# SCRIPT PRESENTASI REVIEW JURNAL

# Computational Vision for Tomato Classification Using a Decision Tree Algorithm

## Slide 1: Cover

Selamat pagi/siang/sore Bapak/Ibu dan teman-teman sekalian. Pada kesempatan ini, saya akan mempresentasikan hasil review jurnal yang berjudul “Computational Vision for Tomato Classification Using a Decision Tree Algorithm” yang ditulis oleh Caroline S. da Fonseca, Bilton G. Nhantumbo, Yuri M. Ferreira, Layana A. da Silva, dan Anderson G. Costa yang dipublikasikan dalam jurnal Engenharia Agrícola pada tahun 2025.

Jurnal ini membahas tentang penggunaan teknologi computer vision dan algoritma decision tree untuk mengklasifikasikan tomat berdasarkan karakteristik kolorimetri yang memengaruhi potensi pembelian konsumen.

## Slide 2: Outline Presentasi

Dalam presentasi ini, saya akan membahas beberapa poin utama, yaitu: 1. Pendahuluan yang meliputi latar belakang masalah dan pentingnya inspeksi kualitas tomat 2. Karakterisasi tomat berdasarkan tingkat kematangan dan warna 3. Detail dataset yang digunakan dalam penelitian 4. Proses akuisisi dan preprocessing data 5. Metodologi penelitian 6. Model Decision Tree yang digunakan 7. Hasil dan analisis 8. Perbandingan dengan pendekatan Machine Learning lainnya 9. Kesimpulan dan aplikasi dari penelitian ini

Mari kita mulai dengan pendahuluan.

## Slide 3: Pendahuluan

Inspeksi kualitas merupakan tahap krusial dalam produksi dan pemasaran tomat. Atribut fisik seperti bentuk dan warna menjadi faktor penentu utama dalam penerimaan konsumen.

Beberapa tantangan dalam penilaian kualitas tradisional meliputi: - Inspeksi manual yang rentan terhadap kesalahan manusia - Inkonsistensi dalam standardisasi - Klasifikasi yang tidak tepat dan keterlambatan proses - Evaluasi subjektif akibat kelelahan

Penyortiran yang tidak tepat berdasarkan tingkat kematangan dapat menyebabkan: - Pematangan dipercepat karena pelepasan etilen - Umur simpan yang berkurang - Peningkatan kerentanan terhadap penyakit - Kerugian ekonomi di seluruh rantai produksi

Penelitian ini bertujuan untuk mengklasifikasikan tomat berdasarkan karakteristik kolorimetri menggunakan algoritma decision tree, dengan tujuan menyediakan solusi otomatis untuk klasifikasi tomat yang dapat meningkatkan kontrol kualitas dan mengurangi kerugian dalam industri pertanian.

## Slide 4: Karakterisasi Tomat

Dalam penelitian ini, tomat diklasifikasikan ke dalam dua kategori berdasarkan potensi pembelian konsumen:

1. Major Purchasing Potential (MPP): 20 buah tomat dengan warna merah matang (>90% warna merah)
2. Low Purchasing Potential (LPP): 40 buah tomat dengan warna baik yang belum matang (hijau) atau menengah (campuran hijau dan merah)

Klasifikasi ini sejalan dengan standar CEAGESP (2003) dan buah yang digunakan berasal dari kultivar Italiano (Solanum lycopersicum L.) yang ditanam di rumah kaca.

Setup untuk akuisisi gambar terdiri dari: - Kamera digital (Canon Powershot G9 X RGB) dipasang 0,24m di atas meja - Dua lampu halogen 100W diposisikan pada ketinggian 0,84m - Latar belakang putih untuk meminimalkan efek bayangan - Gambar diproses menggunakan software ImageJ untuk mengekstrak nilai RGB

Nilai RGB kemudian dikonversi ke model warna tambahan untuk analisis komprehensif: - Model RGB: nilai intensitas Merah, Hijau, Biru - Model HSI: Hue (H) mewakili warna (0° hingga 360°) - Model CIELab: Luminance (L) untuk kecerahan (0 hingga 100), Kromatisitas a (hijau ke merah) dan b (biru ke kuning)

