

Terbit online pada laman web jurnal: <http://jurnal.iaii.or.id>**JURNAL RESTI****(Rekayasa Sistem dan Teknologi Informasi)**

Vol. 5 No. 6 (2021) 1153 – 1160

ISSN Media Elektronik: 2580-0760

**Penerapan Convolutional Neural Networks untuk Mesin Penerjemah  
Bahasa Daerah Minangkabau Berbasis Gambar**Mayanda Mega Santoni<sup>1</sup>, Nurul Chamidah<sup>2</sup>, Desta Sandya Prasvita<sup>3</sup>, Helena Nurramdhani Irmanda<sup>4</sup>, Ria  
Astriratma<sup>5</sup>, Reza Amarta Prayoga<sup>6</sup><sup>1,2,3</sup>Program Studi Informatika, Fakultas Ilmu Komputer, UPN Veteran Jakarta<sup>4,5</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jakarta<sup>6</sup>Pusbanglin BPP Bahasa, Kementerian Pendidikan dan Kebudayaan<sup>1</sup>megasantoni@upnvj.ac.id, <sup>2</sup>nurul.chamidah@upnvj.ac.id, <sup>3</sup>desta.sandya@upnvj.ac.id, <sup>4</sup>helenairmanda@upnvj.ac.id,<sup>5</sup>astriratma@upnvj.ac.id, <sup>6</sup>reza.amarta@kemdikbud.go.id**Abstract**

One of efforts by the Indonesian people to defend the country is to preserve and to maintain the regional languages. The current era of modernity makes the regional language image become old-fashioned, so that most of them are no longer spoken. If it is ignored, then there will be a cultural identity crisis that causes regional languages to be vulnerable to extinction. Technological developments can be used as a way to preserve regional languages. Digital image-based artificial intelligence technology using machine learning methods such as machine translation can be used to answer the problems. This research will use Deep Learning method, namely Convolutional Neural Networks (CNN). Data of this research were 1300 alphabetic images, 5000 text images and 200 vocabularies of Minangkabau regional language. Alphabetic image data is used for the formation of the CNN classification model. This model is used for text image recognition, the results of which will be translated into regional languages. The accuracy of the CNN model is 98.97%, while the accuracy for text image recognition (OCR) is 50.72%. This low accuracy is due to the failure of segmentation on the letters i and j. However, the translation accuracy increases after the implementation of the Leveinstan Distance algorithm which can correct text classification errors, with an accuracy value of 75.78%. Therefore, this research has succeeded in implementing the Convolutional Neural Networks (CNN) method in identifying text in text images and the Leveinstan Distance method in translating Indonesian text into regional language texts.

**Keywords:** Convolutional Neural Networks, Translation, Indonesia Language, Local Language Minangkabau, Optical Character Recognition (OCR)

**Abstrak**

Salah satu upaya bela negara yang dapat dilakukan oleh masyarakat Indonesia yaitu melestarikan dan memertahankan bahasa daerah. Kemajuan zaman membuat penggunaan bahasa daerah dianggap kuno dan ketinggalan zaman, sehingga kebanyakan bahasa daerah tidak dituturkan oleh penuturnya lagi. Jika hal tersebut diabaikan, maka akan terjadi krisis identitas budaya yang menyebabkan bahasa daerah rawan punah. Perkembangan teknologi dapat dimanfaatkan sebagai upaya melestarikan bahasa daerah. Teknologi kecerdasan buatan berbasis citra digital menggunakan metode *machine learning* seperti mesin penerjemah, dapat dimanfaatkan untuk menjawab permasalahan tersebut. Pada penelitian ini akan menggunakan metode *Deep Learning*, yakni *Convolutional Neural Networks* (CNN). Data penelitian yang digunakan sebanyak 1300 citra alfabet, 5000 citra teks dan 200 kosa kata dari bahasa daerah Minangkabau. Data citra alfabet digunakan untuk pembentukan model klasifikasi CNN. Model ini digunakan untuk pengenalan citra teks yang hasilnya akan dilakukan penerjemahan ke bahasa daerah. Akurasi model CNN diperoleh sebesar 98.97%, sedangkan akurasi untuk pengenalan citra teks (*Optical Character Recognition* - OCR) sebesar 50.72%. Akurasi ini rendah disebabkan karena terdapatnya kegagalan segmentasi pada huruf i dan j. Namun akurasi terjemahan meningkat setelah diterapkannya algoritma *Leveinstan Distance* yang dapat memperbaiki kesalahan klasifikasi teks, dengan nilai akurasi sebesar 75.78%. Oleh karena itu pada penelitian ini telah berhasil mengimplementasikan metode *Convolutional Neural Networks* (CNN) dalam mengidentifikasi teks pada citra teks dan metode *Leveinstan Distance* dalam menerjemahkan teks bahasa Indonesia ke dalam teks bahasa daerah.

**Kata kunci:** Convolutional Neural Networks, Penerjemahan, Bahasa Indonesia, Bahasa Daerah Minangkabau, Optical Character Recognition (OCR)

## 1. Pendahuluan

Indonesia dikenal dunia sebagai negara dengan masyarakat majemuk, yakni masyarakat yang memiliki keanekaragaman suku, budaya, bahasa, dan agama. Setiap suku memiliki bahasa sendiri yang biasa disebut sebagai bahasa daerah. Bahasa daerah merupakan salah satu unsur penting dari keberlanjutan lestariannya suatu suku. Jika bahasa daerah terus dilestarikan maka transmisi bahasa daerah akan kekal diwariskan ke generasi-generasi berikutnya. Berdasarkan data yang dikeluarkan oleh Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan, Indonesia memiliki 718 bahasa daerah yang tersebar dari Sabang sampai Merauke [1]. Bahkan *National Geographic* memprediksi dari sekitar 7000 bahasa ibu yang ada saat ini akan punah pada akhir abad ini [2].

