

KLASIFIKASI KENDARAAN BERMOTOR DENGAN MULTI KERNEL SUPPORT VECTOR MACHINE

Muhammad Athoillah

Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas PGRI Adi
Buana Surabaya
athoillah.muhammad@gmail.com

Abstract

Since October 2017, Electronic automatic payments have been applied in all Indonesian toll roads. This payment system needs an extra system which is able to distinguish the types of motor vehicles that will enter the toll road due to the regulation itself which allow four-wheel or more vehicle to pass the toll road. This automatic system can be built by a classification algorithm, and one of the best is Support Vector Machine (SVM), in order to be able to classify non-linear separable data, SVM must be modified by giving kernel function on it. Furthermore, determining the appropriate kernel for every classification problem is a fundamental step, and that's obviously not easy, to solve that problem the researchers found a method that can make this kernel function more flexible, this method called Multi Kernel Learning (MKL). Main idea of this method is formulating some kernel function to be one kernel function. This framework is built an automatic system to classify the types of motor vehicles using Support Vector Machine modified using Multi Kernel Learning method. The experimental result shows that the system has a good average value of accuracy that is 84.60%, the average value of precision is 84.95% and also average value of recall is 84.60%

Keyword: Kendaraan Bermotor, Klasifikasi, Multi Kernel, Support Vector Machine.

PENDAHULUAN

Berdasarkan Peraturan Menteri PUPR Nomor 16/PRT/M/2017, mulai oktober 2017 seluruh transaksi di gerbang jalan Tol di Indonesia berubah menjadi transaksi non-tunai. Itu artinya penggunaan Gerbang Tol Otomatis (GTO) telah diterapkan pada setiap jalan Tol di Indonesia. Dengan Gebang Tol Otomatis ini pengguna jalan Tol tidak lagi dilayani oleh petugas gerbang Tol

melainkan cukup dengan menggunakan kartu *e-money* maka gerbang Tol akan secara otomatis terbuka. Dengan sistem pembayaran otomatis seperti ini, tentunya dibutuhkan sebuah sistem tambahan yang mampu untuk membedakan jenis kendaraan bermotor yang hendak melewati Gerbang Tol Otomatis, hal tersebut berkaitan dengan regulasi bahwa kendaraan bermotor yang diperbolehkan untuk melewati jalan Tol

hanya kendaraan bermotor roda empat atau lebih.

Membangun sistem yang mampu membedakan jenis kendaraan bermotor dapat dilakukan dengan menggunakan algoritma klasifikasi. Salah satu algoritma pengklasifikasian yang sangat baik adalah Support Vector Machine (SVM) (Athoillah, dkk, 2015). Dengan menemukan *hyperplane* terbaik pada *input space*, dimana *hyperplane* tersebut didapatkan dengan menghitung *margin* maksimal dari dua garis vektor pemisah kedua kelas (*support line*), maka hasil pemisahan kelas yang dihasilkan menjadi lebih baik daripada algoritma lain yang hanya memanfaatkan satu garis pemisah yang didapat dengan menghitung jarak terdekat antar data (Clarke, dkk, 2009; Campbell, dkk, 2011).

Meskipun SVM mampu mengklasifikasikan data dengan sangat baik, namun ternyata SVM hanya mampu mengklasifikasikan data yang bersifat linier saja. Untuk dapat mengklasifikasikan data yang bersifat non-linier maka SVM harus dimodifikasi sedemikian rupa, salah satu metode yang dapat digunakan adalah dengan menambahkan fungsi kernel di dalamnya. Namun faktanya, dari beberapa fungsi kernel yang dapat digunakan, tidak semua kernel dapat memberikan hasil yang maksimal pada tiap permasalahan, hal ini dikarenakan setiap data yang digunakan memiliki karakteristik sendiri dan tidak

semua data sesuai dengan fungsi kernel yang digunakan. Oleh karenanya, memilih fungsi kernel yang tepat merupakan hal yang fundamental dalam setiap penggunaan SVM, hal ini tentunya tidaklah mudah, untuk mempermudah para peneliti membuat metode yang menjadikan fungsi kernel tersebut menjadi lebih fleksibel, metode ini biasa disebut dengan *Multi Kernel Learning* (MKL) (Athoillah, dkk, 2015; Rakotomamonjy dkk, 2007). Beberapa penelitian terbaru menunjukkan bahwa pembelajaran multi kernel mampu memberikan hasil yang baik dalam hal mengklasifikasi (Athoillah, dkk, 2015; Athoillah, 2017; Shrivastava dkk, 2015; Li dkk, 2014). Ide dasar dari MKL adalah dengan mengkombinasikan beberapa fungsi kernel menjadi sebuah persamaan, penjelasan lebih lanjut ada pada bab berikutnya.

