

## Terakreditasi SINTA Peringkat 3

Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi Nomor 225/E/KPT/2022 masa berlaku mulai Vol.7 No. 1 tahun 2022 s.d Vol. 11 No. 2 tahun 2026

Terbit online pada laman web jurnal:  
<http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>



Vol. 8 No. 1 (2023) 19 - 26

# JOINTECS

## (Journal of Information Technology and Computer Science)

e-ISSN:2541-6448

p-ISSN:2541-3619

### Klasifikasi Citra Daun Anggur Menggunakan SVM Kernel Linear

Adri Gabriel Sooi<sup>1\*</sup>, Paskalis Andrianus Nani<sup>2</sup>, Natalia Magdalena Rafu Mamulak<sup>3</sup>,  
Corazon Olivia Sianturi<sup>4</sup>, Shine Crossifixio Sianturi<sup>5</sup>, Alicia Herlin Mondolang<sup>6</sup>

<sup>1,2,3</sup>Program Studi Ilmu Komputer, Fakultas Teknik, Universitas Katolik Widya Mandira Kupang

<sup>4</sup>Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Sanata Dharma

<sup>5</sup>Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Sanata Dharma

<sup>6</sup>Program Studi Pendidikan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam (MIPA), Universitas Negeri Malang

<sup>1\*</sup>[adrigabriel@unwira.ac.id](mailto:adrigabriel@unwira.ac.id), <sup>2</sup>[paskalisnani@gmail.com](mailto:paskalisnani@gmail.com), <sup>3</sup>[mamulak.natalia@gmail.com](mailto:mamulak.natalia@gmail.com),  
<sup>4</sup>[173114035@student.usd.ac.id](mailto:173114035@student.usd.ac.id), <sup>5</sup>[185314091@student.usd.ac.id](mailto:185314091@student.usd.ac.id), <sup>6</sup>[aliciaherlinm@gmail.com](mailto:aliciaherlinm@gmail.com)

#### Abstract

*The use of artificial intelligence for the image recognition process has been carried out by many researchers. One of its fields is to recognize diseases of grape leaves. Modeling has been carried out using augmentation before support vector machine classification with kernel cubic, with 97.6% accuracy. Improved performance of image prediction accuracy through modeling can still be improved through various means. Some techniques that can be used include using feature selection, initial processing to find and discard outliers, or selecting classifier algorithms that are specifically able to handle datasets with certain characteristics. Another technique is to pass images in the feature extraction process to obtain models with relatively higher accuracy than previous studies. This study aims to improve the acquisition of accuracy figures using the help of the feature extraction process, as well as comparing the performance of several classifiers, namely k-Nearest Neighbor, Random Forest, Naïve Bayes, Neural Network, and Support Vector Machine. The method used starts from the feature extraction process utilizing the SqueezeNet algorithm to obtain a dataset with a specific composition. Furthermore, the division of training and test data was carried out with a ratio of 60:40. Data training uses a variety of validated classifiers using 2-fold cross-validation. The data used is a secondary dataset of grape leaves, consisting of 7222 leaf images, divided into four validated classes from related studies. The results obtained outperformed the previous study, namely 98.1% on the Support Vector Machine classifier using linear kernels.*

*Keywords: grape leaf dataset; improved accuracy; linear kernel; support vector machine; squeezenet.*

#### Abstrak

Pemanfaatan kecerdasan buatan untuk proses pengenalan citra telah dilakukan oleh banyak peneliti. Salah satu bidangnya adalah mengenali penyakit pada daun anggur. Telah dilakukan pemodelan menggunakan augmentasi mendahului pengklasifikasian *support vector machine* dengan *kernel cubic*, dengan hasil akurasi yang diperoleh adalah 97.6%. Peningkatan kinerja akurasi prediksi citra melalui pemodelan masih dapat ditingkatkan melalui berbagai cara. Beberapa teknik yang bisa digunakan antara lain adalah menggunakan seleksi fitur, pengolahan awal untuk mencari dan membuang outlier, ataupun pemilihan algoritma pengklasifikasi yang secara khusus mampu menangani dataset dengan karakteristik tertentu. Teknik lainnya adalah melewatkan citra pada proses ekstraksi fitur untuk memperoleh dataset yang berkualitas baik dan mampu dilatih untuk memperoleh model dengan akurasi yang relatif lebih tinggi, dibandingkan penelitian sebelumnya. Penelitian ini bertujuan meningkatkan perolehan angka akurasi dengan menggunakan bantuan proses ekstraksi fitur, serta membandingkan kinerja beberapa pengklasifikasi yaitu *k-Nearest Neighbor*, *Random Forest*, *Naïve Bayes*, *Neural Network* dan *Support Vector Machine*. Metode yang digunakan dimulai dari proses ekstraksi fitur memanfaatkan algoritma SqueezeNet untuk mendapatkan dataset dengan komposisi 1000 kolom dan 7222 baris. Selanjutnya dilakukan pembagian data latih dan uji dengan perbandingan

Diterima Redaksi : 06-02-2023 | Selesai Revisi : 11-03-2023 | Diterbitkan Online : 19-03-2023

60:40. Pelatihan data menggunakan ragam pengklasifikasi yang di validasi menggunakan *2-fold cross validation*. Data yang digunakan adalah dataset sekunder daun anggur, yang terdiri dari 7222 citra daun, terbagi dalam empat kelas yang telah tervalidasi dari penelitian terkait. Hasil yang diperoleh mengungguli penelitian sebelumnya yaitu 98.1% pada pengklasifikasi *Support Vector Machine* menggunakan *kernel linear*.

