

SUPPORT VECTOR MACHINE BERBASIS ZERNIKE MOMENTS DALAM KLASIKASI CITRA BAKTERI

Yupie Kusumawati¹, Ibnu Utomo Wahyu Mulyono²

¹Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

²Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

e-mail: ¹yupie@dsn.dinus.ac.id, ²ibnu.utomo.wm@dsn.dinus.ac.id

ABSTRAK

Kontaminasi bakteri patogen dalam produk makanan mahal bagi masyarakat dan industri. Metode tradisional untuk mendeteksi dan mengidentifikasi patogen bawaan makanan seperti *Listeria monocytogenes* biasanya memakan waktu 3-7 hari. Pada makalah ini, sistem klasifikasi dikembangkan untuk mengidentifikasi pengambilan citra bakteri menggunakan alat optical light scattering dan menghasilkan citra berbentuk grayscale. Algoritma klasifikasi yang diusulkan didasarkan pada Invariant Zernike Moment berbasis Support Vector Machine pada kernel Radial Chebyshev Moments yang dihitung dari dataset citra bakteri apda 4 genus yang digunakan sebagai dataset. Sebanyak 400 citra bakteri dengan 100 citra pada masing-masing jenis genus telah di uji dan menghasilkan akurasi pada proses indentifikasi dengan capaian sebesar 78,33% pada 5-fold Cross Validation.

KATA KUNCI: Bakteri pathogen, Support Vector Machine, Invariant Zernike Moment, Radial Chebyshev Moment.

1. PENDAHULUAN

Identifikasi bakteri patogen secara cepat dan spesifik penting untuk aplikasi dalam biodeteksi, terutama dalam mengurangi insiden keracunan makanan atau melawan bioterorisme. Identifikasi konvensional bakteri patogen biasanya memakan waktu 24 hingga 48 jam setelah pertumbuhan bakteri pada cawan agar atau dalam kaldu kultur, karena tingkat pengujian biokimia atau molekuler yang diperlukan untuk memastikan identitas kultur bakteri [1]. Oleh karena itu, teknologi *forward lightscattering* digunakan dalam penelitian ini sebagai metode baru untuk identifikasi koloni bakteri pada pelat agar tanpa memerlukan pemrosesan atau preparasi sampel. Koloni yang dianalisis adalah genus *Listeria*, patogen intraseluler *Gram-positif*. Genus ini terdiri dari enam spesies: *L. monocytogenes*, *L. innocua*, *L. ivanovii*, *L. seeligeri*, *L. welshimeri*, dan *L. grayi*. Spesies patogen adalah *L. monocytogenes* pada manusia dan *L. ivanovii* pada hewan serta sisanya dianggap nonpatogenik [2]–[4].

Masalah penting dalam analisis pola adalah pengenalan otomatis suatu objek dengan menganalisa letak objek, ukuran, dan orientasi. Salah satu bentuk pengenalan pola dalam bentuk dua dimensi yaitu menggunakan *Invariant Zernike Moment* atau dikenal dengan *Zernike Moment* [1], [5], [6]. Hasil proses pengenalan awal pada *Zernike Moment* kemudian digunakan sebagai inputan pada proses klasifikasi dengan tujuan mendapatkan pelabelan dengan akurat meskipun menggunakan fitur *Zernike* bukan satu-satunya factor penentu proses klasifikasi. Dalam penelitian ini, kami menggambarkan kalsifikasi bakteri menggunakan ekstraksi fitur yang efisien, berdasarkan *Invariant Zernike Moment*, yang mampu mengukur perbedaan hamburan bakteri. Identifikasi *scatterograms* bakteri dilakukan melalui analisis bentuk yang diimplementasikan menggunakan teknik pengolahan citra digital. Proses ini melibatkan tiga langkah utama: pra-pemrosesan citra terutama peningkatan citra; misalnya, pemerataan histogram adaptif, ekstraksi fitur yang memanfaatkan ekstraksi *Zernike Moment* dari segmentasi pola, dan klasifikasi [4].