Salah satu bahasa daerah yang ada di Indonesia adalah bahasa Minang. Bahasa minang merupakan bahasa daerah dari suku Minangkabau yang berakar kuat dari daerah Sumatera Barat. Kelestarian bahasa Minangkabau sangat ditentukan oleh sikap penuturnya [3]. Proyeksi dari temuannya kedepan bahasa Minangkabau secara perjalanan waktu mengalami pelapukan. Hal ini disebabkan terjadinya pergeseran penutur khususnya di perkotaan khususnya kalangan remaja dan keluarga muda cenderung meninggalkan penggunaan Bahasa Minangkabau dalam komunikasi sehari-hari. Bahasa Indonesia dan Bahasa Asing terutama Bahasa Inggris menjadi penggeropos bahasa Minangkabau karena terinterferensi dari bahasa Indonesia dan bahasa asing (Inggris). Jika proyeksi ini senyatanya terjadi, maka tidak dapat dihindarkan lagi bahasa Minangkabau di ranah Minang akan mengalami kekaburan identitas bahasa di kalangan remaja dan keluarga muda. Langkah yang tepat perlu diambil untuk mendokumentasikan dan mengalihwahkan bahasa Minangkabau dengan bantuan teknologi sangat berperan penting.

Di era teknologi 4.0 saat ini memengaruhi perubahan gaya hidup dan gaya berkomunikasi masyarakat. Penggunaan bahasa daerah dianggap sebagai gaya komunikasi yang tertinggal, kolot, dan kampungan, sehingga generasi milenial saat ini banyak meninggalkan dan tidak peduli terhadap keberlangsungan bahasa daerah. Menurut temuan Prayoga dan Khatimah [4], ketidakpedulian ini berasal dari keterniatan masyarakat Indonesia sendiri yang menganggap dan meletakkan bahasa daerah sebagai pilihan yang tidak rasional dalam menghadapi era kompetisi global. Maka dari itu, penetrasi bahasa asing menjadi ancaman terhadap eksistensi bahasa daerah, karena tidak lagi menjadi pilihan utama dalam berbagai ranah baik Pendidikan, Ekonomi, dan Pemerintahan. Generasi milenial lebih pragmatis yakni memilih untuk berbahasa asing agar mereka dapat diterima di

lingkungannya. Tetapi, penggunaan bahasa tersebut tidak lah salah, disisi lain bahasa daerah harus tetap dipelajari agar tidak terjadi krisis identitas budaya dan penutur jati yang dapat menyebabkan bahasa daerah rawan punah.

Salah satu perkembangan teknologi di era industri 4.0 adalah kecerdasan buatan. Teknologi kecerdasan buatan berbasis citra digital menggunakan metode machine learning memungkinkan untuk menjawab permasalahan di atas. Kecerdasan buatan merupakan sebuah teknologi yang melatih komputer agar dapat berpikir seperti layaknya manusia. Salah satu produk kecerdasan buatan yang dapat kita manfaatkan untuk melestarikan bahasa daerah adalah mesin penerjemah [5].

Pada umumnya, teks merupakan data input yang digunakan pada mesin penerjemah. Selain teks, mesin penerjemah juga bisa menerima inputan berupa gambar yang berisikan teks. Teknik mengekstrak teks pada gambar biasa disebut *Optical Character Recognition* (OCR). Mesin penerjemah menggunakan OCR akan menerima inputan dari pengguna berupa citra digital yang memiliki teks Bahasa Indonesia. Selanjutnya mesin penerjemah akan menampilkan terjemahan bahasa daerahnya sesuai dengan bahasa daerah yang pengguna tentukan di awal. Pemilihan data citra digital dimaksud untuk memudahkan pengguna dalam menerjemahkan bahasa tanpa harus mengetikkan kata-kata yang ingin diterjemahkan [6].

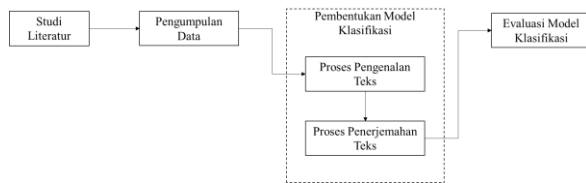
Beberapa penelitian terkait mesin penerjemah Bahasa Indonesia ke Bahasa daerah sudah dilakukan. Penerjemahan Bahasa Indonesia ke Bahasa Daerah Karo [7], Bahasa Indonesia ke Bahasa Melayu dan Bahasa Jawa [8], Bahasa Indonesia ke Bahasa Minangkabau [9], dan Bahasa Indonesia ke Bahasa Sunda [10]. Semua penelitian tersebut masih menggunakan data teks sebagai inputan pada mesin penerjemahnya sehingga mesin penerjemahan dengan memasukkan data teks tidak memudahkan pengguna dalam melakukan penerjemahan. Disisi lain, jika cara tersebut dihadapkan dengan teks yang banyak untuk diterjemahkan, maka akan sangat menyulitkan pengguna. Maka dari itu, penelitian ini mencoba menyempurnakan penelitian sebelumnya dengan lebih memperkuat implementasi basis citra digital dalam penerjemahan Bahasa Indonesia ke Bahasa Daerah. Lebih lanjut, penerjemahan bahasa Indonesia ke Bahasa daerah dengan implementasi basis citra digital diharapkan mampu mempermudah akses pengguna dalam penerjemahan dalam jumlah teks yang banyak. Penerjemahan berbasis citra digital yang sudah dilakukan untuk menerjemahkan Bahasa Indonesia ke Bahasa Inggris atau sebaliknya [6, 11]. Sementara itu untuk Bahasa Indonesia ke Bahasa Daerah telah dilakukan untuk Aksara Lampung [5]. Oleh karena itu, penerjemahan Bahasa berbasis citra digital/gambar masih perlu dilakukan.

