

# Pengenalan Pola Tulisan Tangan Aksara Bima menggunakan Ciri Tekstur dan KNN

*(Handwriting Recognition of Bima Script using Texture Features and KNN)*

Fitri Bimantoro\*, Arik Aranta, Gibran Satya Nugraha, Ramaditia Dwiyanaputra, Ario Yudo Husodo

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Mataram

Jl. Majapahit 62, Mataram, Lombok NTB, Indonesia

Email: [bimo, arikaranta, gibransn, rama, ario]@unram.ac.id

*\*Penulis Korespondensi*

**Abstract** As the fact, that is Bima script did not familiar to bimanese, Bima script as a cultural heritage needs to be preserved. Pattern recognition has been used to recognize several of ancient script. Gray Level Co-occurrence Matrix (GLCM) as features extraction and K Nearest Neighbour (KNN) as a classifier show the good performance to recognize of an ancient script. so, in this research we use GLCM and KNN to recognize Bima script. We use 2640 images of handwriting bima Script that is collected from 10 volunteers. Each volunteer write 22 of Bima script twelve times each script. The experimental result show that the performance of our model is good enough, with 60.86% of accuracy that is obtained by manhattan distances.

**Key words:** Bima Script, handwriting, GLCM, KNN, recognition.

## I. PENDAHULUAN

Indonesia sebagai negara kepulauan terbesar, memiliki lebih dari 17 ribu pulau yang tersebar dari Sabang hingga Merauke[1]. Jumlah pulau yang tersebar tersebut tentu memberikan keragaman yang cukup tinggi di Indonesia, mulai dari suku, Bahasa dan juga Aksara. Nusa Tenggara Barat (NTB) salah satu provinsi yang terdiri dari dua pulau besar memiliki tiga suku utama didalamnya yakni Sasak, Samawa, dan Mbojo (Bima). Dari ketiga suku ini, suku Bima menjadi suku terbesar kedua setelah suku Sasak. Sayangnya, hanya sedikit masyarakat Bima yang mengerti Aksara Bima.

Aksara dan bahasa merupakan salah satu identitas suatu bangsa. Menurut situs endangeredlanguages.net dan endangeredalphabets.net yang menghimpun daftar bahasa dan aksara di seluruh dunia, menyatakan kondisi aksara dan bahasa tersebut dalam dua kategori, yaitu menurun dan muncul/berkembang. Salah satu penyebab aksara dan bahasa dalam kategori menurun adalah karena mulai berkurangnya eksistensi aksara dan bahasa dalam kehidupan sehari[2].

Nusa Tenggara Barat (NTB) salah satu provinsi yang terdiri dari dua pulau besar memiliki tiga suku utama didalamnya yakni Sasak, Samawa, dan Mbojo (Bima). Dari ketiga suku ini, suku Bima menjadi suku terbesar kedua setelah suku Sasak[3]. Namun, berdasarkan survei yang dilakukan terhadap sejumlah responden yang berasal dari Bima, 66,7% menyatakan tidak familiar dengan

aksara bima, dan terdapat 45,7% yang tidak mengetahui keberadaan aksara Bima. Dan bukan tidak mungkin suatu saat nanti masyarakat Bima tidak mengenal aksaranya sendiri apabila tidak ada usaha dalam pelestariannya.

Aksara Bima merupakan salah satu aksara yang sempat hilang keberadaannya, dan setelah penelitian panjang akhirnya Hj. Siti Maryam mendeklarasikannya pada tanggal 28 Juli 2007. Namun, upaya pelestarian ini baru memiliki kekuatan hukum ketika Walikota Bima menuangkannya dalam peraturan Walikota dengan nomor 50 tahun 2019, yang menetapkan mata pelajaran bahasa, sejarah, seni budaya dan keterampilan sebagai mata pelajaran muatan lokal untuk sekolah dasar dan sekolah menengah pertama di kota Bima[4]. Walau demikian, jumlah dan kemampuan tenaga pengajar yang masih sangat terbatas tentunya menjadi salah satu tantangan dalam upaya lanjutan dalam melestarikan aksara Bima. Tercatat pada tahun 2018 hanya 50 Guru SD dan SMP yang mendapat Bimtek Aksara Bima[5]. Tentu saja hal ini tidak akan cukup untuk memenuhi kebutuhan tenaga pengajar dalam upaya melestarikan aksara Bima.

Teknologi informasi saat ini memungkinkan untuk digunakan sebagai salah satu cara untuk mendukung upaya pelestarian aksara. Beberapa penelitian yang memanfaatkan teknologi informasi menerapkan konsep sistem cerdas pada aksara antara lain Bima [6], [7], Jawa [8], [9], Lampung [10], Bali [11], Sasak [12], Thailand [13], Arab [14]–[16], Bangla [13], [17], Devanagiri [13], [17]–[19], Oriya [17], dan Telugu [17], [18], [20]. Penelitian - penelitian tersebut memanfaatkan sistem cerdas untuk mengenali karakter setiap aksara.