Kata kunci: dataset daun anggur; *kernel* linier; peningkatan akurasi; *support vector machine*; *squeeznet*.



## 1. Pendahuluan

Tanaman perkebunan seperti anggur merupakan komoditi unggulan yang memiliki nilai ekonomi tinggi. Selain dapat disajikan segar, olahan menjadi ragam pangan lainnya dapat dilakukan pada buah anggur. Penelitian yang menggunakan objek tanaman anggur seperti buah dan daunnya cukup banyak dilakukan [1]. Salah satu hal yang menjadi perhatian peneliti adalah gangguan penyakit pada daun anggur. Hal ini dapat diamati melalui permukaan daun yang berubah warna dan kondisinya. Beberapa penyakit daun anggur yang dikenal antara lain adalah *Black Rott*, *Isariopsis* dan *Black Measles*. Ketiganya memiliki ciri yang hampir sama yaitu bintik hitam dan coklat yang merusak permukaan daun anggur [2]. Pemanfaatan teknologi kecerdasan buatan [3], [4] telah digunakan untuk berbagai kebutuhan. Secara khusus dapat juga digunakan untuk membantu mengenali gejala penyakit pada tanaman anggur. Salah satu penelitian yang mengangkat pemanfaatan kecerdasan buatan dimaksud adalah pemanfaatan *Convolution Neural Network* (CNN) [5], yang menjadi acuan dalam pengenalan citra untuk berbagai kasus yang membutuhkan bantuan kecerdasan buatan. Namun, apakah pengklasifikasi lainnya dianggap kalah atau tidak mampu menangani permasalahan pengenalan citra [6]–[8].

Penggunaan pengklasifikasi untuk memodelkan citra dengan hasil akurasi yang relatif tinggi dapat dilakukan menggunakan *Random Forest* (RF). Hal ini telah dilakukan dalam kasus pemodelan citra burung jalak yang membandingkan pemanfaatan dua algoritma yaitu jaringan syaraf tiruan dan RF. Penelitian tersebut mampu memodelkan dataset dengan hasil yang mencapai angka akurasi prediksi sebesar 98.6% untuk jaringan syaraf tiruan dan 100% untuk RF [9]. Pengklasifikasi lainnya yang berbasis probabilitas seperti *Naïve Bayes* (NB) pun dapat digunakan untuk melakukan pemodelan dalam pengenalan citra, bahkan yang lebih besar dari sekedar citra yaitu kombinasi ragam dataset yang dikenal dengan Big Data.

Penggunaan NB dibuktikan melalui penelitian untuk klasifikasi manajemen lalu lintas, dimana kombinasi dari citra statis, dan dinamis seperti video digunakan. Pengklasifikasian menggunakan NB dapat meningkatkan akurasi dari angka 49% menjadi 92% [10]. Pengklasifikasian serupa juga dilakukan berbantuan *k-Nearest Neighbor* (*kNN*) [11] terhadap

hasil pertanian, maupun produk olahan pangan dan dapat juga untuk perkebunan anggur. Sebuah penelitian yang menggunakan dataset berisi observasi sejumlah 3419 citra anggur yang terbagi dalam enam kelas, telah dilakukan dan pengolahan data awalnya yaitu mengubah citra RGB menjadi *grayscale*. Selanjutnya dilakukan perbaikan citra berupa penutupan lubang warna yang terdapat pada tiap citra dan membuang latar belakang warna. Sehingga, citra buah anggur benar-benar terbebas dari objek lainnya yang dianggap mengurangi kinerja pengklasifikasi. Namun, hasil yang diperoleh hanya mencapai angka akurasi prediksi yaitu 80.98% [12]. Penelitian terkait lainnya, yang telah dilakukan untuk pemodelan dataset daun anggur menggunakan augmentasi. Proses ini kemudian menghasilkan fitur sejumlah 1000 yang di ekstraksi dari total 2500 data daun anggur. Untuk memperoleh kinerja terbaik maka digunakan ragam kombinasi SVM kernel dalam proses pengklasifikasiannya. Namun, hanya diperoleh hasil akurasi pada angka 97.6% untuk *Support Vector Machine* (SVM) dengan *kernel Cubic* [13].