Salah satu metode *machine learning* yang saat ini sedang menjadi topik hangat di kalangan peneliti yakni metode *Deep Learning*. *Deep Learning* adalah perkembangan dari metode *Neural Networks* untuk menyelesaikan berbagai permasalahan. Salah satu metode deep learning yang digunakan untuk mengklasifikasikan data gambar yakni *Convolutional Neural Networks* (CNN). Banyak permasalahan klasifikasi gambar berhasil diselesaikan dengan menggunakan metode CNN. Teknik CNN menggunakan *high level feature* yang mana fitur yang ada pada gambar tidak perlu didefinisikan secara spesifik (*handcrafted features*) karena di dalam CNN terdapat lapisan yang disebut *feature extraction layer* [12].

Berdasarkan permasalahan yang ada dan penelitian-penelitian sebelumnya yang telah dibahas, maka penelitian ini akan menerapkan *Convolutional Neural Networks* untuk mengklasifikasikan teks pada citra/gambar untuk diterjemahkan dari Bahasa Indonesia ke Bahasa Daerah Minangkabau. Algoritma *Leveinstan Distance* digunakan untuk penerjemahan Bahasa Indonesia ke Bahasa Daerah Minangkabau.

## 2. Metode Penelitian

Tahapan penelitian meliputi beberapa tahapan, yaitu studi literatur, pengumpulan data, pembentukan model klasifikasi OCR CNN dan evaluasi model klasifikasi. Gambaran umum tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

### 2.1. Studi Literatur

Pada tahapan awal penelitian, Penulis mengumpulkan informasi, landasan teori serta hasil pada penelitian sebelumnya yang relevan terkait dengan penelitian yang akan dilakukan. Tahapan ini sangat penting karena akan memperkaya pengetahuan dan landasan teori mengenai penelitian yang akan dilakukan serta untuk mengetahui *state of the art* penelitian *Convolutional Neural Networks* pada proses pengenalan dan klasifikasi citra khususnya pada citra yang mengandung teks.

Penelitian kecerdasan buatan untuk identifikasi teks pada data citra digital sudah banyak dilakukan. Pada penelitian Naseer dan Zafar [13] melakukan analisis komparatif antara metode CNN dan LSTM dalam mengidentifikasi karakter (OCR – *Optical Character Recognition*) dengan *font* ligatures pada bahasa Urdu yang merupakan bahasa nasional Pakistan dengan

performa sekitar 90% sampai dengan 99.8%. Phangtriasu et.al [14] juga melakukan penelitian OCR untuk membandingkan metode *Neural Networks* dan *Support Vector Machine*. Akurasi tertinggi untuk masing-masing metode yakni *Neural Networks* sebesar 93.48%, sedangkan *Support Vector Machine* sebesar 94.43%. Srivastava et.al [15] juga melakukan identifikasi karakter optik pada dokumen cek Bank menggunakan 2D *Convolutional Neural Networks* dengan performa akurasi sebesar 95.71%. Bhunia et.al [16] juga melakukan penelitian untuk mengidentifikasi tulisan pada data citra pemandangan (*natural scene image*) menggunakan *Convolutional-LSTM Network* yang menghasilkan akurasi tertinggi pada dataset MLe2e sebesar 96.70%. Wang et.al [17] melakukan identifikasi pada CAPTCHA menggunakan *Convolutional Neural Networks* dengan akurasi tertinggi 98.96%.

Berdasarkan penelitian penulis sebelumnya [18], Algoritma *Leveinstan Distance* merupakan algoritma yang paling memberikan akurasi yang paling baik dalam pemrosesan data teks khususnya pada mesin penerjemahan Bahasa Indonesia ke Bahasa daerah Minangkabau. Selain itu, Hossain, et.al [19] menggunakan *Leveinstan Distance* dalam melakukan perbaikan otomatis pada hasil terjemahan Bahasa Bengali ke Bahasa Inggris dengan tingkat akurasi sebesar 78.13%. Penelitian Wint, Ducros, dan Arisugi [20] menggunakan *Leveinstan Distance* untuk melakukan perbaikan ejaan pada dataset sosial media dengan tingkat akurasi 90%. Yulianto, Arifudin, dan Alamsyah [21] juga menggunakan *Leveinstan Distance* untuk melakukan *autocomplete* dan *spell checking* dalam proses pencarian data perpustakaan.

### 2.2. Pengumpulan Data

Data citra yang digunakan pada penelitian ini adalah data citra yang mengandung teks cetak (*printed text*). Data penelitian akan dibagi menjadi tiga kelompok data. Pertama, data citra alfabet/karakter/huruf baik itu huruf kecil dan kapital dengan menggunakan lima jenis *font* berbeda yaitu *Times New Roman*, *Calibri*, *Arial*, *Bodoni*, dan *Helvetica* dan lima jenis *device*/perangkat yang berbeda yakni Scanner, Samsung Galaxy J5, Iphone 4, Iphone 5s, dan Iphone 7. Spesifikasi dari setiap *device* dapat dilihat pada Tabel 1.