## Slide 5: Detail Dataset

Berikut adalah karakteristik dataset yang digunakan dalam penelitian:

* Total buah: 60 tomat (kultivar Italiano)
* Distribusi kelas:
  + Major Purchasing Potential (MPP): 20 buah (33,3%)
  + Low Purchasing Potential (LPP): 40 buah (66,7%)
* Fitur yang diekstrak per buah: 7 karakteristik kolorimetri
* Jenis fitur: Semua variabel kontinyu
* Kondisi pertumbuhan: Kultivasi rumah kaca
* Waktu panen: 85 hari setelah penanaman
* Sumber geografis: Wilayah Tengah-Selatan Rio de Janeiro, Brasil (22°25’10” S lintang, 43°25’21” W bujur; ketinggian rata-rata: 610 m)

Karakteristik fitur dalam dataset meliputi: - Red (R): Rentang 80.0-160.0, Rata-rata 121.0, CV 19.0% - Green (G): Rentang 12.0-104.0, Rata-rata 53.5, CV 55.1% - Blue (B): Rentang 5.0-24.0, Rata-rata 11.9, CV 50.0% - Hue (H): Rentang 2.0-61.0, Rata-rata 27.7, CV 76.8% - Luminance (L): Rentang 29.3-45.0, Rata-rata 34.9, CV 9.5% - Chromaticity a: Rentang -2.9-47.5, Rata-rata 24.0, CV 77.3% - Chromaticity b: Rentang 31.2-53.0, Rata-rata 42.4, CV 13.9%

Semua fitur mengikuti distribusi normal berdasarkan uji Shapiro-Wilk (p<0.05).

## Slide 6: Akuisisi & Preprocessing Data

Proses akuisisi gambar terdiri dari beberapa tahap: 1. Persiapan sampel: Tomat dibersihkan dan diposisikan secara individual pada meja pengukuran 2. Pengambilan gambar: Kamera digital Canon Powershot G9 X RGB diposisikan 0,24m di atas meja 3. Kontrol pencahayaan: Dua lampu halogen 100W pada ketinggian 0,84m menyediakan pencahayaan yang konsisten 4. Persiapan latar belakang: Latar belakang putih untuk meminimalkan efek bayangan 5. Penyimpanan digital gambar: Gambar disimpan dalam format JPG resolusi tinggi 6. Kontrol kualitas: Inspeksi visual dari setiap gambar untuk memastikan tidak adanya refleksi, bayangan, atau artefak

Proses pengolahan gambar & ekstraksi fitur: 1. Software: ImageJ open-source digunakan untuk semua pengolahan gambar 2. Pemilihan region of interest: Area buah diisolasi dari latar belakang 3. Ekstraksi RGB: Nilai RGB rata-rata dihitung dari semua piksel dalam region buah 4. Konversi ruang warna: Nilai RGB secara matematis ditransformasi ke: - Model HSI menggunakan transformasi rotasi (Pedrini & Schwartz, 2008) - Model CIELab menggunakan persamaan ruang warna terstandarisasi 5. Organisasi data: Semua fitur yang diekstrak dikompilasi ke dalam dataset terstruktur 6. Jaminan kualitas: Analisis statistik dilakukan untuk memeriksa integritas data

Persiapan data untuk Machine Learning meliputi: - Seleksi fitur menggunakan pengujian signifikansi statistik dengan uji-t Student (p<0.05) - Principal Component Analysis (PCA) untuk penilaian dimensi - Evaluasi multikolinearitas menggunakan matriks korelasi - Indeks ketidakmurnian Gini untuk mengidentifikasi fitur yang paling prediktif - Pembagian data: Training set (75%), Test set (25%) - Tidak diperlukan standardisasi untuk model Decision Tree

## Slide 7: Metodologi

Analisis statistik dari karakteristik kolorimetri menunjukkan: - Semua karakteristik kolorimetri kecuali luminance (L) secara signifikan membedakan antara kelas MPP dan LPP - Dalam model RGB, komponen hijau (G) menunjukkan perbedaan signifikan terbesar antara kelas - Dalam model CIELab, kromatisitas a (variasi antara hijau dan merah) menunjukkan perbedaan yang paling signifikan

