

Sistem Perhitungan Suara Pemilu Otomatis pada Formulir C1 dengan Penerapan Deep Learning

M. Alif Al Hakim¹, Rakha Abid Bangsawan², Muhammad Mariozulfandy³, Adila Alfa Krisnadhi⁴

¹Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, email: malif.al@ui.ac.id

²Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, email: rakha.abid@ui.ac.id

³Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, email: muhammad.mariozulfandy@ui.ac.id

⁴Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, email: adila@cs.ui.ac.id

Corresponding Author: M.Alif Al Hakim

ABSTRAK — Proses perhitungan atau rekapitulasi perhitungan suara pemilihan umum di Indonesia masih dilakukan secara manual. Hal ini tentu saja memakan banyak waktu dan biaya. Walaupun Indonesia telah mulai menerapkan sistem perhitungan suara otomatis menggunakan *optical recognition text* pada Sirekap di tahun 2024 ini, masih terdapat masalah akurasi pada Sirekap tersebut. Oleh karena itu, dibutuhkan sebuah sistem perhitungan suara otomatis yang baru dengan performa yang baik untuk membantu proses perhitungan atau rekapitulasi suara menjadi lebih efektif dan efisien. Penelitian ini bertujuan untuk mengembangkan sistem perhitungan suara otomatis pada Formulir C1 pemilihan umum Presiden dan Wakil Presiden dengan menggunakan metode *fine-tuning* PaddleOCR dan YOLO-NAS. Metode tersebut digunakan karena metode yang digunakan sirekap untuk menentukan *Region of Interest* (ROI) masih dilakukan secara manual. Selain itu akurasi model OCR Sirekap juga diklaim 93% pada data lapangan. Pada penelitian ini, digunakan dataset yang di-*scraping* melalui *website* komisi pemilihan umum. Dataset yang digunakan terdiri dari gambar Formulir C1 yang secara acak diambil dari berbagai Tempat Pemungutan Suara (TPS) dari seluruh kota/kabupaten di Indonesia. Dataset berjumlah 1187 gambar. Setiap gambar pada dataset kemudian dilakukan *preprocessing*, seperti *Auto Orient*, *Gray Scale* dan *Data Augmentation*. Gambar-gambar yang telah diproses kemudian digunakan untuk melatih model deteksi YOLO-NAS yang berguna untuk mendeteksi ROI yaitu kotak hasil perhitungan suara dan model PaddleOCR yang berguna untuk membaca teks pada gambar. Pada evaluasi yang dilakukan, metode ini memiliki performa yang cukup baik dimana memiliki akurasi sebesar 96,9% dalam kebenaran membaca hasil perhitungan suara. Kesalahan prediksi yang terjadi umumnya disebabkan gambar yang sangat *blur* atau tidak konsisten dalam penulisan teks dan *marking* OMR hasil perhitungan suara pada data gambar Formulir C1. Dengan demikian, dari penelitian ini disimpulkan bahwa penggunaan *fine-tuning* PaddleOCR dan YOLO-NAS dapat meningkatkan efisiensi dan akurasi perhitungan suara pemilu. Namun, perlu dicatat bahwa performa model yang dihasilkan bergantung pada kualitas dari gambar.

KATA KUNCI — OCR, Paddle, YOLO, Pemilu, Perhitungan Suara Otomatis

ABSTRACT — The process of calculating or recapitulating the vote count for general elections in Indonesia is still done manually. This of course takes a lot of time and money. Even though Indonesia has started implementing an automatic vote counting system using optical recognition text in Sirekap in 2024, but there are still accuracy problems in Sirekap. Therefore, a new automatic vote counting system with good performance is needed to help the vote counting or recapitulation process become more effective and efficient. This research aims to develop an automatic vote counting system for Form C1 for the general election of President and Vice President using the PaddleOCR and YOLO-NAS fine-tuning methods. This method is used because the method used by sirekap to determine Region of Interest (ROI) is still done manually. Apart from that, the accuracy of the Sirekap OCR model is also claimed to be 93% on field data. In this research, a dataset scraped via the general election commission's website. The dataset used consists of images of Form C1 which were randomly taken from various polling places (TPS) in all cities/districts in Indonesia. The dataset consists of 1187 images. Each image in the dataset is then subjected to preprocessing, such as Auto Orient, Gray Scale and Data Augmentation. The images that have been processed are then used to train the YOLO-NAS detection model which is useful for detecting ROIs, namely boxes resulting from sound calculations and the PaddleOCR model which is useful for reading text on images. In the evaluation carried out, this method performed quite well, having an accuracy of 96.9% in correctly reading the vote count results. Prediction errors that occur are generally caused by images that are very blur or inconsistent in text writing and marking OMR vote calculation results on Form C1 image data. Thus, from this research it is concluded that the use of fine-tuning PaddleOCR and YOLO-NAS can increase the efficiency and accuracy of election vote calculations. However, it should be noted that the performance of the resulting model depends on the quality of the images.