Tabel 1. Spesifikasi device/perangkat pengambilan citra

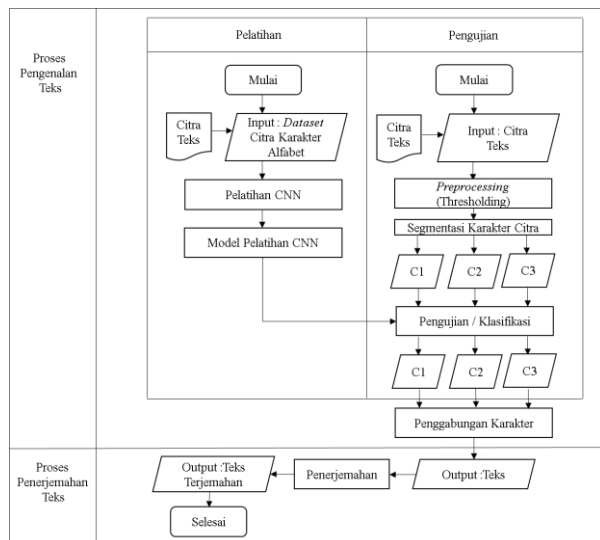
Jenis Perangkat (Kode)	Megapixel/ Scanner Type	Bukaan (f)/ Sensor Type	Ukuran piksel / Optical Resolution
Scanner EPSON (D1)	600 x 1200 dpi	CIS	Flatbed Colour Image Scanner
Samsung Galaxy J5 (D2)	13 MP	f/1.9	1.15 µm
Iphone 4 (D3)	15 MP	f/2.8	1.5 µm
Iphone 5S (D4)	8 MP	f/2.2	1.5 µm
Iphone 7 (D5)	12 MP	f/1.8	1.22 µm

Data citra ini nantinya akan digunakan untuk proses pembentukan model identifikasi karakter dengan menggunakan metode CNN. Kedua, selain data citra alfabet, terdapat juga data citra teks yang terdiri dari kata (bukan kalimat) yang nantinya akan diujikan ke model CNN yang telah dihasilkan sebelumnya. Ketiga, data kamus bahasa Indonesia beserta artinya dalam bahasa daerah Minangkabau.

### 2.3. Pembentukan Model Klasifikasi CNN

Setelah data penelitian diperoleh, selanjutnya dilakukan perancangan metode penelitian. Pada tahapan ini, akan ditentukan constraint penelitian, seperti jumlah data, ukuran gambar dan arsitektur penelitian yang akan digunakan. Setelah itu akan dirancang tahapan implementasi dari metode penelitian yang diusulkan tersebut, sehingga diperoleh hasil pengenalan karakter tulisan print yang terdapat dalam sebuah citra digital. Gambaran umum rancangan metode penelitian dapat dilihat pada Gambar 2.

Proses pengenalan teks bahasa Indonesia pada sebuah citra digital dilakukan melalui dua tahapan. Tahapan pertama adalah tahapan pelatihan (*training*) yang dilakukan untuk mendapatkan model CNN yang dapat mengenali masing-masing karakter alfabet. Data masukan yang digunakan pada tahapan ini adalah data citra karakter alfabet huruf kecil dan kapital dari lima jenis *font* yang berbeda. Selanjutnya data tersebut digunakan untuk pelatihan dengan menggunakan metode CNN, sehingga diperoleh model pelatihan CNN.



Gambar 2. Perancangan Eksperimen Penelitian

Tahapan kedua adalah tahapan pengujian (*testing*) yang dilakukan untuk menguji performa model CNN yang telah diperoleh dari tahapan pelatihan. Pada tahapan ini, data masukan berupa data citra teks yang berisikan kata (bukan kalimat) Bahasa Indonesia. Sebelum dilakukan segmentasi, terlebih dahulu dilakukan praproses data

citra. Hasil segmentasi berupa kumpulan karakter alfabet dari kata teks yang terdapat pada citra sebelumnya. Selanjutnya setiap karakter tersebut, diujikan ke dalam model pelatihan CNN yang diperoleh dari tahapan pelatihan. Hasil identifikasi atau klasifikasi karakter digabungkan kembali untuk membentuk kata. Kata ini nanti akan digunakan untuk diterjemahkan ke dalam bahasa daerah.

Teks yang dihasilkan dari proses pengenalan teks di atas, akan digunakan untuk proses penerjemahan ke dalam bahasa daerah. Proses penerjemahan dilakukan menggunakan algoritma *Levenshtein Distance* dengan cara mengukur jumlah transformasi teks yang diperlukan untuk mengubah teks hasil pengenalan ke teks yang terdapat dalam kamus. Jumlah transformasi terkecil yang akan dipilih sebagai hasil terjemahan.

### 2.4. Evaluasi Model Klasifikasi

Evaluasi dan analisis hasil dilakukan untuk mengetahui pengaruh metode yang diusulkan terhadap hasil akurasi pengenalan karakter teks tercetak pada citra. Hasil yang diperoleh pada penelitian ini akan diujikan apakah identifikasi yang dihasilkan akurat sesuai dengan *groundtruth*. *Groundtruth* merupakan label kelas yang telah diketahui sebelumnya. Pengujian yang akan dilakukan dengan menggunakan nilai akurasi. Terdapat tiga nilai akurasi yang akan diperoleh pada penelitian ini. Pertama adalah nilai akurasi dari metode CNN dalam mengklasifikasikan karakter tercetak pada citra (*Optical Character Recognition*). Nilai evaluasi kedua yakni nilai akurasi klasifikasi kata pada citra teks. Nilai evaluasi ini digunakan untuk mengetahui performa dari model CNN yang diperoleh pada evaluasi sebelumnya untuk mengklasifikasikan kata pada citra teks. Nilai evaluasi ketiga yakni nilai akurasi terjemahan. Nilai evaluasi ini digunakan untuk mengetahui performa dari metode *Levenshtein Distance* dalam menerjemahkan hasil klasifikasi kata pada citra teks.

