

## Klasifikasi Jenis Buah *Cherry* Menggunakan *Support Vector Machine* (SVM) Berdasarkan Tekstur dan Warna Citra

Annisa Rizky Damanik<sup>1</sup>, Siti Annisa<sup>2</sup>, Albesty Islamyati Rafeli<sup>3</sup>, Anggitha Septia Liana<sup>4</sup>, Desta Sandya Prasvita<sup>5</sup>

Informatika / Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

JL. RS Fatmawati No. 1, Pondok Labu, Jakarta Selatan, DKI Jakarta 12450

annisard@upnvj.ac.id<sup>1</sup>, sitia@upnvj.ac.id<sup>2</sup>, albesty@upnvj.ac.id<sup>3</sup>, anggithasl@upnvj.ac.id<sup>4</sup>,

desta.sandya@upnvj.ac.id<sup>5</sup>

**Abstrak.** Buah cherry merupakan buah yang relatif asing bagi masyarakat Indonesia. Buah ini memiliki jenis yang dapat dikonsumsi dan jenis yang tidak dapat dikonsumsi karena beracun. Mengklasifikasi buah cherry yang memiliki jenis beragam membutuhkan metode yang memiliki akurasi yang maksimal sehingga didapat hasil yang akurat agar kita dapat membedakan jenis buah cherry yang dapat dikonsumsi dan tidak. Penelitian ini mengusulkan teknik pengolahan Support Vector Machine (SVM) dalam mengklasifikasi buah cherry. Tekstur dan warna pada buah cherry yang beragam dan dalam posisi dan keadaan kontras cahaya yang berbeda-beda dijadikan data untuk mengklasifikasi jenis buah cherry. Hasil yang didapat algoritma SVM efisien dalam mengklasifikasi tekstur dan warna buah cherry. Akurasi yang dihasilkan dari aplikasi yang menggunakan algoritma SVM sebesar 100%.

**Kata Kunci:** Cherry, Support Vector Machine (SVM), Tekstur, Warna.

## 1 Pendahuluan

### 1.1 Latar Belakang

Buah cherry merupakan buah yang cukup asing bagi masyarakat di Indonesia. Di Indonesia masih banyak petani cherry yang kurang dalam pengetahuan dan sering mengabaikan manakah buah cherry yang beracun atau tidak, hal ini berakhir panen yang tidak maksimal kondisi terburuknya adalah gagal panen. Untuk membantu mengatasi masalah hal tersebut maka dilakukan identifikasi jenis buah cherry menggunakan computer vision. Computer vision telah banyak dipakai dalam mengidentifikasi jenis buah dan tanaman serta penyakit pada tanaman.

Dalam mengidentifikasi jenis buah cherry digunakan informasi tekstur dan warna citra pada buah. Berbagai jenis buah cherry umumnya memiliki tekstur dan warna yang berbeda. Ekstraksi ciri warna menggunakan Support Vector Machine (SVM) dapat mengenali empat jenis buah cherry pada penelitian. Pengenalan jenis buah cherry menggunakan pemrosesan citra yang membutuhkan informasi ciri masing-masing citra yang berwarna tertentu dan memiliki ciri yang unik. Teknik SVM digunakan untuk mendapatkan fungsi pemisah (hyperplane) yang optimal untuk memisahkan observasi yang memiliki nilai variabel target yang berbeda [1]. Peneliti ingin mengidentifikasi jenis buah cherry menggunakan SVM berdasarkan parameter berupa tekstur dan warna dari citra.

Pada penelitian ini terdapat referensi yang digunakan penulis dalam melakukan penelitian. Penelitian [2] bertujuan untuk membandingkan dua metode klasifikasi Naive Bayes dan Support Vector Machine dan mencari metode terbaik berdasarkan hasil akurasi dan error. Dari menggunakan data training dan data testing, hasil yang didapat

tingkat akurasi SVM lebih tinggi yaitu sebesar 96,2704%, sedangkan metode Naive Bayes memiliki hasil akurasi sebesar 92.0746%.