Penelitian tentang aksara Bima yang sudah dilakukan yakni tentang transliterasi latin ke aksara Bima dan Pengenalan pola aksara Bima. Pengenalan pola aksara Bima yang pernah dilakukan menggunakan data yang berasal dari font komputer yang diaugmentasi dan kemudian diklasifikasi menggunakan SVM. Dengan data yang ada, hasil yang diperoleh menunjukkan performa yang baik[6].

Sebelum data diklasifikasi, perlu dilakukan ekstraksi ciri. Ciri yang dapat diekstrak dari sebuah citra berupa ciri warna, bentuk, dan tekstur. Salah satu metode yang cukup handal dalam melakukan ekstraksi ciri tekstur adalah *Gray*

*Level Co-occurrence Matrix* (GLCM). GLCM beberapa kali digunakan dalam mengekstraksi ciri pada tulisan tangan aksara, dan memberikan performa yang cukup baik yakni berkisar antara 80 sampai 95% [18], [21].

*K-Nearest Neighbour* (KNN) merupakan salah satu metode yang digunakan untuk melakukan klasifikasi dengan memanfaatkan kedekatan ciri [22]. Salah satu keunggulan KNN yakni mampu menunjukkan performa yang baik bahkan ketika data memiliki banyak derau [23]. Penggunaan KNN pada pengenalan tulisan tangan memberikan hasil yang baik, dimana pada pengenalan pola tulisan tangan aksara Jawa menghasilkan akurasi sebesar 82,5% [24], dan pada tulisan tangan aksara Bali sebesar 74,6% [23]. Percobaan lain menggunakan tulisan tangan angka yakni dataset MNIST (*Modified National Institute of Standards and Technology*) diperoleh hasil 89.81% [22]. Selain itu pada topik penelitian yang lain KNN digunakan untuk mendiagnosa penyakit ginjal [25] dan juga untuk mengidentifikasi protein [26]. Hasil yang diberikan mampu bersaing dengan mesin pembelajaran yang lainnya, bahkan lebih baik dari pada *Convolutional Neural Network* (CNN) [26].

Untuk itu pada penelitian ini akan dilakukan pengenalan pola tulisan tangan aksara Bima menggunakan GLCM dan KNN karena berdasarkan penelitian yang telah dilakukan sebelumnya, baik GLCM ataupun KNN menunjukkan performa yang baik. Pada bagian selanjutnya akan dijelaskan tentang tinjauan pustaka yang berisi penelitian terkait, teori tentang GLCM dan KNN. Setelah itu dibahas metodologi penelitian dan dilanjutkan dengan hasil dan pembahasan dari pengujian yang telah dilakukan. Dibagian akhir akan disampaikan kesimpulan dari penelitian yang telah dilakukan.

## II. TINJAUAN PUSTAKA

### A. Penelitian Terkait

Penelitian tentang pengenalan aksara cukup menjadi perhatian dalam perkembangan teknologi informasi. Sebelumnya, beberapa penelitian terkait aksara telah dilakukan mulai baik itu aksara yang ada di Indonesia ataupun luar negeri. Penelitian-penelitian tersebut banyak yang berfokus pada bagaimana melakukan ekstraksi ciri hingga mencari tahu performa beberapa metode klasifikasi.

Pengenalan pola aksara Bima pertama kali dilakukan dengan memanfaatkan SVM untuk melakukan klasifikasi aksara. Hasil yang diperoleh menunjukkan performa yang baik yakni akurasi sebesar 90%. Namun, data yang digunakan pada penelitian ini belum menggunakan data tulisan tangan, namun menggunakan data font yang tersedia pada Ms. Word [6]. Walaupun mendapatkan akurasi yang tinggi, variasi data yang digunakan sedikit, sehingga ada kemungkinan model yang dibuat tidak akan baik bila menggunakan data tulisan tangan.

Ekstraksi ciri menggunakan metode zoning pernah dilakukan untuk melakukan pengenalan pola aksara Bali. Sebelum proses ekstraksi ciri arah dan ciri semantik, citra aksara Bali di-zoning kedalam 4 bagian. selanjutnya dari masing-masing zona akan diekstraksi ciri sehingga

didapatkan ciri arah sebanyak 16 dan ciri semantik sebanyak 12 jenis ciri [11]. Penggunaan *integral projection* sebagai metode ekstraksi ciri pada tulisan tangan aksara sasak dan jaringan syaraf tiruan dilakukan. berdasarkan hasil yang didapatkan, penggunaan Jaringan syaraf tiruan terbukti dapat diaplikasikan pada permasalahan ini, walaupun hasil yang didapatkan masih perlu dikembangkan lebih lanjut. tahapan yang dilakukan pada penelitian ini dengan melakukan ekstraksi ciri menggunakan *integral projection* yang menghasilkan 128 ciri. 128 ciri ini kemudian akan direduksi menggunakan PCA sehingga akan tersisa 22 ciri yang selanjutnya menjadi input jaringan syaraf tiruan [12]. Metode lain yang digunakan untuk ekstraksi ciri yakni *histogram of Oriented Gradient* (HOG). Aksara Jawa yang telah didijitalisasi diubah menjadi citra biner, lalu ditentukan *Region of Interest* (ROI) menggunakan *connected component labelling*. setelah didapatkan ROI selanjutnya dengan menggunakan ukuran cell 32x32 piksel, didapatkan ciri histogram sebanyak 9 untuk masing-masing citra. penggunaan SVM dengan RBF kernel menghasilkan klasifikasi mencapai nilai 93.3% [8].