## 3. Hasil dan Pembahasan

### 3.1. Dataset

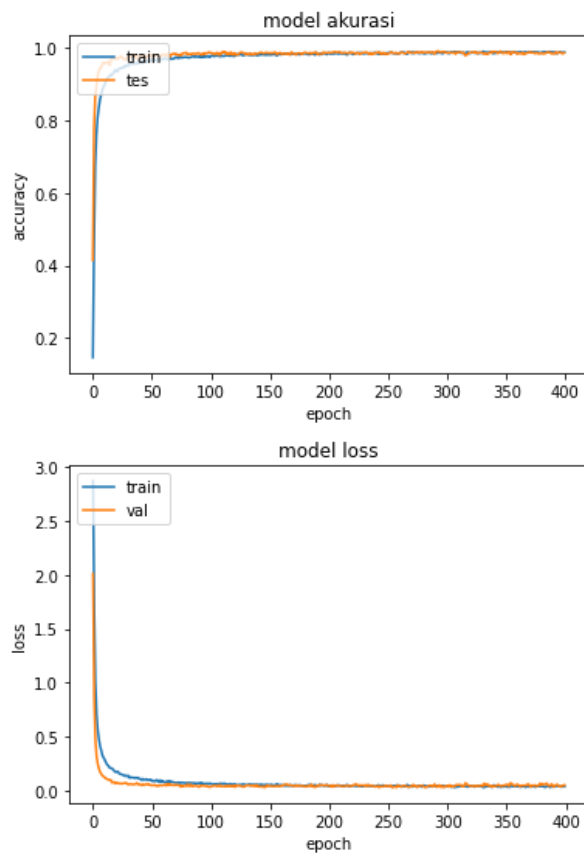
Pada penelitian ini, menggunakan tiga kelompok data. Pertama, citra alfabet yang terdiri dari 26 alfabet baik huruf kecil dan kapital. Data ini digunakan sebagai citra masukan pada pelatihan model CNN dalam melakukan klasifikasi karakter tercetak pada citra (*Optical Character Recognition*). Masing-masing alfabet memiliki 25 citra yang diperoleh dari pengambilan citra pada 5 *device* dan 5 jenis *font*, sehingga total citra pada data ini sebanyak  $5 \text{ device} \times 5 \text{ font} \times 26 \text{ alfabet} \times 2 \text{ (huruf kecil dan huruf kapital)} = 1300$  citra.

Kedua, citra teks yang terdiri dari 200 kata yang berasal dari daftar kata swadesh. Daftar kata ini kami peroleh dari Laboratorium Kebinekaan Bahasa dan Sastra, Kementerian Pendidikan dan Kebudayaan. Masing-

masing kata dilakukan pengambilan gambar pada 5 *device* dan 5 jenis *font*, sehingga total citra pada data ini sebanyak 200 kata x 5 *device* x 5 *font* = 5000 citra. Ketiga, kamus bahasa daerah yakni Bahasa Minangkabau dari 200 kata pada daftar kata swadesh yang diurutkan berdasarkan Bahasa Indonesianya dari huruf A ke Z.

### 3.2. Hasil Eksperimen dan Evaluasi

Pada penelitian ini, terdapat tiga eksperimen yang akan dievaluasi. Pertama adalah performa metode CNN dalam melakukan klasifikasi citra alfabet tercetak (*Optical Character Recognition*). Nilai akurasi terbaik pada eksperimen ini adalah sebesar 98.97%. Nilai *hyperparameter* terbaik yang digunakan pada model yaitu *epoch* sebanyak 200, nilai *batch size* sebesar 50 dan nilai *learning rate* sebesar 0,001.



Gambar 3. Hasil plot model akurasi dan loss function

Gambar 3 menampilkan grafik proses model akurasi dan model loss untuk proses *training* dan *testing*. Pada proses *training* dan *testing* dapat dilihat bahwa dengan bertambahnya epoch, secara umum akurasi juga meningkat, berkorelasi dengan loss yang semakin menurun. Hal ini menunjukkan model dapat belajar dengan baik berdasarkan data *training* dan *testing* yang diberikan.

Nilai akurasi menunjukkan tingkat kebenaran metode CNN dalam melakukan klasifikasi alfabet sangat baik. Nilai akurasi, *precision* dan *recall* tiap alfabet dapat dilihat pada Tabel 2.

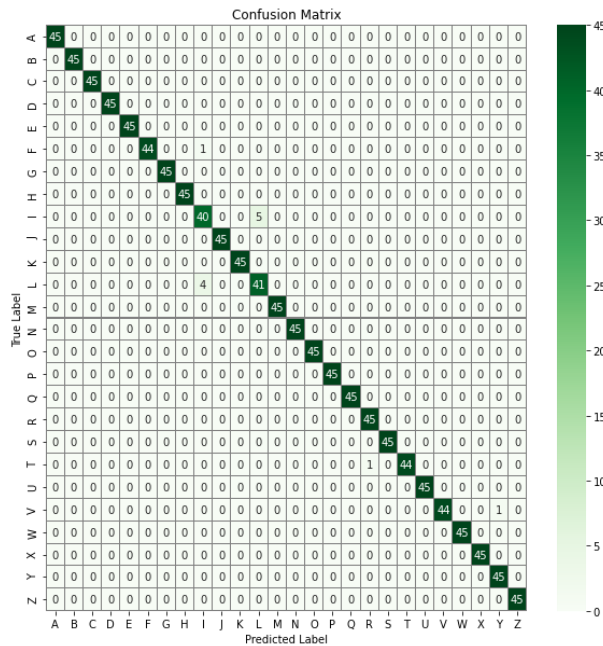
Tabel 2. Nilai akurasi, precision dan recall pada model OCR CNN