Lalu berdasarkan jurnal penelitian [3], penelitian ini bertujuan mencari metode terbaik untuk mengklasifikasi tanaman mangrove dan perkiraan luas area mangrove dengan membandingkan metode SVM dan Decision Tree. Berdasarkan data training dan data testing, hasil yang didapatkan oleh metode SVM yaitu keakuratan sebesar 95% sedangkan Decision Tree sebesar 93%, dimana berdasarkan SVM luas lahan mangrove sebesar 634.62 Ha dan berdasarkan Decision Tree luas lahan sebesar 590.47%.

## 1.2 Data yang Digunakan

Penelitian ini menggunakan dataset 'Fruits 360 dataset: A dataset of images containing fruits and vegetables' yang kami dapat dari situs Kaggle.com (<https://www.kaggle.com/moltean/fruits>). Dataset ini terdiri atas 90308 gambar dari 131 jenis buah dan sayur, diantaranya yang digunakan dalam penelitian ini adalah 2952 gambar dari empat jenis buah cherry, yaitu cherry Rainier sebanyak 984 gambar dimana 738 gambar digunakan untuk training dan 246 gambar untuk testing, cherry Wax Black sebanyak 656 gambar dimana 492 gambar digunakan untuk training dan 164 gambar untuk testing, cherry Wax Red sebanyak 656 gambar dimana 492 gambar digunakan untuk training dan 164 gambar untuk testing, dan cherry Wax Yellow sebanyak gambar dimana 492 gambar digunakan untuk training dan 164 gambar untuk testing.

## 2 Tinjauan Pustaka

### 2.1 Klasifikasi

Klasifikasi data adalah proses pengorganisasian data ke dalam kategori yang membuatnya mudah untuk diambil, disortir, dan disimpan untuk digunakan di masa mendatang. Sistem klasifikasi data yang terencana dengan baik membuat data penting mudah ditemukan dan diambil. Hal ini dapat menjadi sangat penting untuk manajemen risiko, penemuan hukum dan kepatuhan.

Prosedur dan pedoman tertulis untuk kebijakan klasifikasi data harus mendefinisikan kategori dan kriteria apa yang akan digunakan organisasi untuk mengklasifikasikan data dan menentukan peran dan tanggung jawab karyawan dalam organisasi terkait dengan pengelolaan data.

### 2.2 SVM

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang dapat digunakan untuk klasifikasi atau regresi. Namun, sebagian besar digunakan dalam masalah klasifikasi. Dalam algoritma SVM, plot setiap item data sebagai titik dalam ruang n-dimensi, dimana n adalah sejumlah fitur yang dimiliki, dengan nilai setiap fitur menjadi nilai koordinat tertentu[4]. SVM juga merupakan *supervised learning*, SVM dapat mencari fungsi pemisahan antar kelas dengan cara memaksimalkan margin atau garis batas. Pencarian garis batas

mempertimbangkan jarak yang paling jauh untuk memisahkan dua kelas tersebut. Kemudian, klasifikasi dengan menemukan hyperplane dapat membedakan kedua kelas dengan baik. Hyperplane itu sendiri merupakan fungsi yang digunakan untuk pemisah antar kelas atau dapat digunakan untuk klasifikasi di dalam ruang kelas dimensi yang lebih tinggi.

Algoritma SVM seharusnya tidak hanya menempatkan objek ke dalam kategori, tetapi juga memiliki margin di antara grafik seluas mungkin. Beberapa aplikasi SVM antara lain klasifikasi teks dan hypertext, klasifikasi gambar, mengenali karakter tulisan tangan, dan dalam ilmu biologi termasuk klasifikasi protein. Keuntungan dalam menggunakan algoritma SVM ini yaitu dapat menentukan hubungan kompleks antara data tanpa perlu melakukan banyak transformasi sendiri, serta efektif dalam kasus jumlah fitur lebih besar daripada jumlah titik data.

