

**STUDI BERBAGAI METODE PENGENALAN ANGKA DAN
KARAKTER TULISAN TANGAN**

Skripsi



oleh
PAPANG WIDA KRISTIANTO
22053930

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2014

**STUDI BERBAGAI METODE PENGENALAN ANGKA DAN
KARAKTER TULISAN TANGAN**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

PAPANG WIDA KRISTIANTO
22053930

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2014

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

STUDI BERBAGAI METODE PENGENALAN ANGKA DAN KARAKTER TULISAN TANGAN

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar keserjanaan saya.

Yogyakarta, 18 Juni 2014




PARANG WIDA KRISTIANTO
22053930

HALAMAN PERSETUJUAN


Judul Skripsi : STUDI BERBAGAI METODE PENGENALAN
ANGKA DAN KARAKTER TULISAN TANGAN
Nama Mahasiswa : PAPANG WIDA KRISTIANTO
N I M : 22053930
Matakuliah : Skripsi (Tugas Akhir)
Kode : TIW276
Semester : Genap
Tahun Akademik : 2013/2014

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 18 Juni 2014

Dosen Pembimbing I


Ir. Sri Suwarno, M.Eng.

Dosen Pembimbing II


Budi Susanto, SKom., M.T.

HALAMAN PENGESAHAN

STUDI BERBAGAI METODE PENGENALAN ANGKA DAN KARAKTER TULISAN TANGAN

Oleh: PAPANG WIDA KRISTIANTO / 22053930

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 11 Juni 2014

Yogyakarta, 18 Juni 2014
Mengesahkan,

Dewan Penguji:

1. Ir. Sri Suwarno, M.Eng
2. Budi Susanto, SKom., M.T.
3. Kristian Adi Nugraha, S.Kom., M.T.
4. Ignatia Dhian E.K.R., S.Kom



Dekan



(Drs. Wimmie Handiwidjojo, M.T.)

Ketua Program Studi



(Nugroho Agus Haryono, M.Si)

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan anugerah, sehingga penulis dapat menyelesaikan Studi Literatur dengan judul Studi Berbagai Metode Pengenalan Angka dan Karakter Tulisan Tangan dengan baik.

Penulisan laporan Tugas Akhir ini merupakan kelengkapan dan pemenuhan dari salah satu syarat dalam memperoleh gelar Sarjana Komputer. Selain itu bertujuan melatih mahasiswa untuk dapat menghasilkan suatu karya yang dapat dipertanggungjawabkan secara ilmiah, sehingga bermanfaat bagi penggunanya.

Dalam menyelesaikan pembuatan laporan Tugas Akhir ini, penulis telah banyak menerima bimbingan, saran dan masukan dari berbagai pihak, baik secara langsung maupun tidak langsung. Untuk itu dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Ir. Sri Suwarno, M.Eng, selaku Pembimbing I yang telah memberikan bimbingannya dengan sabar dan baik kepada penulis, juga kepada
2. Bapak Budi Susanto, S.Kom., M.T. selaku dosen Pembimbing II, atas bimbingan, petunjuk dan masukan yang diberikan selama pengerjaan Tugas Akhir ini sejak awal hingga akhir.
3. Ayah dan Ibunda tercinta yang memberikan dukungan dan semangat.
4. Semua keluarga besar saya yang telah memberikan semangat.
5. Teman-teman yang telah berjuang bersama.
6. Pihak lain yang tidak dapat penulis sebutkan satu persatu, sehingga Skripsi ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari pembaca sekalian, sehingga suatu saat penulis dapat memberikan karya yang lebih baik lagi.

Akhir kata penulis minta maaf bila ada kesalahan baik dalam penyusunan laporan maupun yang pernah penulis lakukan sewaktu membuat Tugas Akhir ini. Sekali lagi penulis mohon maaf yang sebesar-besarnya. Dan semoga ini dapat berguna bagi kita semua.