Bharathi melakukan pengujian 9 ciri GLCM yaitu *contrast, dissimilarity, entropy, sum of square variance, sum of average, sum variance, sum of entropy, difference variance* dan *difference entropy* pada *Support Vector Machine* dengan data tulisan tangan. Pada tahap preproses, Bharati melakukan proses binarisasi, *morphological Filtering*, segmentasi, dan penyeragaman ukuran. 9 ciri GLCM dicari menggunakan 4 sudut, sehingga didapat total ciri yang akan digunakan adalah 36 ciri. berdasarkan pengujian, hasil yang didapatkan menunjukkan kemampuan pengenalan pola yang cukup baik yaitu berkisar antara 80.76% sampai dengan 95.43% [21]

Singla pada tahun 2014 menulis tentang penggunaan 4 ciri GLCM meliputi *contrast correlation, Energy* dan *Homogeneity* pada Jaringan syaraf tiruan untuk mengenali pola aksara Devanagari. Pengujiannya memanfaatkan 10-fold validation, dengan pembagian dari 30 data 20:10 untuk data latih berbanding data uji mendapatkan hasil yang baik [19]. Selain Singla, penelitian tentang Aksara Devanagari dilakukan dengan memanfaatkan ciri HOG. Ciri yang berhasil diekstrak dengan menggunakan HOG akan menjadi masukan metode klasifikasi *Deep-Learning Feedforward-Backpropagation Neural Network* (DFBNN) dan *Extreme Learning Machine* (ELM). Hasil menunjukkan bahwa performa DFBNN mengungguli ELM, bahkan ketika data aksara lain yang digunakan yaitu aksara Thailand dan aksara Bangla [13]. Ismail menunjukkan penggunaan Jaringan Syaraf Tiruan untuk mendeteksi tulisan Arab. hasil menunjukkan, dibandingkan *decision tree*, jaringan syaraf tiruan menunjukkan performa yang lebih baik dengan rata-rata akurasi yang didapat 96.7% [14].

Pada pengenalan pola tulisan arab, penggunaan *probabilistic neural network* menghasilkan performa yang cukup baik yaitu 90%. ciri yang digunakan pada penelitian ini dengan memanfaatkan intensitas dari citra biner.

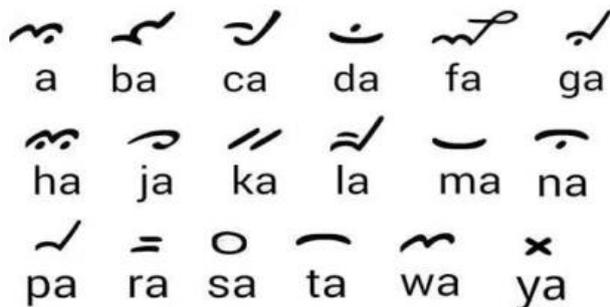
pengujian dilakukan dengan menggunakan 100 citra dan memanfaatkan k-fold validation untuk pengujiannya[16]. Penggunaan Jaringan syaraf tiruan pun dilakukan pada pengenalan pola aksara Lampung. dengan memanfaatkan zoning, 64 ciri hasil ekstraksi selanjutnya digunakan sebagai input Jaringan syaraf tiruan. berdasarkan pengujian yang dilakukan, akurasi yang diperoleh adalah 90,8%[10].

Pada penelitian yang menggunakan KNN sebagai metode klasifikasinya, memberikan hasil yang baik. Pada pengujian data MNIST dengan menggunakan LBP sebagai cirinya, akurasi tertinggi didapat sebesar 89,65% [22]. Pada penelitian yang lainnya yang menggunakan aksara Bali, menunjukkan bahwa KNN memiliki performa yang lebih baik dibandingkan SVM [23]. Pada penggunaannya pada aksara Jawa menunjukan akurasi sebesar 82,5%, didapat dengan menggunakan KNN dengan jumlah k = 3. Namun data yang digunakan hanya sedikit yaitu berjumlah 160 citra aksara Jawa[24]. Pada penelitian dengan topik yang berbeda, KNN digunakan untuk mengidentifikasi jenis protein. Dari hasil yang diperoleh, KNN menunjukkan performa terbaik ketika menggunakan Manhattan sebagai jaraknya, bahkan performa yang diberikan mengungguli CNN[26].

Dengan performa yang diberikan oleh GLCM dan KNN pada penelitian-penelitian sebelumnya, maka dilakukan penelitian yang memanfaatkan GLCM sebagai ekstraksi ciri dan KNN untuk mengklasifikasikan tulisan tangan aksara Bima. Selain itu, data yang digunakan perlu menggunakan data tulisan tangan, bukan data dari aplikasi sehingga dapat meningkatkan variasi data.