## **2.3 Buah Cherry**

Cherry memiliki lebih dari 1000 jenis dan tidak semua jenis buah cherry dapat dimakan, beberapa diantaranya justru mematikan bila dimakan. Buah ini umumnya dibagi menjadi dua jenis, cherry manis dan cherry asam. Adapun jenis-jenis cherry yang akan kami klasifikasi adalah:

### **2.3.1 Cherry Rainier**

Cherry ini merupakan salah satu jenis cherry yang paling terkenal, cherry ini paling dikenal dengan bentuknya yang seperti hati. Cherry ini memiliki warna kuning ke kuning keorenan dengan sapuan warna merah.

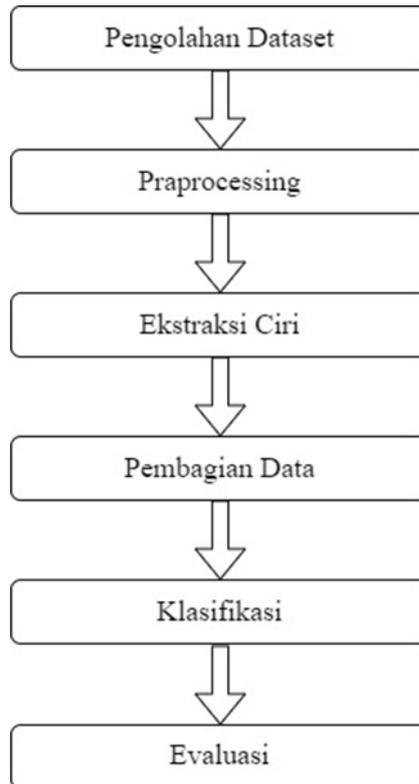
### **2.3.1 Cherry Wax Black, Red and Yellow**

Cherry wax adalah cherry berpohon satu dengan jenis cherry yang beragam yang membedakannya adalah warna dan tekstur. Cherry ini memiliki 3 jenis warna yang berbeda yaitu Black, Red, dan Yellow. Sesuai namanya cherry wax black memiliki warna ungu hampir kehitamannya dapat dilihat pada, cherry wax red memiliki warna merah cerah, dan cherry wax yellow memiliki warna kuning agak ke oranye.

## **3 Metodologi Penelitian**

### **3.1 Implementasi Penelitian**

Penelitian ini mengklasifikasi jenis buah cherry menggunakan data citra yang telah ada berdasarkan ciri tekstur dan warna. Dengan tujuan utama dari proses ini untuk meningkatkan data pada gambar. Hal tersebut terdiri dari pengubahan ukuran pixel pada citra, mengubah warna citra, dan segmentasi citra. Proses pengubahan ukuran pada citra diperlukan seperti tampilan, penyimpanan, dan transmisi citra. Segmentasi citra adalah sebuah proses memisahkan objek yang satu dengan yang lainnya dalam citra menjadi objek-objek berdasarkan karakteristik tertentu. Segmentasi bertujuan menyederhanakan representasi suatu citra lebih mudah untuk dianalisis. Proses ini terdiri dari Identifikasi dan Klasifikasi jenis, preprocessing, ekstraksi ciri, dan Klasifikasi.



**Gambar. 1.** Proses penelitian dimulai dari Pengolahan Dataset, Preprocessing, Ekstraksi Ciri, Pembagian Data, Klasifikasi hingga Evaluasi.