Yogyakarta, Mei 2014

Penulis

©UKYDIN

ABSTRAK

Pengenalan tulisan tangan angka dan karakter merupakan ilmu pengetahuan yang sudah dipelajari sejak bertahun-tahun, terutama pengenalan pola yang dilakukan oleh sistem komputer dalam membaca bermacam-macam tulisan tangan angka dan karakter. Permasalahan yang dimiliki oleh sistem komputer dalam pengenalan pola ini adalah bagaimana caranya supaya sistem komputer dapat mengenali berbagai macam tulisan tangan angka dan karakter dari berbagai input yang ditulis manual oleh bermacam-macam penulis yang berbeda cara penulisannya dapat diotomatisasi ke dalam sistem komputer. Dalam pemecahan permasalahan ini ada beberapa cara atau metode yang dapat digunakan untuk pengenalan tulisan tangan. Di sini penulis akan membahas beberapa metode yang dapat memecahkan permasalahan pengenalan tulisan tangan angka dan karakter, metode-metode tersebut diantaranya adalah metode *K-Nearest Neighbor*(KNN), dan metode *Naïve Bayes*(NB). Melalui studi literatur ini diharapkan akan memberi arahan yang lebih baik dalam hal pengenalan angka dan karakter tulisan tangan bagi penelitian selanjutnya.

Kata kunci: pengenalan pola, *k-nearest neighbor*, *naïve bayes*.

DAFTAR ISI

HALAMAN JUDUL.....	
SAMPUL DALAM.....	
PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
UCAPAN TERIMA KASIH.....	vi
ABSTRAK.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Perumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	2
1.5. Metode/Pendekatan.....	3
1.6. Sistematika Penulisan.....	3
BAB 2 STUDI LITERATUR.....	5
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	6
2.2.1 <i>Preprocessing Image</i>	6
2.2.2 <i>Feature Extraction</i>	8
2.2.3 <i>Machine Learning</i>	9
2.2.4 Sistem Klasifikasi.....	10
2.2.5 Evaluasi Sistem Klasifikasi.....	14

BAB 3 PEMBAHASAN.....	15
3.1 Tahap-tahap Pengenalan Tulisan Tangan	15
3.2 <i>Preprocessing</i>	16
3.3 <i>Feature Extraction</i>	17
3.4 <i>Classification</i>	18
BAB 4 EVALUASI	22
4.1. Konsep Evaluasi Klasifikasi.....	22
4.2. Contoh Evaluasi Klasifikasi	25
BAB 5 KESIMPULAN.....	32
5.1 Kesimpulan	32
DAFTAR PUSTAKA	33

©UKYDWM

DAFTAR GAMBAR

Gambar	Keterangan	Halaman
2.1	Klasifikasi Naïve Bayes	6
2.2	<i>Remove Noise</i>	7
2.3	<i>Remove Noise</i>	7
2.4	<i>Binarizing</i>	7
2.5	<i>Binarizing</i>	7
2.6	<i>Resizing</i>	8
2.7	<i>Resizing</i>	8
2.8	<i>Decision Trees</i>	10
2.9	<i>Neural Networks</i>	11
2.10	<i>Pseudocode kNN</i>	13
3.1	Kerangka pikir pengenalan pola	16
3.2	Transisi Vertikal dan Horisontal	18
3.3	Pembagian citra ke 4 region dan fitur ekstraksi	18
4.1	Gambar Tabel Akurasi antara Klasifikasi BPNN dan kNN	25
4.2	Gambar Tabel Perbandingan BPNN dengan kNN	27

4.3	Prosentase Akurasi Campuran Tulisan Tangan dan Cetak Kannada menggunakan Klasifikasi kNN (k=3) dan SVM	28
4.4	Prosentase Akurasi Tulisan Tangan Kannada menggunakan Klasifikasi kNN (k = 3) dan SVM	29
4.5	Prosentase Akurasi Tulisan Cetak Kannada menggunakan Klasifikasi kNN (k = 3) dan SVM	30
4.6	Rata-rata Prosentase Akurasi Pengenalan Tulisan Tangan dan Cetak Kannada menggunakan metode klasifikasi kNN dan SVM	31

©UKDM

DAFTAR TABEL

Tabel	Keterangan	Halaman
3.1	Ilustrasi Vektor untuk Klasifikasi k-NN	19
3.2	Tabel Penghitungan Jarak Gambar Testing dengan Gambar Dikenali	20
4.1	Tabel Kontingensi	24