KEYWORD — OCR, Paddle, YOLO, Pemilu, Automatic Vote Counting

I. PENDAHULUAN

A. LATAR BELAKANG

Pemilihan Umum (Pemilu) merupakan kegiatan yang dilaksanakan setiap lima tahun di Indonesia dan harus diikuti oleh seluruh warga negara Indonesia yang berusia lebih dari 17 tahun. Pemilu adalah proses pemilihan untuk memilih sebagian besar atau seluruh anggota suatu badan terpilih badan legislatif dan presiden yang dipilih secara langsung oleh masyarakat [1]. Saat ini, proses perhitungan dan rekapitulasi suara Pemilu di Indonesia dilakukan secara manual mulai dari tahap kecamatan, kota, provinsi, dan terakhir nasional sehingga perhitungan ini memakan waktu yang cukup lama hingga minimal satu bulan dari proses pemungutan suara serta biaya yang sangat besar. Kementerian Keuangan (Kemenkeu) mengalokasikan anggaran hingga Rp71,3 triliun untuk Pemilu 2024. Jumlah dana ini naik 57,3 persen dibandingkan anggaran pesta demokrasi serentak pada 2019 yang sebesar Rp45,3 triliun [2].

Sistem perhitungan suara otomatis dapat menangani hal tersebut. Sistem perhitungan suara otomatis ini bekerja dengan menghitung suara setiap paslon secara otomatis dengan data Formulir C1 hasil pemungutan suara. Sistem ini dapat mempercepat proses perhitungan suara karena dilakukan secara otomatis oleh mesin serta mengurangi biaya operasional Pemilu karena semakin sedikit tugas perhitungan suara manual yang dilakukan (semakin mudah tugas semakin kecil biaya) jika sistem ini bekerja dengan baik.

Sistem perhitungan suara otomatis dengan teknologi OCR mulai diterapkan oleh KPU pada Pemilu 2024 dengan nama Sirekap [3]. Namun, masih terdapat beberapa masalah, seperti masalah akurasi yang dimiliki oleh sistem ini dimana utamanya disebabkan oleh kesalahan pembacaan digit dan buruknya pengambilan gambar. Berdasarkan laporan yang dipublikasikan di website Mahkamah Konstitusi, Sirekap memiliki tiga masalah utama yang menghambat dalam proses perhitungan suara [4] yaitu sebagai berikut.

- 1) Tulisan tangan yang gagal dideteksi oleh OCR pada Sirekap. Sirekap mengklaim bahwa akurasi OCR mereka 93% pada lapangan sehingga masih terdapat 7% kesalahan pembacaan gambar.
- 2) Pengambilan gambar yang buruk. Data gambar yang diperoleh Sirekap didapatkan dari *handphone* sehingga kualitas gambar yang diperoleh berbeda-beda dan pastinya ada yang buruk.
- 3) Masalah kertas yang terlipat. Ini menimbulkan kesalahan pada interpretasi model OCR Sirekap. Terlebih lagi, jika bagian hasil perhitungan suara tidak terlihat.

Selain itu, peneliti juga melakukan analisis terhadap proses pembacaan hasil perhitungan suara gambar Formulir C1 pada aplikasi *mobile* Sirekap.¹ Peneliti memperoleh informasi bahwa Sirekap menggunakan model OCR yang telah dilatih dengan MNIST (dataset tulisan tangan digit 0-9). Setelah aplikasi memperoleh gambar, beberapa preprocessing dilakukan

¹diperoleh dari Reverse Engineering Aplikasi *mobile* Sirekap

pada gambar, seperti *grayscale*, menyesuaikan kontras dengan CLAHE, dan menerapkan *adaptive thresholding*. Setelah itu, akan dideteksi kotak hasil perhitungan suara setiap paslon dengan menggunakan koordinat yang telah ditetapkan (konstan) dan ada sedikit penyesuaian dengan ukuran gambar secara matematis untuk dilakukan *cropping*. Setelah kotak hasil perhitungan suara didapatkan, akan dilakukan pembacaan dengan model. Model yang digunakan adalah model MNIST OCR yang digunakan untuk membaca setiap digit pada setiap kotak digit suara. Selain menggunakan OCR, aplikasi juga dapat melakukan OMR (Optical Mark Recognition) dan lebih cenderung akan melakukan itu (jika OMR gagal, model melakukan OCR). OMR diimplementasikan dengan menggunakan OpenCV dengan cara mendeteksi semua bulatan dan menemukan *outlier* bulatan (bulatan hitam yang tidak tepat di area bulatan), lalu dilakukan Optical Mark Recognition pada bulatan tersebut.

Masalah Sirekap ini menimbulkan beberapa masalah lainnya, seperti kegaduhan masyarakat akibat koreksi rekapitulasi suara salah satu paslon Presiden dan Wakil Presiden [5]. Kesalahan ini juga menimbulkan keresahan di media sosial kala itu akan manipulasi suara, padahal hanya kesalahan akibat masalah sistem Sirekap saja. Masalah-masalah tersebut yang mendorong peneliti dalam berkontribusi mengembangkan sistem perhitungan suara otomatis pada Formulir C1 Pemilu. Dari masalah dan analisis Sirekap di atas, peneliti akan mencoba mengimplementasikan sistem perhitungan suara menggunakan deteksi objek (kotak perhitungan suara) dengan model YOLO-NAS dan pembacaan dengan PaddleOCR.