Huruf	Akurasi	Precision	Recall
A	1.000	1.000	1.000
B	1.000	1.000	1.000
C	1.000	1.000	1.000
D	1.000	1.000	1.000
E	1.000	1.000	1.000
F	0.978	1.000	0.978
G	1.000	1.000	1.000
H	1.000	1.000	1.000
I	0.888	0.888	0.888
J	1.000	1.000	1.000
K	1.000	1.000	1.000
L	1.000	0.891	0.911
M	1.000	1.000	1.000
N	1.000	1.000	1.000
O	1.000	1.000	1.000
P	1.000	1.000	1.000
Q	1.000	1.000	1.000
R	1.000	0.978	1.000
S	1.000	1.000	1.000
T	0.978	1.000	0.978
U	1.000	1.000	1.000
V	0.978	1.000	0.978
W	1.000	1.000	1.000
X	1.000	1.000	1.000
Y	1.000	0.978	1.000
Z	1.000	1.000	1.000
A	1.000	1.000	1.000
B	1.000	1.000	1.000
C	1.000	1.000	1.000
D	1.000	1.000	1.000
E	1.000	1.000	1.000
F	0.978	1.000	0.978
G	1.000	1.000	1.000
H	1.000	1.000	1.000
I	0.888	0.888	0.888
J	1.000	1.000	1.000
K	1.000	1.000	1.000
L	1.000	0.891	0.911
M	1.000	1.000	1.000
N	1.000	1.000	1.000
O	1.000	1.000	1.000
P	1.000	1.000	1.000
Q	1.000	1.000	1.000
R	1.000	0.978	1.000
S	1.000	1.000	1.000
T	0.978	1.000	0.978
U	1.000	1.000	1.000
V	0.978	1.000	0.978
W	1.000	1.000	1.000
X	1.000	1.000	1.000
Y	1.000	0.978	1.000
Z	1.000	1.000	1.000

Diperoleh hasil bahwa mayoritas kelas dapat diklasifikasi dengan sempurna yang ditunjukkan dengan nilai akurasi, *precision* dan *recall* sama dengan 1.00. Dari 26 alfabet yang ada, huruf i dan l memiliki nilai *precision* yang paling rendah yakni 0.89. Tingkat *precision* menunjukkan tingkat kebenaran prediksi kelas terhadap keseluruhan data yang diprediksi sebagai kelas tersebut.



Dilihat dari *confussion matrix* pada Gambar 4, jumlah data yang terklasifikasikan benar adalah huruf i (*true positive*) sebanyak 40 data, sedangkan jumlah kelas yang terprediksi sebagai huruf i (*false positive*) sebanyak 45 data. Oleh karena itu *precision* dari huruf i adalah  $40/45 = 0.888$ . Hal ini menunjukkan masih terdapat 5 data yang terklasifikasikan salah sebagai kelas I (*false positive*). Begitu juga dengan huruf l, jumlah data yang terklasifikasikan benar adalah huruf l (*true positive*) sebanyak 41 data. Sementara itu, jumlah kelas yang terprediksi sebagai huruf l sebanyak 46 data, sehingga *precision* dari huruf l adalah  $41/46 = 0.891$ .



Gambar 4. Confusion Matrix Model OCR CNN

Huruf i juga memiliki nilai *recall* paling rendah yakni 0.89. Tingkat *recall* menunjukkan tingkat kebenaran prediksi kelas terhadap keseluruhan data yang sebenarnya kelas tersebut. Pada huruf i jumlah data yang terklasifikasi benar adalah huruf i (*true positive*) sebanyak 40 data, sedangkan jumlah data yang sebenarnya kelas huruf i sebanyak 45 data, sehingga nilai *recall* dari huruf I adalah  $40/45 = 0.888$ . Berbeda dengan huruf l, nilai *recall* pada huruf l adalah 0.91, yang mana jumlah data yang terklasifikasikan benar adalah huruf l (*true positive*) sebanyak 41 data dan sedangkan jumlah data yang sebenarnya kelas huruf l sebanyak 45 data. Pada huruf l, nilai *precision* lebih rendah dibandingkan nilai *recall*. Hal ini menunjukkan bahwa masih terdapat data yang sebenarnya bukan huruf l namun diprediksi sebagai huruf l (*false positive*). Sementara itu, untuk huruf f, t dan v memiliki nilai *precision* yang lebih tinggi daripada nilai *recall*. Hal ini menunjukkan bahwa masih terdapat kejadian *false negative*, yakni kesalahan prediksi pada data kelas yang

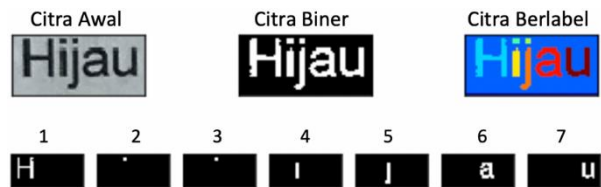
sebenarnya yakni justru diprediksi sebagai data kelas lain.

Eksperimen kedua yang akan dievaluasi adalah hasil klasifikasi kata pada citra teks menggunakan model OCR CNN yang telah dievaluasi sebelumnya. Contoh hasil klasifikasi kata pada citra teks dapat dilihat pada Gambar 5.



Gambar 5. Contoh hasil klasifikasi kata pada citra teks

Dapat dilihat gambar tersebut, masih banyak terjadi kesalahan klasifikasi. Kesalahan sering terjadi diakibatkan kegagalan dalam melakukan segmentasi objek khususnya untuk huruf i dan j. Karena dua huruf ini akan tersegmentasi menjadi 2 objek yang berbeda, yakni huruf i serta tanda titiknya. Contoh kegagalan segmentasi dapat dilihat pada Gambar 6.



Gambar 6. Kesalahan Pelabelan Objek bukan Huruf

Perhitungan akurasi keberhasilan pada proses klasifikasi kata dari masing-masing *device* dan *font* dapat dilihat pada Tabel 3. Dalam tabel akurasi klasifikasi kata ini, dapat dilihat bahwa akurasi tertinggi diperoleh dari data citra teks dengan jenis *font* 1 (*Arial*) dan *font* 4 (*Helvetica*) menggunakan *device* 5 (iPhone 7), yaitu sebesar 58,0%. Selain itu, akurasi terendah diperoleh dari data citra teks dengan jenis *font* 2 (*Bodoni*) dengan menggunakan *device* 3 (iPhone 4), yaitu sebesar 0,5%.