©UKDW

Bab 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Pengenalan tulisan tangan angka dan karakter merupakan ilmu pengetahuan yang sudah dipelajari sejak bertahun-tahun, terutama pengenalan pola yang dilakukan oleh sistem komputer dalam membaca bermacam-macam tulisan tangan angka dan karakter. Permasalahan yang dimiliki oleh sistem komputer dalam pengenalan pola ini adalah bagaimana caranya supaya sistem komputer dapat mengenali berbagai macam tulisan tangan angka dan karakter dari berbagai input yang ditulis manual dapat diotomatisasi ke dalam sistem komputer untuk mempermudah dalam pencatatan dokumen-dokumen penting pada perusahaan atau lembaga perkantoran. Adapun perusahaan atau lembaga yang memerlukan sistem aplikasi pengenalan tulisan tangan angka dan karakter ini diantaranya adalah pada lembaga kantor pos untuk pengenalan tulisan alamat pada surat yang ditulis manual, pengenalan tulisan nomer rekening pada slip bukti pembayaran atau setoran pada bank, pengenalan tulisan nomer polisi pada kendaraan bermotor, dan pada aplikasi-aplikasi lainnya. Pengenalan tulisan tangan ini juga bermanfaat untuk piranti elektronik interaktif seperti pada tablet *Personal Computer*(PC), papan tulis elektronik interaktif, dan lain sebagainya.

Dalam pemecahan permasalahan ini ada beberapa cara atau metode yang dapat digunakan untuk pengenalan tulisan tangan. Di sini penulis akan membahas beberapa metode yang dapat memecahkan permasalahan pengenalan tulisan tangan angka dan karakter, metode-metode tersebut diantaranya adalah metode *K-Nearest Neighbour*(KNN), dan metode *Naïve Bayes*(NB).

Dari beberapa metode pengenalan tulisan tangan angka dan karakter yang ada tersebut penulis akan melakukan identifikasi proses pembelajaran pada masing-masing metode serta memahami metode-metode tersebut untuk membandingkan hasilnya yang didasarkan pada tinjauan pustaka dari beberapa

karya ilmiah yang berhubungan dengan pengenalan pola tulisan tangan angka dan karakter.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah yang telah dikemukakan di atas, pada karya tulisan ilmiah ini akan meninjau bagaimana bentuk penerapan dari metode KNN dan NB terkait dengan pemecahan masalah pada pengenalan citra tulisan tangan angka dan karakter dari suatu citra.

1.3 Batasan Masalah

Untuk lebih memfokuskan pembahasan dalam penelitian ini, maka penulis akan meninjau dari sisi latar belakang algoritma yang dikembangkan, berbagai bentuk penerapan, serta beberapa parameter evaluasi dari kedua algoritma yaitu, KNN dan NB, terkait penerapannya untuk pengenalan angka dan karakter tulisan tangan berdasar pada beberapa rujukan karya ilmiah yang telah dipublikasikan sebelumnya.

Adapun batasan masalah dari studi literatur yang ditulis diantaranya adalah akan diuraikan pada setiap metode beberapa hasil penelitian yang ada, termasuk kinerja pada setiap metodenya. Pembahasan tidak mencakup tentang analisis algoritma yang ada pada setiap metode tersebut.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penulisan studi literatur ini adalah :

1. Menguraikan dua algoritma dasar , yaitu KNN dan NB, dan bentuk penerapannya untuk pengenalan angka dan karakter tulisan tangan.
2. Memberi wawasan kepada siapa saja yang tertarik dengan penerapan metode pengenalan citra tulisan tangan angka dan karakter.

1.5 Metode/Pendekatan

Metode pendekatan dalam penulisan studi literatur ini adalah : studi pustaka untuk mencari informasi mengenai metode-metode yang digunakan untuk pengenalan citra tulisan tangan angka dan karakter melalui buku, jurnal berbahasa Indonesia dan berbahasa asing.

1.6 Sistematika Penulisan

Laporan studi literatur tugas akhir ini disusun menjadi 5 bab, yaitu : Pendahuluan, Studi Literatur, Pembahasan, Evaluasi dan Kesimpulan.

BAB I adalah Pendahuluan, yang berisi latar belakang masalah dilakukan penulisan, perumusan masalah, batasan-batasan masalah, metode penelitian, tujuan serta sistematika penulisan.

