adalah (1) menjalankan algoritma *feature learning* pada sebuah set filter citra yang diambil sebelum training data untuk mempelajari sekumpulan fitur citra, (2) mengevaluasi fitur convolutional dari training image. Mengurangi jumlah fitur yang menggunakan *spatial pooling*, (3) melatih klasifikasi linear baik untuk deteksi teks ataupun pengenalan karakter. Dengan menggunakan dataset ICDAR 2003 metode ini meraih hasil rata-rata akurasi sebesar 82.66%.

Pada percobaan (Neumann L, 2013) di konferensi ICDAR, sebuah metode baru diajaukan, yakni *Oriented Stroke Detection* digunakan untuk melakukan segmentasi pada teks. Metode ini menggabungkan keutungan dari *sliding-window* dan metode *connected component*. Karakter terdeteksi dan dikenali sebagai area gambar yang berisi stroke orientasi spesifik pada posisi tertentu, di mana stroke dideteksi secara efisien dengan menggabungkan bidang gradien gambar dengan satu set filter bar yang berorientasi. Selain itu, representasi karakter baru yang dihitung secara efisien dari nilai yang diperoleh pada tahap deteksi stroke diperkenalkan. Keefektifan representasi ditunjukkan oleh hasil yang dicapai dalam klasifikasi karakter pada dunia nyata menggunakan *euclidian nearest-neighbor* yang dilatih pada data sintetis dalam bentuk polos. Hasil akhir dari penerapan pendekatan ini menunjukkan hasil yang signifikan dibandingkan dengan penelitian sebelumnya, diujikan pada dataset ICDAR 2011, nilai akurasi yang didapatkan untuk *text localization* sebesar 72%, namun untuk *end-to-end text recognition* hanya mendapatkan presentase sebesar 45.2%.

Rujukan berikutnya didapat dari (Yildirim G et al., 2013) dengan jurnalnya yang berjudul "*Text Recognition in Natural Images using Multiclass Hough Forests*". Teknik yang digunakan untuk mendeteksi dan mengenali teks secara terpadu dengan mencari kata secara langsung tanpa membagi gambar menjadi *text region* atau karakter individu. Pada jurnal ini terdapat tiga kontribusi. Pertama, memodifikasi kerangka deteksi objek yang disebut *Hough Forest* (Gall et al., 2011; Razavi et al., 2011) dengan memperkenalkan "Cross-scale Binary Features" yang membandingkan informasi antara image patch pada skala yang berbeda-beda. Kedua, fungsi *word-formation* dan pemetaan lingkungan digunakan untuk mendeteksi dan mengenali teks pada gambar natural. Ketiga, Yildirim dkk mampu menghadirkan genetika data yang realistik untuk karakter teks. Teknik ini diujicobakan pada dataset ICDAR 2003 dan SVHN dan meraih hasil akurasi yang luar biasa, yakni sebesar 85.7%

Pada penelitian yang dilakukan (Baoguang Shi et al., 2016) dengan pendekatan baru, yakni RARE (Robust Scene Text Recognition with Automatic Refrication) dengan judul jurnal yang sama mendapatkan hasil yang sangat baik. Diterapkan pada dua dataset, SVT dan CUTE80, masing-masing mendapatkan nilai akurasi 77% dan 59%. Yang paling menarik dari metode Baoguang adalah kemampuan dalam merefrikasi teks yang tidak proporsional pada sebuah citra. Baoguang menggunakan *Spatial transformer network*, *Grid generator*, dan *sampler* untuk memetakan proporsi dari citra dan menormalisasikan bentuk teks. Baoguang berhasil memetakan kembali kata dengan beragam posisi, baik yang miring ataupun teks yang melingkar.

Hasil penelitian terkait *text recognition* terbaik sejauh ini didapatkan dengan menggunakan metode *Convolutional Neural Network* (CNN). Sebuah penelitian karya (Jaderberg et al., 2016) dengan mengkombinasikan beberapa pendekatan baru untuk menghasilkan nilai *recall* yang tinggi, dan tahap filtering yang cepat untuk meningkatkan presisi. Untuk tahap pengenalan teks, Jaderberg melatih *convolutional neural network*

dengan jaringan yang sangat besar pada seluruh region proposal pada waktu yang bersamaan, sehingga menghasilkan ranking yang tinggi. Untuk teknik *text detection* Jaderberg menggunakan *sliding window* dengan *random forest classifier* yang bekerja pada *HOG features*. Pada tahapan pengenalan teks, tiap kata yang telah di segmentasi akan diinputkan kedalam *deep CNN*. Namun, Jaderberg juga menggunakan model klasifikasi kamus. Pengenalan dicapai dengan melakukan klasifikasi multi-arah di seluruh kamus untuk mendapatkan kata-kata yang potensial. Hasilnya, dari banyak sekali dataset yang diujikan, mulai dari Synthetic Font, ICDAR03, SVT, ICDAR13, dan IIT5k, diperoleh hasil rata-rata 93% untuk akurasi *cropped word recognition*. Namun untuk *end-to-end text spotting* meraih hasil rata-rata 75%.