**B. Aksara Bima**

Berdasarkan penelitian yang dilakukan oleh Hj. Siti Maryam, pada awal penelitiannya menyatakan aksara Bima yang dianggap pernah hilang memiliki dua versi yakni versi Raffles dan versi Zolinger. Namun pada akhirnya aksara yang dikukuhkan sebagai aksara resmi Bima pada tahun 2007 merupakan aksara berdasarkan catatan Zolinger. Aksara ini memiliki kemiripan dengan aksara Bugis, dan sering ditemukan pada naskah-naskah kuno berdampingan dengan aksara Arab Melayu. Bentuk aksara bima seperti pada Gambar 1.



Gambar 1. Jenis aksara Bima

**C. GLCM**

GLCM atau yang disebut juga *gray-level spatial dependece matrix* merupakan metode yang menggunakan pendekatan statistik untuk melihat hubungan spasial dari

piksel. Pada penelitian ini kami menggunakan jarak ketetanggan GLCM d = 1 dan 4 arah sudut  $\Theta = 0^\circ, 45^\circ, 90^\circ,$  dan  $135^\circ$  untuk menghasilkan empat buah matriks GLCM. Selanjutnya, dari empat matriks GLCM kemudian dicari nilai *homogeneity, energy, contrast,* dan *correlation* yang merupakan ciri yang digunakan pada penelitian ini. Untuk mendapatkan nilai dari masing-masing ciri dengan menggunakan Persamaan 1 untuk *homogeneity*, Persamaan 2 untuk *energy*, Persamaan 3 untuk *contrast* dan Persamaan 4 untuk *correlation*.

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \tag{1}$$

$$Energy = \sum_{i,j=0}^{N-1} (P_{1j})^2 \tag{2}$$

$$Contrast = \sum_{i,j=0}^{N-1} P_{1j} (i - j)^2 \tag{3}$$

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu_i)(j-\mu_j)}{\sigma^2} \tag{4}$$

Dimana  $P_{ij}$  merupakan elemen dari setiap matriks GLCM,  $\mu$  merupakan rata-rata matriks GLCM, dan  $\sigma^2$  merupakan Varians masing-masing matriks GLCM yang dapat diperoleh dengan Persamaan 5 dan Persamaan 6.

$$\mu_i = \sum_{i,j=0}^{N-1} iP_{1j} , \mu_j = \sum_{i,j=0}^{N-1} jP_{1j} \tag{5}$$

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{ij} (i - \mu)^2 \tag{6}$$

**D. KNN**

Algoritma Klasifikasi KNN merupakan salah satu metode yang populer. Ide dasar dari KNN adalah dengan membandingkan nilai tetangga terdekat, dimana kesimpulan akhir akan mengikuti kelas dominan dari tetangga terdekat yang dimiliki oleh data tersebut. Salah satu hal penting yang digunakan KNN adalah pengukuran jarak kedua data. Perbedaan jarak yang digunakan akan menghasilkan jumlah k yang berbeda. Pada penelitian ini kami menggunakan empat jenis jarak antara lain euclidean, manhattan, chebyshev dan minkowski.

Jarak Euclidean merupakan salah satu jarak yang paling sering digunakan, dan disajikan pada Persamaan 7.

$$Euclidean = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{7}$$

Selanjutnya jarak Manhattan yang digunakan pada penelitian kami seperti pada Persamaan 8.

$$Euclidean = \sum_{i=1}^n |x_i - y_i| \tag{8}$$

Jarak Chebishev dan jarak Minkowski diberikan pada Persamaan (9) dan Persamaan (10).

$$Chebyshev = \max_i (|x_i - y_i|) \tag{9}$$

$$Minkowski = (\sum_{i=1}^n |x_i - y_i|^p)^{\frac{1}{p}} \quad (10)$$

### III. METODOLOGI PENELITIAN

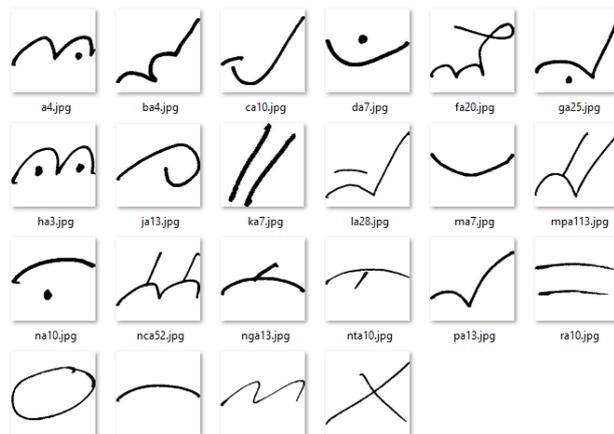
#### A. Data

Pada penelitian ini, pengumpulan data melibatkan 10 orang responden. 10 orang responden ini akan diminta untuk menuliskan 22 karakter aksara Bima sebanyak 12 kali penulisan untuk setiap karakter. Setiap karakter ditulis pada kertas berukuran F4 dengan menggunakan ballpoint ukuran 1.0 mm. Selanjutnya dilakukan proses digitalisasi menggunakan scanner dengan pengaturana dpi sebesar 600 dpi. Citra hasil scan tulisan tangan aksara Bima ditampilkan pada Gambar 2.