Support Vector Machine (SVM) dan *Jaringan Syarat Tiruan (JST)* adalah metode populer di bidang pengenalan multi-target [2]. Algoritma ini umumnya didasarkan pada karakteristik morfologi objek sehingga dapat memperoleh hasil yang lebih baik. Namun, JST sulit untuk

mencapai hasil yang diinginkan ketika data sampel terbatas dalam aplikasi praktis. SVM berbeda dengan JST, karena diusulkan berdasarkan teori prinsip minimisasi risiko struktural dalam sistem statistik yang sesuai dengan sampel yang terbatas [4]. SVM adalah algoritma pembelajaran terawasi berdasarkan konsep *hyperplane* yang bertujuan untuk memisahkan sekumpulan objek dengan margin maksimum [3]. SVM menentukan beberapa support vector dari *feature space* yang berguna untuk menentukan *hyperplane* yang optimal. SVM memiliki kemampuan dan keunggulan yang dapat diprediksi dengan baik dan cocok diterapkan dalam pengujian makanan untuk sampel terbatas [7]–[9].

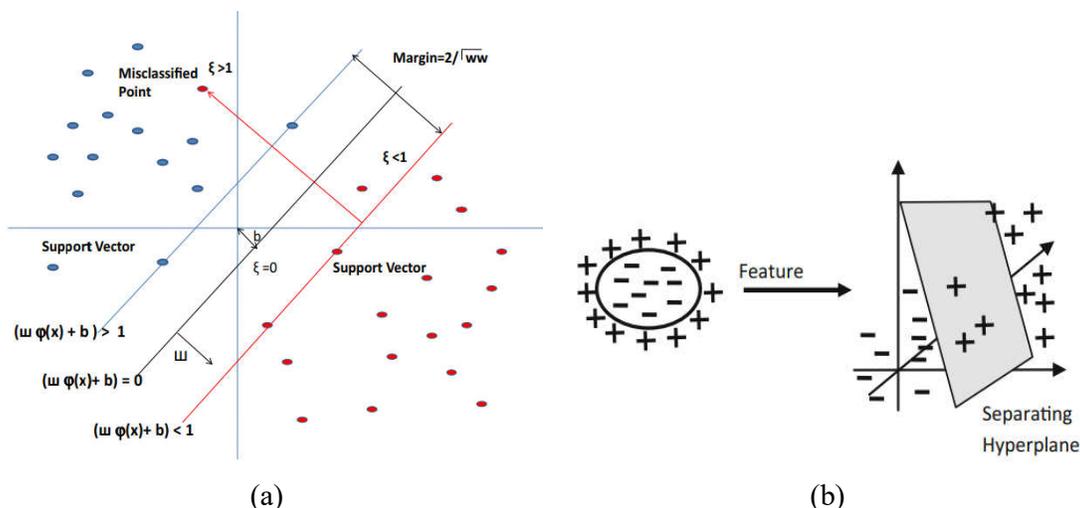
Beberapa pendekatan populer lainnya seperti KNN, distribusi *Gaussian*, histogram fitur template, CNN, fitur SIFT dan *bag of visual words* sudah pernah digunakan untuk mengklasifikasi bakteri *pathogen* [1], [10]. Momen ortogonal dapat bersifat diskrit maupun kontinu. Momen diskrit utama yang sering digunakan adalah *Krawtchouk moments*, *Tchebichef moments*, *dual Hahn moments*, *Racah moments* serta *Orthogonal moments* seperti *Zernike moments*, *Legendre moments*, *pseudo-Zernike moments*, *orthogonal Fourier–Mellin moments* dan *Gaussian–Hermite moments*.