Principal Component Analysis (PCA): - PC1 (79,43% dari varians): Terutama membedakan tomat matang dari yang belum matang - Fitur yang memengaruhi kelas MPP: R (intensitas merah), a (kromatisitas merah), b (kromatisitas kuning) - Fitur yang memengaruhi kelas LPP: H (hue) dan G (intensitas hijau) - PC2 (17,73% dari varians): Memengaruhi variasi dalam kelas kematangan menengah - Pemisahan yang jelas antara kelas ditunjukkan dalam bidang dua dimensi

Pengembangan Decision Tree: - Algoritma: Part (Recursive Partitioning and Regression Trees) mengimplementasikan teknik CART - Kriteria pemisahan: Indeks Gini untuk memaksimalkan homogenitas kelas - Pruning: Diterapkan untuk mengoptimalkan parameter kompleksitas (cp) dan menghasilkan model yang paling akurat - Pembagian data: - Training set: 75% (15 buah MPP dan 30 buah LPP) - Test set: 25% (5 buah MPP dan 10 buah LPP) - Metrik kinerja: Akurasi, presisi, recall, dan F1-score

Nilai parameter kompleksitas (cp) optimal 0,019 sesuai dengan tree dengan empat node, mewakili struktur terbaik dengan menyeimbangkan kesederhanaan dan akurasi.

## Slide 8: Implementasi Model Decision Tree

Implementasi Algoritma: - Lingkungan pemrograman: Software R (versi 4.3.2) - Algoritma: Part (Recursive Partitioning and Regression Trees) - Teknik implementasi: CART (Classification and Regression Trees) - Kriteria split: Pengukuran ketidakmurnian indeks Gini - Gini index = 1 - Σ(pi)² - Di mana pi adalah probabilitas kelas i dalam suatu node - Package R yang digunakan: - rpart - Untuk implementasi decision tree - caret - Untuk pelatihan dan evaluasi model - e1071 - Untuk utilitas statistik

Optimasi Hyperparameter: - Parameter kompleksitas (cp) mencegah overfitting dengan mengontrol pertumbuhan tree - Nilai cp yang lebih kecil menciptakan tree yang lebih besar, sementara cp yang lebih besar menciptakan tree yang lebih kecil - Nilai cp optimal 0,019 dipilih karena memberikan error validasi silang terendah sekaligus mempertahankan struktur tree yang sederhana dengan hanya 4 node

Rule Set Decision Tree: - IF (G < 71) THEN - IF (R < 86) THEN - PREDICTION = “LPP” # 7% dari sampel pelatihan - ELSE - IF (B < 19) THEN - PREDICTION = “MPP” # 2% dari sampel pelatihan - ELSE - PREDICTION = “LPP” # 31% dari sampel pelatihan - END IF - END IF - ELSE - PREDICTION = “LPP” # 67% dari sampel pelatihan - END IF

Interpretasi aturan: - Titik keputusan pertama (node root) bergantung pada intensitas hijau (G) - Intensitas hijau < 71 mengindikasikan buah yang berpotensi matang (perlu pengecekan lebih lanjut) - Intensitas hijau ≥ 71 langsung mengklasifikasikan buah sebagai LPP (belum matang) - Untuk buah yang berpotensi matang, intensitas merah (R) menjadi faktor keputusan kedua - Jika intensitas merah < 86, buah diklasifikasikan sebagai LPP (kemerahan tidak cukup) - Untuk buah dengan intensitas merah yang memadai, intensitas biru (B) memberikan klasifikasi akhir - Intensitas biru < 19 mengindikasikan MPP (tomat matang sepenuhnya)

## Slide 9: Model Decision Tree

Struktur Decision Tree yang dihasilkan menunjukkan:

* Node pertama: G < 71 (Intensitas hijau)
  + Jika G > 71, klasifikasikan sebagai LPP (33% dari buah)
  + Jika G < 71, lanjut ke node berikutnya
* Node kedua: R < 86 (Intensitas merah)
  + Jika R < 86, klasifikasikan sebagai LPP (7% dari buah)
  + Jika R > 86, lanjut ke node berikutnya
* Node ketiga: B < 19 (Intensitas biru)
  + Jika B < 19, klasifikasikan sebagai MPP (2% dari buah)
  + Jika B > 19, klasifikasikan sebagai LPP (27% dari buah)
* 67% dari buah dalam training set langsung diklasifikasikan sebagai MPP