Gambar 2. Citra hasil scan tulisan tangan aksara Bima

Setelah proses akuisisi dilakukan, selanjutnya dilakukan *cropping* sehingga masing-masing responden menghasilkan 12 buah tulisan tangan untuk satu karakter. Sehingga didapatkan dataset aksara Bima sejumlah 2640 citra. Setelah itu ukuran citra akan diubah menjadi 64 x 64 piksel dan disimpan dalam bentuk jpg seperti pada Gambar 3.

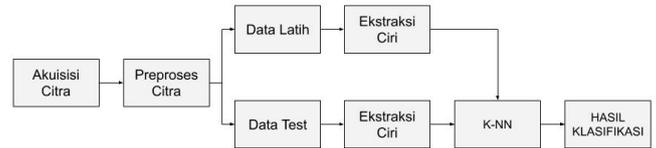


Gambar 3. Hasil Citra Tulisan Tangan Aksara Bima.

#### B. Perancangan Sistem

Perancangan sistem yang kami ajukan sesuai diagram pada Gambar 4. Berdasarkan Gambar 4, hasil akuisis citra

akan dilakukan preproses data dimana citra akan dipotong dan diubah ukurannya menjadi 64x64 dan disimpan dalam bentuk jpg. Selanjutnya, pada citra tersebut dilakukan ekstraksi ciri menggunakan GLCM, hasilnya kemudian digunakan sebagai dasar untuk melakukan klasifikasi menggunakan KNN



Gambar 4. Diagram kerja sistem

#### C. Skenario Pengujian

Beberapa skenario pengujian yang dilakukan antara lain:

- Menentukan nilai rasio data latih dan uji paling optimal. Pada skenario ini akan digunakan rasio data latih dan uji sebanyak 5 jenis rasio yaitu 90:10, 80:20, 70:30, 60:40 dan 50:50. Pada pengujian ini, nilai k yang digunakan yaitu 1 sampai 9 dan jarak yang digunakan merupakan jarak yang paling sering digunakan yaitu euclidean.
- Skenario kedua dari penelitian ini dengan membandingkan keempat jarak yang berdeda yaitu euclidean, manhattan, chebyshev dan minkowski. Pada skenario ini, rasio yang digunakan adalah nilai rasio terbaik berdasarkan skenario pengujian sebelumnya.
- Skenario terakhir yang digunakan adalah dengan melihat nilai presisi dan recall dari parameter terbaik yang telah didapatkan dari skenario sebelumnya. Hal ini dilakukan untuk menganalisis performa yang didapatkan, sehingga didapatkan kesimpulan penyebab performa tersebut terjadi.

Untuk metrik performa yang digunakan adalah akurasi, presisi dan recall. Akurasi yang digunakan mengikuti Persamaan 11, presisi mengikuti persamaan 12 dan recall sesuai dengan Persamaan 13.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

$$Presisi = \frac{TP}{TP+FN} \quad (12)$$

$$Recall = \frac{TP}{TP+FP} \quad (13)$$

Dimana :

- TP = data yang benar yang diklasifikasikan dengan benar
- TN = data yang salah yang diklasifikasikan dengan salah
- FN = data yang benar namun diklasifikasikan dengan salah

- FP = data yang salah namun diklasifikasikan dengan benar

IV. HASIL DAN PEMBAHASAN

Percobaan dilakukan dengan melakukan beberapa skenario antara lain (1) dengan melakukan percobaan pada variasi rasio data latih dan uji, dan (2) dengan memberikan variasi nilai *k* serta variasi jarak yang digunakan pada KNN.

A. Percobaan dengan variasi rasio data

Penggunaan rasio jumlah data latih dan data uji pada pembuatan model, akan menghasilkan performa yang berbeda[27], untuk itu perlu dilakukan percobaan dengan menggunakan rasio jumlah data latih dan data uji yang berbeda-beda. Dari rasio terbaik yang didapatkan, akan digunakan sebagai model pengujian selanjutnya.

Tabel I merupakan hasil percobaan yang dilakukan dengan menggunakan variasi rasio jumlah data latih dan data uji. Data latih yang digunakan minimal harus sejumlah data uji, sehingga rasio data latih dan data uji yang digunakan yaitu 90:10, 80:20, 70:30, 60:40 dan 50:50. Pada Tabel I KNN menggunakan jarak euclidean, dengan jumlah *k* dari 1 hingga 9.

Berdasarkan Tabel I, jumlah *k* yang digunakan mampu meningkatkan performa dari KNN, performa paling tinggi didapatkan ketika nilai *k* 7,8, dan 9. Sedangkan Rasio data latih dan data uji yang menghasilkan performa paling baik pada rasio 70:30, dan performa paling rendah pada rasio data 50:50. Performa terendah dikarenakan variasi data latih dianggap kurang jika dibandingkan dengan variasi data uji, sehingga kemungkinan salah klasifikasi semakin tinggi. Dari segi waktu komputasi, waktu yang diperlukan masih terbilang cepat, hal ini karena ciri dan data yang digunakan sedikit.