Paojja et al (2016) mencoba melakukan perbandingan dari berbagai metode yang digunakan untuk mengenali text pada video, atau OCR dari Video. Dimana dataset yang digunakan adalah text scrolling atau news ticker pada siaran berita. Terdapat tiga langkah dalam pengenalan teks pada tiap metode, yakni deteksi teks, ekstraksi teks dan pengenalan teks. Dari hasil **survey Text Extracting and Recognition from videos** diperoleh kesimpulan bahwa melakukan ekstraksi frame dari video saja sudah merupakan tantangan tersendiri. Teks harus diekstrak berdasarkan segmentasi kata dan karakter di ekstrak dari *Punjabi Text Line*. Tujuan utama dari penelitian ini adalah untuk mengembangkan metodologi untuk mengekstrak text dari video dengan peningkatan recognition rate, dengan resolusi video yang rendah, pengenalan karakter menjadi bertambah sulit. Sehingga diharapkan dapat menjadi daya tarik bagi modifikasi berikutnya dalam di waktu selanjutnya.

Haojin Yang et al (2016) mencoba mengangkat teknik OCR video secara *real-time* menggunakan metode *Maximally Stable Extremal Regions* (MSERs), dan untuk menghilangkan kesalahan dalam pembacaan region diverifikasi dengan metode *Convolutional Neural Network* (CNN). Untuk pengenalan kata penulis membangun sebuah metode kerangka dasar untuk mensegmentasi region teks dari background, kemudian menggunakan metode Pengenalan kata CNN untuk membaca pengenalan teks. Hasil penelitian diujikan menggunakan dataset dari ICDAR 2015 Robust Reading Competition Challenge untuk Task 3 "*Focused Scene Word Recognition*". Kategori evaluasi adalah "*Word Recognition Rate (Uppercase)*" (WRR-U) karena hasil output metode merupakan case-insensitif.

Prajit Ramachandran (2016) membahas implementasi algoritma Faster R-CNN untuk mendeteksi objek pada Video. Hasil implementasi ini dipresentasikan pada ICCV 2015 ImageNet and MS COCO Visual Recognition Challenges Joint Workshop. Penelitian ini menggunakan video clip dari the ImageNet 2015 Object Detection from Video Challenge. Dalam jurnal ini disampaikan bahwa penulis menggunakan VGGNet sebagai Convolutional backend dari Faster-RCNN. VGGNet memberikan perbaikan yang substansial melebihi arsitektur yang dijelaskan Zeiler dan Fergus. Dari hasil validasi set, penulis mendapatkan nilai akurasi mAP 53%.

Baoguang Shi et al (2016) membahas tentang penerapan metode Neural Network untuk pengenalan scene text, dimana proporsi maupun orientasi teksnya beragam dan tidak ideal. Dibandingkan dengan metode yang pernah diajukan sebelumnya, metode ini memiliki sifat khas sebagai berikut : (1) menggunakan data latih secara end-to-end, (2) secara natural menangani urutan yang tidak beraturan, tanpa segmentasi karakter atau normalisasi skala horizontal. (3) Tidak terbatas pada leksikon yang telah ditetapkan dan mencapai kinerja yang luar biasa di kedua leksikon dan pengenalan scene teks berdasarkan leksikon. (4) Menghasilkan model yang efektif untuk dapat dipraktikan di dunia nyata. Dari hasil pengujian novel neural network architecture, yang disebut Convolutional Recurrent Neural Network (CRNN), yang mengintegrasikan keuntungan metode Deep Convolutional Neural Networks (DCNN) dan Recurrent Neural Networks (RNN). CRNN mampu mengambil input gambar dari berbagai dimensi dan menghasilkan prediksi dengan panjang yang berbeda. Pada pengujian eksperimen pengenalan scene text dengan CRNN diraih hasil yang superior atau performa yang

sangat tinggi, dibandingkan dengan metode konvensional seperti DCNN dan RNN. Hal ini membuktikan keuntungan dari algoritma yang ditawarkan pada makalah ini.

U. Bhattacharya et al (2016) membahas mengenai teknik pendeteksi Scene Text tanpa batasan dengan menggunakan sebuah metode yang mudah di implementasikan. Berdasarkan deteksi stroke pada pixel secara efisien dan membandingkan dengan pixel disekitarnya. Stroke-spesifik keypoints mendeteksi secara efisien dan *text fragments* secara berurutan di ekstrak menggunakan *local thresholding* mengikuti keypoint properties. Metode klasifikasi dengan efektif mengkalkulasi fitur untuk mengeliminasi area non-text. Stroke-specific keypoints menghasilkan segmentasi region dua kali lebih kecil dan tetap bisa mendeteksi 25% karakter lebih dari yang dihasilkan MSER secara umum dan 4 kali lebih cepat dalam proses. Setelah tahapan klasifikasi, jumlah region tereduksi sebanyak 7 kali dari metode standard dan lebih cepat hampir 3 kali lipat.

