

Implementasi Metode Convolutional Neural Network Untuk Identifikasi Pola Aksara Batak

Tumiar Butar – Butar, Berto Nadeak, Imam Saputra

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: tumiarbutar21@gmail.com

Abstrak—Pola aksara batak merupakan tulisan-tulisan bahasa batak yang merupakan bagian penting dalam pengenalan aksara batak. Aksara batak terdiri dari lima dialek besar yaitu dialek Angkola-Mandailing, Karo, Pakpak-Dairi, Simalungun, dan Toba. Tulisan-tulisan yang ada pada aksara batak khususnya aksara batak Toba, tentu sangat sulit untuk dipahami dan dipelajari oleh generasi milenial saat ini. Hal ini disebabkan oleh jaranginya ditemui orang yang ahli dalam bidang aksara batak untuk memberikan pengajaran mengenai aksara batak khususnya aksara batak Toba. Convolutional neural network (CNN) merupakan bagian dari jaringan saraf tiruan yang digunakan sebagai tempat mengolah data dalam bentuk dua dimensi (multy layer perceptron) yakni suara dan gambar atau dengan kata lain sebagai bentuk pengenalan pola. Convolutional neural network (CNN) digunakan pada pola aksara batak sebagai alat untuk pengenalan pola yang akan membantu mengenali pola aksara batak kedalam bentuk komputer dengan menggunakan MatLab sehingga lebih efektif dan efisien untuk dipelajari. Dengan demikian, penggunaan metode Convolutional neural network (CNN) pada aksara batak akan kembali memunculkan rasa keingintauan untuk mempelajari aksara batak toba. Dan juga dengan adanya metode Convolutional neural network (CNN) pada aksara batak akan lebih mempermudah proses pengidentifikasian.

Kata Kunci: Pola Aksara Batak; Convolutional Neural Network (CNN); MatLab

Abstract—The pattern of the Batak script is the writings of the Batak language which is an important part in the introduction of the Batak script. The Batak script consists of five major dialects, namely Angkola-Mandailing, Karo, Pakpak-Dairi, Simalungun, and Toba dialects. The writings in the Batak script, especially the Toba Batak script, are of course very difficult to understand and study by the current millennial generation. This is due to the lack of people who are experts in the field of Batak script to teach about the Batak script, especially the Toba Batak script. Convolutional neural network (CNN) is part of an artificial neural network that is used as a place to process data in two-dimensional form (multi-layer perceptron), namely sound and images or in other words as a form of pattern recognition. Convolutional neural network (CNN) is used in Batak script patterns as a tool for pattern recognition that will help recognize Batak script patterns into computer form using MatLab so that it is more effective and efficient to learn. Thus, the use of the Convolutional neural network (CNN) method in the Batak script will again bring up a sense of curiosity to learn the Batak script, especially the Toba Batak script. And also with the Convolutional Neural Network (CNN) method in Batak script, it will make the identification process easier.

Keywords: Batak Script Pattern; Convolutional Neural Network (CNN); MatLab

1. PENDAHULUAN

Setiap suku daerah yang berada di pulau Indonesia, pasti memiliki warisan budaya tersendiri yang sampai saat ini masih berdiri kokoh. Salah satu warisan budaya yang masih di pegang erat oleh bangsa Indonesia adalah warisan bahasa daerah. Bahasa daerah ini umumnya selalu digunakan oleh nenek moyang di zaman dulu sebagai alat komunikasi yang sah baik itu secara lisan maupun non lisan. Alat komunikasi non lisan ini biasanya digunakan untuk menulis simbol-simbol tertentu yang dijadikan sebagai alat komunikasi.