Proses pemangkasan pohon mengevaluasi parameter kompleksitas (cp), yang mencegah overfitting dan membatasi pertumbuhan pohon: - Nilai cp optimal: 0,019 - Nilai ini sesuai dengan pohon dengan 4 node - Mewakili keseimbangan terbaik antara kesederhanaan dan akurasi - Mencegah overfitting sekaligus mempertahankan kekuatan klasifikasi

Berdasarkan indeks Gini, tingkat kepentingan fitur adalah: 1. Green (G): 48,3% - Diskriminator utama pada node pertama 2. Red (R): 31,7% - Diskriminator sekunder pada node kedua 3. Blue (B): 12,5% - Diskriminator final pada node ketiga 4. Hue (H): 4,8% 5. Chromaticity a: 2,1% 6. Chromaticity b: 0,6%

Menariknya, meskipun signifikansi statistik dari Hue (H) dan kromatisitas (a, b), model memilih komponen RGB sebagai prediktor yang paling efektif, kemungkinan karena korelasi langsung mereka dengan tampilan visual tomat.

## Slide 10: Evaluasi & Analisis Kinerja Model

Strategi Evaluasi menggunakan beberapa metrik kinerja:

* Akurasi: 86,7% (Proporsi prediksi benar di antara semua prediksi)
  + Akurasi = (TP + TN) / (TP + TN + FP + FN)
* Presisi: 83,3% (Proporsi prediksi positif yang benar)
  + Presisi = TP / (TP + FP)
* Recall (Sensitivitas): 100,0% (Proporsi positif aktual yang diidentifikasi dengan benar)
  + Recall = TP / (TP + FN)
* F1-Score: 90,9% (Rata-rata harmonik dari presisi dan recall)
  + F1 = 2 × (Presisi × Recall) / (Presisi + Recall)

Analisis Matriks Konfusi: - True Positives (TP): 3 tomat MPP yang benar diklasifikasikan sebagai MPP - True Negatives (TN): 10 tomat LPP yang benar diklasifikasikan sebagai LPP - False Positives (FP): 0 tomat LPP yang salah diklasifikasikan sebagai MPP - False Negatives (FN): 2 tomat MPP yang salah diklasifikasikan sebagai LPP - Analisis kesalahan: Model cenderung salah dengan mengklasifikasikan beberapa tomat matang sebagai belum matang, bukan sebaliknya - Implikasi praktis: Bias konservatif ini lebih disukai untuk keamanan pangan dan kontrol kualitas

Analisis Pentingnya Fitur: - Green (G): 48,3% - Fitur terpenting yang memainkan peran kunci dalam klasifikasi - Red (R): 31,7% - Kontributor penting kedua untuk klasifikasi - Blue (B): 12,5% - Memberikan penyempurnaan tambahan dalam klasifikasi - Meskipun semua karakteristik kolorimetri (kecuali L) menunjukkan perbedaan signifikan secara statistik antar kelas, algoritma decision tree hanya memilih tiga (G, R, B) sebagai prediktor yang paling efektif - Decision tree lebih memilih fitur RGB daripada ruang warna alternatif (HSI, CIELab), menunjukkan bahwa intensitas warna langsung memberikan kekuatan diskriminatif yang cukup

## Slide 11: Hasil & Analisis

Matriks Konfusi menunjukkan: - Kelas LPP: 10 dari 10 buah diklasifikasikan dengan benar (100%) - Kelas MPP: 3 dari 5 buah diklasifikasikan dengan benar (60%) - 2 buah MPP salah diklasifikasikan sebagai LPP - Akurasi keseluruhan 86,7%

Model menunjukkan kinerja yang sangat baik dalam mengidentifikasi tomat LPP, dengan pengenalan sempurna. Untuk tomat MPP, kinerja baik tetapi menunjukkan beberapa ruang untuk peningkatan. Akurasi yang lebih tinggi untuk LPP bermanfaat dari perspektif keamanan pangan, karena memastikan bahwa tomat belum matang atau matang sebagian tidak salah dipasarkan sebagai matang sepenuhnya.