Menurut penelitian yang telah dilakukan oleh [11], prediksi urutan protein virulen bakteri memiliki implikasi untuk identifikasi dan karakterisasi faktor-faktor terkait virulensi baru, menemukan target obat/vaksin baru terhadap protein yang sangat diperlukan untuk patogenitas, dan memahami mekanisme virulensi yang kompleks pada patogen. Dalam penelitian ini, SVM mampu mencapai akurasi 81,8% yang mencakup area 86% di *Receiver Operator Characteristic (ROC)*, lebih baik daripada salah satu pengklasifikasi SVM lapisan satu berdasarkan fitur urutan tunggal atau ganda. Menurut [2], SVM menunjukkan akurasi tinggi dalam membedakan bakteri patogen dari non-patogen, dalam kumpulan data yang relevan secara klinis, yang hanya terdiri dari bakteri yang diinangi manusia. Di antara gen yang menerima bobot positif tertinggi dalam pengklasifikasi yang dihasilkan, kami menemukan gen yang diketahui terkait dengan patogenitas bakteri, selain kandidat baru, yang keterlibatannya dalam virulensi bakteri tidak pernah dilaporkan. Sedangkan menurut penelitian yang dilakukan oleh [3], telah membuktikan bahwa, dibandingkan dengan hasil yang dikenali oleh mata manusia, SVM dapat secara efektif membedakan bakteri dari non-bakteri dalam gambar, dan sangat mengurangi waktu deteksi setiap sampel

2. TINJAUAN PUSTAKA

2.1 Support Vector Machine (SVM)

Saat ini tujuan penelitian SVM difokuskan pada penerapannya pada ranah praktis. Klasifikasi citra adalah tugas penting untuk analisis citra di berbagai bidang seperti studi lingkungan, aplikasi militer, ilmu hidrologi, sektor pertanian, penginderaan jauh. Menurut [2], [8], [12]–[14] memperkenalkan skema inovatif untuk klasifikasi citra hiper-spektral seperti tampak pada Gambar 1. Pada penelitian ini, SVM dapat meningkatkan akurasi klasifikasi dengan menggunakan tekstur dan fitur spektral. Metode ini hanya memberikan akurasi yang lebih baik dengan dataset yang lebih kecil. Invarian momen digunakan untuk mengekstrak fitur spasial dari suatu gambar. Fitur tekstur dan fitur spasial keduanya digunakan untuk melakukan tugas klasifikasi menggunakan SVM. Jona Inisai [2], menggunakan konsep metode kontur aktif dan SVM. Penelitian ini melakukan pengelompokan gambar 3D otak untuk analisis volumetrik. Segmen pertama jaringan otak dari jaringan intrakranial pada gambar MR. Kontur aktif VFC memodifikasi citra otak dan mengekstrak fitur untuk klasifikasi. Pengklasifikasi SVM dilatih untuk setiap jaringan. Hasil penelitian menunjukkan bahwa algoritma ini menghasilkan hasil yang baik dibandingkan dengan pendekatan lain [8].



Gambar 1. (a) Model Pemisahan Masalah pada SVM, (b) Pemisahan Hyperplane SVM

SVM sangat bagus ketika pengguna tidak mengetahui mengenai data yang digunakan. SVM dapat bekerja dengan baik bahkan dengan data tidak terstruktur dan semi terstruktur seperti teks, citra, dan *tree*. Dengan fungsi kernel yang sesuai, pengguna dapat memecahkan masalah yang kompleks. Tidak seperti di jaringan saraf, SVM tidak diselesaikan untuk optima lokal. Skala SVM relatif baik untuk data dimensi tinggi. Model SVM memiliki generalisasi dalam praktiknya sehingga risiko over-fitting lebih sedikit di SVM. Jika dibandingkan dengan model ANN, SVM memberikan hasil yang lebih baik.