B. RUMUSAN MASALAH

Berdasarkan latar belakang yang telah dijelaskan, pertanyaan penelitian yang akan tim peneliti ajukan adalah sebagai berikut.

- 1) Bagaimana efektivitas dan akurasi dari sistem perhitungan suara pada formulir C1 dengan optimasi PaddleOCR dan YOLO ?
- 2) Apa tantangan dan hambatan yang dihadapi oleh sistem perhitungan suara pada formulir C1 dengan optimasi PaddleOCR dan YOLO ?
- 3) Bagaimana performa kedua metode, yaitu YOLO dan PaddleOCR dalam mengenali digit suara pada gambar kualitas yang buruk ?

C. TUJUAN PENELITIAN

Berikut adalah tujuan yang ingin dicapai dari penelitian ini.

- 1) Mengetahui bagaimana efektivitas dan akurasi dari sistem perhitungan suara pada formulir C1 dengan optimasi PaddleOCR dan YOLO.
- 2) Mengetahui tantangan dan hambatan yang dihadapi oleh sistem perhitungan suara pada formulir C1 dengan optimasi PaddleOCR dan YOLO.
- 3) Mengetahui bagaimana performa PaddleOCR dan YOLO pada gambar dengan kualitas yang buruk.

D. MANFAAT PENELITIAN

Penelitian ini diharapkan dapat memberikan manfaat bagi pihak-pihak yang membutuhkan, baik secara teoritis maupun praktis, diantaranya sebagai berikut.

1) Manfaat Teoritis

- Sebagai *blueprint* untuk membuat sistem perhitungan suara pada formulir C1 secara otomatis.
- Menambah wawasan dan pengetahuan pembaca terkait sistem perhitungan suara dari dataset yang diberikan.

2) Manfaat Praktis

- Membantu pemerintah dalam upaya untuk menghitung perhitungan suara sah dari formulir C1.
- Meningkatkan efisiensi tenaga dan waktu dalam melakukan rekapitulasi suara sah.
- Mendeteksi data gambar Formulir C1 yang memiliki kualitas (ketajaman, kecerahan, pengambilan) yang buruk dengan performa model perhitungan suara pada gambar tersebut.

E. BATASAN PENELITIAN

Batasan penelitian pada metode yang dipaparkan dalam makalah adalah sebagai berikut.

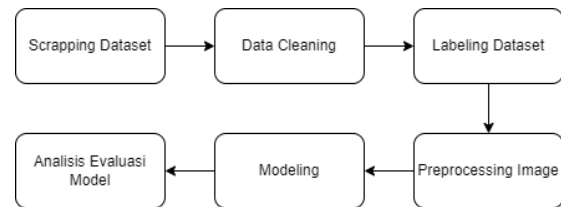
- 1) Model yang dibuat hanya untuk deteksi suara pada Formulir C1 pemilihan paslon Presiden dan Wakil Presiden.
- 2) Performa model bergantung pada kualitas (seperti ketajaman, kecerahan, dan jarak pengambilan) gambar Formulir C1 yang diberikan.
- 3) Penelitian hanya berfokus pada pengembangan model sistem perhitungan suara otomatis pada Formulir C1 Pemilu.
- 4) Jumlah dataset gambar Formulir C1 yang digunakan dalam pengembangan model sebanyak 1187 data.
- 5) Bagian kotak perolehan suara pada data gambar Formulir C1 harus terlihat jelas untuk menjalankan sistem perhitungan suara yang dibuat.

II. METODOLOGI

Bab ini menjelaskan alur penelitian singkat yang dilakukan oleh tim peneliti, *dataset* yang digunakan untuk melatih dan validasi model, penjelasan mengenai model dan pemrosesan data yang dilakukan, serta metrik yang digunakan untuk mengevaluasi performa model dalam mendeteksi digit pada perhitungan suara sah pada formulir C1.

A. RINGKASAN ALUR PENELITIAN

Penelitian dilakukan untuk membuat sistem perhitungan suara otomatis pada setiap Paslon Formulir C1. Berikut adalah diagram yang dapat dilihat pada gambar 1 dan penjelasan singkat alur penelitian untuk mencapai hal tersebut.