Penelitian yang menggunakan teknik membandingkan pengklasifikasi, banyak digunakan untuk mengukur kinerja sebuah model. Salah satunya adalah penelitian yang menggunakan dataset hasil ternak unggas yaitu berupa Ayam Petelur, yang membandingkan kinerja pengklasifikasi *Neural Network* (NN) dan *Decision Tree* (DT). Digunakan rasio 50:50 untuk membagi dataset dan sukses memperoleh hasil akurasi prediksi tertinggi pada angka 92.5% [14]. Pengukuran kinerja antara dua pengklasifikasi juga dilakukan pada ragam penelitian, salah satunya adalah menggunakan dataset logo mobil. Seperti penelitian sebelumnya tentang ayam petelur, penelitian logo mobil juga menggunakan perbandingan dataset 50:50 untuk proses pelatihan dan pengujiannya. Dengan menggunakan dua pengklasifikasi yaitu NN dan DT, penelitian berhasil memperoleh angka akurasi pada 92.50% sebagai yang terbaik [15]. Dari sejumlah penelitian terkait, proses peningkatan akurasi menggunakan data sekunder masih dapat dilakukan. Beberapa teknik yang lazim dipakai untuk meningkatkan akurasi pengklasifikasi antara lain melalui ekstraksi fitur, seleksi fitur, pencarian dan pembuangan *outlier*, pemilihan *n-fold cross validation* yang tepat dan dataset yang tidak *overfitting*. Namun, karena data yang digunakan adalah data sekunder dan telah tervalidasi dengan hasil yang relatif baik, maka *overfitting* tidak terjadi.

Pada penelitian ini, dataset yang digunakan adalah data sekunder yang telah tervalidasi menggunakan pengklasifikasi SVM dengan *kernel cubic* [13]. Selanjutnya digunakan ekstraksi fitur berbasis SqueezeNet [16][17] yang memungkinkan SVM dengan *kernel linear* menghasilkan akurasi yang lebih tinggi dari penelitian sebelumnya. Perangkat yang digunakan untuk menunjang penelitian adalah *Central Processing Unit Core i7* dengan *memory 8GB*, tanpa *Graphic Processor Unit*.

Tujuan penelitian ini bermaksud memperbaiki kinerja pengklasifikasi. Perbandingan beberapa antara pengklasifikasi yaitu *kNN*, RF, NB, SVM dan NN ditujukan untuk mencari algoritma dan konfigurasi terbaiknya. Dilakukan proses pelatihan data untuk empat kelas dan pengujian data guna memberikan model untuk pengenalan citra daun anggur yang mengalami gangguan penyakit *Black Measles (BM)*, *Black Rott (BR)*, *Isariopsis (Isa)* dan daun sehat *Healthy Leaf (HL)*.

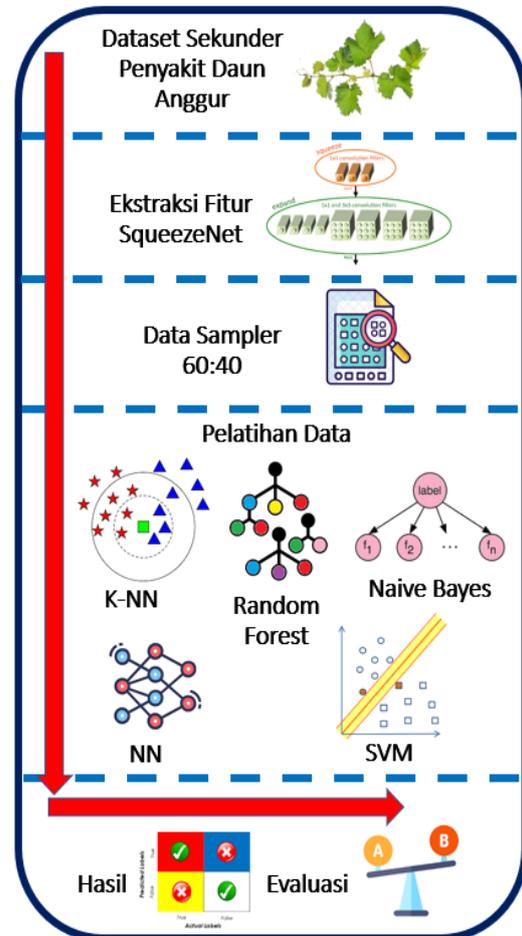
## 2. Metode Penelitian

Percobaan yang telah dilakukan disusun dalam beberapa tahap mulai dari perolehan data hingga akhir klasifikasi menghasilkan matriks confusion. Seluruh tahap percobaan dapat dilihat pada Gambar 1. Tahap pertama adalah akuisisi dataset sekunder yang telah tervalidasi oleh penelitian sebelumnya mengenai tiga jenis penyakit pada daun anggur [18]. Citra penyakit daun anggur dikelompokkan dalam tiga kelas yaitu BM, Isa, BR dan satu kelas daun sehat yaitu HL. Setiap kelas berisi citra sebanyak 1000 sampai 1144 daun, dengan total 7222 citra berukuran 256 x 256 pixel, kedalaman 24bit dengan kerapatan *pixel 96dpi*.