Pola Aksara Batak merupakan tulisan-tulisan bahasa Batak yang digunakan dalam *Surat Batak*[1]. Pola Aksara Batak atau nama *Surat Batak* yaitu Aksara yang digunakan untuk menulis bahasa Batak yang terbagi dalam lima dialek besar yaitu dialek Angkola-Mandailing, Karo, Pakpak-Dairi, Simalungun, dan Toba. Dalam bahasa Batak,, aksara Batak sering disebut dengan *si sia-sia* atau *surat na sampulu sia* karena aksara (*ina ni surat*) berjumlah Sembilan belas dan tergantung pada daerahnya.

Pada perkembangan zaman sekarang ini, untuk dapat menulis aksara batak yang merupakan salah satu warisan budaya Indonesia, tentunya sangat sulit untuk dipelajari dan bahkan dilakukan. Hal ini tentunya dikarenakan para generasi sekarang sudah banyak meninggalkan warisan budaya yang dinilai kurang *instant* dan beralih ke teknologi yang lebih canggih yaitu komputer. Apalagi aksara batak sekarang hanya ada di museum luar negeri, sehingga hal ini membuat semakin redupnya keingintauan atau niat para generasi muda untuk mengetahui dan mempelajari pola aksara batak tersebut. Salah satu upaya yang dapat dilakukan agar dapat terus melestarikan warisan budaya yang termasuk aksara batak dengan menyesuaikan perkembangan teknologi yang kekinian, yaitu dengan cara memanfaatkan teknologi komputer untuk dapat mengembangkan sistem yang bisa mengenali masukan (*input*) gambar yang berupa tulisan tangan atau bisa disebut dengan *handwriteten character recognition*. Adapun salah satu tahapan penting pengenalan pola aksara batak adalah *Feature extraction*. Tahapan ini ditujukan untuk mengekstraksi ciri khusus dari suatu karakter sehingga dapat dijadikan suatu pembeda dengan karakter yang lain.

Handwriteten character recognition merupakan suatu bagian dari OCR (*Optical Character Recognition*) atau juga termasuk dalam jaringan saraf tiruan (JST) yang mana karakter merupakan *input* yang berasal dari tulisan tangan yang akan dipresentasikan kedalam komputer yang dapat berupa sebuah citra[2]. *Optical Character Recognition* merupakan salah satu sistem komputer yang bisa membaca huruf, teks maupun tulisan tangan manusia. *Optical Character Recognition* ini tentunya berupa teks yang sesuai dengan gambar *output scanner* yang mana tingkat akurasi penerjemahan karakternya tergantung pada kejelasan gambar yang berkaitan serta metode yang digunakan.

Perkembangan teknologi di zaman milenial ini tentunya akan dimanfaatkan untuk membantu para generasi sekarang agar lebih mudah mengakses atau mempelajari pola aksara batak dengan menggunakan komputer sebagai media pengenalannya. Hal ini dilakukan agar para generasi milenial sekarang tidak mudah melupakan warisan budayanya. Nantinya komputer akan membantu mengidentifikasi pola aksara batak tersebut sehingga lebih mudah diidentifikasi dan dipelajari sehingga warisan budaya berupa bahasa daerah termasuk aksara batak dapat terus di budayakan.

Convolutional neural network (CNN) merupakan bagian dari jaringan saraf tiruan yang digunakan sebagai tempat mengolah data dalam bentuk dua dimensi yakni suara dan gambar atau dengan kata lain sebagai bentuk pengenalan pola[3]. *Convolutional neural network* (CNN) digunakan pada pola aksara batak sebagai alat untuk pengenalan pola yang akan membantu mengenali pola aksara batak kedalam bentuk komputer sehingga lebih efektif dan efisien untuk dipelajari.

Penelitian ini bertujuan sebagai pengembangan suatu aplikasi yang mampu mengidentifikasi pola aksara batak dengan memanfaatkan teknologi komputer sebagai pengenalannya, khususnya dalam pengolahan citra untuk membantu dalam pengenalan pola aksara batak dan menulis aksara batak dengan mengimplementasikan metode *Convolutional neural network* (CNN). Dalam penelitian ini, metode *Convolutional neural network* (CNN) difungsikan untuk pengenalan pola aksara batak dengan mengenali ciri-ciri yang membangun pola aksara batak tersebut.