Berdasarkan latar belakang yang telah dijabarkan, pada penelitian ini dibangun sebuah sistem yang mampu untuk mengklasifikasikan dan mengidentifikasi kendaraan bermotor berupa roda dua maupun roda empat dengan menggunakan algoritma SVM yang dimodifikasi dengan metode *Multi Kernel Learning* (MKL).

METODE PENELITIAN

1. Pra-Proses

Penelitian ini diawali dengan pra-proses, yaitu dengan mengubah file-file

gambar yang menjadi dataset penelitian menjadi bentuk data matriks yang kemudian dilakukan perhitungan untuk mendapatkan nilai histogram dari citra tersebut. Histogram adalah perhitungan frekuensi kemunculan nisbi (*relative*) dari intensitas sebuah citra. Secara matematis nilai histogram sebuah citra dapat dihitung dengan persamaan berikut:

$$h_i = \frac{n_i}{n}, \quad i = 0, 1, 2, \dots, L - 1$$

Dengan, n_i merupakan jumlah *pixel* yang mempunyai derajat keabuan i , kemudian n merupakan jumlah seluruh *pixel*, sedangkan L merupakan nilai dari derajat keabuan (Athoillah, 2017). Jadi, misalkan terdapat sebuah citra *input* seperti Gambar 1 dibawah ini:



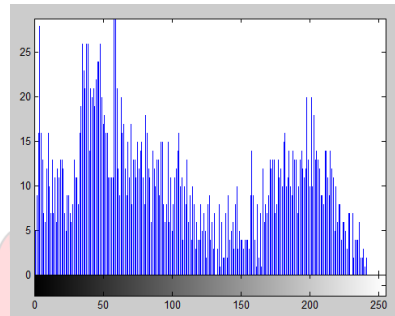
Gambar 1 Contoh citra *input*

Kemudian setelah diubah menjadi *greyscale* dan dilakukan *resize* sehingga menjadi citra dengan ukuran 50 x 50 maka didapatkan data sebagai berikut.

6	5	4	2	...	162
7	7	4	4	...	173
7	5	3	2	...	174
21	21	19	19	...	177

⋮	⋮	⋮	⋮		⋮
34	33	33	36	...	36

Dari data tersebut didapatkan histogram sebagai berikut:



Gambar 2. Contoh histogram

Gambar 2 menunjukkan histogram dari Gambar 1 dimana sumbu horizontal menunjukkan nilai *grey level* sedangkan sumbu vertikal menunjukkan nilai jumlah dari *grey level* tersebut. Dari hasil ini maka didapat vektor *input* untuk setiap citra berupa $X = (x_1, x_2, x_3, \dots, x_{256})$ dengan x_i merupakan *feature* yang didapat dari jumlah pada setiap *grey level*.

2. Proses Klasifikasi

Nilai $X = (x_1, x_2, x_3, \dots, x_{256})$ yang telah didapat dari proses sebelumnya, kemudian diklasifikasikan dengan menggunakan algoritma SVM atau lebih dikenal dengan proses *training SVM*. Pada dasarnya SVM merupakan algoritma yang hanya dapat mengklasifikasikan data yang bersifat linier. Untuk dapat mengklasifikasikan data non-linier,

algoritma SVM harus dimodifikasi dengan memasukkan fungsi kernel. Fungsi kernel disini berfungsi untuk memetakan data x dari *input space* ke ruang vektor yang berdimensi lebih tinggi (*feature space*) dengan fungsi Φ sehingga $\Phi : x \rightarrow \Phi(x)$. Dengan demikian, maka pada ruang vektor yang baru ini, *hyperplane* dapat dikonstruksikan dan data dapat dipisahkan (Scholkopf & Smola, 2001). Karena sulitnya menemukan fungsi transformasi dari Φ , maka perhitungan *dot product* tersebut dapat digantikan dengan fungsi kernel $K(x_i, x_j)$ dimana fungsi tersebut mendefinisikan transformasi Φ secara implisit. Inilah yang disebut dengan “*kernel trick*” (Hamel, 2009) dirumuskan dengan:

$$K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (1)$$

Dengan penambahan fungsi kernel tersebut, hasil klasifikasi dari data x dapat diformulasikan sebagai berikut:

$$\begin{aligned} f(x_d) &= \sum_{i=1}^{ns} \alpha_i y_i x_i x_d + b \\ &= \sum_{i=1, x_i \in SV}^{ns} \alpha_i y_i \langle \Phi(x_i), \Phi(x_d) \rangle + b \\ &= \sum_{i=1, x_i \in SV}^{ns} \alpha_i y_i K(x_i, x_d) + b \end{aligned} \quad (2)$$

Dengan $SV = Support Vector$ ialah subset dari *training set* yang terpilih sebagai *support vector*, dengan kata lain data x yang berkorespondensi pada $\alpha_i \geq 0$.

