

Penentuan Metode Terbaik Dalam Menentukan Jenis Pohon Pisang Menurut Tekstur Daun (Metode *K-NN* dan *SVM*)

Ahmad Hudawi As.¹, Wali Ja'far Shudiq², M. Fadhilur Rahman³

Fakultas Teknik Universitas Nurul Jadid, Paiton, Probolinggo

Info Artikel

Riwayat Artikel

Diterima: 13-12-2020

Direvisi: 21-12-2020

Disetujui: 30-12-2020

Kata Kunci

Pohon pisang;

Data Mining;

K-Nearest Neighbor (K-NN);

Support Vector Machine

(*SVM*);

✉ Corresponding Author

Wali Ja'far Shudiq,

Tel. +62 85257767603

wali.jafar@unuja.ac.id

ABSTRAK

Di sejumlah masyarakat banyak ditemui berbagai jenis pohon pisang. Tidak hanya buahnya yang rasanya manis, tapi juga pohonnya bisa membantu penghijauan alam. Seringkali masyarakat kecewa saat pohon pisang yang ditanamnya tidak sesuai dengan yang diharapkan saat menanamnya. Hal ini bisa dimaklumi mengingat pohon pisang membutuhkan waktu yang lama untuk tumbuh sebelum berbuah. Maka akan lebih baik jika dapat diketahui sejak awal jenis pohon pisang tersebut berdasarkan komponen pohon yang mudah diamati yaitu tekstur daun. Metode yang digunakan adalah dua metode data mining yang klasifikasi yaitu *K-Nearest Neighbor (K-NN)* dan *Support Vector Machine (SVM)*, yang akan mencari model terbaik dari kedua metode tersebut, dalam mencari tingkat keakurasian yang paling tinggi. Hasil dari penelitian ini menunjukkan kinerja metode *K-Nearest Neighbor (K-NN)* dengan nilai akurasi mencapai 74,00% lebih baik dari hasil kinerja metode *Support Vector Machine (SVM)* dengan nilai akurasi mencapai 67,89%.

PENDAHULUAN

Pisang merupakan tumbuhan yang memiliki daun besar memanjang. Beberapa jenisnya menghasilkan buah yang dinamakan sama. Buah pisang tersusun dalam tandan dengan kelompok tersusun menjari. Hampir semua buah pisang memiliki kulit yang sama ketika matang, namun ada beberapa yang berwarna jingga, merah, hijau, ungu, atau bahkan berwarna hitam. Buah pisang juga merupakan sumber energi atau karbohidrat dan mineral. Selain sebagai sumber energi, buah pisang juga bermanfaat untuk kesehatan tubuh karena mengandung banyak vitamin [1].

Selain buah yang manis rasanya, pohonnya sendiri merupakan aset penghijauan alam. Di area tersebut seringkali masyarakat kecewa dengan pohon pisang yang ditanamnya karena tidak sesuai dengan yang diharapkan. Akan lebih bagus lagi jika masyarakat mengetahui jenis pohon pisang yang di tanamnya sebelum pohon pisang tersebut berbuah melalui tekstur daun pisang. Oleh karena itu, dibutuhkan suatu metode untuk menentukan jenis pohon pisang berdasarkan tekstur daun.

Pada penelitian ini diusulkan penggunaan dua metode, yaitu *K-NN* dan *SVM* untuk mengetahui metode yang terbaik dalam menentukan jenis pohon pisang menurut tekstur daun. Metode *k-nearest neighbor (k-NN* atau *K-NN*) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Ketepatan algoritma *K-NN* ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Metode *Support Vector Machine (SVM)* adalah sistem pembelajaran

Penentuan Metode Terbaik Dalam Menentukan Jenis Pohon Pisang Menurut Tekstur Daun (Metode K-NN dan SVM)

Ahmad Hudawi As, Wali Ja'far Shudiq, M. Fadhilur Rahman

yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan *learning* bias yang berasal dari teori pembelajaran statistik. Dari metode-metode di atas bertujuan untuk menentukan metode terbaik dalam menentukan jenis pisang menurut daun.