B. Percobaan dengan variasi jarak KNN

Pada skenario percobaan ini, akan dibandingkan KNN dengan menggunakan parameter jarak yang berbeda yakni *Euclidean*, *Manhattan*, *Chebyshev*, dan *Minkowski*. Berdasarkan Tabel II, performa dari masing-masing jarak yang digunakan menunjukkan hasil yang mirip. *k* pada masing-masing jarak KNN memberi pengaruh performa yang sedikit berbeda namun memberikan tren yang sama, yakni semakin performa akan bernilai semakin tinggi seiring bertambahnya nilai *k*. Pada jarak *Euclidean* dan *Minkowski* memberikan hasil yang sama karena nilai probabilitas pada *minkowski* diberikan nilai 1. Jarak *Manhattan* memberikan hasil paling tinggi pada nilai *k* = 9 dengan akurasi 60,86%. Jika dilihat berdasarkan waktu komputasi, waktu komputasi tidak terlalu berpengaruh dengan variasi jarak yang digunakan pada KNN.

TABLE I. HASIL PERCOBAAN KNN DENGAN JARAK EUCLIDEAN

Rasio=90:10				
k	Acc	Prec	Rec	time
1	52.27	52.27	52.27	0.04378
2	53.03	53.03	53.03	0.03659
3	54.17	54.17	54.17	0.04194
4	55.3	55.3	55.3	0.05002
5	56.44	56.44	56.44	0.0433
6	57.95	57.95	57.95	0.04113
7	59.09	59.09	59.09	0.04324
8	59.09	59.09	59.09	0.04277
9	58.71	58.71	58.71	0.04308
Rasio=80:20				
k	Acc	Prec	Rec	time
1	51.7	51.7	51.7	0.06049
2	51.89	51.89	51.89	0.05256
3	53.98	53.98	53.98	0.05282
4	54.92	54.92	54.92	0.05648
5	56.44	56.44	56.44	0.05116
6	57.2	57.2	57.2	0.06812
7	59.28	59.28	59.28	0.05536
8	58.33	58.33	58.33	0.07201
9	57.2	57.2	57.2	0.05524
Rasio=70:30				
k	Acc	Prec	Rec	time
1	53.41	53.41	53.41	0.08218
2	51.77	51.77	51.77	0.08725
3	54.17	54.17	54.17	0.07678
4	54.42	54.42	54.42	0.07209
5	56.82	56.82	56.82	0.08368
6	57.32	57.32	57.32	0.07773
7	58.59	58.59	58.59	0.06457
8	59.97	59.97	59.97	0.08493
9	<b>60.35</b>	<b>60.35</b>	<b>60.35</b>	0.06903
Rasio=60:40				
k	Acc	Prec	Rec	time
1	52.65	52.65	52.65	0.09089
2	51.04	51.04	51.04	0.07431
3	53.88	53.88	53.88	0.09626
4	54.92	54.92	54.92	0.08122
5	56.82	56.82	56.82	0.0867
6	57.77	57.77	57.77	0.08516
7	57.39	57.39	57.39	0.0881
8	58.33	58.33	58.33	0.12499
9	58.81	58.81	58.81	0.11018
Rasio=50:50				
k	Acc	Prec	Rec	time
1	51.59	51.59	51.59	0.0987
2	50.23	50.23	50.23	0.09979
3	52.65	52.65	52.65	0.10139
4	55	55	55	0.09508
5	54.02	54.02	54.02	0.10558
6	55.38	55.38	55.38	0.10888
7	55.38	55.38	55.38	0.13597
8	56.14	56.14	56.14	0.09292
9	55.91	55.91	55.91	0.12791

TABLE II. HASIL PERCOBAAN KNN DENGAN VARIASI JARAK

EUCLIDEAN DISTANCE				
k	Acc	Prec	Rec	time
1	53.41	53.41	53.41	0.08218
2	51.77	51.77	51.77	0.08725
3	54.17	54.17	54.17	0.07678
4	54.42	54.42	54.42	0.07209
5	56.82	56.82	56.82	0.08368
6	57.32	57.32	57.32	0.07773
7	58.59	58.59	58.59	0.06457
8	59.97	59.97	59.97	0.08493
9	60.35	60.35	60.35	0.06903
MANHATTAN DISTANCE				
k	Acc	Prec	Rec	time
1	51.52	51.52	51.52	0.07632
2	51.26	51.26	51.26	0.06681
3	54.29	54.29	54.29	0.05868
4	56.06	56.06	56.06	0.08104
5	57.32	57.32	57.32	0.0879
6	59.34	59.34	59.34	0.07941
7	57.83	57.83	57.83	0.07124
8	59.6	59.6	59.6	0.06768
9	<b>60.86</b>	<b>60.86</b>	<b>60.86</b>	0.06795
CHEBYSHEV DISTANCE				
k	Acc	Prec	Rec	time
1	51.39	51.39	51.39	0.06774
2	52.53	52.53	52.53	0.07232
3	53.91	53.91	53.91	0.07434
4	54.42	54.42	54.42	0.07392
5	57.07	57.07	57.07	0.0708
6	57.45	57.45	57.45	0.09585
7	57.58	57.58	57.58	0.08401
8	56.94	56.94	56.94	0.07738
9	59.34	59.34	59.34	0.08835
MINKOWSKI DISTANCE				
k	Acc	Prec	Rec	time
1	53.41	53.41	53.41	0.09374
2	51.77	51.77	51.77	0.0757
3	54.17	54.17	54.17	0.07464
4	54.42	54.42	54.42	0.08218
5	56.82	56.82	56.82	0.06891
6	57.32	57.32	57.32	0.0792
7	58.59	58.59	58.59	0.07475
8	59.97	59.97	59.97	0.0741
9	60.35	60.35	60.35	0.08147