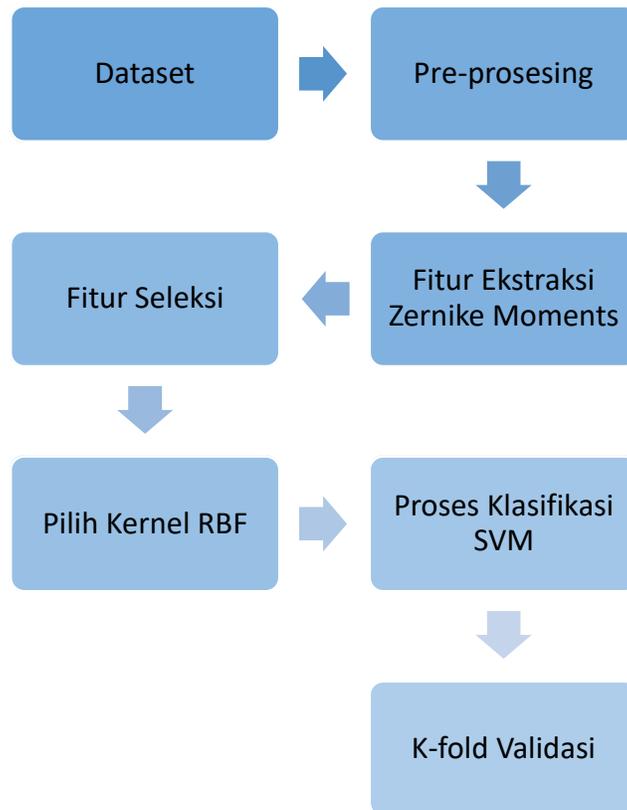
2.2 Invariant Zernike Moment

Fitur mewakili properti skalar dari objek yang dicitrakan. Dalam banyak kasus, sejumlah kecil fitur tidak cukup untuk menangkap informasi yang cukup tentang suatu objek untuk memungkinkan klasifikasi otomatis. Fitur seperti area, perimeter, kekompakan, dan deskriptor *Fourier* terlalu sederhana dan terlalu terbatas agar berguna untuk banyak jenis gambar. Oleh karena itu, fitur yang lebih kompleks misalnya geometrik dan *invariant* momen Zernike [1], [6]. Secara umum, invarian momen adalah alat untuk mengukur karakteristik batas dan konten dalam gambar. Sifat invarian momen yang diinginkan adalah ortogonalitas, sensitivitas rendah terhadap noise, kompleksitas komputasi rendah, konten informasi tinggi, dan invarians di bawah fotometrik dan transformasi geometris misalnya, transformasi affine yang mencakup penskalaan, rotasi, translasi. Pemilihan yang tepat dari berbagai fitur yang akan digunakan untuk pengenalan objek adalah salah satu faktor kunci dalam keberhasilan keseluruhan proses identifikasi. Polinomial yang berbeda menyebabkan berbagai jenis momen. *Zernike Moment* diwakili oleh satu set polinomial kompleks, yang membentuk set ortogonal lengkap di bagian dalam lingkaran [4], [5]. Untuk citra digital, integral pada persamaan (1) dapat diubah menjadi penjumlahan dimana $f(x,y)$ merupakan fungsi citra. Selain itu, koordinat dari citra akan dinormalisasi menjadi [0,1] dengan transformasi pemetaan.

$$\begin{aligned}
 Znm &= \frac{n+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) [V^*nm(x,y)] \\
 &= \frac{n+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) R_{n,m}(\rho_{xy}) \exp(-jm\theta_{xy}) \quad (1)
 \end{aligned}$$

3. METODE PENELITIAN

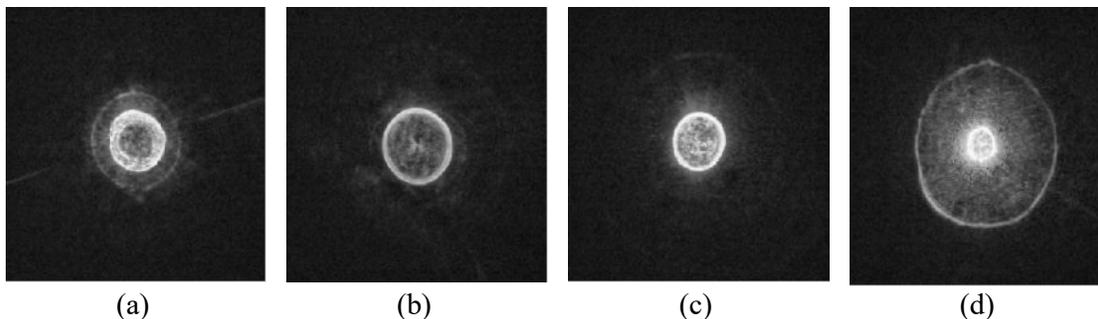
Pada bagian ini, proses klasifikasi citra bakteri patogen dapat di lihat pada Gambar 2.