2. METODOLOGI PENELITIAN

2.1 Citra Digital

Kejadian yang menjadi tanda untuk perkembangan pengolahan citra adalah peristiwa pengiriman photo secara digital pada tahun 1921 yang berhasil dilakukan. Peristiwa photo tersebut akan dikirim melalui samudra Atlantik dengan menggunakan kabel bawah laut sebagai medianya dari New York menuju London. Sistem pengiriman ini ditemukan oleh Harry G.Bartholomew dan Maynard D.McFarlane yang dikenal sebagai *Bartlane cable picture transmission system*. Kelebihan menggunakan sistem *Bartlane cable picture transmission system* ini yaitu reduksi waktu yang dilakukan pada saat pengiriman photo yang biasanya memerlukan waktu yang cukup lama (beberapa minggu) dapat lebih efisien yaitu cuma 3 jam saja. Adapun kekurangan system *Bartlane cable picture transmission system* ini adalah terletak dalam proses pengiriman data secara digital dan teknik pencetakan kembali untuk sisi penerima jika ingin mendapatkan satu resolusi gambar yang baik[1].

Dalam sistem ini, foto yang akan dikirim dalam bentuk digital dan pada penerimanya yaitu printer telegraph yang akan diubah kembali menjadi sebuah gambar tercetak. Walaupun dengan menggunakan gambar digital saja, photo belum dapat dianggap sebagai hasil dari pengolahan gambar digital. Hal ini dikarenakan dalam proses pembuatan atau rekayasanya tidak menggunakan komputer sebagai medianya[1].

2.2 Convolutional Neural Network (CNN)

Convolutional neural network (CNN) merupakan bagian dari jaringan saraf tiruan yang digunakan sebagai tempat mengolah data dalam bentuk dua dimensi (*multy layer perceptron*) yakni suara dan gambar atau dengan kata lain sebagai bentuk pengenalan pola[20]. *Convolutional neural network* (CNN) juga bagian dari metode pengklasifikasian dengan *backpropagation* dan *feedward*[20] [21].

Convolutional neural network (CNN) pada perkembangannya memiliki tipe lapisan yaitu, pada lapisan pertama merupakan lapisan klasifikasi. Lapisan klasifikasi merupakan sebuah lapisan yang tersusun dari beberapa lapisan yang mana pada setiap lapisannya terdiri dari *neuron* yang memiliki koneksi penuh antara lapisan satu dengan lapisan lainnya. Sedangkan layer kedua merupakan ekstraksi fitur. Ekstraksi fitur ini terletak pada awal arsitektur yang terdiri atas lebih dari satu lapisan dan setiap lapisannya juga dilapisi dengan *neuron* yang terkoneksi hanya pada daerah local saja. Pada saat data input pada *layer* bisa dilakukan perhitungan dengan menggunakan bobot yang ada. Nantinya hasil dari perhitungan itu ditransformasikan dengan menggunakan perhitungan yang non-linear yang bisa disebut sebagai fungsi aktivasi[20] [22].

Adapun proses-proses yang akan dilakukan oleh *Convolutional neural network* (CNN) adalah sebagai berikut [23]:

1. *Convolutional network*

Adapun rumus yang dapat digunakan pada tahap ini adalah sebagai berikut :

$$Z_i^t = B^1 + \sum_{j=1}^{k^1-1} W_j^1 * Z_j^{1-1}, \text{ dimana } i \in [1, K^1], B^1 \quad (1)$$