Metrik Kinerja: - Akurasi: 86,7% - Ketepatan keseluruhan klasifikasi tinggi - Presisi: 83,3% - Keandalan prediksi positif baik - Recall: 100% - Identifikasi sempurna dari semua sampel LPP - F1-Score: 90,9% - Keseimbangan yang sangat baik antara presisi dan recall

Skor F1 yang tinggi menunjukkan bahwa model secara efektif menyeimbangkan false positive dan false negative, membuatnya cocok untuk aplikasi praktis dalam kontrol kualitas.

Perbandingan dengan Aplikasi Lain: - Habib et al. (2020) menggunakan decision tree untuk mendeteksi penyakit pada pepaya, mencapai tingkat akurasi melebihi 75% - Özaltın (2024) menerapkan algoritma machine learning untuk klasifikasi buah kurma berdasarkan fitur morfologi dan kolorimetri - Raihen & Akter (2024) mengklasifikasikan kismis berdasarkan ukuran, bentuk, warna, dan tekstur menggunakan decision tree dengan akurasi tinggi

Kinerja yang dicapai dalam penelitian ini (akurasi 86,7%, skor F1 90,9%) dibandingkan dengan baik dengan aplikasi pertanian serupa, menunjukkan efektivitas pendekatan ini untuk klasifikasi buah.

## Slide 12: Perbandingan dengan Pendekatan ML Lainnya

Decision Tree vs. Metode Machine Learning Lainnya:

| Aspek | Decision Tree | Support Vector Machine | Neural Network |
| --- | --- | --- | --- |
| Interpretabilitas | Tinggi - Aturan keputusan transparan | Rendah - Model black box | Sangat Rendah - Black box kompleks |
| Skalabilitas Fitur | Tidak diperlukan | Diperlukan | Diperlukan |
| Kebutuhan Ukuran Sampel | Bekerja baik dengan dataset kecil | Ukuran sampel sedang diperlukan | Ukuran sampel besar lebih disukai |
| Hubungan Non-linear | Secara alami menangani non-linearitas | Membutuhkan fungsi kernel | Membutuhkan arsitektur dalam |
| Kecepatan Pelatihan | Cepat | Sedang | Lambat |
| Kecepatan Prediksi | Sangat cepat | Cepat | Cepat |
| Tuning Hyperparameter | Sederhana (sedikit parameter) | Kompleks | Sangat kompleks |
| Penanganan Nilai Hilang | Dukungan bawaan | Membutuhkan preprocessing | Membutuhkan preprocessing |
| Kompleksitas Implementasi | Rendah | Sedang | Tinggi |

Mengapa Decision Tree untuk Aplikasi Ini? - Kompatibilitas dataset kecil: Dengan hanya 60 sampel tomat, decision tree menghindari masalah overfitting yang umum dalam model yang lebih kompleks - Interpretabilitas untuk pakar domain: Insinyur pertanian dan ilmuwan pangan dapat memahami dan mempercayai aturan klasifikasi - Implementasi dalam sistem real-time: Pemrosesan cepat cocok untuk lini sortir otomatis - Batas keputusan non-linear: Dapat menangkap hubungan kompleks antara fitur warna dan kematangan - Umpan balik pentingnya fitur: Memberikan wawasan tentang karakteristik kolorimetri mana yang paling relevan

Perbandingan dalam Aplikasi Pertanian Serupa: - Deteksi Penyakit Pepaya (Habib et al., 2020): SVM mencapai akurasi >90% - Klasifikasi Buah Kurma (Özaltın, 2024): Decision Tree: 86,59%, KNN: 92,18%, Neural Networks: 93,85%, SVM: 92,74% - Deteksi Pemalsuan Kimia pada Buah (Sattar et al., 2024): Decision Tree: 81%, Naïve Bayes: 82%, Deep Learning: 96,71%