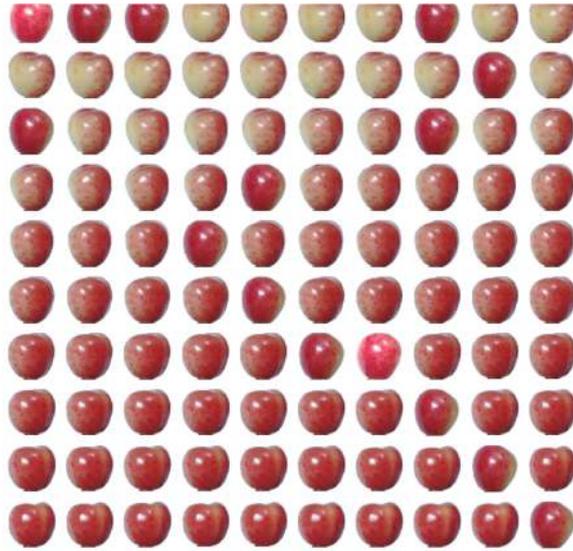
### 3.1.1 Preprocessing

Data yang digunakan adalah data yang bersumber dari kaggle, data terdiri dari empat kelas atau empat jenis buah cherry yaitu cherry Rainier sejumlah 984, cherry wax black sejumlah 656, cherry wax red sejumlah 656, dan cherry wax yellow sejumlah 656 dengan total data sebanyak 2952. Gambar akan di-*scaling* terlebih dahulu agar tidak ada perbedaan mencolok pada data, atau menormalisasikan data pada rentang nilai tertentu. Hal ini juga membantu dalam mempercepat kahlkulasi dalam suatu algoritma[5].

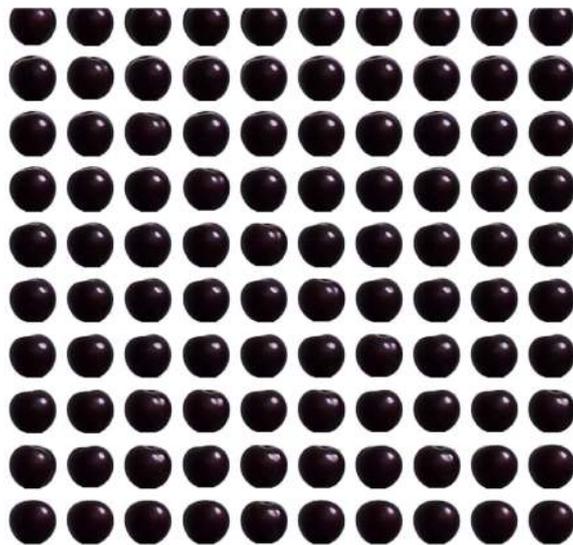
### 3.1.2 Ekstraksi Ciri

Ekstraksi ciri adalah mengambil informasi yang dibutuhkan dalam membedakan jenis dari buah cherry. Ekstraksi ciri yang dibutuhkan adalah tekstur dan warna. Tahapan dari proses ini adalah segmentasi objek dan segmentasi RGB, Konversi RGB ke BGR, dan Scaling[5].

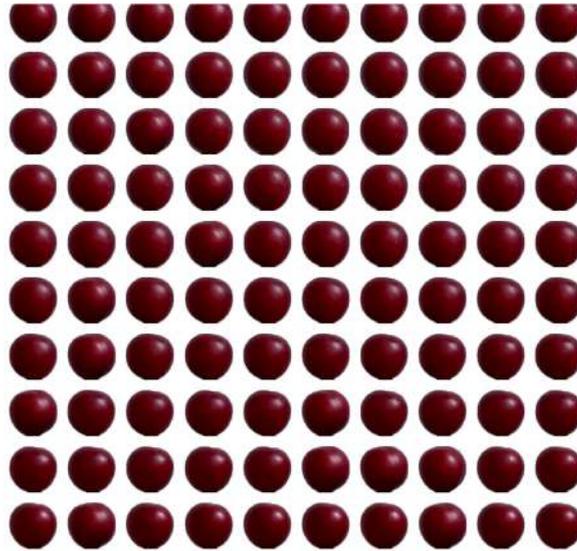
Dalam program ini kami memvisualisasi dari keempat jenis cherry dengan sampel 100 gambar yang akan ditampilkan untuk melihat perbedaannya dengan indra mata sebelum menggunakan sistem metode yang akan kami gunakan. Visualisasi sebagai berikut