2. *Non-linearity layer*

Pada tahap ini akan digunakan rumus yang dapat membantu perhitungan agar lebih mudah, yaitu :

$$a^1 = f(Z^1) \quad (2)$$

3. *Pooling layer*

Poling layer memiliki rumus sebagai berikut :

$$p^1 = \max_{1 \in r} a_i^1 \quad (3)$$

4. *Fully-connected layer*

Pada tahap ini akan dilakukan perubahan data dari data matriks x-dimensi menjadi matriks satu dimensi.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Pola aksara batak merupakan tulisan-tulisan bahasa batak yang merupakan bagian penting dalam pengenalan suatu aksara batak. Aksara batak ini tentu terdiri dari lima dialek besar yaitu dialek Angkola-Mandailing, Karo, Pakpak-Dairi, Simalungun, dan Toba. Tulisan-tulisan yang ada pada aksara batak khususnya aksara batak Toba, tentu sangat sulit untuk dipahami dan dipelajari oleh generasi milenial saat ini. Hal ini tentu disebabkan oleh jaranginya ditemui orang yang ahli dalam bidang aksara batak untuk memberikan pengajaran mengenai aksara batak khususnya aksara batak Toba.

Pada saat ini, untuk dapat mempelajari aksara batak secara langsung tentu sangat sulit dilakukan. Hal ini dikarenakan generasi milenial telah meninggalkan hal yang kurang modern sehingga membuat semakin redupnya keingin tahuan para generasi sekarang untuk dapat mempelajarinya. Adapun salah satu upaya yang dapat dilakukan agar dapat mengembalikan minat untuk mempelajari aksara batak yaitu dengan mengimbangi tingkat modernisasi yang ada pada saat ini.

Saat ini, generasi muda sekarang tentu mengetahui alat komunikasi modern yang sering dipakai diseluruh pelosok dunia yaitu komputer. komputer inilah yang nantinya akan membantu proses pengenalan pola aksara batak dengan mudah dengan memanfaatkan metode yang ada. Metode yang dapat digunakan untuk dapat memproses pola aksara batak dengan mudah yaitu metode *Convolution Neural Network* (CNN) sebagai pembantu pengenalan pola pada aksara batak, sehingga untuk dapat mengidentifikasi proses aksara batak khususnya aksara batak Toba dapat menjadi lebih mudah.

3.1.1. Penerapan *Convolution Neural Network*

Convolutional neural network (CNN) merupakan bagian dari jaringan saraf tiruan yang digunakan sebagai tempat mengolah data dalam bentuk dua dimensi yaitu berupa suara dan gambar atau dengan kata lain sebagai bentuk pengenalan pola yang telah banyak digunakan dalam berbagai pengenalan pola. *Convolutional neural network* (CNN) digunakan pada pola aksara batak adalah sebagai alat untuk pengenalan pola yang akan membantu mengenali pola aksara batak kedalam bentuk komputer sehingga lebih efektif dan efisien untuk dipelajari atau dipahami.

Berikut ini adalah pada penerapan metode *Convolutional neural network* (CNN) pada aksara batak dengan jumlah total aksara batak 18 yang tidak memiliki persamaan antara satu dengan yang lainnya dan yang akan dilakukan pengujian adalah sebesar 5 aksara saja. Pada contoh yang akan diimplementasikan adalah aksara batak toba (a). Adapun bentuk yang dapat membedakan aksara batak satu dengan aksara yang lainnya dapat dilihat dalam tabel berikut ini :

Tabel 1. Perbedaan bentuk pada aksara batak toba

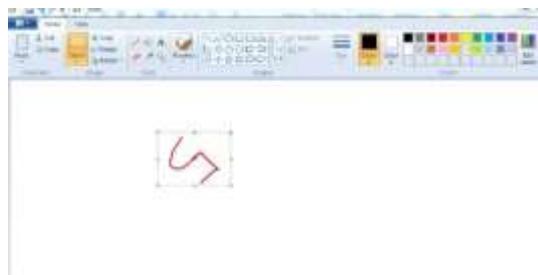
Simbol aksra	Arti	Keterangan ciri
5	A	Berbentuk vertikal
6	Ba	Berbentuk melekung busur
7	Da	Berbentuk diagonal kanan
8	Ga	Berbentuk melekung kubah
9	Ha	Berbentuk gelombang

