Implementasi Shape Feature dan K-Nearest Neighbor untuk Klasifikasi Tanda Tangan

Muchamad Kurniawan[1], Naili Saidatin[2], dan Hendro Nugroho[3]
Institut Teknologi Adhi Tama Surabaya[1, 2, 3]
e-mail: muchamad.kurniawan@itats.ac.id

ABSTRACT

One proof that is used to verify a person's identity is through signatures that contain special characters and additional forms. The use of signatures is often found in several activities, especially in the field of administration. In this study, signature recognition using two segmentation patterns, square and triangle, aims to determine the effect of pattern recognition on the level of signature authenticity and signature classification. In making this system the off-line signature classification is applied by taking a signature image as input that will be used in the next process. The first digital image processing is preprocessing on digital images, followed by the feature extraction process and finally the classification process using the K-Nearest Neighbor (K-NN) method. Based on the classification results obtained by using two approaches to image cutting results obtained are not much different, whether square or triangle has a pretty good accuracy above 95%. Better results are obtained from the triangle approach, this mapping consistently produces an accuracy of 98.25%. For further research is with high accuracy in classification data it is necessary to conduct research to identify the authenticity of signatures.

Keyword: *Signature, digital image, K-Nearest Neighbor (K-NN) method, square and triangle.*

ABSTRAK

Salah satu bukti yang digunakan untuk verifikasi identitas seseorang adalah melalui tanda tangan yang mengandung karakter khusus dan bentuk-bentuk tambahan. Penggunaan tanda tangan sering dijumpai pada beberapa kegiatan khususnya dalam bidang administrasi. Pada penelitian ini dilakukan pengenalan tanda tangan dengan menggunakan dua pola segmentasi yaitu square dan triangle, yang bertujuan untuk mengetahui pengaruh pengenalan pola terhadap tingkat akurasi keaslian tanda tangan dan klasifikasi tanda tangan. Dalam pembuatan sistem ini diterapkan klasifikasi tanda tangan secara off-line dengan mengambil sebuah image tanda tangan sebagai input yang akan digunakan dalam proses selanjutnya. Proses pengolahan citra digital diawali dengan proses prepocesing pada citra digital , dilanjutkan dengan proses fitur ekstraksi dan terakhir proses klasifikasi dengan metode K-Nearest Neighbor (K-NN). Berdasarkan hasil klasifikasi didapatkan dengan melakukan dua pendekatan pemotongan citra hasil yang didapatkan tidak jauh berbeda, baik persegi atau segitiga mempunyai akurasi yang cukup bagus diatas 95%. Hasil lebih bagus didapatkan dari pendekatan segitiga, pendetakan ini secara konsisten menghasilkan akurasi 98.25%.

Kata kunci: Tanda tangan, citra digital, metode K-NN, square dan triangle

PENDAHULUAN

Salah satu bukti yang digunakan untuk verifikasi identitas seseorang adalah melalui tanda tangan [1] yang mengandung karakter khusus dan bentuk-bentuk tambahan [2]. Penggunaan tanda tangan sering dijumpai pada beberapa kegiatan khususnya dalam bidang administrasi. Namun, sampai saat ini masih ditemui kegiatan memalsukan tanda tangan yang dilakukan oleh orang yang bukan merupakan pemilik tanda tangan tersebut. Beberapa kegiatan pemalsuan tanda tangan dapat dijumpai dalam kehadiran proses belajar mengajar, undangan resmi tertentu, pemalsuan surat keterangan dokter, pemalsuan surat pernyataan orang tua, dan lain sebagainya yang memerlukan tanda tangan. Penentuan keaslian tanda tangan semakin didukung dengan adanya perkembangan teknologi informasi. Penentuan keaslian tanda tangan adalah salah satu bagian dari kecerdasan teknologi informasi berbasis