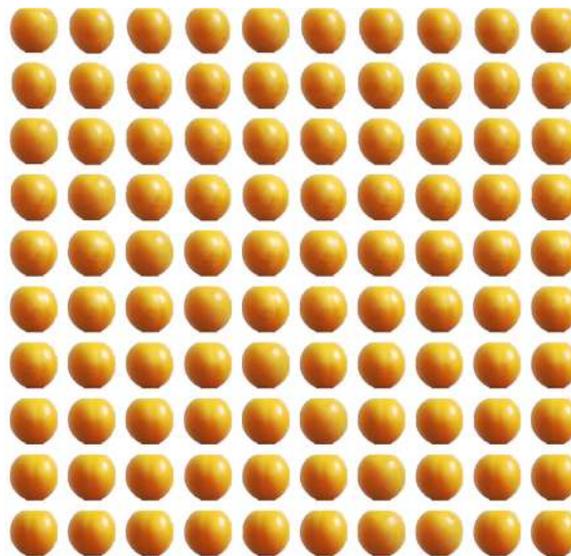
**Gambar. 2.** Cherry Rainer, salah satu jenis buah cherry yang ingin diklasifikasi jenisnya.



**Gambar. 3.** Cherry Wax Black, salah satu jenis buah cherry yang ingin diklasifikasi jenisnya.



**Gambar. 4.** Cherry Wax Red, salah satu jenis buah cherry yang ingin diklasifikasi jenisnya.



**Gambar. 5.** Cherry Wax Yellow, salah satu jenis buah cherry yang ingin diklasifikasi jenisnya.

### 3.1.3 Pembagian Data

Data akan dibagi menjadi data training dan data testing, training bertujuan untuk membangun model, sementara testing yang dilakukan untuk menemukan keakuratan dari model. Data training sejumlah 75%, atau 738 data cherry rainier, 492 cherry wax black, 492 cherry wax red, dan 492 cherry wax yellow untuk diuji coba, dan 25% untuk data testing, atau 246 data cherry rainier, 164 cherry wax black, 164 cherry wax red, dan 164 cherry wax yellow, atau totalnya berupa 2214 data training dan 738 data testing.

### 3.1.4 Klasifikasi

Support Vector Machine (SVM) merupakan salah satu teknik terbaik dalam melakukan klasifikasi. SVM berfungsi mencari hyperplane terbaik dalam memaksimalkan jarak antar kelas. Hyperplane sendiri merupakan pemisah antar kelas [6]. Cara kerja SVM klasifikasi yaitu dengan cara mencoba memisahkan dari setiap kelas dengan cara membuat marginnya seluas mungkin. Awalnya prinsip SVM hanya dapat mengklasifikasi dua kelas, namun dengan pattern recognition mampu menangani lebih dari dua kelas. Pattern recognition sendiri adalah mentransformasikan data pada ruang input, input space, ke ruang yang berdimensi lebih tinggi (feature space), dan optimasi dilakukan pada ruang vektor yang baru tersebut.

### 3.1.5 Evaluasi

Tahapan ini bertujuan untuk mencari tahu apakah model klasifikasi yang telah diciptakan sudah berjalan dengan maksimal atau tidak. Pada tahapan ini kami menggunakan confusion matrix, confusion matrix merupakan alat ukur performa dari model klasifikasi yang dibuat dimana outputnya bisa terdiri dari dua atau lebih kelas [7]. Confusion matrix berisi dari kombinasi yang berbeda dari nilai prediksi dan nilai yang sebenarnya.