Dengan metode ini mendukung variasi skrip secara luas (Latin, Hebrew, Chinese, etc). Hasil pengujian stroke detector dengan FASText pada dataset ICDAR 2013 menunjukkan hasil akurasi sebesar 82% dengan peningkatan kecepatan proses 4 kali lebih tinggi ketimbang MSER.

Annmaria Cherian (2016) membahas mengenai penerapan metode Hough Transform untuk memperbaiki orientasi dari scene image dan menggunakan deteksi karakter yang efisien dan teknik lokalisasi. Metode yang digunakan untuk mem-filter komponen non-text adalah SVM. Setelah dilakukan proses filtering, pengenalan karakter digunakan untuk mengenali text secara akurat. Dataset yang digunakan pada penelitian ini adalah Scene Text-Perpective dataset yang berikan gambar oaoan iklan atau billboard yang terletak di pinggir jalan raya dengan perpekstif gambar yang miring. Dan hasil dari pengujian menunjukkan bahwa metode yang digunakan cukup simple dan melebihi metode yang ada.

Lukas Neumann et al (2016) mencoba melakukan pendekatan Extremal Regions digunakan pada jurnal ini melakukan pembacaan Scene Text localization and recognition atau disebut dengan istilah Photo OCR. Pada tahap pertama, probabilitas dari tiap ER menjadi karakter diestimasi menggunakan kalkulasi fitur Dengan pendekatan algoritma dalam waktu yang konstan. Selanjutnya ER dengan probabilitas local dipilih untuk tahap kedua, dimana akurasi klasifikasi telah ditingkatkan menggunakan komputasi dengan fitur yang baik. Probabilitas karakter yang paling tinggi kemudian dipilih pada tahap akhir ketika konteks dari tiap karakter telah diketahui. Metode ini diujikan pada tiga dataset publik. Pertama menggunakan ICDAR 2013 dataset dan meraih hasil yang luar biasa pada text localization. Kedua menggunakan SVT dataset dan hasilnya secara signifikan melebihi metode lain. Dataset yang ketiga menggunakan ICDAR 2015 Robust Reading Competition, dan mencapai nilai akurasi yang luar biasa kompetitif.

Raja Vikramdeep Singh et al (2016) membahas tentang implementasi metodologi OCR untuk membantu mengkoleksi data kendaraan yang masuk melalui jalan toll dan monitoring lalu lintas di India. Tahapan dalam proses meliputi empat fase, fase pertama adalah image acquisition, kedua Extraction, ketiga pixilation and segmentation, dan keempat adalah Recognition. Proses Automobile Plate Recognition and Extraction System pada jurnal ini telah sukses diimplementasikan dengan adanya optimasi dari kecepatan serta efisiensi sistem dengan mengurangi algoritma image recognition untuk alphabet dan pembacaan nomor menjadi satu algoritma saja, yang mudah dan efisien mengurangi kompleksitas dari pembagian dan penggabungan noise dari gambar output akhir dan menyediakan output yang bebas noise dan pencocokkan background dengan warna font dari nomor plat kendaraan. Hasilnya diperoleh angka akurasi sebesar 98%.

Gokhan Yildirim et al (2016) menjabarkan teknik pengenalan text pada gambar natural menggunakan metode multiclass hough forest. Jurnal ini menawarkan sebuah teknik untuk mendeteksi dan mengenali text dalam beragam kondisi dengan mencari kata secara langsung tanpa mengurangi gambar menjadi region atau karakter individu. Kami mempersembahkan tiga kontribusi. Pertama, penulis memodifikasi object detection framework yang disebut Hough Forest (Gall et al., 2011) dengan memperkenalkan "Cross-Scale Binary Features" yang membandingkan informasi antar patch gambar yang

sama dengan skala berbeda. Kami menggunakan teknik yang sudah dimodifikasi untuk menghasilkan pemetaan bagi setiap karakter teks. Kedua, dalam fungsi formasi kata dan pemetaan lingkungan yang terkomputasi yang kami buat, digunakan untuk mendeteksi dan mengenali teks dalam gambar natural. Hasil pengujian metode ini menggunakan dataset Street View House Number (Netzer et al., 2011) dan ICDAR 2003 (Lucas et al., 2003). Untuk SVHN dataset algoritma kami telah melampaui metode sebelumnya dan telah dibandingkan perfrmormanya menggunakan beberapa sampel data training. Dan dengan menggunakan dataset ICDAR 2003 kami mendapatkan hasil yang luar biasa dalam pengenalan kata dalam hal performa, yakni mencapai 85.7%.