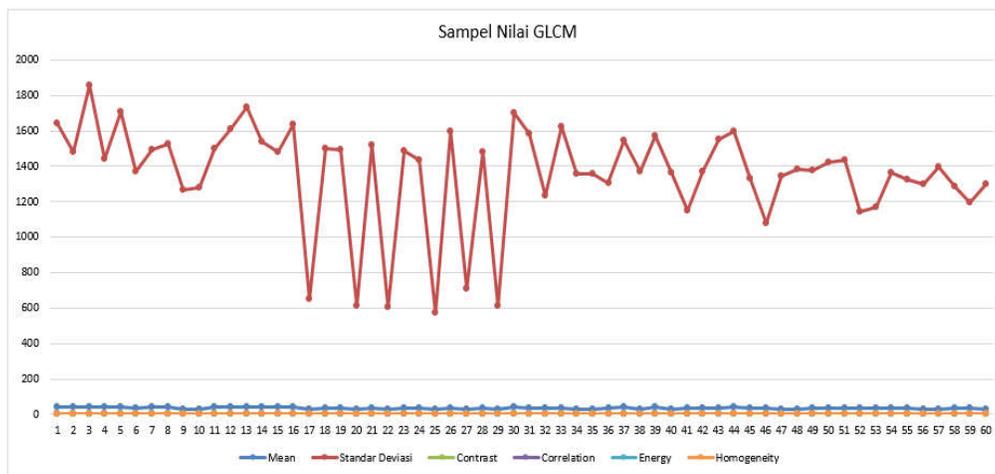
Gambar 2. Skema Usulan Metode Klasifikasi Bakteri Patogen Menggunakan SVM

Pada Gambar 2, SVM dikenal beberapa jenis kernel antara lain linear, polynomial dan Radial Basis Function (RBF). Pada penelitian ini, kami menggunakan RBF. RBF umumnya mampu memberikan tingkat akurasi klasifikasi yang sama dengan SVM. Salah satu keuntungan penting dari algoritma pembelajaran yang diusulkan, dibandingkan dengan mesin vektor dukungan, adalah bahwa algoritma pembelajaran yang diusulkan biasanya membutuhkan waktu yang jauh lebih sedikit untuk mengetahui nilai parameter yang optimal dengan validasi silang. Perbandingan dengan SVM menarik, karena telah ditunjukkan dalam sejumlah penelitian terbaru bahwa SVM secara umum mampu memberikan tingkat akurasi yang lebih tinggi daripada algoritma klasifikasi data lain yang ada.

Pada makalah ini, dataset awal sejumlah 1200 citra yang terdiri dari 4 kelas bakteri *Listeria*, dimana masing-masing kelas terdiri dari 400 data yang diambil dari <https://www.kaggle.com>. Pada penelitian ini digunakan 300 data pada setiap kelas sebagai data testing seperti tampak pada Gambar 3 berikut.