Image processing adalah ilmu yang mempelajari proses pengolahan gambar dimana baik masukan dan keluarannya berbentuk berkas digital. *Image processing* dilakukan untuk memperbaiki kesalahan data gambar yang terjadi akibat transmisi dan selama akuisisi data, serta untuk meningkatkan kualitas gambar supaya lebih mudah dilihat oleh manusia baik dengan melakukan manipulasi gambar atau analisa terhadap gambar [2].

Berdasarkan permasalahan di atas, maka judul penelitian ini adalah “Penentuan Metode Terbaik Dalam Menentukan Jenis Pohon Pisang Menurut Tekstur daun (Metode *K-NN* dan *SVM*)”. Dengan penelitian ini diharapkan dapat membantu masyarakat dalam menentukan metode terbaik untuk menentukan jenis pohon pisang, Sehingga didapat suatu metode dengan nilai yang akurat.

METODE

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [3].

Data mining adalah campuran dari “statistik, kecerdasan buatan, pengenalan pola, dan basis data” yang masih terus berkembang. Data dalam jumlah besar dapat diolah sehingga menjadi informasi untuk menyelesaikan persoalan yang tidak dapat diatasi secara tradisional. Pekerjaan data mining antara lain: Model prediksi, analisis *cluster*, analisis asosiasi, dan deteksi anomali [4].

a. Metode K-NN

K-Nearest Neighbor (K-NN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada *K-NN* [5]. Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel latih. Pengklasifikasian tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik uji, akan ditemukan sejumlah K objek (titik *training*) yang paling dekat dengan titik uji. Klasifikasi menggunakan *voting* terbanyak di antara klasifikasi dari K objek. Algoritma *K-NN* menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari sample uji yang baru. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak *Euclidian*. Algoritma metode *KNN* sangatlah sederhana, bekerja dengan berdasarkan pada jarak terpendek dari sample uji ke sample latih untuk menentukan *KNN* nya. Setelah mengumpulkan *KNN*. Kemudian diambil mayoritas dari *KNN* untuk dijadikan prediksi dari sample uji [6]. Jarak *Euclidian* dapat dicari dengan menggunakan persamaan (1):

$$d_1 = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (1)$$

b. Metode SVM

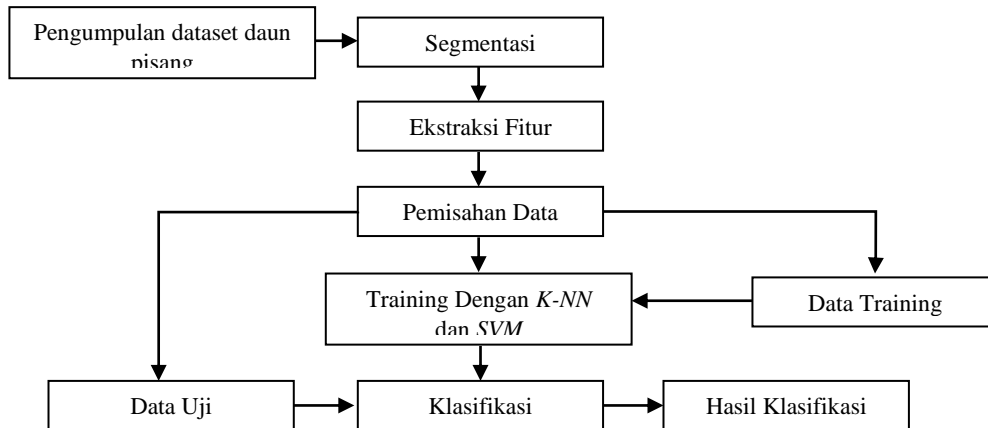
SVM adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization (SRM)* dengan tujuan menemukan hyperplane terbaik yang memisahkan dua buah class pada *input space* [7]. Tulisan ini membahas teori dasar *SVM* dan aplikasinya dalam bio informatika, khususnya pada analisa ekspresi gen yang diperoleh dari analisa *microarray*.