## 4 Hasil dan Pembahasan

### 4.1 Datasets

Dataset yang kami gunakan dalam penelitian ini berupa citra RGB beberapa jenis buah cherry pada dataset. Jumlah citra yang terdapat pada dataset yaitu 2952. Dataset tersebut dibagi menjadi data training 75% dan data test 25% untuk setiap kelas, dengan rincian data training sejumlah 738 data cherry rainier, 492 cherry wax black, 492 cherry wax red, dan 492 cherry wax yellow untuk diuji coba, dan untuk data testing sejumlah 246 data cherry rainier, 164 cherry wax black, 164 cherry wax red, dan 164 cherry wax yellow, atau totalnya berupa 2214 data training dan 738 data testing.

```
Ada 738 TRAINING gambar dari CHERRY RAINIER
Ada 492 TRAINING gambar dari CHERRY WAX BLACK
Ada 492 TRAINING gambar dari CHERRY WAX RED
Ada 492 TRAINING gambar dari CHERRY WAX YELLOW
Ada 246 TEST gambar dari CHERRY RAINIER
Ada 164 TEST gambar dari CHERRY WAX BLACK
Ada 164 TEST gambar dari CHERRY WAX RED
Ada 164 TEST gambar dari CHERRY WAX YELLOW
-----
Cherry Rainier
Cherry Wax Black
Cherry Wax Red
Cherry Wax Yellow
-----
```

**Gambar. 6.** Aplikasi melakukan klasifikasi terhadap jenis buah cherry. Jenis buah cherry dibagi menjadi empat kelas, yaitu cherry rainier, cherry wax black, cherry wax red, dan cherry wax yellow.

### 4.2 Hasil Penelitian

Tahap penelitian merupakan tahapan pengujian metode yang digunakan dapat mengetahui secara keseluruhan sistem mengenali dengan baik atau tidak terhadap suatu citra buah. Proses mengklasifikasi jenis buah cherry menggunakan metode Support Vector Machine dengan bantuan Python. Parameter untuk membedakan keempat jenis buah cherry tersebut adalah tekstur dan warna.

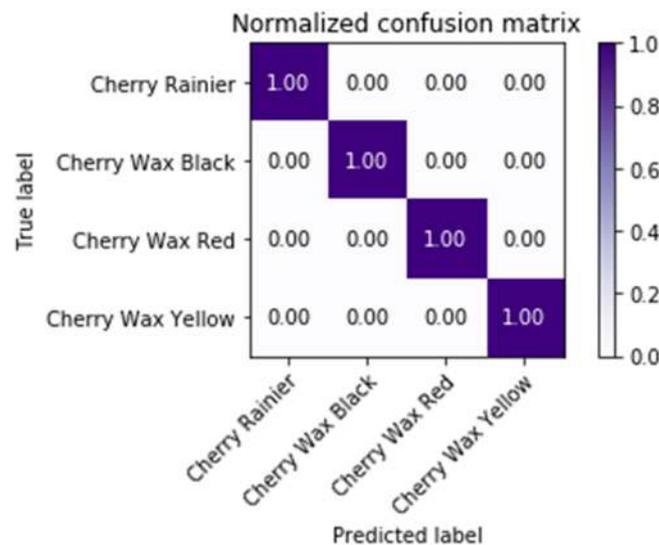
Pertama data akan dibaca dalam keadaan asli berwarna, selanjutnya data akan di-resize, dan warna citra akan di-convert dari RGB menjadi BGR. Selanjutnya, data akan di-scalling, metode ini mengubah nilai pixel menjadi 0-1 hal ini guna mempercepat kalkulasi yang akan dilakukan.

Maka peneliti membangun model menggunakan metode Support Vector Machine (SVM) dari data yang sudah ditentukan. Kemudian hasil yang didapat akan dievaluasi, apakah telah mengerjakan tugasnya secara akurat atau tidak. Hasil yang didapat juga akan dievaluasi menggunakan Confusion matrix guna mengetahui seberapa besar prediksi yang benar dan yang salah.

Uji penelitian menggunakan algoritma Support Vector Machine (SVM) menghasilkan akurasi 100%. Hal ini menunjukkan hasil fitur dapat merepresentasikan permukaan buah melalui citra.

Akurasi menggunakan algoritma SVM: 100.00%

**Gambar. 7.** Algoritma SVM menghasilkan akurasi 100%



**Gambar. 8.** Normalized confusion matrix dari klasifikasi buah cherry menggunakan SVM, dimana didapat hasil 1.0 di setiap jenis buah cherry yang tepat.