C. Analisis Performa

Selain Akurasi, hal yang menjadi pertimbangan dalam menentukan performa sebuah sistem adalah presisi dan recall. Berdasarkan Tabel III, Presisi dan Recall masing-masing Aksara cukup berbeda, performa terbaik pada aksara ke 9 (ka), artinya aksara ini memiliki karakteristik yang berbeda dengan aksara lainnya sehingga dapat dikenali dengan baik. Sedangkan untuk aksara paling sulit dibedakan adalah aksara ke 7 (ha) karena sangat mirip dengan aksara ke-1 (a) dan Aksara ke 21 (wa). Bentuk Aksara ka ditampilkan pada Gambar XXX, dan kemiripan

bentuk Aksara ha dan a seperti yang ditampilkan pada Gambar 5.



Gambar 5. Bentuk Aksara ka



(a)

(b)



(c)

Gambar 6. Kemiripan Aksara (a) Ha dengan aksara (a) A dan (c) Wa

TABLE III. PRESISI DAN RECALL DENGAN RASIO 70:30 DAN JARAK MANHATTAN

id aksara	Prec	Rec
1	36	36
2	39	31
3	38	39
4	47	44
5	60	58
6	58	58
7	33	33
8	59	61
9	100	100
10	41	58
11	61	64
12	70	72
13	55	58
14	77	75
15	56	53
16	47	50
17	69	86
18	97	94
19	81	72
20	68	69
21	62	42
22	91	83

## V. KESIMPULAN

Berdasarkan percobaan yang telah dilakukan, performa terbaik didapatkan ketika parameter  $k$  bernilai 9 dan dengan menggunakan rasio dataset 70:30 dengan tingkat akurasi, presisi dan recall sebesar 61%. Hasil yang diperoleh pada penelitian kami masih kurang optimal, ini disebabkan karena penggunaan GLCM yang bersifat *global*, sedangkan pada citra yang digunakan memiliki latar belakang seragam, selain itu ada beberapa karakter aksara Bima yang memiliki bentuk serupa. Untuk penelitian selanjutnya perlu dipertimbangkan penggunaan ciri-ciri yang bersifat lokal atau perlu dipertimbangkan untuk melakukan proses seleksi ciri untuk mengurangi ciri yang tidak dominan. Perlakuan pre-proses perlu dilakukan juga untuk mengurangi derau yang berlebihan.