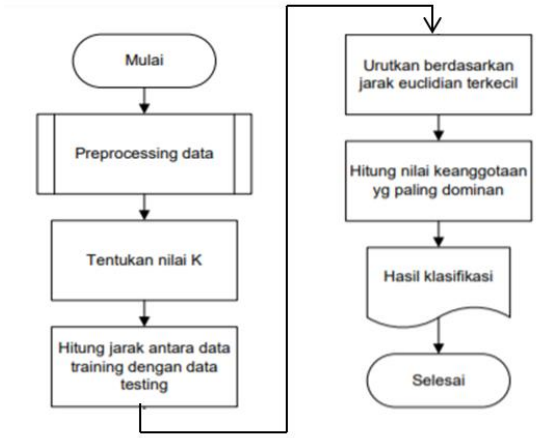
citra digital. Sistem verifikasi tanda tangan dapat dibagi menjadi dua bagian yaitu secara offline [3-5] dan online [6-8]. Verifikasi tanda tangan secara offline merupakan metode otentikasi yang menggunakan dinamika tulisan tangan seseorang dan menganalisa bentuk fisik dari tanda tangan tersebut [9]. Verifikasi tanda tangan secara off-line dilakukan dengan mengambil sebuah *image* tanda tangan sebagai *input* yang akan digunakan dalam proses selanjutnya. Sedangkan *input* untuk verifikasi tanda tangan secara on-line diambil dari tanda tangan yang didapatkan langsung dari digitizer yang dapat menghasilkan nilai– nilai dinamik, seperti nilai koordinat, lama tanda tangan, dan kecepatan tanda tangan .

Pengenalan tanda tangan dapat dilakukan dengan menerapkan metode ekstraksi ciri bentuk. Beberapa penelitian telah menerapkan berbagai macam metode dan ekstraksi ciri pada signature verification diantaranya dilakukan oleh penelitian [10] menggunakan Feed Forward Back Propagation Error Neural Network dengan ekstraksi ciri Discrete Wavelet Transform untuk melakukan verifikasi pada tanda tangan online, penelitian [11] menggunakan Euclidean Distance Model dengan ekstraksi ciri Geometric Center dan penelitian [12] dengan judul "Automatic Medical Image Classification and Abnormality using K-Nearest Neighbour". Penelitian lain [13] juga membahas mengenai verifikasi tanda tangan namun menggunakan algoritma "Siamese" Deep Neural Network. Penelitian ini memberikan performa Equal Error Rate (ERR) sebesar 4,5%. Dari penelitian-penelitian tersebut dapat dilihat bahwa para peneliti melalui berbagai macam penelitian berusaha untuk menemukan metode klasifikasi dan ekstraksi ciri yang menghasilkan tingkat akurasi yang tinggi. Tingkat akurasi yang tinggi penting dalam verifikasi tanda tangan untuk menghindari ancaman pemalsuan tanda tangan . Namun, dari beberapa penelitian tersebut belum ada yang menjelaskan dengan beberapa perbedaan pengenalan bentuk pola citra.

Oleh karena itu, pada penelitian ini dilakukan pengenalan tanda tangan dengan menggunakan dua pola segmentasi yaitu square dan triangle, yang bertujuan untuk mengetahui pengaruh pengenalan pola terhadap tingkat akurasi keaslian tanda tangan serta dilakukan uji klasifikasi tanda tangan dengan metode pengelompokan menggunakan K-Nearest Neighbor (K-NN). Metode pengelompokan K-NN mampu melakukan klasifikasi dengan baik untuk citra yang rumit. Contohnya, pada penelitian [12] mampu melakukan klasifikasi gambar medis dengan akurasi 80% dan merupakan yang terbesar dibandingkan dengan SVM linear dan RBF kernel. Dalam pembuatan sistem ini diterapkan klasifikasi tanda tangan secara off-line dengan mengambil sebuah image tanda tangan sebagai input yang akan digunakan dalam proses selanjutnya. Proses pengolahan citra digital diawali dengan teknik grayscale, thresholding, segmentation, size normalization, thinning, dan region. Selanjutnya dilakukan penelusuran piksel untuk menyimpan semua posisi piksel yang mengandung foreground dan dicari piksel tengah dari citra. Kemudian semua piksel yang tersimpan akan dihitung centroid distance function-nya dan kemudian akan dikelompokkan sesuai dengan banyak segmen yang membagi citra sehingga setiap kelompok memiliki jumlah data yang sama dengan menggunakan klasifikasi K-Nearest Neighbor (KNN).