Faktanya, setiap data memiliki karakteristik yang berbeda-beda, sehingga tidak setiap fungsi kernel dapat digunakan dan memberikan hasil yang maksimal. Untuk itu diperlukan pendekatan lain yang dapat membuat fungsi kernel tersebut menjadi lebih fleksibel dan mampu digunakan pada tiap data yang ada, pendekatan ini disebut dengan metode *Multi Kernel Learning* (MKL) (Athoillah, dkk, 2015; Rakotomamonjy dkk, 2007),

Jika pada Persamaan (1) fungsi kernel didefinisikan dengan $K(x, x')$ maka pada pembelajaran dengan metode multi kernel, fungsi tersebut dapat digantikan oleh fungsi kombinasi dari beberapa fungsi kernel dengan $K_1(x, x'), K_2(x, x'), \dots, K_M(x, x')$ yang dirumuskan sebagai berikut:

$$K(x, x') = \sum_{m=1}^M d_m K_m(x, x') \quad (3)$$

dengan kendala

$$d_m \geq 0, \sum_{m=1}^M d_m = 1$$

Dimana M adalah jumlah dari fungsi kernel (Rakotomamonjy dkk, 2008). Pada penelitian ini fungsi kernel yang digunakan dan dikombinasikan diantaranya adalah fungsi kernel Polinomial order tiga, fungsi kernel RBF dan fungsi kernel Linier, berikut ini adalah persamaan dari ketiga fungsi tersebut:

Kernel Linier:

$$K(x_i, x_j) = x_i x_j^T \quad (4)$$

Kernel Polynomial:

$$K(x_i, x_j) = (x_i x_j^T + 1)^P \quad (5)$$

Kernel RBF:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (6)$$

PEMBAHASAN

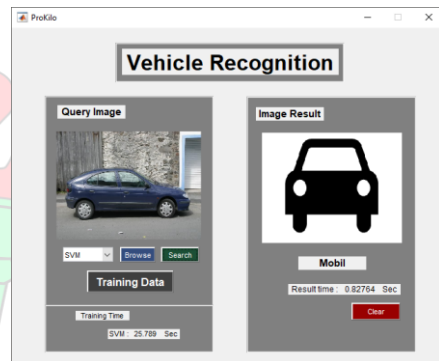
1. Dataset

Dataset yang digunakan dalam penelitian ini adalah dataset citra yang berupa citra kendaraan roda dua dan citra kendaraan roda empat, dimana masing-masing kelompok citra berjumlah 500 citra. Sehingga total dari citra dataset yang digunakan berjumlah 1000 citra. Dari jumlah total tersebut 90% digunakan untuk proses *training* sedangkan 10% sisanya digunakan sebagai data *testing*.

Proses validasi dari penelitian ini dilakukan dengan menggunakan metode *k-fold cross validation* (Mohri, dkk, 2012). Dimana dataset yang digunakan dibagi menjadi sepuluh bagian ($k = 10$), dari sepuluh bagian tersebut sembilan bagiannya digunakan sebagai kelompok data *training* dan satu kelompok digunakan sebagai kelompok data *testing*. Kemudian percobaan dilakukan sebanyak sepuluh kali ($k = 10$) dengan kombinasi kelompok data *training* dan *testing* yang berbeda pada tiap percobaannya.

2. Hasil Uji Coba

Hasil uji coba didapatkan dengan menghitung nilai akurasi, nilai presisi dan nilai *recall* dari sistem dalam menentukan citra *input* sebagai kendaraan beroda dua maupun beroda empat. Berikut ini adalah *user interface* dari sistem klasifikasi kendaraan kendaraan bermotor yang dibangun pada penelitian ini.