Pengertian yang lainnya adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan

mengimplementasikan *learning* bias yang berasal dari teori pembelajaran statistik [8]. Berikut penghitungan nilai *SVM* dapat dilihat pada persamaan (2) :

$$f(x_t) = \sum_{s=1}^{ns} \alpha_s y_s x_s \cdot x_t + b \quad (2)$$

Adapun alur dari tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Rancangan Penelitian

1. Pengumpulan Dataset

Dataset daun pisang terdiri dari 25 citra daun pisang dari masing – masing jenis. Citra tersebut diambil menggunakan kamera *smartphone* dengan resolusi gambar 1024 x 512. Untuk mempermudah proses implementasi, setiap daun diletakkan pada *background* berwarna putih dan difoto dari atas dengan jarak sekitar 1 meter. Dataset yang akan diteliti terdiri dari 25 foto dari masing-masing jenis pisang. Terdapat 2 jenis pisang yang akan diteliti, sehingga total foto dalam dataset tersebut berjumlah 50 foto. Hal ini dilakukan agar tingkat akurasi metode yang dikembangkan semakin akurat. Setiap 25 foto untuk satu jenis pisang, akan diambil dari 5 pohon yang berbeda-beda. Sehingga akan diambil 10 jenis dari setiap pohon pisang, untuk 5 pohon pisang yang sejenis. Jadi total pohon pisang yang harus dikunjungi pada tahap ini adalah 10 pohon pisang dari 2 jenis pohon pisang yang ada di kalangan masyarakat.

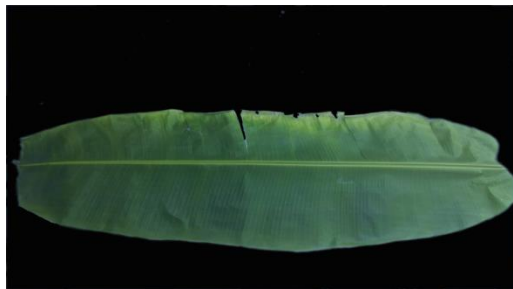
2. Segmentasi

Dalam pengolahan citra, terkadang menginginkan pengolahan hanya pada obyek tertentu. Oleh sebab itu, perlu dilakukan proses segmentasi citra yang bertujuan untuk memisahkan antara objek (*foreground*) dengan *background*. Pada umumnya keluaran hasil segmentasi citra adalah berupa citra biner di mana objek (*foreground*) yang dikehendaki berwarna putih (1), sedangkan *background* yang ingin dihilangkan berwarna hitam (0). Sama halnya pada proses perbaikan kualitas citra, proses segmentasi citra juga bersifat eksperimental, subjektif, dan bergantung pada tujuan yang hendak dicapai.

Pada segmentasi ini peneliti menggunakan aplikasi Adobe Photoshop untuk mengganti *background* menjadi hitam dengan menggunakan citra warna RGB yaitu #000000. Berikut hasil gambar sebelum di segmentasi dan sesudah di segmentasi.



Gambar 2. Daun Pisang yang Belum di Segmentasi



Gambar 3. Daun Pisang yang Sudah di Segmentasi

3. Ekstraksi Fitur

Pada tahap ini, penelitian ini melakukan otomatisasi klasifikasi 2 jenis pisang yaitu: pisang raja dan pisang raja sereh berdasarkan tekstur daun. Hal ini karena umumnya warna daun adalah hijau, maka peneliti menggunakan fitur warna hijau dari sistem warna RGB. Hasil analisis setiap fitur dari 2 pendekatan analisis tekstur dengan *Principal Component Analysis*(PCA). Didapatkan ada 25 fitur yang tepat untuk digunakan. Analisis citra dengan pendekatan statistik memberikan fitur rata-rata, *smoothness*, dan *entropy* nilai intensitas warna hijau pada daun yang merupakan karakter morfologi alami daun. Pendekatan *moment invariant* sudah terbukti *robust* terhadap pergeseran, perputaran dan penskalaan obyek dalam citra, maka tepat jika digunakan dalam penelitian ini untuk peletakan obyek dalam citra yang bisa dipengaruhi oleh 3 hal tersebut. Dari 7 fitur moment invariant, ada 5 yang mempunyai pengaruh yang signifikan yaitu moment: 1, 2, 4, 6, dan 7. Sedangkan pendekatan matrik *co-occurrence* memberikan alternatif teknik analisis citra selain pendekatan statistik, fitur yang diekstrak dari matrik ini adalah energi, dan kontras. Energi untuk mengukur kekonstanan intensitas, sedangkan kontras untuk mengukur tingkat perbedaan nilai intensitas warna hijau pada daun. Dari analisis dengan PCA, maka paramater tekstur yang digunakan adalah: rata-rata intensitas, *smoothness*, *entropy*, 5 komponen *moment invariants*, *energy*, kontras. Pengujian dilakukan pada 2 jenis daun pisang, yaitu: pisang raja dan pisang raja sereh. Masing – masing jenis diambil sampel 25 fitur daun dari 2 jenis daun pisang, sehingga total ada 50 fitur daun yang digunakan sebagai data set dalam penelitian.