TINJAUAN PUSTAKA

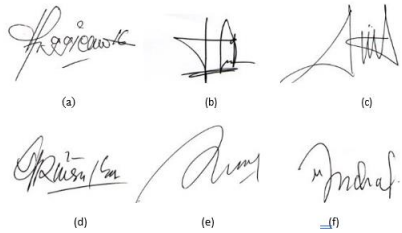
Metode pengelompokan K-NN mampu melakukan klasifikasi dengan baik untuk citra yang rumit. Contohnya, pada penelitian [12] mampu melakukan klasifikasi gambar medis dengan akurasi 80% dan merupakan yang terbesar dibandingkan dengan SVM linear dan RBF kernel. Algortima K-NN merupakan algoritma klasifikasi yang cukup sederhana [14], tetapi mempunyai hasil akurasi yang cukup bagus [15]. Cara kerja algoritma ini digambarkan seperti gambar 1. Pada penelitian ini menggunakan teknik one leave out cross validation, dikarekan dataset yang kami miliki tidaklah banyak. Cara kerja teknik ini dengan mengeluarkan satu data dijadikan data testing dan sisa datanya dijadikan data training. Langkah ini diulang terus sampai semua data menjadi data testing.



Gambar 1. Algoritma K-NN

METODE

Klasifikasi tanda tangan pada penelitian ini dimulai dengan proses prepocesing pada citra digital , dilanjutkan dengan proses fitur ekstraksi dan terakhir proses klasifikasi. Pada proses prepocesing kami menggunakan tiga langkah : grayscale, resize dan segmentasi. Pada proses fitur ekstraksi dilakukan dengan dua cara: segitiga dan persegi. Proses terakhir kami menggunakan K-Neirest Neighbor untuk klasifikasi. Dataset yang digunakan pada penelitian ini dilakukan dengan mengambil contoh citra tanda tangan sebanyak 6 koresponden dengan 1 koresponden sebanyak 8 kali, contoh citra tanda tangan seperti pada gambar 2 dan 3. Pengambilan jumlah dataset atau tanda tangan terhadap satu koresponden bertujuan untuk membuat pola yang lebih bagus, dikarenakan setiap kali tanda tangan tidak mugkin sama persis. Teknik pengambilan tanda tangan yang kami gunakan adalah koresponden memberikan contoh tanda tangan dalam kertas putih dengan ketebalan pena yang hampir sama.



Gambar 2. Contoh tanda tangan setiap koresponden



Gambar 3. Contoh 4 tanda tanggan koresponen ke-1

HASIL DAN PEMBAHASAN

Pada penelitian ini pada proses prepocesing diawali dengan konversi RGB ke grayscale dan resize citra digital menjadi 1 : 2, 1 untuk baris dan 2 untuk kolom. Formulasi yang digunakan untuk merubah RGB ke grayscale menggunakan persamaan 1.

$$grayscale = R * 0.2898 + G * 0.5870 + B * 0.1140 \quad ....(1)$$

Pada proses *resize image* kami mengggunakan *library* yang telah digunakan oleh matlab dengan memanggil fungsi imresize(). Proses selanjutnya adalah segmentasi yaitu mengkonversi citra grayscale menjadi dua level nilai, 0 dan 1 atau hitam dan putih. Pembentukan citra biner dibatasi oleh *threshold* t seperti pada persamaan 2. Dimana x adalah nilai pixel yang didapatkan. Untuk mendapatkan nilai t kami menggunakan algoritma otsu, yang dipangil dengan fungsi matlab *im2bw().* Hasilnya dapat dilihat pada gambar 4.

$$biner = (1, t \geq x), (0, x < t) \quad ....(2)$$



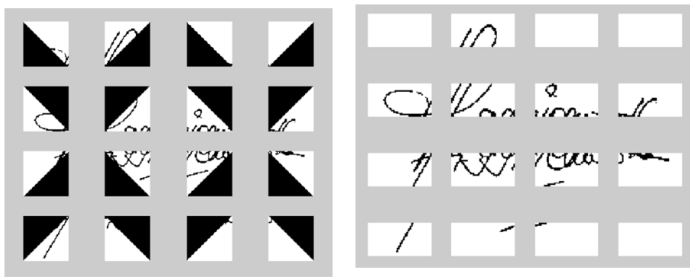
(a) (b) (c) (d)