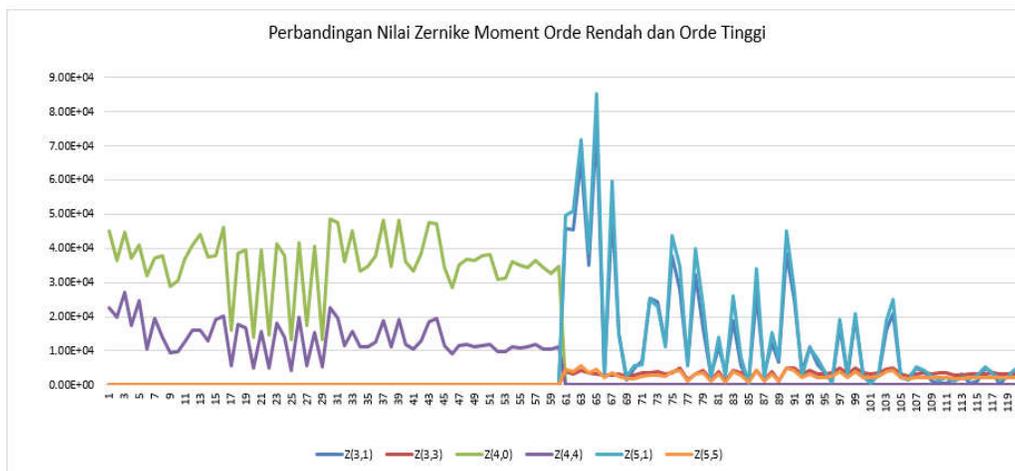


(a) (b) (c) (d)
 Gambar 3. Representasi citra bakteri *Listeria*: (a) *L. monocytogenes* GG8, (b) *L. welshimeri* 105(2)-2L, (c) *L. innocua* C3-3(L), (d) *L. ivanovii* V12



Gambar 4. Sampel nilai GLCM

Pada Gambar 4, dapat dilihat nilai GLCM pada sampel citra dataset bakteri. Dari 1200 data yang digunakan, kami tampilkan 60 sebagai gambaran hasil pemrosesan GLCM. Pada ekstraksi ciri tekstur, digunakan mean dan standard deviasi pada ekstraksi ciri tekstur orde pertama. Perhitungan GLCM menggunakan dengan jarak $d = 1$ dan menghitung rata-rata pada sudut 0° , 45° , 90° dan 135° untuk menghitung nilai contrast, correlation, energy dan homogeneity.



Gambar 5. Sampel Perbandingan Nilai Zernike Moment Orde Rendah dan Orde Tinggi

Berdasarkan Gambar 5, tampak bahwa nilai Zernike Moment pada orde rendah dan tinggi cukup berbeda. Pemilihan masukan (input) pada Zernike moment dipilih secara manual, yaitu dengan mengambil nilai dengan order rendah dan nilai dengan order tinggi. Pemilihan nilai order dilakukan karena mengingat Zernike moment memiliki komputasi tinggi. Pada orde tinggi nilai, Zernike Moment diimplementasikan pada nilai (18,0), (18,18), (19,1), (19,19), (20,0), dan (20,20), sedangkan pada orde rendah diimplementasikan pada nilai (3,1), (3,3), (4,0), (4,4), (5,1), (5,5). Nilai Zernike moment dengan orde rendah, dapat digunakan untuk mendapatkan fitur global suatu bentuk, sedangkan polinomial dengan orde tinggi dapat memperoleh rincian dari fitur lokal. Perhitungan ekstraksi ciri tekstur dilakukan dengan citra biner sebagai masukannya.