Pada penelitian milik (Hartanto dan Kartikasari, 2016) melakukan penelitian HGR dengan menggunakan SIBI sebagai objek penelitiannya. Pendekatan yang digunakan yaitu pendekatan non-contact dengan memanfaatkan kamera webcam yang telah tertanam, metode yang digunakan yaitu mencoba menangkap atau mendeteksi pergerakan tangan dengan menggunakan Skin Color Segmentation model YCbCr color space dan klasifikasi Back Propagation Neural Network untuk mencari pola pada pergerakan tangan yang terdeteksi pada webcam secara real-time. Pada percobaan yang mereka lakukan langkah yang digunakan untuk melakukan HGR ialah, pertama menangkap HGR pada frame. Penangkapan citra pada frame menggunakan kamera yang telah tertanam pada komputer, selanjutnya user harus melakukan pendeteksian warna kulit. Tahap kedua ialah hasil citra RGB yang telah di tangkap, dilakukan perubahan dengan menggunakan YCbCr color space untuk mereduksi variasi cahaya. Selanjutnya menggunakan CbCr untuk mendeteksi pixel warna kulit. Tahap ketiga, melakukan ekstraksi terhadap citra YCbCr hal ini dilakukan untuk mengetahui radian agar dapat mendeteksi jari. Tahap terakhir yaitu proses klasifikasi dengan Back Propagation Neural Network untuk mencari pola dari input dan mencocokkannya dengan dataset alfabet pola SIBI dan menampilkan hasil klasifikasi kedalam sebuah text. Dari percobaan yang telah mereka uji didapatkan akurasi sebesar 62.6% terhadap 26 huruf alfabet SIBI.

Rujukan selanjutnya yaitu penelitian milik (Garcia dan Viesca, 2016), mereka melakukan percobaan terhadap HGR untuk menerjemahkan American Sign Language (ASL) secara real-time. Penelitian ini sama dengan penelitian sebelumnya dengan memanfaatkan Skin Color Segmentation untuk mengetahui keberadaan HGR. Model HSV color space dan klasifikasi Convolutional Neural Networks (CNN) digunakan dalam pengujian ASL. Namun mereka menggunakan dataset ILSVRC2012 pola ASL dari huruf a sampai k (tanpa j) yang telah dikeluarkan oleh Surrey University dan Massey University. Dari hasil pengujian yang telah mereka lakukan, CNN digunakan untuk mengetahui presentase keakuratan dari objek yang diteliti. Didapatkan presentase sebesar 98% untuk huruf ASL dari a sampai dengan e. Sedangkan dari huruf ASL a sampai dengan k (tanpa huruf j) didapatkan sebesar 70%.

Penelitian selanjutnya milik (Swamy et al, 2014) memanfaatkan HGR untuk Indian Sign Language (ISL) dengan menggunakan Skin Color Segmentation model HSV color space dan Local Binary Pattern (LBP). Pada penelitian ini mereka hanya melakukan percobaan terhadap alfabet ISL huruf m. Pengujian huruf m dilakukan dengan berbagai kondisi cahaya yang berbeda dan latar belakang yang berbeda-beda. Dari hasil pengujian didapatkan nilai akurasi 52.6% sampai dengan 62%.

Penelitian milik (Hossian et al, 2014) mengangkat penggunaan bahasa isyarat ASL sebagai media pengontrol sebuah mobil robot yang dinamakan Moto-Robo. Untuk mendeteksi sebuah skin segmentation pada citra yang tertangkap pada kamera menggunakan YIQ color space, sedangkan untuk klasifikasi ASL menggunakan Backpropagation Neural Network. Hasil yang dicapai dalam mengenali 5 huruf ASL sebagai pengontrol gerak Moto-Robo yaitu 80%-92.0%.

Menerjemahkan bahasa isyarat kembali diteliti oleh (Shinde dan Kagalkar, 2015). Bahasa isyarat Marathi menjadi objek penelitian mereka. Penggunaan Skin Color Segmentation dengan model HSV color space digunakan untuk mendeteksi kontur pixel kulit manusia pada citra. Sedangkan klasifikasi dalam hal mendeteksi arti dari citra masukan menggunakan Euclidean Distance. Nilai akurasi dalam mendeteksi 6 huruf

Marathi Sign Language bisa mencapai 30%-100%. Namun waktu proses klasifikasi yang digunakan cukup memakan waktu yang cukup lama.

Thailand Sign Language(TSL) menjadi objek penelitian (Pariwat dan Seresangtakul,2017) untuk menerjemahkan bahasa isyarat kedalam text. Teknik yang digunakan yaitu menggunakan SVM teknik untuk mendapatkan nilai akurasi dalam mengklasifikasikan TSL menerapkan metode Linear Kernel, Polynomial Kernel, dan RBF Kernel. Nilai akurasi dalam mendeteksi 15 huruf TSL dicapai sebesar 91.20%. Automasi penerjemah Arabic Sign Language(ArSL) diteliti oleh (Saleh et al,2016) dengan menggunakan sensor SOFTKINECT untuk mendeteksi pergerakan tangan manusia, sedangkan untuk klasifikasi dari hasil yang diambil oleh sensor menggunakan PCANet. Nilai akurasi terhadap pengenalan 28 ArSL mencapai 96%-99.5%.