Sebelum membahas tentang tentang penerapan metode *Convolutional neural network* (CNN), terlebih dahulu lakukan langkah-langkah penelusuran penyelesaiannya adalah sebagai berikut :

1. Lakukan tahap pemusatan objek dengan *cropping* pada citra RGB

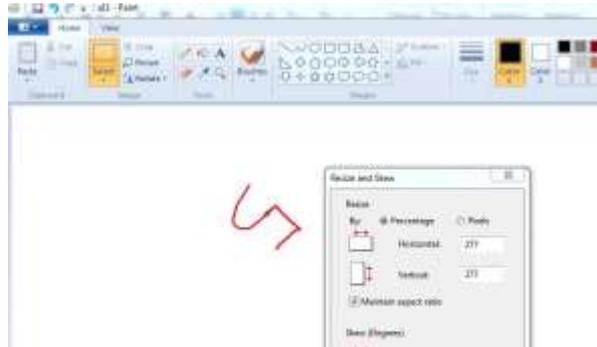


Gambar 1. pola aksara batak (a)



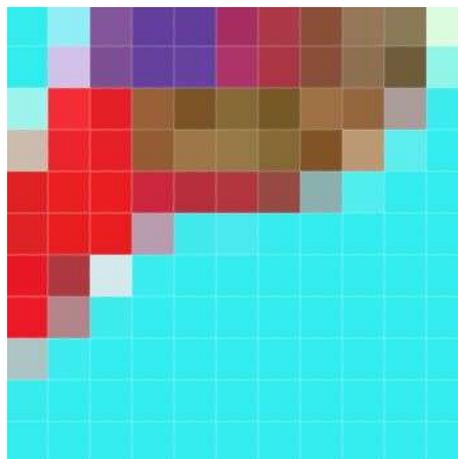
Gambar 2. hasil *cropping* pada pola aksara batak (a)

- Langkah selanjutnya, lakukan proses resize sebesar 227 x 227 piksel. Hal ini dilakukan agar lebih mempermudah proses training dengan metode CNN.



Gambar 3. Hasil Proses Resize Piksel Pola Aksara Batak (a)

- Langkah selanjutnya adalah proses konvolusi yang dilakukan dengan cara memisahkan data set training untuk testing dengan menggunakan Alexnet pada matlab. Maka didapatkan hasil dari proses training tersebut dengan ukuran 11 x 11 x 3. Maksud dari angka 3 adalah nilai RGB (*red, green, and blue*) dengan stride 4 dan padding (lapisan) 0. Adapun proses perhitungannya adalah sebagai berikut :



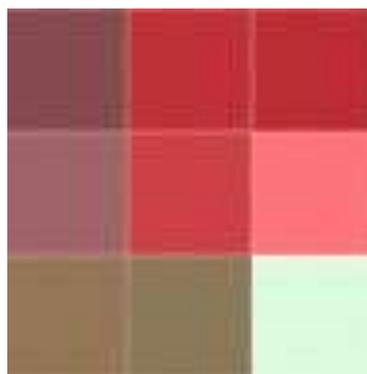
Gambar 4. Piksel Ukuran 11 x 11

$$\text{Konvolusi} = \frac{\text{nilai piksel} + 2 \times \text{padding} - \text{konvolusi}}{\text{stride}}$$

$$\text{Konvolusi} = \frac{528 + 2 \times 0 - 11}{4} + 1$$

$$\text{Konvolusi} = 129$$

- Setelah proses konvolusi selesai, maka lakukan proses *max pooling* ukuran 3 x 3, padding 0 dan dengan stride 2. Maka perhitungannya adalah sebagai berikut :