Eksperimen ketiga yang akan dievaluasi adalah hasil penerjemahan klasifikasi kata pada citra teks yang akan diterjemahkan dari Bahasa Indonesia ke Bahasa Daerah menggunakan algoritma *Leveinstan Distance*.

Tabel 3. Akurasi Klasifikasi Kata Berdasarkan Device dan Font

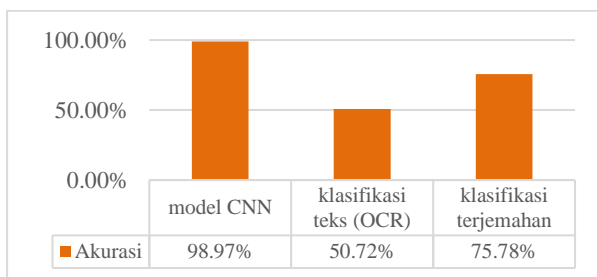
Font/Device	Akurasi (%)				
	D1	D2	D3	D4	D5
F1	85.0	81.0	25.0	82.5	96.0
F2	13.0	5.0	1.0	7.5	17.5
F3	50.5	80.0	4.0	48.0	89.5
F4	85.5	88.5	46.0	85.5	96.0
F5	43.5	45.5	4.5	39.5	48.0

Perhitungan akurasi keberhasilan pada proses penerjemahan dari masing-masing *device* dan *font* dapat dilihat pada Tabel 4. Dalam tabel akurasi penerjemahan ini, dapat dilihat bahwa akurasi mengalami peningkatan. Akurasi penerjemahan tertinggi diperoleh dari data citra teks dengan jenis *font* 1 (*Arial*) dan *font* 4 (*Helvetica*) dengan menggunakan *device* 1 (Scanner), jenis *font* 3 (*Calibri*) dengan *device* 2 (Samsung Galaxy J5) dan jenis *font* 1 (*Arial*) dengan *device* 4 (iPhone 5s) yaitu sebesar 99,5%.

Tabel 4. Akurasi Terjemahan Berdasarkan Device dan Font

Font/Device	Akurasi (%)				
	D1	D2	D3	D4	D5
F1	95.5	95.0	73.0	99.5	98.5
F2	47.5	24.0	10.0	34.0	42.0
F3	97.5	99.5	52.5	95.0	98.0
F4	99.5	98.0	85.5	99.0	98.5
F5	80.0	77.0	42.5	77.5	71.5

Dari perhitungan akurasi di atas, dapat dilihat bahwa hasil akurasi mengalami peningkatan pada saat proses penerjemahan. Rangkuman hasil evaluasi dari semua eksperimen yang telah dilakukan dapat dilihat pada Gambar 7. Akurasi dari model CNN sebesar 98.97%, kemudian mengalami penurunan pada proses klasifikasi kata (OCR) menjadi 50.72%. Penurunan ini terjadi karena banyaknya kesalahan dalam melakukan segmentasi, sehingga mempengaruhi performa model OCR CNN dalam melakukan klasifikasi kata pada citra teks. Namun penurunan akurasi ini, dapat ditingkatkan dengan metode *Leveinstan Distance* pada tahap penerjemahan dengan akurasi terjemahan sebesar 75.78%.



Gambar 7. Grafik Perbandingan Hasil Evaluasi Seluruh Eksperimen

#### 4. Kesimpulan

Penelitian ini telah berhasil mengimplementasikan metode *Convolutional Neural Networks* (CNN) dalam mengidentifikasi teks pada citra teks dan metode *Leveinstan Distance* dalam menerjemahkan teks bahasa

Indonesia ke dalam teks bahasa daerah. Hasil penelitian dievaluasi menggunakan tiga nilai akurasi pada tiga eksperimen. Pertama yakni akurasi model CNN pada eksperimen klasifikasi alfabet tercetak pada citra. Akurasi model CNN untuk keseluruhan alfabet diperoleh sebesar 98.97%, dengan mayoritas alfabet dapat dikenali secara sempurna sebesar 100%, namun terdapat beberapa kelas alfabet seperti huruf “f, i, t, v”.

Kedua yakni akurasi klasifikasi teks (OCR) pada eksperimen klasifikasi kata pada citra teks. Model CNN digunakan untuk mengidentifikasi teks pada citra teks dengan tingkat akurasi klasifikasi teks (OCR) sebesar 50.72%. Pada eksperimen ini masih banyak terjadi kesalahan klasifikasi. Kesalahan sering terjadi diakibatkan kegagalan dalam melakukan segmentasi objek khususnya untuk huruf i dan j. Karena dua huruf ini akan tersegmentasi menjadi 2 objek yang berbeda, yakni huruf i serta tanda titiknya.

Ketiga adalah akurasi terjemahan. Hasil identifikasi teks pada eksperimen kedua, akan digunakan untuk diterjemahkan dari teks Bahasa Indonesia ke Bahasa Daerah menggunakan algoritma *Leveinstan Distance*. Tingkat akurasi pada eksperimen ini 75.78%. Algoritma *Leveinstan Distance* dapat meningkatkan performa terjemahan dengan memperbaiki kesalahan dalam proses identifikasi sebelumnya.

#### Ucapan Terimakasih

Terima kasih kepada Universitas Pembangunan Nasional Veteran Jakarta yang telah mendanai penelitian ini pada skema hibah internal Riset Dosen Pemula tahun 2021.