Pada tahap kedua, proses yang dilakukan adalah ekstraksi fitur untuk mentransformasi data citra menjadi numerik berupa 7222 observasi/baris dan 1000 fitur/kolom. Proses ini menggunakan algoritma SqueezeNet [19], yang bekerja melalui pengenalan gambar yang mencapai akurasi relatif tinggi dengan parameter 50x lebih sedikit. Model dilatih pada himpunan data ImageNet. Selanjutnya pada tahap ketiga, data dipisahkan dengan komposisi 60:40. Tujuannya adalah menyediakan 60% dari data untuk melakukan pelatihan data, sedangkan sisanya 40% digunakan untuk pengujian terhadap model hasil pelatihan data. Proses pelatihan data melalui sebuah langkah validasi yang menggunakan *2-Fold Cross Validation*. Dipilihnya angka terkecil yaitu angka 2 didasarkan pada keterujian dataset sekunder yang digunakan [20].

Memasuki tahap ketiga adalah proses pelatihan data menggunakan beberapa pengklasifikasi yaitu RF, dengan jumlah pohon sebanyak 50. Pengklasifikasi kedua yang dipakai dalam penelitian ini adalah *kNN* dengan nilai  $k = 1$ , menggunakan metrik jarak *Euclidean*, dan bobot yang digunakan adalah jarak.

Selanjutnya digunakan juga pengklasifikasi NB dan SVM dengan *kernel linier*. Parameter yang digunakan dalam SVM yaitu  $Cost=3$ , *Regression Loss Epsilon* pada angka 0.1. Pengklasifikasi terakhir yang digunakan adalah NN menggunakan 9 *neuron* dalam *hidden layer*. Fungsi aktivasi yang digunakan adalah *Logistic*, dan fungsi *solver* menggunakan *Stochastic Gradient Descent (SGD)* [21].



Gambar 1 Alur Penelitian, Evaluasi dan Pengukuran Kinerja

### 2.1. Random Forest

Pengklasifikasi RF bekerja menggunakan teknik kombinasi atau dikenal berbasis *ensemble*. Prinsip utamanya adalah penggabungan berbagai pohon keputusan dalam operasinya [21]. Dalam prosesnya, algoritma ini menghasilkan banyak pohon keputusan yang digabung untuk membuat prediksi. Keuntungan dari penggabungan banyak pohon keputusan adalah akurasi yang tinggi. Memanfaatkan agregasi *bootstrap* yang mengolah seluruh observasi yang tertotasi dalam  $X = x_1 \dots x_n$  dipasangkan dengan kelas yaitu  $Y = y_1 \dots y_n$ . Prosesnya diulangi sebanyak  $B$  selanjutnya diambil sampel acak untuk pelatihan data, menggunakan rumus (1).

$$\hat{f} = \frac{1}{B} \sum_{b=1}^B f_b(x') \quad (1)$$

## 2.2. $k$ NN

Sebagai pengklasifikasi non-parametrik terbimbing,  $k$ NN menghasilkan luaran yaitu keanggotaan kelas. Setiap observasi diklasifikasikan berdasarkan kedekatan antar objek dan direpresentasi dalam nilai  $k$  positif menggunakan rentang angka yang kecil seperti 1,3,5,10 dan 20 [22]. Algoritma ini dapat memanfaatkan ragam metrik jarak untuk proses klasifikasi, antara lain *Mahalanobis*, *Chebyshev*, *Manhattan*, *Minkowski*, *Hamming* serta *Euclidean*. Kedekatan antar objek dalam observasi dihitung menggunakan metrik jarak *Euclidean* pada rumus (2).

$$d(e, f) = \sqrt{\sum_{i=1}^n (e_i - f_i)^2} \quad (2)$$

$d$  merepresentasi jarak antara dua objek yang diwakili notasi  $e$  dan  $f$

## 2.3. Naïve Bayes

Menggunakan pengklasifikasi berbasis probabilitas, NB juga termasuk dan dapat dipakai untuk pengolahan citra. Pemodelan data di dalam NB mengolah observasi yang telah dikelompokkan dan tersusun dalam kelas [23], [24]. Lazimnya pengelompokkan dilakukan melalui tahapan *ground truth*. Tersedia dua model NB, yaitu jaringan bayes sederhana dan estimasi kerapatan kernel. Jaringan bayes sederhana yang dipakai dalam penelitian ini dapat dilihat pada rumus (3).

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)} \quad (3)$$

## 2.4. Support Vector Machine

Pengklasifikasi selanjutnya adalah SVM dengan *kernel* linear [25]. Digunakan dataset latih dalam notasi  $(x_1, y_1), \dots, (x_n, y_n)$  kelas yang diwakili oleh  $y_i$  yang beranggotakan  $x_i$ . Pengelompokkan menggunakan *hyperplane* linier seperti terlihat pada rumus (4). Parameter  $w$  menotasikan vector terhadap *hyperplane*,  $b$  adalah offset.

$$w^T x - b = 0 \quad (4)$$

$$\emptyset(z) = \frac{1}{1+e^{-z}} \quad (5)$$

$$Q(w) = \frac{1}{n} \sum_{i=1}^n Q_i(w) \quad (6)$$

## 2.5. Neural Network

*Multilayer Perceptron* dengan propagasi mundur atau yang dikenal dengan *Neural Network* disiapkan sebagai pengklasifikasi terakhir. Berbantuan fungsi aktivasi *logistic* sesuai rumus (5) dan SGD sesuai rumus (6) dalam bentuk estimasi stokastik. NN bekerja dengan sejumlah *neuron* dalam sejumlah *hidden layer*.