Fitur pertama yang dihitung secara statistis adalah rerata intensitas. Komponen fitur ini dihitung berdasar persamaan (3)

$$m = \sum_{i=0}^{L-1} i \cdot p(i) \quad (3)$$

Dalam hal ini, i adalah aras keabuan pada citra f dan $p(i)$ menyatakan probabilitas kemunculan i dan L menyatakan nilai aras keabuan tertinggi. Rumus di atas akan menghasilkan rerata kecerahan objek.

Fitur kedua berupa deviasi standar. Perhitungannya menggunakan persamaan (4) sebagai berikut:

$$\sigma = \sqrt{\sum_{i=0}^{L-1} (i - m)^2 p(i)} \quad (4)$$

Dalam hal ini, σ^2 dinamakan varians atau momen orde dua ternormalisasi karena $p(i)$ merupakan fungsi peluang. Fitur ini memberikan ukuran kekontrasan.

Fitur *skewness* merupakan ukuran ketidaksimetrisan terhadap rerata intensitas. Perhitungannya menggunakan persamaan (5) :

$$skewness = \sum_{i=1}^{L-1} (i - m)^3 p(i) \quad (5)$$

Skewness sering disebut sebagai momen orde tiga ternormalisasi. Nilai negatif menyatakan bahwa distribusi kecerahan condong ke kiri terhadap rerata dan nilai positif menyatakan bahwa distribusi kecerahan condong ke kanan terhadap rerata. Dalam praktik, nilai *skewness* dibagi dengan $(L-1)^2$ supaya ternormalisasi.

Deskriptor energi adalah ukuran yang menyatakan distribusi intensitas piksel terhadap jangkauan aras keabuan. Perhitungannya menggunakan persamaan (6) :

$$energi = \sum_{i=0}^{L-1} [p(i)]^2 \quad (6)$$

Citra yang seragam dengan satu nilai aras keabuan akan memiliki nilai energi yang maksimum, yaitu sebesar 1. Secara umum, citra dengan sedikit aras keabuan akan memiliki energi yang lebih tinggi dari pada yang memiliki banyak nilai aras keabuan. Energi sering disebut sebagai keseragaman.

Entropi mengindikasikan kompleksitas citra. Perhitungannya menggunakan persamaan (7):

$$entropi = - \sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (7)$$

Semakin tinggi nilai entropi, semakin kompleks citra tersebut. Perlu diketahui, entropi dan energi berkecenderungan berkebalikan. Entropi juga merepresentasikan jumlah informasi yang terkandung di dalam sebaran data.

Properti kehalusan biasa disertakan untuk mengukur tingkat kehalusan/kekasaran intensitas pada citra. Perhitungannya menggunakan persamaan (8) :

$$R = 1 - \frac{1}{1 + \sigma^2} \quad (8)$$

Pada rumus di atas, σ adalah deviasi standar. Berdasarkan rumus di atas, Nilai R yang rendah menunjukkan bahwa citra memiliki intensitas yang kasar. Perlu diketahui, di dalam menghitung kehalusan, varians perlu dinormalisasi sehingga nilainya berada dalam jangkauan [0 1] dengan cara membaginya dengan $(L-1)^2$.

4. Pemisahan Data

Dari 50 data citra dibagi menjadi 2 kelompok yaitu 40 data training dan 10 data uji, komposisi yang digunakan menyesuaikan hasil analisis dengan teknik *K-fold Cross Validation*. Kenapa harus ada pemisahan data di penelitian ini, yaitu karena untuk membuktikan ke akurasian *output* dari *image processing* dan ekstraksi fitur apakah hasil dari keduanya benar dengan data *testing* yang sudah ada. Dalam pemisahan data ini terdapat data training dan data testing. Data training adalah data percobaan kembali untuk nilai yang muncul untuk memastikan apakah data itu sesuai dengan *output* nilai yang muncul. Sedangkan data *testing* adalah data yang nilai *output* sudah benar untuk proses pengklasifikasian daun.