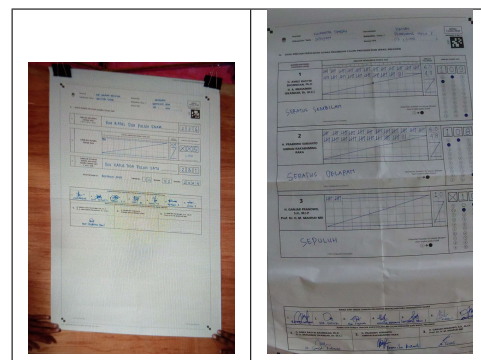
Gambar 1. Alur Penelitian

1) *Scrapping* Dataset Formulir C1 Pemilu

Data yang akan dideteksi hasil suaranya dan digunakan untuk *training* model adalah gambar Formulir C1. Data gambar ini diperoleh dari laman <https://pemilu2024.kpu.go.id/pilpres/hitung-suara/>. Tim peneliti melakukan *automation scraping* dengan *package* selenium 4.22.0 Python untuk memperoleh tiga gambar Formulir C1 secara acak pada TPS-TPS di seluruh kota/kabupaten di Indonesia. Kami memperoleh 1194 gambar Formulir C1 dari hasil *scrapping* ini.

2) Pembersihan Dataset

Dari semua data hasil *scrapping*, masih terdapat beberapa data yang salah format akibat proses *scrapping* otomatis atau sangat kotor untuk di-*training* ke model sehingga akan dilakukan pembersihan untuk memperoleh dataset Formulir C1 yang diinginkan oleh tim peneliti. Dataset yang sangat kotor ini adalah data yang dinyatakan pada batasan model. Namun, tim peneliti tetap menyisakan dataset yang kotor, seperti pada gambar 2 yang akan diproses terlebih dahulu dan di-*training* pada model. Gambar 2 ini adalah contoh data yang dihapus karena kesalahan format (kiri) dan sangat kotor karena kotak hasil perhitungan suara salah satu Paslon tidak terlihat (kanan). Banyak dataset setelah dilakukan pembersihan ada 1187 data.

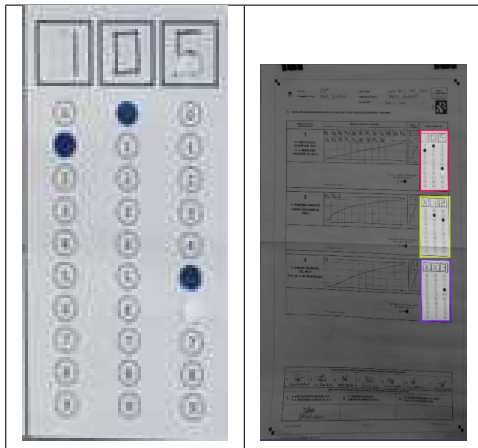


Gambar 2. Kesalahan Format Saat *Scrapping* dan Data Kotor (Terpotong)

3) Labeling Dataset

Setelah melakukan pembersihan pada data, akan dilakukan *labeling* pada dataset yang tersisa. Terdapat dua proses *labeling* pada dataset yang akan dilakukan pada data. Labeling pertama adalah *labeling* untuk menentukan *output* dari model YOLO-NAS untuk deteksi objek *region of interest* (ROI) dalam hal ini kotak

hasil perhitungan suara seperti yang diilustrasikan pada gambar 3.



Gambar 3. Region of Interest (ROI) pada Dataset

Labeling kedua adalah untuk menentukan *output* dari model PaddleOCR dalam menentukan jumlah suara yang dibaca dari ROI.

4) *Preprocessing* Dataset

Preprocessing dilakukan pada dataset sebelum di-*training* ke model. Dataset yang akan dideteksi dan digunakan pada proses *training* tidak selalu bersih sehingga perlu dilakukan pemrosesan terlebih dahulu pada dataset agar data lebih bersih dan dapat diterima oleh model. Setelah melakukan *preprocessing*, dataset akan dibagi menjadi 70% data *training*, 20% data *validation*, dan 10% data *testing*.

5) Modeling

Terdapat dua model yang digunakan oleh tim peneliti pada penelitian ini, yaitu YOLO-NAS dan PaddleOCR. Penjelasan mengenai kedua model tersebut terdapat pada bagian MODEL di bawah. Secara singkat, model YOLO-NAS akan digunakan untuk deteksi objek ROI. Sementara, PaddleOCR untuk pembacaan hasil perhitungan suara. Data *training* yang telah diperoleh sebelumnya yang akan digunakan pada proses *training* kedua model ini.

6) Evaluasi Model

Setelah melakukan *training* pada model, performa model tersebut akan dievaluasi dengan Mean Average Precision (YOLO-NAS) dan Counter Level Accuracy, Digit Level Accuracy, Average Underprediction Error, dan Average Overprediction Error.

B. DATASET

Dataset yang digunakan dalam penelitian berupa kumpulan foto formulir C1 yang di-*scraping* dari laman KPU. Data yang telah dikumpulkan dibersihkan terlebih dahulu untuk memperoleh dataset yang valid sesuai dengan format formulir C1. Jumlah dataset yang didapatkan setelah *scraping* berjumlah 1194.

Dari dataset yang telah dikumpulkan terdapat beberapa data yang kotor, seperti berwarna, blur, gelap, diambil dari jarak jauh, miring, dan terang. Berikut adalah beberapa contoh dataset dengan kondisi tersebut di gambar 4.