$\emptyset$  notasi fungsi logistic,  $z$  notasi asimtot horizontal,  $Q_i$  adalah hasil penjumlahan dari tiap iterasi  $i$ , sedangkan  $w$  adalah parameter untuk meminimalkan  $Q$ .

## 2.6. Mengukur Kinerja Pengklasifikasi

Beberapa besaran yang digunakan untuk mengetahui kinerja pengklasifikasi adalah *Area Under the receiver operating characteristic Curve (AUC-ROC)* atau disingkat AUC [26], *Classification Accuracy(CA)*, *F1*, *Precision* dan *Recall*. Keseluruhan variabel ini bekerja berdasarkan beberapa hasil prediksi dan observasi aktual yaitu, hasil di mana model memprediksi kelas positif dengan benar yaitu *True Positive(TP)*, hasil di mana model memprediksi kelas negatif dengan benar yaitu *True Negative(TN)*, hasil di mana model salah memprediksi kelas positif yaitu *False Positive(FP)*, dan hasil di mana model salah memprediksi kelas negatif *False Negative(FN)* [27]. Dari ke empat besaran ini membentuk beberapa rumus yaitu:

$$CA = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Recall(True Positive Rate) = \frac{TP}{TP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$F1 = 2 * \frac{Recall*Precision}{Recall+Precision} \quad (10)$$

$$False Positive Rate = \frac{FP}{FP+TN} \quad (11)$$

Nilai AUC diperoleh dengan cara memetakan rumus (8) dan (11) untuk membentuk *Receiver Operator Characteristic Curve (ROC)*, di mana nilai tertinggi atau terbaik adalah mendekati 1.0 yang berarti tercipta pengukuran terbaik.

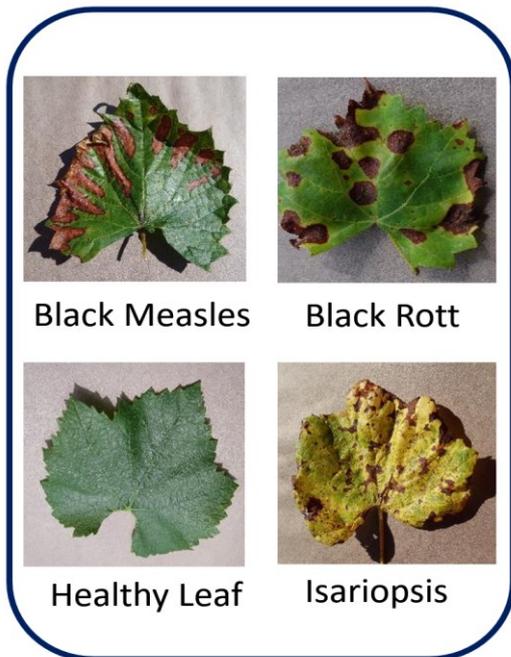
## 3. Hasil dan Pembahasan

Keseluruhan tahapan diawali oleh proses ekstraksi fitur, menggunakan bantuan *SqueezeNet* bekerja dalam proses *image embedding* yang membaca citra secara lokal dan melewatkannya pada model *deep learning*. Setiap citra dihitung fitur vektornya masing-masing dan disimpan dalam 1000 kolom dan 7222 baris sesuai banyaknya citra yang dibaca. Proses pemodelan dan pengenalan data didahului dengan pemisahan dataset menjadi dua kelompok yaitu untuk tujuan pelatihan dan pengujian. Dataset dibagi dalam dua kelompok dengan perbandingan 60:40, yaitu 4334 observasi untuk pelatihan sedangkan sisanya 2888 digunakan untuk data uji.

Setelah keseluruhan proses pelatihan data dan pengujian dilakukan menggunakan beberapa pengklasifikasi, diperoleh hasil akurasi tertinggi yaitu 98.1% dengan menggunakan SVM dengan *kernel* linier. Validasi yang digunakan adalah *2-fold cross validation*, dengan toleransi numerik yang digunakan adalah 0.001. Pengklasifikasi lainnya memperoleh angka akurasi prediksi berturut-turut adalah NN sebesar 96%, RF 94.7%,  $k$ NN 94.3% dan NB 89%, serta SVM dengan *kernel cubic* 879%. Secara lebih terperinci besaran ukuran nilai pengklasifikasi dapat dilihat pada Tabel 1.