Gambar 5. piksel 3 x 3

$$\text{Max pooling} = \frac{\text{nilai piksel} + 2 \times \text{padding} - \text{konvolusi}}{\text{stride}}$$

$$\text{Max pooling} = \frac{289 + 2 \times 0 - 3}{2} + 1$$

$$\text{Max pooling} = 145$$

4. KESIMPULAN

Kesimpulan yang dapat diambil dari hasil penelitian implementasi metode *convolutional neural network*, yaitu Proses identifikasi pola aksara batak yaitu dengan memanfaatkan algoritma yang sesuai dengan identifikasi sebuah objek. Proses identifikasi pola aksara batak menggunakan *convolutional neural network* yaitu dengan sesuai dengan tahapan pada metode tersebut, sehingga identifikasi dapat dilakukan dengan mudah dan efektif

REFERENCES

- [1] “ชุมชน ฉายะจินดา,” *No Title ผู้สื่อข่าว*.
- [2] O. N. Shpakov and G. V. Bogomolov, “Technogenic activity of man and local sources of environmental pollution,” *Stud. Environ. Sci.*, vol. 17, no. C, pp. 329–332, 1981, doi: 10.1016/S0166-1116(08)71924-1.
- [3] J. Utama, “Akuisisi Citra Digital menggunakan Pemrograman MATLAB,” *Maj. Unikom*, vol. 9, no. 1, pp. 71–80, 2011.
- [4] Widyatama & Suprpty, “Bab II Landasan Teori,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2018, doi: 10.1017/CBO9781107415324.004.
- [5] A. K. Tarigan, S. D. Nasution, and A. Karim, “Aplikasi Pembelajaran Citra Dengan Menggunakan Metode Computer Assisted Instruction (Cai),” vol. 3, no. 4, pp. 1–4, 2016.
- [6] N.Nafi'iyah-S.Mujilawati, *Buku Ajar Citra Binarisasi Dan Enhancement*. Yogyakarta, 2018.
- [7] Darma Putra, *Pengolahan Citra Digital*. Yogyakarta: ANDI, 2010.
- [8] P. I. Hidayati *et al.*, “No Title.”
- [9] L. De Russis *et al.*, “OpenCV Java Tutorials Documentation,” *Semin. Nas. Teknol. Inf. Komun. Terap. 2011*, vol. 1, no. 1, pp. 187–198, 2011, [Online]. Available: http://repo.pens.ac.id/1324/1/Paper_TA_MBAH.pdf.
- [10] A. B. S and H. Maulana, “Pengenalan Citra Wajah Sebagai Identifier Menggunakan Metode Principal Component Analysis (PCA),” vol. 9, no. 2, pp. 166–175, 2016.
- [11] K. Z. Transform and D. A. N. Bray, “DISTANCE UNTUK PENGENALAN POLA HURUF,” pp. 153–159.
- [12] U. Catur and D. Yudhatama, “IDENTIFIKASI LAHAN TAMBANG TIMAH MENGGUNAKAN METODE KLASIFIKASI TERBIMBING MAXIMUM LIKELIHOOD PADA CITRA LANDSAT 8 (Identification of Tin Mining Area Using Maximum Likelihood Supervised Classification on Landsat 8 Image),” vol. 8, 2015.
- [13] F. Muhlis, *Aplikasi Data Penginderaan Jauh Untuk Pengelolaan Sumber Daya Alam Dan Lingkungan*. Pasuruan, Jawa Timur: Qiara Media, 2020.
- [14] M. Evi, D. I. P. Rovinsi, and J. A. W. A. B. Arat, “E STIMASI P ERUBAHAN L AHAN S AWAH DENGAN K LASIFIKASI T IDAK T ERBIMBING C ITRA,” vol. 11, no. 2, pp. 55–66, 2017.
- [15] F. Hadi Amalia, T. Agung Budi, and K. N. Ramadhani, “Pengenalan Angka Tulisan Tangan Dengan Penerapan Freeman Chain Code yang Dimodifikasi,” *E-Proceeding Eng.*, vol. 2, no. 2, pp. 6101–6109, 2015.