Gambar 4. Sampel Gambar Kotor pada Dataset

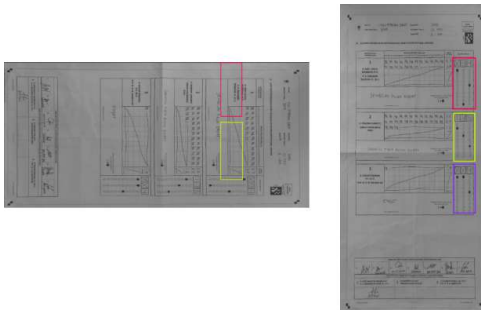
C. PREPROCESSING

Dataset yang kotor akan ditangani dengan melakukan *preprocessing* pada dataset terlebih dahulu sebelum di-*training* ke model. Berikut adalah beberapa metode *preprocessing* yang dilakukan.

1) *Grayscale*: Proses Grayscale ini mengubah gambar dari RGB pixel (berwarna) ke Grayscale pixel (hitam putih) dengan mengkonversi warna Red (R), Green (G), dan Blue (B) ke skala intensitas cahaya (terang/gelap). Pemrosesan ini dapat membantu memperjelas gambar yang berwarna, terang, atau gelap sehingga dapat lebih mudah dikenali model.

2) *Auto Orientation*: Metode ini dilakukan untuk mengubah orientasi gambar secara otomatis menjadi *portrait* dengan format yang benar. Ini diperlukan untuk menangani data yang salah orientasi seperti *landscape* atau terbalik sebelum di-*training* ke model. Pemrosesan ini dapat membantu model deteksi objek seperti model YOLO. Gambar 5 merupakan contoh dari penerapan *preprocessing auto orientation*.

3) *Data Augmentation*: Data Augmentation adalah proses menghasilkan data baru dari data yang sudah ada dengan menambahkan beberapa pemrosesan ke data baru tersebut. Dalam hal ini, tim peneliti melakukan beberapa pemrosesan, yaitu Rotation (mengatur agar rotasi gambar beragam), Shear (mengatur agar perspektif gambar beragam), Brightness (mengatur agar kecerahan gambar beragam), dan Blur (mengatur agar keburaman gambar beragam). Ini dilakukan agar model



Gambar 5. Auto Orientation pada Gambar Kotor

dapat dilatih untuk lebih terbiasa dengan berbagai kondisi data gambar walaupun data gambar tersebut 'kotor'.

D. MODEL

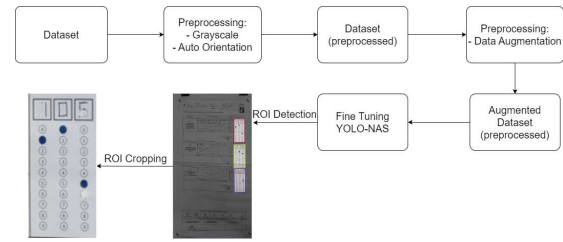
Proses *modeling* data akan dilakukan pada dua model. Model pertama adalah YOLO-NAS yang akan digunakan untuk deteksi objek ROI, yaitu kotak hasil perhitungan suara pada gambar Formulir C1. Setelah itu, akan dilakukan *cropping* pada kotak hasil perhitungan suara. Jumlah gambar kotak hasil perhitungan suara yang diperoleh dari setiap gambar Formulir C1 akan sesuai dengan jumlah Paslon. Gambar kotak hasil perhitungan suara tersebut kemudian akan dibaca dengan model kedua yaitu Paddle-OCR sehingga akan diperoleh sistem yang secara otomatis menghitung suara dari setiap Paslon pada Formulir C1.

1) YOLO-NAS

YOLO merupakan model yang umumnya digunakan dalam proses deteksi objek [6]. YOLO-NAS merupakan peningkatan dari model YOLO dimana dilakukan optimisasi pada arsitektur *neural network* YOLO untuk meningkatkan akurasi dan kecepatan. Tim peneliti menggunakan model ini untuk mendeteksi Region of Interest (ROI) pada gambar Formulir C1 dalam hal ini kotak hasil perhitungan suara. Untuk melatih model ini, tim peneliti terlebih dahulu melakukan anotasi ROI pada data. Setelah itu, tim peneliti akan melatih model dengan data yang telah dianotasi tersebut. Model yang telah dilatih tersebut dapat melakukan prediksi koordinat ROI kotak hasil perhitungan suara pada gambar Formulir C1 baru. Dengan memperoleh koordinat ROI dari gambar Formulir C1, dapat dilakukan *cropping* pada gambar untuk memperoleh gambar ROI tersebut. Hal ini dilakukan karena tim peneliti hanya ingin memfokuskan model PaddleOCR yang digunakan untuk membaca hasil suara pada ROI tersebut. Tujuannya agar data yang akan dibaca oleh model PaddleOCR lebih jelas dan tidak kompleks. Visualisasi Modeling YOLO-NAS ini terdapat pada Gambar 6.