Tabel 1 Rincian Hasil Perbandingan Evaluasi Kinerja Klasifikasi Daun Anggur

Model	AUC	CA	F1	Precision	Recall
SVM Linier	<b>0.999</b>	<b>0.981</b>	<b>0.981</b>	<b>0.981</b>	<b>0.981</b>
NN	0.996	0.960	0.960	0.960	0.960
RF	0.993	0.947	0.947	0.947	0.947
kNN	0.962	0.943	0.943	0.944	0.943
NB	0.975	0.890	0.890	0.890	0.890
SVM Cubic[13]	0.990	0.879	0.881	0.901	0.879



Gambar 2 Contoh Beberapa Citra pada Dataset Daun Anggur [13]

Sebagai hasil dari percobaan yang telah dilakukan terhadap dataset daun anggur, diperoleh nilai AUC terbesar pada angka 99.9%, diikuti F1, CA, Precision dan Recall sebesar 98.1% untuk SVM dengan kernel linier. Hasil tersebut mengungguli penelitian sebelumnya pada angka 96.7% dan telah dilakukan pemodelan ulang menggunakan SVM dengan kernel cubic pada angka AUC 99% diikuti CA dan Recall sebesar 87.9%, serta 88.1% untuk F1 dan 90.1% untuk Precision. Selanjutnya dari keseluruhan proses pengklasifikasian menggunakan ragam algoritma mulai dari kNN, RF, NB, SV dan NN diperoleh yang terbaik yaitu SVM yang menggunakan kernel linier. Dengan hasil akurasi prediksi pada angka 98.1%. Secara terperinci dapat terlihat pada Tabel 2. Mayoritas diagonal matrix konfusi terisi hampir maksimum dengan rincian BR sebanyak 1091 observasi atau 95%, diikuti Bm sebesar 96.7%, kemudian Isa 99.90% dan HL 100%. Beberapa observasi yang gagal diprediksi adalah BR sebesar 45 dan dideteksi sebagai BM dan 1 Isa. Sedangkan BM sejumlah 37 gagal dideteksi sebagai BR. Untuk Isa sebanyak satu observasi gagal diprediksi dan dikenali sebagai BM. Sedangkan untuk HL semuanya berhasil diprediksi.

Tabel 2 Confusion Matrix Pengklasifikasi SVM dengan Kernel Linier

		Predicted				
		BR	BM	Isa	HL	Σ
Actual	BR	<b>1091</b>	45	1	0	1137
	BM	37	<b>1107</b>	0	0	1144
	Isa	0	1	<b>1037</b>	0	1038
	HL	0	0	0	<b>1015</b>	1015
	Σ	1128	1153	1038	1015	<b>4334</b>

Sebagai perbandingan terhadap penelitian sebelumnya di mana SVM dengan kernel cubic menghasilkan angka prediksi di bawah SVM kernel linier, dan ini terlihat pada Tabel 3. Kesalahan prediksi yang cukup besar terlihat pada kelas BR dimana terdapat sejumlah 225 observasi yang salah terprediksi sebagai BM. Nilai kesalahan prediksi yang cukup besar juga terlihat pada kelas Isa, dimana 61 observasi terprediksi sebagai BM.

Tabel 3 Hasil Pengulangan Pengklasifikasi menggunakan SVM Kernel Cubic terhadap dataset sekunder Daun Anggur [13]

		Predicted				
		BR	BM	Isa	HL	Σ
Actual	BR	<b>898</b>	225	6	12	1141
	BM	21	<b>1115</b>	2	12	1150
	Isa	31	143	<b>842</b>	4	1020
	HL	7	61	1	<b>954</b>	1023
	Σ	957	1544	851	982	<b>4334</b>

Adapun kekurangan dalam proses prediksi pengenalan citra yang telah dilakukan dapat terlihat secara umum melalui ilustrasi pada Gambar 2. Terlihat jelas bahwa Isariopsis memiliki perbedaan warna hampir di atas 99% dibandingkan dengan HL dan juga daun anggur berpenyakit BR dan BM. Hal ini sangat memudahkan pengklasifikasi untuk mengelompokkannya dalam kelas yang unik, sehingga terlihat pada tabel konfusi memperoleh angka tertinggi yaitu 99.90%. Dilain pihak, untuk citra BR dan BM, keduanya masih memiliki irisan corak sehingga dalam proses pengklasifikasian terjadi kesalahan prediksi sebesar 3.9% dan 3.2% masing-masing untuk BR dan BM.

#### 4. Kesimpulan

Beberapa fakta yang dapat disimpulkan dari penelitian ini adalah bahwa proses ekstraksi fitur dapat meningkatkan perolehan akurasi prediksi. Terjadi proses peningkatan akurasi prediksi dari 97.6% menggunakan SVM dengan kernel cubic pada penelitian sebelumnya, berhasil ditingkatkan menjadi 98.1% menggunakan SVM dengan kernel linier dan ekstraksi fitur berbantuan Squeezenet. Hal ini terbukti dari penggunaan dataset sekunder yang telah tervalidasi dari penelitian sebelumnya, dapat ditingkatkan akurasi pemodelannya. Kualitas dan keragaman dataset yang digunakan dalam pemodelan pun turut mempengaruhi tingkat akurasi yang dihasilkan oleh pengklasifikasi. Pada keadaan tertentu ditemukan adanya irisan antara dua kelas yang memiliki kemiripan atau kedekatan sebaran data, mampu membuat kesalahan prediksi walaupun relatif kecil. Implementasi dari model yang terbentuk dapat diaplikasikan pada ragam kebutuhan praktis lainnya.