- [16] R. Gunawan, S. Suwarno, and W. Hapsari, “PENERAPAN OPTICAL CHARACTER RECOGNITION (OCR) UNTUK PEMBACAAN METERAN LISTRIK PLN,” vol. 10, no. 2, pp. 127–134, 2014.
- [17] M. E. R. Bezerra and A. L. I. de Oliveira, “Optical character recognition,” *Digit. Doc. Anal. Process.*, vol. 1, no. 1, pp. 207–230, 2013.
- [18] D. H. Tanjung, “Jaringan Saraf Tiruan dengan Backpropagation untuk Memprediksi Penyakit Asma,” *Creat. Inf. Technol. J.*, vol. 2, no. 1, p. 28, 2015, doi: 10.24076/citec.2014v2i1.35.
- [19] E. Hartato, D. Sitorus, and A. Wanto, “Analisis Jaringan Saraf Tiruan Untuk Prediksi Luas Panen Biofarmaka Di Indonesia,” *SemanTIK*, vol. 4, no. 1, pp. 49–56, 2018.
- [20] S. Ilahiyah and A. Nilogiri, “Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network,” *JUSTINDO (Jurnal Sist. dan Teknol. Inf. Indones.)*, vol. 3, no. 2, pp. 49–56, 2018, doi: 10.32528/JUSTINDO.V3I2.2254.
- [21] V. Maha, P. Salawazo, D. Putra, J. Gea, F. Teknologi, and U. P. Indonesia, “Implementasi Metode Convolutional Neural Network (CNN) Pada Peneganaan Objek Video Cctv,” *J. Mantik Penusa*, vol. 3, no. 1, pp. 74–79, 2019.
- [22] G. Wicaksono, S. Andryana, and B. -, “Aplikasi Pendeteksi Penyakit Pada Daun Tanaman Apel Dengan Metode Convolutional Neural Network,” *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 1, p. 9, 2020, doi: 10.31328/jointecs.v5i1.1221.
- [23] M. T. Stefanus Christian Adi Pradhana, Untari Novia Wisesty S.T.,M.T., Febryanthi Stehvanie S.T., “Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network,” *e-Proceeding Eng.*, vol. 7, no. 1, pp. 2558–2567, 2020.
- [24] J. Sitorus, “Perancangan Aplikasi Pengenalan Pola Huruf Aksara Batak Toba Menerapkan Metode Direction Feature Extraction (Dfe),” *JURIKOM (Jurnal Ris. Komputer)*, vol. 2, no. 6, pp. 48–55, 2015.
- [25] C. Nasoichah, “Aksara Batak Dalam Kebhinnekaan Nusantara,” *Kebudayaan*, vol. 11, no. 1, p. 1, 2018, doi: 10.24832/jk.v11i1.13.
- [26] R. A.S-M.Salahuddin, *rekayasa perangkat lunak dan implementasinya*. Yogyakarta: Informatika, 2011.
- [27] A.Nugroho, *Rekayasa Perangkat Lunak Menggunakan UML dan Java*. Yogyakarta: ANDI, 2009.
- [28] J. HM, *Rekayasa Perangkat Lunak dan Implementasinya*. Yogyakarta: ANDI, 2005.
- [29] J. Ilmiah and F. Exacta, “Identifikasi objek berdasarkan citra warna menggunakan matlab,” vol. 4, no. 2, pp. 181–190, 2011.
- [30] B. Cahyono, “Penggunaan Software Matrix Laboratory (Matlab) Dalam Pembelajaran Aljabar Linier,” *Phenom. J. Pendidik. MIPA*, vol. 3, no. 1, p. 45, 2016, doi: 10.21580/phen.2013.3.1.174.
- [31] F. Alfian, “Pengaruh Perbedaan Kernel Edge Detection Kirsch Pada Sketching Citra Digital Dengan Bahasa Pemrograman Matlab,” *J. Teknol. Inf. RESPATI*, vol. XIII, no. November, pp. 1–8, 2017.