## DAFTAR PUSTAKA

- [1] "Indonesia Archipelago - Indonesian Archipelago." <http://www.indo.com/indonesia/archipelago.html> (accessed Feb. 27, 2020).
- [2] "Atlas of Endangered Alphabets – Indigenous and minority writing systems, and the people who are trying to save them." <https://www.endangeredalphabets.net/> (accessed Oct. 28, 2020).
- [3] N. T. B. BPS Provinsi, "Penduduk Kabupaten/Kota (Jiwa), 2018-2020," 2020. <https://ntb.bps.go.id/indicator/12/29/1/penduduk-kabupaten-kota.html> (accessed Oct. 28, 2020).
- [4] Walikota Bima, *Peraturan Walikota Bima No. 50 Tahun 2019*. 2019.
- [5] "Tingkatkan Nilai Kearifan Lokal, Dikbud Bimtek Aksara Mbojo | Kahaba.net." <https://kahaba.net/berita-bima/58385/tingkatkan-nilai-kearifan-lokal-dikbud-bimtek-aksara-mbojo.html> (accessed Nov. 16, 2020).
- [6] M. Rizky, I. Nurhania, and I. S. Areni, "Mbojo Character Recognition Using Shearlet Transform and Support Vector Machine," *Proceeding - 2018 Int. Semin. Intell. Technol. Its Appl. ISITIA 2018*, pp. 339–344, 2018, doi: 10.1109/ISITIA.2018.8710976.
- [7] A. Aranta, F. Bimantoro, and I. P. T. Putrawan, "Penerapan Algoritma Rule Base dengan Pendekatan Hexadesimal pada Transliterasi Aksara Bima Menjadi Huruf Latin," *J. Teknol. Informasi, Komputer, dan Apl. (JTika)*, vol. 2, no. 1, pp. 130–141, 2020, doi: 10.29303/jtika.v2i1.96.
- [8] Y. Sugianela and N. Suciati, "Ekstraksi Fitur Pada Pengenalan Karakter Aksara Jawa Berbasis Histogram of Oriented Gradient," *JUTI J. Ilm. Teknol. Inf.*, vol. 17, no. 1, p. 64, 2019, doi: 10.12962/j24068535.v17i1.a819.
- [9] C. K. Dewa, A. L. Fadhilah, and A. Afiahayati, "Convolutional Neural Networks for Handwritten Javanese Character Recognition," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 12, no. 1, p. 83, 2018, doi: 10.22146/ijccs.31144.
- [10] A. Aryantio and R. Munir, "Pengenalan Aksara Lampung Menggunakan Jaringan Syaraf Tiruan," in *Konferensi Nasional Informatika (KNIF)*, 2015, pp. 34–38.
- [11] I. W. A. S. Darma, I. K. G. D. Putra, and M. Sudarma, "Ekstraksi Fitur Aksara Bali Menggunakan Metode Zoning," *Teknol. Elektro*, vol. 14, no. 2, 2015.
- [12] E. Dina Juliani U M, I. G. P. S. Wijaya, and F. Bimantoro, "Pengenalan Pola Tulisan Tangan Suku Kata Aksara Sasak Menggunakan Metode Integral Projection dan Neural Network," *J. Comput. Sci. Informatics Eng.*, vol. 3, no. 1, p. 19, Jun. 2019, doi: 10.29303/jcosine.v3i1.222.
- [13] S. Iamsa-At and P. Horata, "Handwritten character recognition using histograms of oriented gradient features in deep learning of artificial neural network," in *2013 International Conference on IT Convergence and Security, ICITCS 2013*, 2013, no. 1, doi: 10.1109/ICITCS.2013.6717840.
- [14] S. M. Ismail and S. Norul Huda Sheikh Abdullah, "Geometrical-Matrix Feature Extraction For On-Line Handwritten Characters Recognition," *J. Theor. Appl. Inf. Technol.*, vol. 10, no. 1, 2013, Accessed: Feb. 26, 2020. [Online]. Available: [www.jatit.org](http://www.jatit.org).
- [15] F. Alotaibi, M. T. Abdullah, R. B. H. Abdullah, R. W. B. O. K. Rahmat, I. A. T. Hashem, and A. K. Sangaiah, "Optical Character Recognition for Quranic Image Similarity Matching," *IEEE Access*, vol. 6, pp. 554–562, Nov. 2017, doi: 10.1109/ACCESS.2017.2771621.
- [16] Vidia, "Pengenalan Tulisan Tangan Bahasa Arab Menggunakan Metode Probabilistic Neural Network," *J. Ilmu Komput. dan Desain Komun. Vis.*, vol. 4, no. 1, pp. 28–35, 2019.
- [17] D. Sen Maitra, U. Bhattacharya, and S. K. Parui, "CNN based common approach to handwritten character recognition of multiple scripts," in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2015*, pp. 1021–1025, doi: 10.1109/ICDAR.2015.7333916.
- [18] Vijayalaxmi M B and B. V Dhandra, "Script Recognition using GLCM and DWT Features," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 1, 2015, doi: 10.17148/IJARCCCE.2015.4157.
- [19] L. Singla and S. Singh, "Offline Handwritten Devanagari Numerals Recognition using GLCM Features & Neural Networks," *Int. J. Eng. Res. Technol.*, vol. 3, no. 6, pp. 1476–1480, 2014, Accessed: Feb. 26, 2020. [Online]. Available: [www.ijert.org](http://www.ijert.org).
- [20] N. Prameela, P. Anjusha, and R. Karthik, "Off-line Telugu handwritten characters recognition using optical character recognition," in *Proceedings of the International Conference on Electronics, Communication and Aerospace Technology, ICECA 2017*, Apr. 2017, vol. 2017-January, pp. 223–226, doi: 10.1109/ICECA.2017.8212801.
- [21] V. C. Bharathi and M. Kalaiselvi Geetha, "Segregated Handwritten Character Recognition using GLCM Features," 2013.
- [22] N. Ilmi, T. A. B. Wiharja, and K. N. Ramadani, "Pengenalan Angka Tulisan Tangan Dengan Menggunakan Local Binary Pattern Variance Dan Klasifikasi K-nearest Neighbour," *eProceedings Eng.*, vol. 2, no. 3, 2015.
- [23] I. A. D. P. Sari, B. Hidayat, and U. Sunarya, "Pengenalan Aksara Bali Dengan Metode Local Binary Pattern," *eProceedings Eng.*, vol. 2, no. 2, 2015.
- [24] A. Susanto, D. Sinaga, E. H. Rachmawanto, and others, "Unjuk Kerja K-Nearest Neighbors Pada Pengenalan Karakter Jawa Berbasis Local Binary Pattern," *SNATIF*, vol. 5, no. 1, 2018.
- [25] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease," *IEEE Access*, vol. 8, pp. 20991–21002, 2020, doi: 10.1109/ACCESS.2019.2963053.
- [26] X. Gao and G. Li, "A KNN Model Based on Manhattan

- Distance to Identify the SNARE Proteins,” *IEEE Access*, vol. 8, pp. 112922–112931, 2020, doi: 10.1109/ACCESS.2020.3003086.
- [27] R. Medar, V. S. Rajpurohit, and B. Rashmi, “Impact of Training and Testing Data Splits on Accuracy of Time Series Forecasting in Machine Learning,” in *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*, Sep. 2018, pp. 1–6, doi: 10.1109/ICCUBEA.2017.8463779.