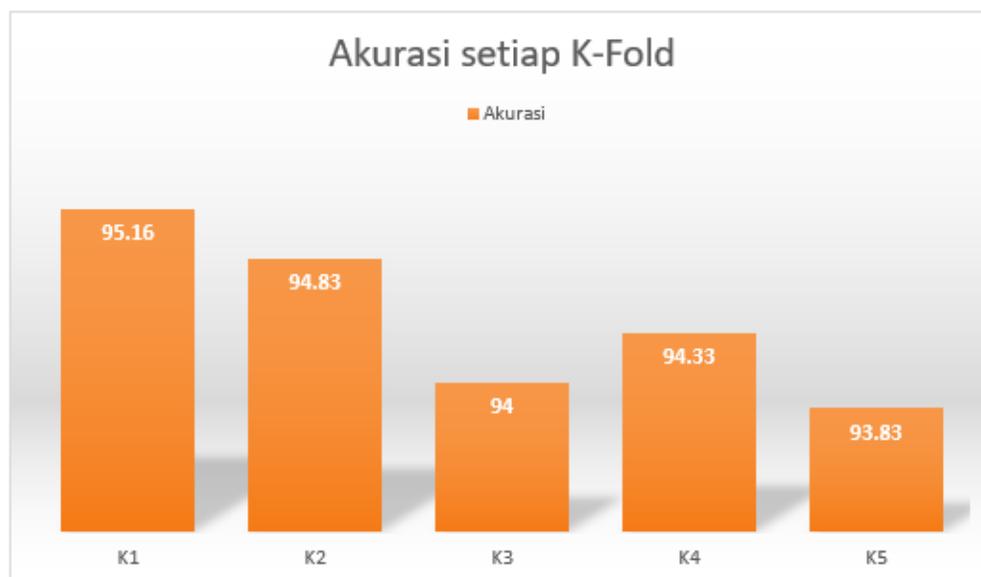
Untuk mengetahui performa dari usulan metode yang telah kami gunakan, digunakan perhitungan *confusion matrix* pada 5 *fold* validasi dari *fold* ke 1 sampai *fold* ke 5. *Confusion matrix* dapat di lihat pada Tabel 1 sedangkan akurasi hasil dapat di lihat pada Gambar 4. Berdasarkan Tabel 1, diketahui bahwa tingkat pencocokan citra pada proses klasifikasi memperoleh nilai cukup tinggi. Di sisi lain, masih terdapat kesalahan pada proses klasifikasi antara 1 sampai 7 buah citra. Pada citra *L. monocytogenes GG8* dan *L. welshimeri 105(2)-2L* saling terbalik dalam proses pengenalan citra dan menghasilkan nilai kesalahan yang sama yaitu 5.

Tabel 1. Hasil perhitungan Confusion Matrix

	Hasil Klasifikasi			
	<i>L. monocytogenes GG8</i>	<i>L. welshimeri 105(2)-2L</i>	<i>L. innocua C3-3(L)</i>	<i>L. ivanovii V12</i>
<i>L. monocytogenes GG8</i>	286	1	6	5
<i>L. welshimeri 105(2)-2L</i>	7	292	3	5
<i>L. innocua C3-3(L)</i>	5	7	290	3
<i>L. ivanovii V12</i>	2	0	1	287

Selanjutnya untuk akurasi dari setiap *fold*-validasi menurut persamaan (2) dapat dilihat pada Gambar 4.

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{2} \\
 &= \frac{286 + 292 + 290 + 287}{286 + 292 + 290 + 287 + 7 + 5 + 2 + 1 + 7 + 6 + 3 + 1 + 5 + 5 + 3} \times 100\% \\
 &= \frac{1155}{1200} \times 100\% = 96,25\%
 \end{aligned}$$



Gambar 6. Hasil akurasi pada setiap K-Fold

Berdasarkan Gambar 6, diketahui bahwa semua akurasi fold menghasilkan nilai cukup tinggi yaitu diatas 90%. Nilai fold terendah ada pada k3 sedangkan fold tertinggi pada k1. Nilai fold dipengaruhi oleh penggunaan fitur ekstraksi ciri yaitu Zernike Moment dan penggunaan kernel RBF pada SVM.

5. KESIMPULAN

Pada penelitian ini telah dilakukan proses klasifikasi citra bakteri patogen *Listeria* pada 4 buah spesies yaitu *L. monocytogenes GG8*, (*b*) *L. welshimeri 105(2)-2L*, (*c*) *L. innocua C3-3(L)*, (*d*) *L. ivanovii V12*. Sebanyak 1200 citra dataset kemudian dipilih menjadi 900 data testing pada SVM. Kernel SVM yang digunakan yaitu RBF yang dikombinasikan dengan preprosesing ekstraksi ciri *Zernike Moment*. Dengan menggunakan akurasi, dihasilkan nilai tinggi pada 96,25% sedangkan hasil uji pada *5-fold* menghasilkan akurasi tertinggi yaitu pada $k1 = 95,16\%$ dan terendah pada $k3 = 94\%$.