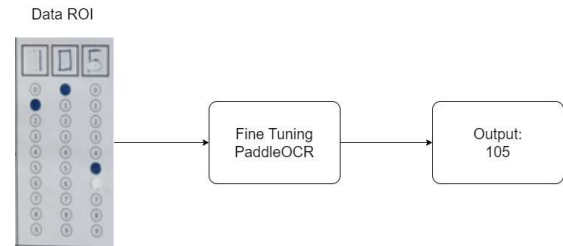
2) PaddleOCR

PaddleOCR adalah salah satu *pretrained* model OCR yang telah dikembangkan untuk deteksi dan pengenalan teks pada gambar [7]. Model ini dikenal dengan akurasi dan efisiensinya dalam melakukan tugas OCR, seperti



Gambar 6. Alur Modeling YOLO-NAS

pengenalan plat kendaraan, pengenalan tulisan tangan, dan lainnya. Tim peneliti menggunakan model ini untuk membaca jumlah suara dari setiap gambar ROI yang diterima. Model ini dilatih dengan data-data gambar ROI dan labelnya. Output dari model ini adalah jumlah suara dari ROI tersebut. Tim peneliti menggunakan model PaddleOCR dengan *epoch* 220, *optimizer* Adam, *learning rate* Cosine 0.001, dan *loss function* CTC Loss. Visualisasi Modeling PaddleOCR ini terdapat pada Gambar 7.



Gambar 7. Alur Modeling PaddleOCR

E. METRIK EVALUASI MODEL

Berikut adalah penjelasan beberapa metrik yang digunakan untuk mengevaluasi performa model dalam mendeteksi objek ROI dan pembacaan hasil suara Paslon.

1) Mean Average Precision

Mean Average Precision adalah metrik untuk mengevaluasi model YOLO-NAS dalam mendeteksi objek (ROI). Metrik ini melibatkan nilai *precision* dan *recall*.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Average Precision (AP) = Area dibawah kurva *precision-recall*. Kurva *precision-recall* melakukan *plotting precision* terhadap *recall* untuk nilai ambang batas yang berbeda. Ketika ambang deteksi bervariasi, nilai *precision* dan *recall* berubah sehingga menciptakan kurva. Kurva ini di-plot untuk setiap kelas. Pada kasus ini, terdapat tiga kotak perhitungan suara (ROI) pada setiap data gambar Formulir C1 sehingga terdapat tiga kelas dan masing-masing kelas tersebut membentuk AP.

$$\text{Mean AP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i$$

N adalah jumlah kelas dan i adalah kelas ke- i .

- 2) **Counter Level Accuracy** Counter Level Accuracy adalah salah satu metrik yang digunakan untuk mengevaluasi model PaddleOCR dalam membaca hasil suara pada ROI. Pada model PaddleOCR, data yang dimasukkan adalah ROI dari gambar Formulir C1. Metrik ini bekerja dengan menghitung persentase jumlah prediksi yang benar (True Positive (TP)) dari jumlah seluruh data *testing*. Berikut adalah rumus matematika yang menyatakan metrik ini.

$$\text{Count Level Accuracy (CLA)} = \frac{\text{TP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- 3) **Digit Level Accuracy**

Digit Level Accuracy adalah salah satu metrik yang digunakan untuk mengevaluasi model PaddleOCR dalam membaca hasil suara pada ROI. Metrik ini dihitung dengan melakukan ekstraksi digit pada hasil prediksi PaddleOCR dan membandingkan masing-masing digit tersebut dengan digit yang bersesuaian pada label. Persentase jumlah digit bersesuaian yang benar dari jumlah seluruh digit pada data (ROI) *testing* inilah yang menjadi Digit Level Accuracy (DLA) dari data.

- 4) **Average Overprediction Error dan Average Underprediction Error**

Ini adalah *custom* metrik yang dibuat oleh tim peneliti untuk mengukur seberapa jauh simpangan kelebihan (*overprediction*) dan kekurangan (*underprediction*) dari jumlah suara setiap paslon yang diprediksi oleh model PaddleOCR. Kedua metrik ini dihitung dengan mengurangi hasil prediksi PaddleOCR pada data dengan label data tersebut. Jika hasil pengurangan bernilai negatif, tim peneliti akan mengkategorikan sebagai *underprediction*. Jika hasilnya bernilai positif, tim peneliti akan mengkategorikan sebagai *overprediction*. Masing-masing *underprediction* dan *overprediction* ini akan dikumpulkan dan dirata-ratakan pada seluruh data *testing* sehingga diperoleh Average Overprediction Error dan Average Underprediction Error.