5. Klasifikasi

Proses klasifikasi dilakukan dengan memproses satu persatu data uji untuk diketahui keluaran kelas yang diberikan oleh sistem. Pada *K-NN*, masing-masing data uji dilakukan

Penentuan Metode Terbaik Dalam Menentukan Jenis Pohon Pisang Menurut Tekstur Daun (Metode K-NN dan SVM)

Ahmad Hudawi As, Wali Ja'far Shudiq, M. Fadhilur Rahman

pengujian 3 kali yaitu: 1-NN, 3-NN, dan 5-NN untuk setiap *K-fold*, sedangkan SVM dilakukan pengujian 3 kali untuk masing-masing K fold. Kemudian hasilnya dilakukan pencocokan dengan kelas yang sesungguhnya sehingga diketahui akurasi sistem dalam melakukan klasifikasi.

HASIL DAN PEMBAHASAN

Dataset yang diambil berjumlah 50 *record* terhitung dari daun yang di foto, dan dataset keseluruhan akan dibagi 2 yaitu variable X sebagai atribut hasil ekstraksi dan variable Y sebagai label jenis pisang. Daun pisang terdiri dari daun pisang raja dan daun pisang raja sereh yang masing-masing 25 daun, sehingga total keseluruhan dataset berjumlah 50 daun pisang. Untuk lebih detailnya, perhatikan tabel 1 tentang karakteristik dataset daun pisang.

Tabel 1. Karakteristik Dataset Daun Pisang

No	Nama Atribut	Variable X	Keterangan
1	Mu	X1	Numerik
2	Deviasi	X2	Numerik
3	Skewness	X3	Numerik
4	Energi	X4	Numerik
5	Entropy	X5	Numerik
6	Smoothness	X6	Numerik
7	Label atau Variable Y	Pisang Raja, Pisang Raja Sereh	Nominal

a. Implementasi Metode K-Nearest Neighbor (K-NN)

Pada implementasi ini digunakan *Number Of Validation* untuk mencari parameter yang paling baik untuk sebuah algoritma klasifikasi, sedangkan nilai k yang tinggi akan mengurangi efek *noise* pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter, dengan menggunakan *cross validation*. Dengan demikian hasil dari validasi ke-2 sampai ke-10 dan k1 sampai k15. Ditunjukkan pada tabel 2.

Tabel 2. Hasil Akurasi dari Metode K-NN

	Metode K-NN							
	K1	K3	K5	K7	K9	K11	K13	K15
Ke-2	60,00%	36,00%	32,00%	32,00%	26,00%	28,00%	20,00%	24,00%
Ke-3	67,77%	64,09%	61,89%	57,84%	63,85%	57,97%	63,97%	57,84%
Ke-4	53,53%	53,53%	49,84%	45,99%	51,92%	37,98%	39,74%	36,06%
Ke-5	74,00%	68,00%	58,00%	60,00%	62,00%	70,00%	58,00%	62,00%
Ke-6	62,04%	53,94%	56,02%	48,15%	43,98%	43,98%	37,96%	47,69%
Ke-7	65,31%	59,69%	55,61%	49,74%	55,87%	49,74%	58,16%	58,16%
Ke-8	54,46%	54,46%	52,38%	48,21%	46,43%	42,26%	48,21%	50,60%
Ke-9	70,00%	61,85%	56,67%	47,04%	43,33%	49,26%	47,78%	58,15%
Ke-10	72,00%	66,00%	56,00%	54,00%	62,00%	54,00%	50,00%	56,00%

Pada metode K-NN, pengujian ini mencari akurasi dari validasi ke-10, dimana nilai k masing – masing validasi adalah 15, dan nilai k harus ganjil, karena mencari tetangga terdekat. Akurasi tertinggi adalah 74.00%, ada di validasi ke 5 dan nilai k 1.

b. Implementasi Metode SVM

Pada implementasi ini digunakan *Number Of Validation* untuk mencari parameter yang paling baik dari sebuah algoritma klasifikasi, sedangkan akurasi menunjukkan kedekatan hasil pengukuran dengan nilai sesungguhnya. Dengan demikian hasil dari validasi ke – 2 sampai ke – 10 ditunjukkan pada tabel 3.