Model decision tree kami (akurasi 86,7%) dibandingkan dengan baik dengan aplikasi pertanian serupa. Meskipun neural network dan model hybrid dapat mencapai akurasi yang lebih tinggi dalam beberapa kasus, pendekatan kami menawarkan beberapa keunggulan: - Implementasi sederhana dengan persyaratan tuning hyperparameter minimal - Interpretabilitas yang jelas dari aturan keputusan berdasarkan ambang batas kolorimetri - Pendekatan praktis untuk klasifikasi real-time dalam sistem sortir - Tidak perlu preprocessing atau normalisasi fitur yang kompleks - Secara efektif menangkap hubungan antara karakteristik warna dan preferensi konsumen

## Slide 13: Kesimpulan & Aplikasi

Temuan Utama: - Semua karakteristik kolorimetri kecuali luminance (L) secara signifikan membedakan antara kelas MPP dan LPP - Komponen hijau (G) menunjukkan perbedaan signifikan terbesar antara kelas dalam model RGB - Kromatisitas a (mengukur variasi merah-hijau) adalah karakteristik paling andal dalam model CIELab - Principal component analysis secara efektif memisahkan kelas tomat pada bidang dua dimensi - Algoritma decision tree mencapai kinerja klasifikasi yang sangat baik dengan akurasi tinggi (86,7%) dan skor F1 (90,9%) - Parameter model warna RGB (R, G, B) dipilih sebagai prediktor paling efektif oleh indeks Gini - Struktur tree optimal berisi hanya 4 node, menyeimbangkan kesederhanaan dengan akurasi

Aplikasi Praktis: 1. Sistem Penyortiran Otomatis: - Integrasi dengan sistem conveyor belt di rumah pengepakan untuk klasifikasi tomat real-time - Mengurangi biaya tenaga kerja dan meningkatkan standardisasi

1. Pemanenan Robotik:
   * Implementasi dalam sistem pemanenan robotik untuk mengidentifikasi tomat matang di lapangan
   * Memungkinkan pemanenan selektif dan mengurangi kerugian pasca panen
2. Kontrol Kualitas Ritel:
   * Aplikasi di titik ritel untuk memastikan kualitas yang konsisten
   * Memperpanjang umur simpan dengan pengelompokan yang tepat
   * Meningkatkan kepuasan konsumen
3. Aplikasi Mobile:
   * Pengembangan aplikasi smartphone yang memungkinkan produsen dan konsumen untuk menilai kualitas tomat menggunakan kamera perangkat

Arah Masa Depan: Integrasi computer vision dengan algoritma decision tree mengoptimalkan klasifikasi buah dan meningkatkan proses produksi pertanian otomatis. Dengan memastikan pemilihan produk berkualitas tinggi, teknologi ini menguntungkan industri pertanian dan secara bersamaan meningkatkan kepuasan dan keamanan konsumen. Penelitian selanjutnya harus fokus pada memperluas pendekatan ini ke buah dan sayuran lain, memasukkan parameter kualitas tambahan, dan mengembangkan model machine learning yang lebih canggih untuk lebih meningkatkan akurasi klasifikasi.

## Slide 14: Referensi

1. Fonseca, C. S., Nhantumbo, B. G., Ferreira, Y. M., Silva, L. A., & Costa, A. G. (2025). “Computational vision for tomato classification using a decision tree algorithm.” *Engenharia Agrícola*, 45, e20240124.
2. Andreuccetti, C., Ferreira, M. D., & Tavares, M. (2005). “Perfil dos compradores de tomate de mesa em supermercados da região de Campinas.” *Horticultura Brasileira*, 23(1), 148-153.
3. Brandão, C. P., Oliveira, M. S. P., Santos, J. C., & Oliveira Junior, M. C. M. (2021). “Perfil e preferências do consumidor de frutos de pupunha da cidade de Belém, Pará.” *Research, Society and Development*, 10(7), e28810716502.
4. Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). “Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables.” *Food and bioprocess technology*, 4, 487-504.
5. Habib, M. T., Majumder, A., Jakaria, A. Z. M., Akter, M., Uddin, M. S., & Ahmed, F. (2020). “Machine vision based papaya disease recognition.” *Journal of King Saud University-Computer and Information Sciences*, 32(3), 300-309.

Terima kasih atas perhatiannya!