III. HASIL DAN PEMBAHASAN

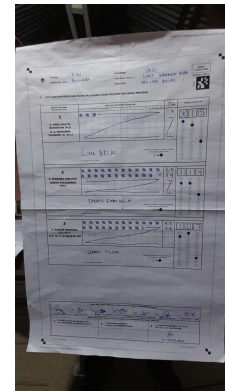
Bab ini menjelaskan hasil dan evaluasi terhadap sistem yang telah dibuat pada penelitian ini. Sebelumnya, telah dijelaskan bahwa 1187 dataset gambar Formulir C1 dibagi menjadi 70% data *training*, 20% data *validation*, dan 10% data *testing* dan telah dijelaskan juga bahwa data gambar Formulir C1 akan di-*training* pada model YOLO-NAS untuk mendeteksi koordinat ROI kotak hasil perhitungan suara. Tabel I menunjukkan hasil evaluasi performa model YOLO-NAS pada data *testing* dengan metrik yang telah dijelaskan sebelumnya.

Dari hasil tersebut dapat dikatakan bahwa performa model YOLO-NAS yang di-*training* untuk mendeteksi ROI kotak

Metrik	Hasil
Precision	99.7%
Recall	100.0%
Mean AP	100.0%

TABEL I
HASIL EVALUASI MODEL YOLO-NAS

hasil perhitungan suara sangat baik. Walaupun pada dataset yang digunakan hanya dilakukan dua *preprocessing* untuk pembersihan, yaitu Grayscale dan Auto Orientation, serta *preprocessing* Data Augmentation untuk pengotoran, performa YOLO-NAS ini sudah sangat baik dalam mendeteksi ROI. Ini merupakan suatu perbedaan atau bahkan peningkatan dari Sirekap dimana deteksi *cropping* ROI dilakukan pada koordinat yang telah ditetapkan (konstan) dengan sedikit penyesuaian sesuai ukuran gambar yang dilakukan secara matematis dengan OpenCV. Menurut peneliti, kekurangan dari pendekatan Sirekap dalam mendeteksi ROI adalah rentannya terhadap data gambar yang kotor, seperti miring, pengambilan dari jarak jauh, pengambilan dari jarak dekat, dan kondisi kertas yang bisa saja kusut atau sedikit terlipat. Masalah tersebut dapat ditangani dengan model YOLO-NAS karena dapat melakukan deteksi objek (ROI) secara otomatis. Gambar 8 adalah contoh gambar dengan masalah tersebut yang dapat dideteksi ROI-nya dengan baik oleh YOLO-NAS.



Gambar 8. Contoh Gambar Kotor untuk Mendeteksi ROI

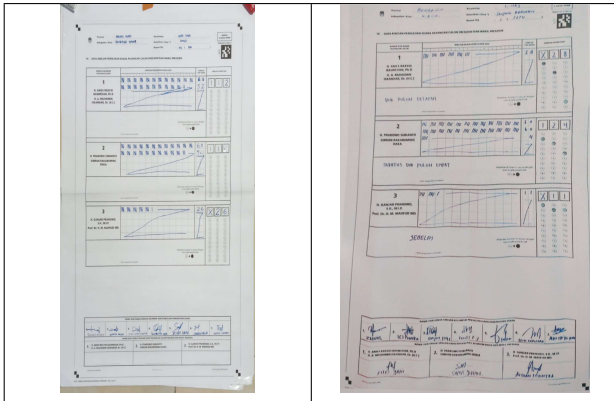
Sebelumnya, tim peneliti telah menjelaskan bahwa data gambar ROI yang telah diperoleh akan dimasukkan ke model PaddleOCR. PaddleOCR ini akan menghasilkan *output* berupa jumlah hasil perhitungan suara yang dibaca dari data gambar ROI. Tabel II menunjukkan hasil evaluasi performa model PaddleOCR pada data *testing* dengan metrik yang telah dijelaskan sebelumnya.

Hasil tersebut dapat dikatakan cukup baik walaupun masih ada beberapa kesalahan dalam model PaddleOCR melakukan pembacaan. Setelah dilakukan analisis lebih lanjut pada data yang gagal dibaca, kesalahan pembacaan secara umum disebabkan karena pengisian yang tidak konsisten (antara bulatan OMR dan teks) atau gambar yang terlalu blur. Contohnya, seperti kedua gambar pada gambar 9 yang gagal dideteksi dimana pada gambar kiri tidak konsisten antara teks dan

Metrik	Hasil		
Count Level Accuracy	0.969		
Digit Level Accuracy	0.983		
	Paslon 1	Paslon 2	Paslon 3
Average Overprediction Error	0	52.0	1.0
Average Underprediction Error	-61.0	-114.0	-44.25

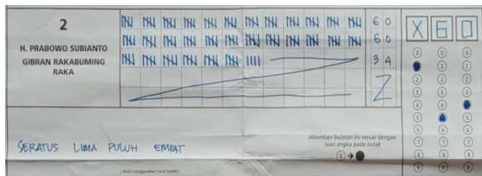
TABEL II
HASIL EVALUASI MODEL PADDLEOCR

bulatan OMR, serta gambar kanan yang sangat blur untuk dibaca oleh model.