Gambar 4. a) citra asli, b) citra grayscale, c) citra resize, d) citra biner

Ekstraksi fitur pada penelitian ini dibagi menjadi dua cara, segitiga dan persegi. Semua jenis ekstraksi akan membagi citra menjadi 16 bagian. Hasil ekstraksi segitiga dapat dilihat pada gambar 5(a), pada gambar tersebut terdapat 16 gambar parsial dari citra contoh 4 (d). Cara kerja untuk membagi data seperti itu sebagai berikut : bagi citra menjadi seperempat bagian dari citra asli secara berurutan, dari setiap bagian tersebut bagi lagi menjadi dua bagian segita atas dan bawah. Akhirnya didapatkanlah 16 bagian citra parsial yang ukuranya sama rata. Dari setiap bagian hitung jumlah nilai pixel hitam. Bagian yang hitam itu tidak termasuk dalam perhitungan total pixel, hal itu terjadi karena citra yang ditampikan harus berbentuk persegi sehingga citra sisa dari segitiga secara otomatis diberikan tanda hitam oleh Bahasa pemrograman matlab. Penjumlahan nilai hitam ini merupakan nilai ekstraksi fitur citra untuk setiap dataset. Gambar 5(b) merupakan hasil pemecahan citra dengan pendekatan persegi, citra dibagi langsung menjadi 16 bagian sama rata. Tabel 1 dan tabel 2 merupakan hasil ekstraksi fitur untuk pembagian secara segitiga dan persegi. Terdapat perbedaan nilai antara dua nilai ekstraksi tersebut, terutama pada bagian ujung citra.

Tabel 1. Hasil ekstraksi fitur segitiga citra contoh koresponden 1 citra ke-1

partial ke-	sum pixel	partial ke-	sum pixel	partial ke-	sum pixel	partial ke-	sum pixel
1	36	5	144	9	6	13	2
2	93	6	1141	10	144	14	77
3	45	7	187	11	214	15	88
4	83	8	24	12	15	16	9



(a) (b)

Gambar 5. a) Ekstraksi fitur segitiga, b) Ekstraksi fitur persegi

Tabel 2. Hasil ekstraksi fitur persegi citra contoh koresponden 1 citra ke-1

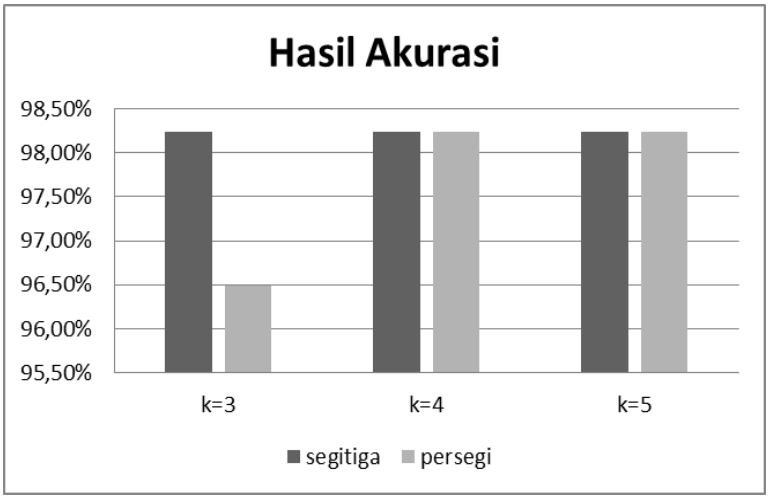
partial ke-	sum pixel	partial ke-	sum pixel	partial ke-	sum pixel	partial ke-	sum pixel
1	0	5	93	9	0	13	0
2	125	6	185	10	147	14	79
3	92	7	182	11	216	15	96
4	36	8	20	12	13	16	0

Tabel 3. Hasil klasifikasi

	K=3		K=4		K=5	
	segitia	persegi	segitia	persegi	segitia	persegi
n data	57		57		57	
TP + TN	56	55	56	56	56	56
FP + FN	1	2	1	1	1	1

Hasil percobaan dengan nilai K :3,4,5 dan hasil yang diperoleh dapat dilihat pada tabel 3. Berdasarkan tabel 3 total keseluruhan data adalah 57 data, TP merupakan notasi untuk *true positif*, TN untuk *true negative*, FP adalah false positif, dan FN merupakan singkatan *false negative*. Hampir seluruh percobaan mempunyai hasil yang sama yaitu 56 data terklasifikasi dengan benar dan 1 data tidak terklasifikasi. Hanya terdapat satu data yang berbeda yaitu pada pengujian K=3 dengan pendekatan persegi, perbedaanya terdapat tambahan 1 data yang tidak terklasifikasi dengan benar. Nilai akurasi yang didapatkan dari persamaan 3, digambarkan pada gambar 6, dengan nilai modus akurasi 98.25%, sedangkan nilai terendah akurasi yang diperoleh 96.50%.