Penelitian lanjutan yang dapat dilakukan berbasis temuan ini adalah masih dimungkinkan peningkatan akurasi dengan menyesuaikan ragam parameter yang ada.

### Ucapan Terimakasih

Disampaikan terima kasih yang sebesar besarnya kepada Rektor UNIKA Widya Mandira yang telah memberikan program peningkatan kuantitas dan kualitas publikasi nasional bagi seluruh dosen dan mahasiswa untuk berpartisipasi di dalamnya. Tidak lupa kepada LPPM yang telah menyelenggarakan berbagai program pelatihan penulisan artikel untuk dipublikasikan pada berbagai jurnal nasional terakreditasi SINTA Kemdikbud.

### Daftar Pustaka

- [1] Y. Qi, R. Wang, Q. Qin, and Q. Sun, "Soil Affected the Variations in Grape and Wine Properties Along the Eastern Foot of Helan Mountain, China," *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science*, vol. 69, no. 6, pp. 494–502, Aug. 2019, doi: 10.1080/09064710.2019.1611914.
- [2] G. Gopal and A. P. J., "Identification of Plant Leaf Diseases Using a Nine-Layer Deep Convolutional Neural Network," *Computers & Electrical Engineering*, vol. 76, pp. 323–338, Jun. 2019, doi: 10.1016/j.compeleceng.2019.04.011.
- [3] M. Imanullah, E. M. Yuniarno, and A. G. Sooi, "A Novel Approach in Low-cost Motion Capture System using Color Descriptor and Stereo Webcam," *IJTech*, vol. 10, no. 5, pp. 942–952, Oct. 2019, doi: 10.14716/ijtech.v10i5.2789.
- [4] D. Purwitasari, C. Fatichah, A. G. Sooi, S. Sumpeno, and M. H. Purnomo, "Productivity-based Features from Article Metadata for Fuzzy Rules to Classify Academic Expert," in *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)*, Morioka, Japan, Oct. 2019, pp. 1–6. doi: 10.1109/ICAWS.2019.8923316.
- [5] R. Marani, A. Milella, A. Petitti, and G. Reina, "Deep Neural Networks for Grape Bunch Segmentation in Natural Images from a Consumer-Grade Camera," *Precision Agric*, vol. 22, no. 2, pp. 387–413, Apr. 2021, doi: 10.1007/s11119-020-09736-0.
- [6] E. Talahaturuson *et al.*, "Exploring Indonesian Netizen's Emotional Behavior Through Investment Sentiment Analysis Using TextBlob-NLTK (Natural Language Toolkit)," in *Technology 4.0 for Smart Ecosystem: A New Way of Doing Digital Business*, Semarang, Indonesia, Oct. 2022. doi: <https://doi.org/10.1109/iSemantic55962.2022.9920431>.
- [7] A. B. Gumelar *et al.*, "Human Voice Emotion Identification Using Prosodic and Spectral Feature Extraction Based on Deep Neural Networks," in *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*, Kyoto, Japan, Aug. 2019, pp. 1–8. doi: 10.1109/SeGAH.2019.8882461.
- [8] A. B. Gumelar, E. M. Yuniarno, D. P. Adi, A. G. Sooi, I. Sugiarto, and M. H. Purnomo, "BiLSTM-CNN Hyperparameter Optimization for Speech Emotion and Stress Recognition," in *Wireless Technologies and Intelligent Systems for Better Human Lives*, Surabaya, Indonesia, Sep. 2021, pp. 220–225. doi: 10.1109/IES53407.2021.9594024.
- [9] A. Y. Rahman, "Klasifikasi Citra Burung Jalak Menggunakan Artificial Neural Network dan Random Forest," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 2, Art. no. 2, Aug. 2022, doi: 10.26418/jp.v8i2.53480.
- [10] H. Chen, S. Hu, R. Hua, and X. Zhao, "Improved Naive Bayes Classification Algorithm for Traffic Risk Management," *EURASIP J. Adv. Signal Process.*, vol. 2021, no. 1, p. 30, Jun. 2021, doi: 10.1186/s13634-021-00742-6.
- [11] I. Istiadi, A. Y. Rahman, and A. D. R. Wisnu, "Identification of Tempe Fermentation Maturity Using Principal Component Analysis and K-Nearest Neighbor," *Sinkron : jurnal dan penelitian teknik informatika*, vol. 8, no. 1, Art. no. 1, Jan. 2023, doi: 10.33395/sinkron.v8i1.12006.
- [12] W. Saputro and D. B. Sumantri, "Implementasi Citra Digital Dalam Klasifikasi Jenis Buah Anggur Dengan Algoritma K-Nearest Neighbors (KNN) Dan Data Augmentasi," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 5, no. 2, pp. 248–253, Dec. 2022, doi: 10.31539/intecom.v5i2.4337.
- [13] M. Koklu, M. F. Unlarsen, I. A. Ozkan, M. F. Aslan, and K. Sabanci, "A Cnn-Svm Study Based on Selected Deep Features for Grapevine Leaves Classification," *Measurement*, vol. 188, p. 110425, Jan. 2022, doi: 10.1016/j.measurement.2021.110425.
- [14] F. Nurdiansyah and F. Marisa, "Klasifikasi Ayam Petelur Menggunakan Artificial Neural Network dan Decision Tree," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 7, no. 3, Art. no. 3, Oct. 2022, doi: 10.31328/jointecs.v7i3.4053.
- [15] S. W. Iriananda, R. P. Putra, F. Nurdiansyah, Fitri Marisa, and I. Istiadi, "Klasifikasi Logo Mobil Menggunakan Jaringan Syaraf Tiruan Backpropagation dan Decision Tree," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 7, no. 1, Art. no. 1, Feb. 2022, doi: 10.31328/jointecs.v7i1.3464.
- [16] C. Y. Jerandu *et al.*, "Image Classification of Decapterus Macarellus Using Ridge Regression," in *2022 8th International Conference on Education and Technology (ICET)*, FMIPA Universitas