BAB II adalah Studi Literatur, berisi tinjauan pustaka yang berisi berbagai referensi mengenai teori-teori yang digunakan.

BAB III adalah Pembahasan, pada bab ini akan membahas secara lebih terperinci dari prosedur-prosedur metode KNN, NB, dan metode lain yang dipakai untuk mendukung metode-metode tersebut. Akan dijelaskan secara terperinci juga fungsi masing-masing metode tersebut. Dari hasil yang didapat dari implementasi metode penelitian-penelitian orang lain akan dilakukan antara lain :

1. Mengidentifikasi proses pembelajaran pengenalan pola.
2. Membandingkan dan mencermati hasil dari masing-masing metode pada tulisan karya ilmiah yang ada diantaranya adalah :
 - KNN – NB

BAB IV adalah Evaluasi, berisi konsep evaluasi klasifikasi dalam pengenalan pola tulisan tangan angka dan karakter serta contoh evaluasi klasifikasi dari karya-karya orang lain yang tertuang dalam literatur yang ada.

BAB V adalah Kesimpulan, berisi rangkuman kesimpulan keseluruhan dari teori masing-masing metode pengenalan citra tulisan tangan angka dan karakter.

©UKDW

BAB 5

KESIMPULAN

5.1. Kesimpulan

Pada penulisan karya ilmiah yang telah dilakukan ini dapat ditarik kesimpulan bahwa pengenalan tulisan tangan angka dan karakter dapat dilakukan dengan menggunakan metode k-Nearest Neighbor dan metode Naïve Bayes. Dari kedua metode yang telah dibahas di atas k-Nearest Neighbor lebih banyak dipakai untuk pengenalan tulisan tangan daripada metode Naïve Bayes.

©UKDW

DAFTAR PUSTAKA

- Albashiti, A. H., & Tamimi, H. (2012). A Lexicon Based Offline Arabic Handwritten Recognition Using Naive Bayesian Classifier with Gaussian Distribution. *image*, 15(16), 17.
- Bischoff, A., & Wang, P. S. (1992). *Handwritten digit recognition using neural networks*.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Clarendon Press.
- Cano, J., Pérez-cortes, J. C., Arlandis, J., & Llobet, R. (2002). *Training Set Expansion in Handwritten Character Recognition*. doi:10.1007/3-540-70659-3_57
- Dhandra, B., Mukarambi, G., & Ngangarge, M. (2011). Zone Based Features for Handwritten and Printed. *International Journal of Computer Applications* , 5-9
- Dhandra, B., Hangarge, M., & Mukarambi, G. (2010). Spatial Features for Handwritten Kannada and English Character Recognition. *Recent Trends in Image Processing and Pattern Recognition*, 146-151.
- Elglaly, Y., & Quek, F. (2011). Isolated Handwritten Arabic Characters Recognition using Multilayer Perceptrons and K Nearest Neighbor Classifiers. 1-6.
- Fausett, L. V. (1994). *Fundamentals of neural networks: Architectures, algorithms,*

and applications. Englewood Cliffs, NJ ; Delhi: Prentice-Hall ; Dorling Kindersley.

Han, J., Kamber, M., & Pei, J. (2006). *Data mining: concepts and techniques*. Morgan kaufmann.

Holambe, A. N., & Thool, D. C. (2010). Printed and Handwritten Character & Number Recognition of Devanagari Script using SVM and KNN. *International J. of Recent Trends in Engineering and Technology*, 3(2), 163-166.

Hotta, S., Kiyasu, S., & Miyahara, S. (2004). *Pattern recognition using average patterns of categorical k-nearest neighbors*. doi:10.1109/ICPR.2004.1333790

Iofina, G. V. (2010). A study of metrics in finite sets for application in classification and recognition problems. *Pattern Recognition and Image Analysis*. doi:10.1134/S1054661810040036

Larrañaga, P., Calvo, B., Santana, R., Bielza, C., Galdiano, J., Inza, I., ... & Robles, V. (2006). Machine learning in bioinformatics. *Briefings in bioinformatics*, 7(1), 86-112.

Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill.

Provost, F., & Kohavi, R. (1998). Glossary of terms, *Machine Learning*. *Machine learning*, 30(2), 271-274.

Putra, D. (2010). *Pengolahan citra digital*. Penerbit Andi.