#### Daftar Rujukan

- [1] KEMDIKBUD, “Bahasa dan Peta Bahasa di Indonesia,” 2019. <https://petabahasa.kemdikbud.go.id/> (accessed Nov. 08, 2021).
- [2] National geographic, “How many of the world’s languages are endangered National Geographic,” 2018. <https://www.nationalgeographic.co.uk/travel/2018/07/how-many-worlds-languages-are-endangered> (accessed Nov. 10, 2021).
- [3] Lindawati, “Bahasa Minangkabau Di Masa Depan Sebuah Proyeksi,” in *International Seminar on Language Maintenance and Shift (LAMAS)* 7, 2017, pp. 348–352.
- [4] P. Reza Amarta and K. Husnul, “Pola Pikir Penggunaan Bahasa Inggris Pada Masyarakat Perkotaan Di Jabodetabek,” *SIMULACRA / Pusat Studi Sosiologi dan Pengembangan Masyarakat*, vol. 2, no. 1, pp. 39–52, 2018.
- [5] M. Agus, S. Erlina, R. Farli, Wajiran, and B. Rohmat Indra, “Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR),” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 7, no. 1, pp. 52–57, 2021.
- [6] W. Antonius Kevin, S. Nanik, and K. Wijayanti Nurul, “Aplikasi Penerjemah Gambar Teks Berbahasa Inggris Menggunakan Teknologi Realitas Tertambah Pada Perangkat Berbasis Android,” 2018.
- [7] A. Ginting and A. Nazori, “Penerjemah Dua Arah Bahasa Indonesia Ke Bahasa Daerah (Karo) Menggunakan Teknik Statistical Machine Translation (Smt) Sebagai Fitur Pada Situs

- Web Untuk Meningkatkan Web Traffic,” Jurnal TELEMATIKA MKOM vol.4 no.1, Maret 2012.
- [8] H. Sujaini, “Peningkatan Akurasi Penerjemah Bahasa Daerah dengan Optimasi Korpus Paralel,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTEI)*, vol. 7, no. 1, pp. 7–12, 2018.
- [9] D. Soyusiawaty, “E-Translator With Rule Based Indonesia-Minang Dan Minang-Indonesia,” *Jurnal Informatika*, vol. 2, no. 2, 2008, [Online]. Available: <http://cimbuak.net>.
- [10] R. Darwis, H. Sujaini, R. Dwi Nyoto, “Peningkatan Mesin Penerjemah Statistik dengan Menambah Kuantitas Korpus Monolingual (Studi Kasus : Bahasa Indonesia-Sunda),” vol. 7, no. 1, 2019.
- [11] U. Anisa Eka, N. Oky Dwi, and M. Kurniawan Teguh, “Aplikasi Penerjemah Bahasa Inggris – Indonesia dengan Optical Character Recognition Berbasis Android,” *Jurnal Teknologi dan Sistem Komputer*, vol. 4, no. 1, pp. 167–177, 2016.
- [12] Y. Jia *et al.*, “Caffe: Convolutional architecture for fast feature embedding,” in *MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia*, Nov. 2014, pp. 675–678. doi: 10.1145/2647868.2654889.
- [13] A. Naseer and K. Zafar, “Comparative Analysis of Raw Images and Meta Feature based Urdu OCR using CNN and LSTM,” 2018. [Online]. Available: [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- [14] M. R. Phangtriatu, J. Harefa, and D. F. Tanoto, “Comparison between Neural Network and Support Vector Machine in Optical Character Recognition,” in *Procedia Computer Science*, 2017, vol. 116, pp. 351–357. doi: 10.1016/j.procs.2017.10.061.
- [15] S. Srivastava, J. Priyadarshini, S. Gopal, S. Gupta, and H. S. Dayal, “Optical character recognition on bank cheques using 2D convolution neural network,” in *Advances in Intelligent Systems and Computing*, vol. 697, Springer Verlag, 2019, pp. 589–596. doi: 10.1007/978-981-13-1822-1\_55.
- [16] A. Kumar Bhunia, A. Konwer, A. Kumar Bhunia, A. Bhowmick, P. P. Roy, and U. Pal, “Script Identification in Natural Scene Image and Video Frame using Attention based Convolutional-LSTM Network.”
- [17] J. Wang, J. Qin, X. Xiang, Y. Tan, and N. Pan, “CAPTCHA recognition based on deep convolutional neural network,” *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 5851–5861, 2019, doi: 10.3934/mbe.2019292.
- [18] M. Mega Santoni, N. Chamidah, D. Sandya Prasvita, R. Amarta Prayoga, and B. Permana Sukma, “Penerjemahan Bahasa Indonesia ke Bahasa Minang dari Optical Character Recognition dengan Menggunakan Algoritme Edit Distance Translating Indonesian into Minang Languages from Optical Character Recognition Using the Edit Distance Algorithm,” *Jurnal Ilmu Komputer Agri-Informatika*, vol. 7, no. 2, pp. 105–113, 2020, [Online]. Available: <http://journal.ipb.ac.id/index>.
- [19] H. Md. Mosabbir, L. Md. Farhan, R. Ahmed Sady, D. Amit Kumar, and Monira Mukta, “Auto-correction of English to Bengali Transliteration System using Levenshtein Distance,” *International Conference on Smart Computing & Communications (ICSCC)*, 2019.
- [20] W. Zar Zar, Th’eo Ducros, and A. Masayoshi, “Spell Corrector to Social Media Datasets in Message Filtering Systems,” 2017.
- [21] M. Maulana Yulianto and R. Arifudin, “Autocomplete and Spell Checking Levenshtein Distance Algorithm to Getting Text Suggest Error Data Searching in Library,” *Scientific Journal of Informatics*, vol. 5, no. 1, pp. 2407–7658, 2018, [Online]. Available: <http://journal.unnes.ac.id/nju/index.php/sji>