Penelitian milik (Shweta et al,2016) mencoba menerjemahkan bahasa isyarat ASL kedalam sebuah suara atau *text to speech*. Teknik yang digunakan dalam mendeteksi HGR hanya merubah citra kedalam bentuk binary dengan menggunakan Geometric operations, Neighborhood dan block operations. Untuk klasifikasi menggunakan Mel Frequency Cepstrum Coefficients yang disediakan oleh Matrix Laboratory. Hasil yang dicapai dalam mendeteksi 9 huruf ASL didapatkan nilai akurasi sebesar 91%. Alfabet Bangla Sign Language diteliti oleh (Azher dan Chwdhury,2016) untuk menerjemahkannya kedalam text dengan menggunakan teknik Skin Color Segmentation model HSV color space untuk mendeteksi keberadaan kontur kulit manusia dan klasifikasi untuk pengenalan menggunakan SVM. Dari 6 huruf Bangla Sign Language didapatkan nilai akurasi bisa mencapai 97.7%. (George et al,2016) melakukan penelitian tentang bagaimana mendeteksi sebuah penyakit autonium pada kulit manusia dengan menggunakan obyek penelitian dari citra *psoriasis*. Untuk mengetahui *Skin Map* (SM) pada citra menggunakan dua pendekatan yaitu *Otsu's Thresholding* dan *Decision Tree* (DT) dengan menggunakan operasi *Morphological*. Dan untuk mengetahui kulit pada citra menggunakan *Skin Color Segmentation* dengan model YCbCr, HSV, RGB, XYZ and CIELab color space. Klasifikasi *Histogram-based Bayesian* digunakan untuk melakukan percobaan terhadap 100 citra *psoriasis*. Dimana klasifikasi ini sangat efektif untuk menguji metode yang disajikan karena *Skin Probability Maps* (SPMs) yang dihasilkan dari ruang warna yang disajikan dapat menghasilkan SM terbaik untuk gambar *psoriasis*. Dari hasil pengujian didapatkan untuk ruang warna YCbCr dengan Otsu's thresh presentase tertinggi pada TNR yaitu sebesar 98.81% sedangkan untuk DT nilai presentase terbesar yaitu di TNR sebesar 97.61%. Ruang warna HSV dari Otsu's thresh presentase terbesar yaitu TNR 98.81% dan DT sebesar 97.91%. Dan ruang warna CIELab nilai presentase terbesar berada di TPR sebesar 92.49% dengan Otsu's Thresh dan DT sebesar 94.60% di TPR.

Penelitian milik (Ting dan Lin, 2016) mencoba melakukan pengujian teknik CIELab color space terhadap 20 gambar dengan nilai kontras yang minim. Mereka ingin membuktikan bahwa dengan menggunakan teknik CIELab dan Fuzzy Automatic Contrast Enhancement (FACE) dapat meningkatkan nilai kontras yang lebih baik sehingga mendapatkan gambar yang lebih jelas dan baik dilihat. Dari hasil percobaan yang telah mereka lakukan didapatkan nilai peningkatan kontras dari 20 gambar didapatkan nilai RMSC sebesar 12.6 sampai 243.7. Penelitian yang dilakukan (Pandey et al, 2014) melakukan penelitian untuk mengetahui kualitas dari sebuah mangga. Tujuan mereka yaitu untuk mendapatkan standard mutu mangga yang bagus berdasarkan warna mangga dan ukuran mangga. Teknik yang digunakan yaitu CIELab color space dan Fuzzy classifier digunakan untuk mengklasifikasikan mangga. Dari hasil penelitian yang mereka lakukan didapatkan akurasi dari 3 buah mangga sebesar 85.71% - 96.87%. (Mejda et al, 2-15) melakukan penelitian terhadap pendeteksi wajah pada sebuah citra. Teknik yang digunakan yaitu menggunakan Skin color segmentation model YCbCr color space dengan menggunakan Gabor Filter untuk mengetahui kontur wajah pada citra. Sedangkan untuk klasifikasi yang digunakan ialah Neural Network. Mereka bertujuan untuk mendapatkan berapa lama waktu yang dibutuhkan dalam mendeteksi wajah pada

citra dengan metode yang mereka tawarkan. Dari hasil penelitian didapatkan bahwa lama waktu eksekusi sebesar 12.46500 – 77.77300 perdetik.

PENUTUP

Dari hasil perbandingan diatas dapat disimpulkan bahwa dalam mendeteksi sebuah obyek hal yang paling diutamakan yaitu kecepatan dalam pengolahan sebuah proses sehingga menghasilkan informasi yang lebih baik dan akurat. Berdasar tabel 1, banyak yang membahas mengenai pengenalan teks, rata-rata untuk tahapan object detection hingga text segmentation metode yang paling banyak digunakan dan akurat adalah Convolutional Neural Network dengan pencapaian akurasi mencapai 93%. Meski umumnya para peneliti melakukan modifikasi terhadap metode CNN tersebut, sehingga diraih hasil yang beragam, namun secara umum dapat dikatakan bahwa penggunaan CNN adalah metode yang terbaik untuk fase awal pengenalan dan segmentasi teks. Sementara pencapaian terbaik berikutnya diperoleh metode OCR, yakni mencapai 98% pada pembacaan plat nomor yang memiliki font yang sama pada tiap kendaraan, dilihat dari besarnya nilai memang tampak sangat menjanjikan, namun hasil tersebut tidak dapat diraih untuk pembacaan teks secara umum dengan jenis font yang berbeda-beda. Hasil terbaik yang diperoleh metode OCR adalah 88% dengan kondisi pembacaan image yang stabil dan pencahayaan yang cukup baik serta jenis font standar sebuah buku, bukan pembacaan teks secara real-time yang ditemui di jalan dengan variasi yang sangat beragam.