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad ....(3)$$



Gambar 6. Hasil Akurasi

KESIMPULAN

Kesimpulan yang dapat ditarik pada penilitian klasifikasi tanda tangan ini adalah pentingnya pemilihan fitur pada citra tanda tangan. Dengan melakukan dua pendekatan pemotongan citra hasil yang didapatkan tidak jauh berbeda, baik persegi atau segitiga mempunyai akurasi yang cukup bagus diatas 95%. Artinya tingkat keberhasilan metodologi yang kami usulkan sangat tinggi. Hasil lebih bagus didapatkan dari pendekatan segitiga, pendetakan ini secara konsisten menghasilkan akurasi 98.25%. Saran untuk penelitian selanjutnya adalah membandingkan berbagai macam fitur ekstraksi yang telah dikerjakan pada penelitian-penelitian sebelumnya baik dengan tekstur atau dengan pembagian citra untuk membuat klasifikasi tanda tangan lebih akurat lagi.

DAFTAR PUSTAKA

[1] Isha, *et al*. "Offline Signature Verification based on Euclidean distance using Support Vector Machine." International Journal of Advanced Engineering, Management and Science, vol. 2, no. 8, Aug. 2016.

[2] K. A. Al-Dulaimi, "Handwritten Signature Verification Technique based on Extract Features", International Journal of Computer Application, Vol.30, No.2, 2011.

[3] I.A. Ismail, M.A.Ramadan, T. S. El-Danaf and A.H. Samak, "An Efficient Off line Signature Identification Method Based On Fourier Descriptor and Chain Codes", 2010, (IJCSNS-2010),VOL.10 No.5, pp.29-35.

[4] J.O.Pinzon-Arenas, R. Jimenez-Moreno, C.G. Pachon-Suescun, "Offline Signature Verification using DAG-CNN", International Journal of Electrical and Computer Engineering, Vol.9, No.4, 2019.

[5] Z.J.Xing, F. Yin, Y.C. Wu, C.L. Liu, "Offline Signature Verification Using Convolution Siamese Network", Ninth International Conference on Graphic and Image Processing, 2017.

[6] R. Munir, " Pengolahan Citra Digital dengan Pendekatan Algoritmatik", Bandung, 2004.

[7] V. Iranmanesh, *et al*., "Online Handwritten Signature Verificaion Using Neural Network Classifier Based on Principal Component Analysis", Hindawi Publishing Corporation, The Scientific World Journal, 2014.

[8] V. G. Yogesh, "Online Signature Verification Systems : A Survey", International Journal on Emerging Technologies, Vol.6, No.2, 2015.

[9] I. Bhattac, P. Ghosh, S. Biswas, "Offline Signature Verification Using Pixel Matching Technique", Procedia Technology, Vol.10, p.970-977, 2013.

[10] M. Fahmi, " Online Handwritten Signature Verification System Based On DWT Features Extraction and Neural Network", Ain Shams Engineering Journal, Vol 1, h 59-70,2010.

[11 B. Mahji, *et al*., " Novel Feature for Off-line Signature Verification", International Journal of Computers, Vol. 1, No 1, h.17-24, 2006.

[12] R.J.Ramkete *et al.*, "Automatic Medical Image Classification and Abnormality Detection Using K-Nearest Neighbour", International Journal of Advance Computer Research, Vol. 2, No. 4, p.190-196, 2012.

[13] Y.Yin, X.Zhou, "End-to-end Online Handwriting Signature Verification", Tenth International Conference on Graphic and Image Processing (ICGIP), 2018.

[14] C. Li, et al., "Using the K-Nearest Neigbor Algorithm for the Classification of Lymph Node Metastasis in Gastric Cancer", Hindawi Publishing Corporation, Computational and Mathematical Methods in Medicine, 2012.

[15]Suguna, K.Thanushkodi, "An Improved k-Nearest Neighbor Classification Using Genetic Algorithm", International Journal of Computer Science Issues, Vol. 7, No. 2, 2010.