Gambar 3. User Interface Sistem

Nilai presisi merupakan perhitungan dari tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Adapun *recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi, sedangkan nilai akurasi didefinisikan dengan tingkat kedekatan antara nilai prediksi dengan nilai sebenarnya (Abe, 2010). Untuk lebih jelasnya perhatikan tabel dari *Confusion matrix* nilai klasifikasi yang dihasilkan sistem dengan nilai pengamatan yang sebenarnya berikut:

Tabel 1. *Confusion matrix*

Klasifikasi sistem	
<i>Benar</i>	<i>Salah</i>

Sebenarnya	Benar	True Positif (TP)	False Negative (FN)
	Salah	False Positive (FP)	True Negative (TN)

$$\text{Akurasi} = \frac{TP + TN}{\text{Jumlah keseluruhan data}} \times 100\%$$

$$\text{Presisi} = \frac{TP}{TP + FP} \times 100\%$$

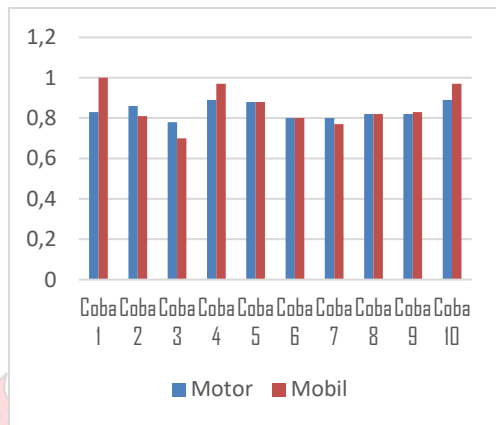
$$\text{Recall} = \frac{TP}{TP + FN} \times 100\%$$

Berikut adalah tabel yang menunjukkan nilai rata-rata dari hasil percobaan pada penelitian ini:

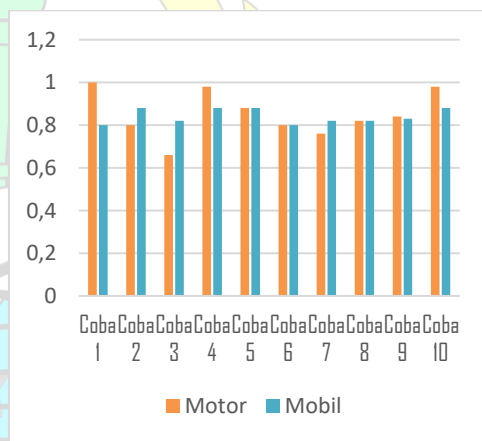
Tabel 2. Nilai rata-rata hasil klasifikasi

Kelas	Presisi	Recall	Akurasi
Motor	84.02%	85.20%	84.60%
Mobil	85.88%	84.00%	84.60%
Rata2	84.95%	84.60%	84.60%

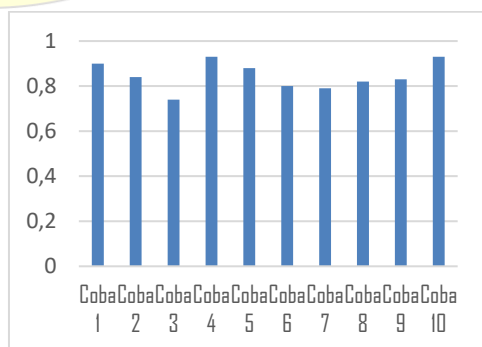
Berdasarkan Tabel 2 terlihat bahwa sistem mampu mengklasifikasikan kendaraan bermotor dengan baik. Hal ini dibuktikan dengan nilai akurasi yang mencapai 84.60%, sedangkan nilai rata-rata presisi dan recall yang mencapai 84.95% dan 84.60% menunjukkan konsistensi dari tingkat ketepatan sistem dalam mengenali kendaraan bermotor pada seluruh percobaan adalah baik. Sebagai pemahaman lebih lanjut bagaimana kemampuan sistem dalam mengklasifikasikan kendaraan bermotor dalam penelitian ini, berikut adalah data hasil uji coba sistem secara keseluruhan pada setiap percobaan.



Gambar 4. Grafik nilai presisi sistem secara keseluruhan



Gambar 5. Grafik nilai recall sistem secara keseluruhan



Gambar 6. Grafik nilai akurasi sistem secara keseluruhan

Terlihat pada Gambar 4, Gambar 5, dan Gambar 6, grafik nilai presisi, *recall*, dan akurasi sistem secara keseluruhan pada tiap percobaan konsisten menunjukkan nilai di persekitaran 80%. Hal ini menunjukkan bahwa sistem secara stabil mampu mengklasifikasikan semua dataset citra kendaraan bermotor dengan sangat baik pada tiap percobaannya.