Table 1. Table comparison of Methodology Text Detection

No	Title	Author	Year Publish	Methodology	Result
Penelitian Text Recognition pada Scene Text Video Footage (dataset)					
1	Text recognition in natural images using multiclass hough forests	Gokhan Yildirim, Radhakrishna Achanta, Sabine Susstrunk	2013	Hough Forest	85.7%
2	OCR from video stream of book flipping	Dibayan Chakraborty, Partha Pratim Roy, Jose M. Alvarez, Umapada Pal	2013	Region of Interest (ROI)	88%
3	FASText: Efficient Unconstrained Scene Text Detector	U. Bhattacharya, S.K. Parui and S. Mondal	2016	MSER & FASText	76%
4	SceneTextReg: A Real-Time Video OCR System	Haojin Yang, ChengWang, Christian Bartz, Christoph Meinel	2016	Convolutional Neural Network (CNN)	82%
5	An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition	Baoguang Shi, Xiang Bai	2016	Convolutional Recurrent Neural Network (CRNN)	83.4%
Penelitian Text Recognition pada Video Real-Time (Streaming)					
6	Real-Time Scene Text Localization and Recognition	Lukas Neumann, Jiri Matas	2012	Extremal Regions (ER) & syntetic fonts	68.7%
7	Robust Scene Text Recognition with Automatic Rectification	Baoguang Shi, Xinggang Wang, Pengyuan Lyu, Cong Yao, Xiang Bai	2016	Robust text recognizer with Automatic Rectification (RARE)	59%
8	Object Detection in Video using Faster R-CNN	Prajit Ramachandran	2015	Faster R-CNN	53%
9	Real-Time Lexicon-Free Scene Text Localization and Recognition	Lukas Neumann, Jiri Matas	2016	Extremal Regions (ER)	45.2% on ICDAR13 68.1% on SVT 41.8% on ICDAR15
Penelitian Text Recognition pada Image					

10	Automobile Number Plate Recognition And Extraction Using Optical Character Recognition	Raja Vikramdeep Singh, Navneet Randhawa	2016	Optical Character Recognition Methodology	98%
11	Automatic Localization and Recognition of Perspectively Distorted Text in Natural Scene Images	Annmaria Cherian	2016	Hough Transform & SVM	72%
12	A Method for Text Text Localization and Recognition in Real-World Image	Lukas Neumann, Jiri Matas	2011	SVM	67%
13	Estimating hidden parameters for text localization and recognition	Lukas Neumann, Jiri Matas	2011	MSER	60%
14	Text Detection and Character Recognition in Scene Images with Unsupervised Learning	Adam Coates et all	2011	Feature Learning	81%
15	Scene Text Localization and Recognition with Oriented Stroke Detection	Lukas Neumann, Jiri Matas	2013	Oriented Stroke Detection	Text-localization: 72.3% End-to-end: 45.2%
16	Text Recognition in Natural Images Using Multiclass Hough Forest	Gokhan Yildirim, Radhakrishna Achanta, Sabine Susstrunk	2013	Hough Forest	84.5%
17	Efficient Scene text localization and recognition with local character refinement	Lukas Neumann, Jiri Matas	2015	MSER with Character Refinement	77.1%
18	Reading Text in the Wild with Convolutional Neural Networks	Jaderberg et al	2016	CNN	AVG: 93%
19	Real-Time Lexicon-Free Scene Text Localization and Recognition	Lukas Neumann, Jiri Matas	2016	External Region(ER)	AVG: 73%
20	Robust Scene Text Recognition with Automatic Rectification	Baoguang Shi et al	RARE	RARE	SVT: 77% CUTE80: 59%

Sedangkan Dari tabel 2 didapatkan bahwa model skin color segmentation yang baik digunakan yaitu CIELab color space lebih baik dari model warna lainnya yaitu dengan nilai akurasi bisa mencapai 96.87%. Dan untuk metode klasifikasi didapatkan Convolutional Neural Network memiliki hasil akurasi sebesar 98%.

Table 2. Table comparion of methodology Human Skin Detection

No	Title	Author	Year Publish	Methodology	Result
Research about Skin Human for Color Space					
21	A Sign Language Recognition Approach For Human-Robot Symbiosis	Afzal Hossian, Shahrin Chowdhury, and Asma-ull-Hosna	2014	Skin Color Segmentation : YIQ Back Propagation Neural Network	Accuracy : 80%-92%
22	Advanced Marathi Sign Language Recognition using Computer Vision	Amitkumar Shinde dan Ramesh Kagalkar	2015	Skin Color Segmentation : HSV Euclidean Distance	Accuracy : 30% - 100% But the executing to long enough
23	Indian Sign Language Interpreter with Android Implementation	Shanmukha Swamy, Chethan M P, dan Mahantesh Gatwadi	2014	Skin Color Segmentation model HSV color space dan Local Binary Pattern	Accrucary : 52.6% - 62%
24	Android Based Real-Time Static Indonesian Sign Language Recognition System Prototype	Rudy Hartanto, Annisa Kartikasari	2016	Skin Color Segmentation : YCbCr color space; Back Propagation Neural Network	Accuracy : 62.6%
25	Real-time American Sign Language Recognition with Convolutional Neural Networks	Brandon Garcia dan Sigberto Alarcon Viesca	2016	Convolutional Neural Networks	Accuracy : 70%-98%