6. SARAN

Pada penelitian lebih lanjut, sebaiknya digunakan jenis data misalnya data original, data CLAHE, dan data hasil preprosesing. Dapat pula di uji dengan varian proporsi dataset sehingga dapat menghasilkan akurasi yang lebih optimal. Algoritma dapat dikombinasikan menggunakan preprosesing, fitur ekstraksi ciri atau warna dan metode segmentasi citra lain misalnya morfologi atau deteksi tepi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada **LPPM Universitas Dian Nuswantoro** yang telah memberi “**dukungan financial**” terhadap penelitian ini khususnya pada skema penelitian Penelitian Terapan Perguruan Tinggi (PTPT) Tahun Anggaran Genap 2020/2021.

DAFTAR PUSTAKA

- [1] B. Bayraktar, P. P. Banada, E. D. Hirleman, A. K. Bhunia, J. P. Robinson, and B. Rajwa, “Bacterial phenotype identification using Zernike moment invariants,” in *Advanced Biomedical and Clinical Diagnostic Systems IV*, 2006, vol. 6080, no. 765, p. 60800V.
- [2] J. I. R. P, S. C, and V. S, “AN IMPLEMENTATION OF SVM CLASSIFIER FOR BACTERIAL BLIGHT IDENTIFICATION,” *Eur. J. Mol. Clin. Med.*, vol. 07, no. 11, pp.

- 5586–5595, 2020.
- [3] R. Zhang, S. Zhao, Z. Jin, N. Yang, and H. Kang, “Application of SVM in the food bacteria image recognition and count,” in *2010 3rd International Congress on Image and Signal Processing*, 2010, vol. 4, pp. 1819–1823.
 - [4] W. M. Ahmed, B. Bayraktar, A. K. Bhunia, E. D. Hirleman, J. P. Robinson, and B. Rajwa, “Classification of Bacterial Contamination Using Image Processing and Distributed Computing,” *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 1, pp. 232–239, Jan. 2013.
 - [5] I. Gogul and V. S. Kumar, “Flower species recognition system using convolution neural networks and transfer learning,” *2017 4th Int. Conf. Signal Process. Commun. Networking, ICSCN 2017*, pp. 1–6, 2017.
 - [6] A. Khotanzad and Y. H. Hong, “Invariant Image Recognition by Zernike Moments,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, 1990.
 - [7] M. A. Chandra and S. S. Bedi, “Survey on SVM and their application in image classification,” *Int. J. Inf. Technol.*, Jan. 2018.
 - [8] J. Kim, B.-S. Kim, and S. Savarese, “Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines,” *Appl. Math. Electr. Comput. Eng.*, pp. 133–138, 2012.
 - [9] G. M. Foody and A. Mathur, “A relative evaluation of multiclass image classification by support vector machines,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1335–1343, Jun. 2004.
 - [10] C. E. Cabrera Ardila, L. Alberto Ramirez, and F. A. Prieto Ortiz, “Spectral analysis for the early detection of anthracnose in fruits of Sugar Mango (*Mangifera indica*),” *Comput. Electron. Agric.*, vol. 173, no. March, p. 105357, Jun. 2020.
 - [11] Y. Chen *et al.*, “Variety identification of orchids using Fourier transform infrared spectroscopy combined with stacked sparse auto-encoder,” *Molecules*, vol. 24, no. 13, 2019.
 - [12] S. K. Mishra and V. H. Deepthi, “Brain image classification by the combination of different wavelet transforms and support vector machine classification,” *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, Jul. 2020.
 - [13] S. Zhang, X. Wu, S. Zhang, Q. Cheng, and Z. Tan, “An effective method to inspect and classify the bruising degree of apples based on the optical properties,” *Postharvest Biol. Technol.*, vol. 127, pp. 44–52, 2017.
 - [14] P. B. Padol and A. A. Yadav, “SVM classifier based grape leaf disease detection,” *Conf. Adv. Signal Process. CASP 2016*, pp. 175–179, 2016.
 - [15] Y.-L. Bei, S. Qiao, M.-X. Liu, X.-R. Zhu, and Q. Zhang, “A Color Image Watermarking Scheme Against Geometric Rotation Attacks Based on HVS and DCT-DWT,” in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2018, pp. 343–347.