PENUTUP

Simpulan

Pada penelitian ini dibangun sebuah sistem yang mampu mengklasifikasikan kendaraan bermotor roda dua dan roda empat dengan menggunakan algoritma Support Vector Machine (SVM) yang dimodifikasi dengan menambahkan fungsi multi kernel didalamnya, fungsi kernel tersebut dibentuk dengan memformulasikan beberapa fungsi kernel diantaranya kernel Polinomial order tiga, kernel RBF dan kernel Linier menjadi sebuah fungsi kernel tunggal, metode ini disebut dengan metode *Multi Kernel Learning* (MKL). Pada penelitian ini, percobaan dilakukan dengan menggunakan dataset berupa citra kendaraan bermotor roda dua dan roda empat sebanyak masing-masing 500 citra, yang mana dari jumlah tersebut 90% darinya digunakan sebagai data *training* dan 10% menjadi data *testing*. Validasi dari penelitian ini dilakukan dengan

menggunakan metode *k-fold cross validation* dimana data dibagi menjadi 10 bagian ($k=10$) yang kemudian dilakukan percobaan sebanyak 10 kali pula dengan kombinasi data *testing* dan *training* yang berbeda tiap percobaannya.

Hasil dari percobaan menunjukkan bahwa pada penelitian ini, sistem klasifikasi yang dibangun dengan menggunakan algoritma SVM multi kernel memiliki performa yang baik dengan rata-rata nilai akurasi yang mencapai 84.60%, rata-rata nilai presisi adalah 84.95% serta rata-rata nilai *recall* yang mencapai 84.60%.

Saran

Untuk penelitian selanjutnya dapat dilakukan dengan menggunakan dataset yang berbeda, misalnya dengan mengklasifikasikan suara, video, maupun data yang lainnya, atau penelitian dapat dilakukan dengan mengkombinasikan fungsi kernel yang lain selain dari fungsi yang telah digunakan dalam penelitian ini.

DAFTAR PUSTAKA

- Abe, S. (2010), *Support Vector Machines for Pattern Classification*, Springer-Verlag, London, UK.
- Athoillah, M. (2017). Pengenalan Wajah Menggunakan SVM Multi Kernel dengan Pembelajaran yang Bertambah. *Jurnal Online*

- Informatika (JOIN)*, Vol 2, No 2, hal-84-91
- Athoillah, M., Irawan, M.I., dan Imah, E.M. (2015). Study Comparison of SVM-, K-NN- and Backpropagation-Based Classifier for Image Retrieval. *Journal of Computer Science and Information*, Vol 8, No 1, hal-11-19
- Athoillah, M., Irawan, M.I., and Imah, E.M, (2015). Support Vector Machine with Multiple Kernel Learning for Image Retrieval. *Proceeding of International Conference on Information, Communication Technology and System (ICICTS)*, ITS, Surabaya.
- Campbell, C dan Ying, Y. (2011). Learning with Support Vector Machines, Buku seri *Synthesis Lectures on Artificial Intelligence and Machine Learning*, Morgan & Claypool Publisher, UK
- Clarke, B., Fokoue, E., dan Zhang, H.H. (2009). Principles and Theory for Data Mining and Machine Learning, *Springer Science + Business Media*, New York, USA.
- Hamel, L. (2009), *Knowledge Discovery with Support Vector Machine Learning*, John Wiley & Sons, Inc, New Jersey, USA
- Li, D., Wang, J., Zhao, X., Liu, Y., dan Wang, D. (2014). Multiple Kernel-Based Multi-Instance Learning Algorithm for Image Classification, *Jurnal Visual Communication and Image Representation*, Vol 25, No 5, Hal 1112-1117
- Mohri, M., Rostamizadeh, A. dan Talwalkar, A. (2012), *Foundations of Machine Learning*, Buku seri *The Adaptive Computations and Machine Learning*, MIT Press, Massachusetts Institute of Technology, USA.
- Rakotomamonjy, A., Bach, F.R., Canu, S. Dan Grandvalet, Y. (2008), "SimpleMKL", *Journal Machine Learning Research* 9, hal 2491-2521
- Rakotomamonjy, A., Bach, F.R., Canu, S. Dan Grandvalet, Y. (2007), "More Efficiency in Multiple Kernel Learning", *Proceesings of 24th International Conference on Machine Learning*, Corvallis, OR, USA. hal 775-782
- Scholkopf, B dan Smola, A.j. (2001), *Learning with Kernels*, Buku seri *The Adaptive Computations and Machine Learning*, MIT Press, Massachusetts Institute of Technology, USA.
- Shrivastava, A., Pillai, J.K., dan Patel, V.M. (2015). Multiple Kernel-based Dictionary Learning for Weakly Supervised Classification, *Jurnal Pattern Recognition*, Vol 48, No 8, Hal 2667-2675.