Berdasarkan keseluruhan hasil penelitian bahwa fitur tekstur maupun ruang warna mempresentasikan permukaan buah melalui citra dan menghasilkan keakuratan yang maksimal dan efisien. Hal ini menunjukkan bahwa untuk

mengenalinya suatu buah melalui citra memerlukan fitur tekstur dan ruang warna, kemungkinan bila tidak menggunakan fitur tersebut terjadi kesalahan klasifikasi berdasarkan informasi fitur pada dataset.

## 5 Kesimpulan dan Saran

### 5.1 Kesimpulan

Akurasi yang didapat dari mengklasifikasi jenis buah cherry dengan menggunakan metode Support Vector Machine (SVM) menembus angka 100%, hal ini menunjukkan bahwa model yang dibentuk menggunakan metode ini memiliki keakuratan yang tinggi. Hasil Confusion matrixnya pun menunjukkan angka 1.00 disetiap jenisnya pada baris dan kolom untuk jenis yang sama yang menunjukkan bahwa nilai prediksi bernilai benar terhadap nilai keasliannya, dan menunjukkan angka 0.00 pada baris dan kolom untuk jenis yang sama menunjukkan bahwa tidak terdapat nilai prediksi yang bernilai salah terhadap nilai keasliannya. Berdasarkan akurasi yang didapat, model klasifikasi SVM layak digunakan untuk mengklasifikasi jenis buah cherry.

### 5.2 Saran

Hasil dari penelitian kami masih dapat dikembangkan dengan membandingkannya dengan teknik klasifikasi berbeda maupun dengan ekstraksi ciri dengan metode yang berbeda untuk dapat menghasilkan hasil yang lebih baik. Selain itu, data testing dapat diberikan lebih banyak agar testing yang dilakukan dapat lebih sempurna kedepannya. Serta dapat dilakukan penelitian dengan objek penelitian yang sama namun dikembangkan dengan penggunaan dataset yang lebih kompleks dan beragam jenis sehingga dapat menjadikan model yang lebih bervariasi pada penelitian selanjutnya.

## 6 Referensi

- [1] Williams, G. (2011). Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery. Diunduh dari <http://users.umiaccs.umd.edu/~oard/teaching/301/spring16/readings/Williams.pdf>
- [2] Apriyani, H., & Kurniati. (2020). Perbandingan Metode Naive Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus. Diunduh dari <https://conference.binadarma.ac.id/index.php/BDCCS/article/view/1692>
- [3] Firmansyah, S., Gaol, J., & Susilo, S. B. (2019). Perbandingan Klasifikasi SVM dan Decision Tree untuk Pemetaan Mangrove Berbasis Objek Menggunakan Citra Satelit Sentinel-2B di Gili Sulat, Lombok Timur. Diunduh dari <https://journal.ipb.ac.id/index.php/jpsl/article/view/21873>
- [4] Lestari, F. R., Purnama, I. P. N., Sajiah, A. M., & Aksara, L. M. B. (2019). Identifikasi Penyakit Tanaman Jeruk Siam Menggunakan Metode M-Svm. Diunduh dari <http://publikasi.dinus.ac.id/index.php/semnastik/article/view/2889>
- [5] Mihai Oltean. (2019). Fruit Classification: PCA, SVM, KNN, Decision Tree. Diunduh dari <https://www.kaggle.com/waltermaffy/fruit-classification-pca-svm-knn-decision-tree/execution>.
- [6] Widodo B. K., Prasetyo H., Ahmad G., Wibisono, B. A., & Prasvita, D. S. (2021). Analisis Perbandingan Algoritma SVM dan CNN untuk Klasifikasi Buah. Diunduh dari <https://conference.upnvj.ac.id/index.php/senamika/article/view/1564/1336>
- [7] Athallah M. A., Arkadia A., Rifqi S. N., Trianto, & Prasvita D. S. (2021). Klasifikasi Kematangan Buah Pisang Berdasarkan Fitur Warna dengan Metode SVM. Diunduh dari <https://conference.upnvj.ac.id/index.php/senamika/article/download/1781/1344>