26	<i>Thai Finger-Spelling Sign Language Recognition Using Global and Local Features with SVM</i>	Thongpan Pariwat dan Pusadee Seresangtakul	2017	<i>SVM Teknik : Linear Kernel Accuracy, Polynomial Kernel, RBF Kernel, Sigmoid Kernel</i>	Accuracy : 91.20%
27	Arabic Sign Language Fingerspelling Recognition from Depth and Intensity Images	Salch Alyl, Basma Osman , Walaa Aly dan Mahmoud Saber	2016	<i>Sensor SOFTKINECT dan PCANet</i>	Accuracy : 96%-99.5%
28	Real Time Two Way Communication Approach for Hearing Impaired and Dumb Person Based on Image Processing	Shweta S. Shinde, Rajesh M. Autee dan Vitthal K. Bhosale	2016	<i>Mel Frequency Cepstrum Coefficients dan Geometric operations, Neighborhood and block operations.</i>	Accuracy : 91%
29	Hand Sign Language Recognition for Bangla Alphabet using Support Vector Machine	Md Azher Uddin dan Shayhan Ameen Chowdhury	2016	<i>Skin Color Segmentation : HSV color space dan SVM</i>	Accuracy : 97.7%
30	Pixel-based Skin Segmentation in Psoriasis Images	Y. George, M. Aldeen, and R. Garnavi	2016	<i>Skin Color Segmentation : CIELabs color space dan Bayesian Classifier</i>	Accuracy : 90%-97%
31	Fuzzy Automatic Contrast Enhancement Based on Fuzzy C-Means Clustering in CIELAB Color Space	Po Ting Lin dan Boting Rex Lin	2016	<i>Skin Color Segmentation : CIELab color space dan Fuzzy Automatic Contrast Enhancement (FACE)</i>	root-mean-square contrast (RMSC) enhancement : 12.6-243.7
32	Non-destructive quality grading of mango (<i>Mangifera Indica L</i>) based on CIELab colour model and size	Rashmi Pandey, Nikunj Gamit dan Sapan Naik	2014	<i>Skin Color Segmentation : CIELab color space dan FUZZY</i>	Accuracy : 85.71 % - 96.87%
33	Implementation of skin color selection prior to Gabor filter and neural network to reduce execution time of face detection	Mejda Chihaoui, Akram Elkefi, Wajdi Bellil dan Chokri Ben Amar	2015	<i>Skin Color Segmentation : YCbCr color space dan Sd-Gabor-NN</i>	Execution time (sec): 12.46500 – 77.77300

DAFTAR PUSTAKA

- [1] Hossian, Afzal., Chowdhury, Shahrin dan Asma-ull-Hosna. "A Sign Language Recognition Approach For Human-Robot Symbiosis"., International Journal of Computer Science, Engineering and Information Technology (IJCEIT), Vol. 4, No.4, August 2014.
- [2] Shanmukha Swamy, Chethan M P, dan Mahantesh Gatwadi. "Indian Sign Language Interpreter with Android Implementation", International Journal of Computer Applications, Volume 97– No.13, July 2014.
- [3] Shinde, Amitkumar dan Kagalkar, Ramesh. "Advanced Marathi Sign Language Recognition using Computer Vision"., International Journal of Computer Applications, Volume 118 – No. 13, pp.1-7. 2015.
- [4] Hartanto, Rudy dan Kartikasari, Annisa. "Android Based Real-Time Static Indonesian Sign Language Recognition System Prototype"., International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 1-6. Oct. 2016.
- [5] Garcia, Brandon., dan Viesca, Sigberto Alarcon. "Real-time American Sign Language Recognition with Convolutional Neural Networks", Stanford University., Stanford, CA, 2016.
- [6] Pariwat, Thongpan. dan Seresangtakul, Pusadee. "Thai Finger-Spelling Sign Language Recognition Using Global and Local Features with SVM"., 9th International Conference on Knowledge and Smart Technology (KST), pp. 116-120. March 2017.