Gambar 9. Data yang Gagal Dibaca oleh PaddleOCR

Walaupun demikian, akurasi model yang dikembangkan oleh tim peneliti lebih baik dibandingkan model MNIST OCR Sirekap yang diklaim memiliki performa akurasi 93% pada lapangan. Menurut peneliti, pendekatan yang dilakukan oleh Sirekap masih sangat rentan terhadap tulisan yang tidak jelas, seperti terdapat bagian yang dicoret/ditip-ex pada digit sehingga menimbulkan kesalahan dalam pembacaan digit. Pendekatan yang tim peneliti lakukan untuk menangani hal tersebut adalah membuat model untuk juga meninjau bulatan OMR pada bagian bawah sehingga selain fokus ke digit, bulatan tersebut juga dapat membantu mengenali digit/teks. Gambar 10 adalah contoh data dengan tulisan tidak konsisten yang berhasil dibaca oleh model PaddleOCR karena bulatan juga dipertimbangkan.



Gambar 10. Contoh Gambar Berhasil Dideteksi

Selain dari akurasi, tim peneliti juga melakukan evaluasi terhadap performa kecepatan model. Tim peneliti melakukan pengukuran terhadap kecepatan prediksi model dan diperoleh hasil sebagai berikut.

Waktu prediksi model dapat dikatakan cukup cepat. Ini disebabkan arsitektur model YOLO-NAS dan PaddleOCR itu sendiri yang optimum dan efisien.

Model	Inference Time
YOLO-NAS	0.13 s
PaddleOCR	0.017 s

TABEL III
HASIL KECEPATAN MODEL

IV. KESIMPULAN

Secara umum, penerapan *Deep Learning* menggunakan teknik *fine tuning* YOLO-NAS dan PaddleOCR memiliki performa akurasi dan efektivitas yang baik. Akan tetapi, terdapat tantangan dimana sistem hanya menggunakan sebagian dari formulir C1 sebagai *region of interest* sehingga jika terdapat inkonsistensi pada *region of interest* tersebut, model akan mengalami kesulitan untuk memprediksi. Sistem ini juga terbukti dapat mengenali digit dengan cukup baik pada gambar dengan kualitas buruk, seperti sudut foto yang kurang baik, jarak foto yang terlalu jauh, dan foto yang tergolong gelap. Selain itu, sistem ini juga memiliki performa yang lebih baik dibandingkan dengan Sirekap. Model ini dapat memberikan fleksibilitas yang lebih baik pada berbagai kondisi dibandingkan dengan sistem Sirekap. Sistem ini diharapkan dapat membantu proses perhitungan rekapitulasi suara tiap TPS dengan lebih baik dan efisien.

V. UCAPAN TERIMA KASIH

Tim peneliti mengucapkan banyak terima kasih kepada Tuhan Yang Maha Esa, serta yang terhormat para pemimpin Universitas Indonesia, pemimpin Fakultas Ilmu Komputer, dan Dosen Pendamping kami atas bantuan dan dukungannya untuk menyelesaikan karya tulis ilmiah kami, diantaranya.

- 1) Badrul Munir, Ph.D. (Direktur Kemahasiswaan Universitas Indonesia)
- 2) Muhammad Hafizhuddin Hilman, Ph.D. (Manajer Kemahasiswaan dan Hubungan Alumni Fakultas Ilmu Komputer)
- 3) Alfian Farizki Wicaksono, Ph.D. (Koordinator Program Studi Sarjana Ilmu Komputer dan Kelas Internasional)
- 4) Adila Alfa Krisnadhi, Ph.D. (Dosen Pendamping)

REFERENSI

- [1] "Election political science," accessed: 2024-05-14. [Online]. Available: <https://www.britannica.com/topic/election-political-science>
- [2] "Details on specific regulatory document," accessed: 2024-05-14. [Online]. Available: <https://mkri.id/index.php?page=web.Berita&id=20196>
- [3] KPU, "Indonesia: Sirekap app helps to monitor election recapitulation," *Asia News Monitor*, Jan 2024, accessed: 2024-05-14. [Online]. Available: <https://www.proquest.com/newspapers/indonesia-sirekap-app-helps-monitor-election/docview/2916439354/se-2>
- [4] U. Argawati, "Pakar it ungkap tiga sumber masalah sirekap," April 2024, accessed: 2024-06-28. [Online]. Available: <https://mkri.id/index.php?page=web.Berita&id=20196>
- [5] "Kpu koreksi publikasi: Anies dapat 3 juta suara di 1 tps, tepis isu turunkan," accessed: 2024-05-14. [Online]. Available: <https://news.detik.com/pemilu/d-7196395/kpu-koreksi-publikasi-anies-dapat-3-juta-suara-di-1-tps-tepis-isu-turunkan>
- [6] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, p. 1680-1716, Nov. 2023. [Online]. Available: <http://dx.doi.org/10.3390/make5040083>

- [7] C. Li, W. Liu, R. Guo, X. Yin, K. Jiang, Y. Du, Y. Du, L. Zhu, B. Lai, X. Hu, D. Yu, and Y. Ma, "Pp-ocrv3: More attempts for the improvement of ultra lightweight ocr system," 2022. [Online]. Available: <https://arxiv.org/abs/2206.03001>