Tabel 3. Hasil Akurasi Metode SVM

<i>Number Of Validation</i>	Akurasi
Ke-2	56.00%
Ke-3	67.89%
Ke-4	57.69%
Ke-5	58.00%
Ke-6	59.95%
Ke-7	50.00%
Ke-8	56.55%
Ke-9	53.70%
Ke-10	60.00%

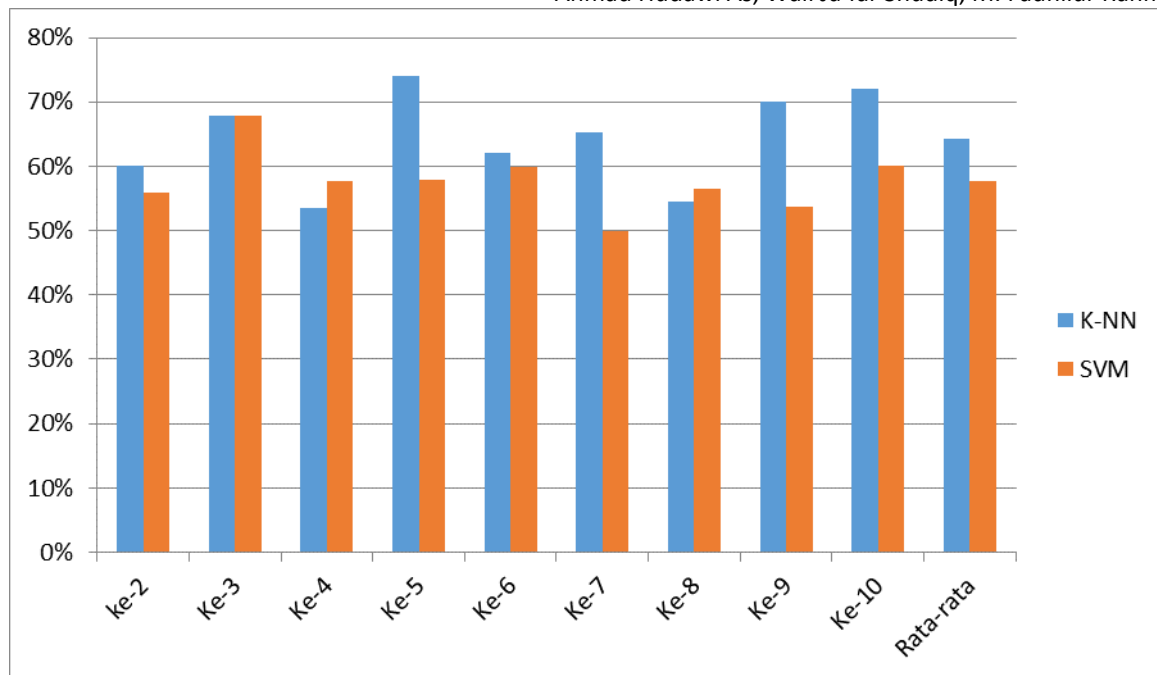
c. Analisis Hasil Komparasi

Setelah nilai akurasi sudah didapatkan dari dua metode *K-NN* dan *SVM*, maka dari penelitian ini akan dibandingkan akurasinya , hasil akurasi tertinggi pada dataset daun pisang. Dan akurasi tertinggi pada dataset daun pisang adalah metode *K-NN*, pada K1 dan validasi ke-5 dengan nilai akurasi 74,00%. Untuk lebih detailnya perhatikan tabel 4 perbandingan kedua metode yang sudah dipaparkan di atas

Tabel 4. Perbandingan Metode Sesuai Nilai Akurasi

Eksperimen	Metode	
	K-NN dan pada K1	SVM
Ke-2	60,00%	56,00%
Ke-3	67,77%	67,89%
Ke-4	53,53%	57,69%
Ke-5	74,00%	58,00%
Ke-6	62,04%	59,95%
Ke-7	65,31%	50,00%
Ke-8	54,46%	56,55%
Ke-9	70,00%	53,70%
Ke-10	72,00%	60,00%
Rata-rata	64,35%	57,75%

Dari hasil diatas akan dibuat model grafik dengan tujuan lebih diperjelas perbandingan dari metode *K-NN* dan *SVM* yang digunakan dalam penelitian ini. Kronologi gambar 3 perbandingan grafik dari kedua metode yang sudah dipaparkan di atas, sebagai berikut :



Gambar 4. Perbandingan Akurasi

Dari gambar 4 perbandingan akurasi menunjukkan warna biru memiliki nilai yang lebih bagus dari pada warna merah. Warna biru adalah metode *K-NN*.