- Negeri Malang, Oct. 2022, pp. 81–86. doi: 10.1109/ICET56879.2022.9990820.
- [17] A. J. Lado *et al.*, “Comparison of Neural Network and Random Forest Classifier Performance on Dragon Fruit Disease,” in *Wireless Technologies and Intelligent Systems for Better Human Lives (IES-IEEE) 2021*, Politeknik Negeri Surabaya, Sep. 2021, pp. 351–355. doi: 10.1109/ IES53407. 2021.9593992.
- [18] M. Shantkumari and S. V. Uma, “Grape Leaf Image Classification Based on Machine Learning Technique for Accurate Leaf Disease Detection,” *Multimed Tools Appl*, vol. 82, no. 1, pp. 1477–1487, Jan. 2023, doi: 10.1007/s11042-022-12976-z.
- [19] Y. P. Purbanugraha, A. Fatchur Rochim, and I. Setiawan, “Improvement Accuracy Identification and Learning Speed of Offline Signatures Based on SqueezeNet with ADAM Backpropagation,” in *2022 9th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Aug. 2022, pp. 248–253. doi: 10.1109/ ICITACEE55701. 2022. 9924082.
- [20] A. G. Sooi, K. Yoshimoto, H. Takahashi, S. Sumpeno, and M. H. Purnomo, “Dynamic Hand Gesture Recognition on 3D Virtual Cultural Heritage Ancient Collection Objects Using k-Nearest Neighbor,” *Engineering Letters*, vol. 26, no. 3, pp. 356–363, 2018.
- [21] A. B. Gumelar, A. Yogatama, D. P. Adi, F. Frismanda, and I. Sugiarto, “Forward Feature Selection for Toxic Speech Classification Using Support Vector Machine and Random Forest,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 2, Art. no. 2, Jun. 2022, doi: 10.11591/ijai.v11.i2.pp717-726.
- [22] A. Nugroho, S. Sumpeno, and M. H. Purnomo, “A Decision Guidance for Solving Success Rate Political Campaign Using Distance Weighted kNN in Nassi-Shneiderman Framework,” *IJIES*, vol. 14, no. 2, pp. 410–420, Apr. 2021, doi: 10.22266/ijies2021.0430.37.
- [23] N. Z. Fanani *et al.*, “Two Stages Outlier Removal as Pre-processing Digitizer Data on Fine Motor Skills (FMS) Classification Using Covariance Estimator and Isolation Forest,” *IJIES*, vol. 14, no. 4, pp. 571–582, Aug. 2021, doi: 10.22266/ijies2021.0831.50.
- [24] N. Z. Fanani, A. G. Sooi, S. Sumpeno, and M. H. Purnomo, “Penentuan Kemampuan Motorik Halus Anak dari Proses Menulis Hanacaraka Menggunakan Random Forest,” *JNTETI*, vol. 9, no. 2, Art. no. 2, May 2020, doi: 10.22146/jnteti.v9i2.153.
- [25] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, “Human Motion Recognition Based on Svm in Vr Art Media Interaction Environment,” *Hum. Cent. Comput. Inf. Sci.*, vol. 9, no. 1, p. 40, Nov. 2019, doi: 10.1186/s13673-019-0203-8.
- [26] O. A. Montesinos López, A. Montesinos López, and J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-89010-0.
- [27] - Istiadi *et al.*, “Classification of Tempeh Maturity Using Decision Tree and Three Texture Features,” *JOIV: International Journal on Informatics Visualization*, vol. 6, no. 4, Art. no. 4, Dec. 2022, doi: 10.30630/joiv.6.4.983.

*Halaman ini sengaja dikosongkan*