- [7] Saleh Alyl, Basma Osman , Walaa Aly dan Mahmoud Saber."Arabic Sign Language Fingerspelling Recognition from Depth and Intensity Images"., International Computer Engineering Conference (ICENCO), pp. 99-104. 2016
- [8] Shweta S. Shinde, Rajesh M. Autee dan Vitthal K. Bhosale., "Real Time Two Way Communication Approach for Hearing Impaired and Dumb Person Based on Image Processing"., International Conference on Computational Intelligence and Computing Research (ICCIC), pp.1-5. 2016.
- [9] Md Azher Uddin dan Shayhan Ameen Chowdhury., "Hand Sign Language Recognition for Bangla Alphabet using Support Vector Machine"., International Conference on Innovations in Science, Engineering and Technology (ICISSET). Pp 1-4, 2016.
- [10] Y. George, M. Aldeen, and R. Garnavi., "Pixel-based Skin Segmentation in Psoriasis Images"., nternational Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Pp 1352- 1356, 2016.
- [11] Po Ting Lin dan Boting Rex Lin., "Fuzzy Automatic Contrast Enhancement Based on Fuzzy C-Means Clustering in CIELAB Color Space"., International Conference on Mechatronic and Embedded Systems and Applications (MESA). Pp 1-10, 2016.
- [12] Rashmi Pandey, Nikunj Gamit dan Sapan Naik., "Non-destructive quality grading of mango (*Mangifera Indica* L) based on CIELab colour model and size"., International Conference on Advanced Communications, Control and Computing Technologies., pp 1246-1251, 2014.
- [13] Mejda Chihaoui, Akram Elkefi, Wajdi Bellil dan Chokri Ben Amar., "Implementation of skin color selection prior to Gabor filter and neural network to reduce execution time of face detection"., International Conference on Intelligent Systems Design and Applications (ISDA). Pp 341-346, 2015.
- [14] Y. George, M. Aldeen, and R. Garnavi, "Pixel-based Skin Segmentation in Psoriasis Images", 2016 IEEE 38th Annual International Conference, Engineering in Medicine and Biology Society (EMBC), 2016.
- [15] Lukas Neumann, Jiri Matas, "Real-Time Lexicon-Free Scene Text Localization and Recognition". IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016
- [16] Haojin Yang, ChengWang, Christian Bartz, Christoph Meinel, "SceneTextReg: A Real-Time Video OCR System Haojin", Hasso Plattner Institute (HPI), University of Potsdam, Germany. 2016
- [17] Baoguang Shi, Xiang Bai. "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. 2016
- [18] Baoguang Shi, Xinggang Wang, Pengyuan Lyu, Cong Yao, Xiang Bai. "Robust Scene Text Recognition with Automatic Rectification". IEEE Conference on Computer Vision and Pattern Recognition Robust. 2016
- [19] Raja Vikramdeep Singh, Navneet Randhawa. "Automobile Number Plate Recognition And Extraction Using Optical Character Recognition". INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 3. 2014
- [20] Gokhan Yildirim, Radhakrishna Achanta, Sabine Susstrunk. "Text Recognition in Natural Images Using Multiclass Hough Forest". 2016
- [21] Prajit Ramachandran. "Object Detection in Video using Faster R-CNN". ICCV Workshop, India. 2015
- [22] Dibayan Chakraborty, Partha Pratim Roy, Jose M. Alvarez, Umapada Pal. "OCR from video stream of book flipping". Proceedings - 2nd IAPR Asian Conference on Pattern Recognition, ACPR. 2013
- [23] U. Bhattacharya, S.K. Parui and S. Mondal. "FASText: Efficient Unconstrained Scene Text Detector". Proceedings of the IEEE International Conference on Computer Vision. 2016

- [24] Annmaria Cherian. "Automatic Localization and Recognition of Perspectively Distorted Text in Natural Scene Images". IEEE International Conference on Computer Vision. 2016
- [25] Pooja, Renu Dhir, "Video Text Extraction and Recognition: A Survey", Proceeding IEEE WISPNET, 2016
- [26] Busta, M., Neumann, L., dan Matas, J. FASText: Efficient Unconstrained Scene Text Detector Michal. Proceedings of the IEEE International Conference on Computer Vision. 2016
- [27] Farinella G. Maria, Battiato S, dan Cipolla R., Advanced Topics in Computer Vision. Springer London Heidelberg New York Dordrecht. New York, US. 2013.
- [28] Jaderberg, M., Simonyan, K., Vedaldi, A., Zisserman, A., Reading Text in the Wild with Convolutional Neural Networks. International Journal of Computer Vision. 2016.
- [29] Liao, M., Shi, B., Bai, X., Wang, X., Liu, W. 2016. TextBoxes: A Fast Text Detector with a Single Deep Neural Network. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence.
- [30] Neumann, L., Matas, J. A Method for Text Text Localization and Recognition in Real-World Image. Center for Machine Perception, Czech Technical University in Prague, Czech Republic. 2011.
- [31] Neumann, L., Matas, J. Estimating hidden parameters for text localization and recognition. Computer Vision Winter Workshop (CVWW). 2011.
- [32] Neumann, L., Matas, J. Scene text localization and recognition with oriented stroke detection. Proceedings of the IEEE International Conference on Computer Vision. 2013.
- [33] Neumann, L., Matas, J. Scene text localization and recognition with oriented stroke detection. Proceedings of the IEEE International Conference on Computer Vision. 2013.
- [34] Neumann, L., Matas, J., Real-Time Lexicon-Free Scene Text Localization and Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016.