SIMPULAN (PENUTUP)

Berdasarkan uraian di atas, tentang penentuan metode terbaik dalam menentukan jenis pohon pisang berdasarkan tekstur daun. Pisang yang diteliti ada dua jenis pisang yaitu pisang raja dan pisang raja seroh. Dimana peneliti menggunakan 2 metode, metode tersebut adalah *K-NN* dan *SVM*, yang tujuannya mencari akurasi dan akurasi yang tertinggi yang akan digunakan dalam penelitian ini sebagai acuan dari penentuan jenis pisang. Penelitian ini menggunakan k-fold validasi dari ke-2 sampai 10. Nilai akurasi tertinggi pada metode *K-NN* dengan akurasi **74,00%** terdapat pada k-fold validasi ke-5 dan nilai k 1, sedangkan nilai tertinggi dari metode *SVM* dengan nilai akurasi **67,89%**. Di k-fold validasi ke-3. Dan nilai rata – rata metode *K-NN* adalah **64,35%** dan metode *SVM* adalah **57,75%**. Jadi nilai akurasi tertinggi terdapat pada metode *K-NN*.

Secara matematis, penelitian ini menemukan metode *K-Nearest Neighbor (K-NN)* yang memiliki akurasi yang paling tinggi dibandingkan dengan metode *Support Vector Machine (SVM)*, mungkin pada peneliti selanjutnya bisa menemukan metode yang lain dimana akurasinya lebih tinggi dari sebelumnya. perlu dikembangkan atau ditingkatkan. Peningkatan yang perlu dilakukan bisa dalam mendapatkan nilai akurasi yang lebih tinggi dengan menambah variable atau indikator, biasanya dapat menggunakan optimasi atau *feature selection* atau yang lainnya.

DAFTAR PUSTAKA

- [1] D. A. Dewi, "Pengertian Pisang dan Manfaatnya," 2014. [Online]. Available: <http://dinaalizadewi.blogspot.com/2014/01/pengertian-pisang-dan-manfaatnya.html>.
- [2] A. Fajar, *Pengolahan Citra Digital : Konsep & Teori*. Andi Offset, 2013.
- [3] M. Riadi, "Pengertian, Fungsi, Proses dan Tahapan Data Mining," 2017. [Online]. Available: <https://www.kajianpustaka.com/2017/09/data-mining.html>. [Accessed: 18-

- Mar-2018].
- [4] A. Pamungkas, “Data Mining,” 2018. [Online]. Available: <https://pemrogramanmatlab.com/data-mining-menggunakan-matlab/>.
 - [5] W. J. Shudiq, “Penerapan K-Nearest Neighbor Berbasis Algoritma Genetika untuk Klasifikasi Mutu Padi Organik,” in *Prosiding SANTIF*, 2017.
 - [6] Y. Anggoro, B. D. Setiawan, and P. P. Adikara, “Implementasi Metode Fuzzy K-Nearest Neighbor Untuk Klasifikasi Penyakit Tanaman Kedelai Pada Citra Daun,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 6, pp. 2381–2389, 2018.
 - [7] S. AULIA, S. HADIYOSO, and D. N. RAMADAN, “Analisis Perbandingan KNN dengan SVM untuk Klasifikasi Penyakit Diabetes Retinopati berdasarkan Citra Eksudat dan Mikroaneurisma,” *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 3, no. 1, p. 75, Jan. 2015.
 - [8] S. . Permana, “Pengertian, Kelebihan dan Kekurangan SVM,” 2012. [Online]. Available: <http://cgeduntuksemua.blogspot.co.id/2012/03/pengertian-kelebihan-dan-kekurangan.htm>.