

Identifikasi Biji-Bijian Berdasarkan Ekstraksi Fitur Warna, Bentuk dan Tekstur Menggunakan Random Forest

Luthfi Alwi, *Akademi Pariwisata Majapahit*, Arya Tandy Hermawan, *Teknologi Informasi Institut Sains dan Teknologi Terpadu Surabaya*, dan Yosi Kristian, *Teknologi Informasi Institut Sains dan Teknologi Terpadu Surabaya*

Abstrak - Proses identifikasi atau pengenalan biji-bijian merupakan aspek penting dalam dunia industri pengolahan pangan. Sebuah industri pangan berskala besar, proses pencampuran beberapa macam biji-bijian dalam pengolahan sebuah produk pangan sangat memperhatikan ketepatan dalam memilih bahan. Agar tidak terjadi kesalahan yang fatal, diperlukan sebuah proses identifikasi dari bahan yang digunakan. Proses pengidentifikasian terhadap beberapa varian biji-bijian dapat dilakukan dengan cara mengekstraksi fitur dari citra (image) dengan menganalisa melalui parameter warna, bentuk dan tekstur serta melakukan proses pengklasifikasian untuk mengukur tingkat keakuratan. Penelitian ini melakukan identifikasi terhadap varian biji-bijian (padi, jagung, kacang tanah dan kedelai) dengan melakukan ekstraksi fitur warna menggunakan RGB dan HSV, ekstraksi fitur bentuk menggunakan Morphological Threshold dan ekstraksi fitur tekstur menggunakan Grey Level Co-occurrence Matrix (GLCM) dan Local Binary Pattern (LBP). Untuk proses pengklasifikasian, peneliti menggunakan metode Random Forest Classifier (RF) untuk mendapatkan tingkat akurasi yang tinggi dengan batasan-batasan yang mempengaruhi keakuratan dalam proses pengklasifikasian untuk dikembangkan dalam proses selanjutnya. Peneliti menggunakan tools MATLAB R2015b untuk proses identifikasi mulai dari proses ekstraksi fitur sampai proses klasifikasinya. Berdasarkan hasil pengujian yang dilakukan didapatkan tingkat akurasi sebesar 99.8 %. Dapat disimpulkan bahwa pengambilan dataset berupa gambar atau image biji-bijian yang diteliti dapat dijadikan patokan untuk pengidentifikasian dan dapat dikembangkan dalam proses selanjutnya.

Kata kunci: *Identifikasi, Ekstraksi Fitur, RGB, HSV, Morphological Threshold., GLCM, LBP, RF, MATLAB*

I. PENDAHULUAN

Proses identifikasi atau pengenalan biji-bijian merupakan aspek penting dalam dunia industri pengolahan pangan. Sebuah industri pangan berskala besar, proses pencampuran beberapa macam biji-bijian dalam pengolahan sebuah produk pangan sangat memperhatikan ketepatan dalam memilih bahan agar tidak terjadi kesalahan dalam proses produksi karena berpengaruh pada hasil akhir dari sebuah produksi. Agar tidak terjadi kesalahan yang fatal, diperlukan sebuah proses identifikasi dari bahan yang digunakan. Dengan sebuah sensor (intelligent camera) yang digunakan dari hasil

sebuah proses identifikasi maka sebuah proses produksi produk pangan dapat berjalan dengan baik dan tidak terjadi kesalahan dalam pencampuran bahan.

Proses pengidentifikasian terhadap beberapa varian biji-bijian dapat dilakukan dengan cara mengekstraksi fitur dari citra (image) dengan menganalisa melalui parameter warna, bentuk dan tekstur serta melakukan proses pengklasifikasian untuk mengukur tingkat keakuratan. Proses identifikasi berdasarkan parameter fitur warna, bentuk dan tekstur dengan metode computer vision diharapkan dapat menjamin tingginya tingkat akurasi dari sebuah varian biji-bijian dan bisa dijadikan acuan untuk penelitian selanjutnya.

Citra digital merupakan gambar dua dimensi yang bisa ditampilkan pada layar komputer sebagai himpunan atau diskrit nilai digital yang disebut pixel/ picture elements.[1] Citra sebagai salah satu komponen multimedia memegang peranan sangat penting sebagai bentuk informasi visual. Citra mempunyai karakteristik yang tidak dimiliki oleh data teks, yaitu citra kaya akan informasi. Maksud dari “citra kaya akan informasi” adalah citra dapat memberikan informasi yang lebih banyak dibandingkan dengan informasi yang disajikan dalam bentuk teks[2]

Pengolahan citra merupakan proses pengolahan dan analisis citra yang banyak melibatkan persepsi visual. Proses ini mempunyai ciri data masukan dan informasi keluaran yang berbentuk citra. Istilah pengolahan citra digital secara umum didefinisikan sebagai pemrosesan citra dua dimensi dengan komputer. Dalam definisi yang lebih luas, pengolahan citra digital juga mencakup semua data dua dimensi.[3]

Ekstraksi fitur adalah proses pengindeksan suatu database berupa citra (image) dengan isinya. Salah satu proses ekstraksi fitur adalah menganalisa berdasarkan isi visual seperti warna, bentuk dan tekstur. Setelah proses ekstraksi fitur, dilakukan proses klasifikasi untuk menentukan tingkat keakuratan dari proses identifikasi yang telah dilakukan. [4]

II. TINJAUAN PUSTAKA

A. Image Acquisition

Berdasarkan penelitian yang dilakukan oleh Dongming Li, Yongfu Liu And Liang Gao [5] yang menerangkan bahwa sebuah gambar (image) dapat dipelajari dan dieksplorasi dengan berbagai parameter yang dapat diekstraksikan mulai dari fitur bentuk, warna, ukuran dan tekstur dengan melakukan teknologi pengenalan (recognition) karakteristik dari sebuah gambar (image). Recognition adalah proses untuk memberikan label pada objek sesuai dengan fitur-fitur yang dimiliki oleh objek tersebut.

Recognition dilakukan dimulai dengan proses pengambilan citra (image) yang diperlukan sebagai dataset

Luthfi Alwi, Akademi Pariwisata Majapahit, Surabaya, Jawa Timur, Indonesia (e-mail: luthfi.alwi@yahoo.com)

Arya Tandy Hermawan, Departemen Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya, Jawa Timur, Indonesia (e-mail: arya@stts.edu)

Yosi Kristian, Departemen Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya, Jawa Timur, Indonesia (e-mail: yosi@stts.edu)

dengan menggunakan beberapa perangkat seperti kamera, scanner, dan lain-lain atau yang sering disebut dengan Image acquisition. Proses pengambilan gambar (image) biji yang dilakukan peneliti dilakukan dengan menggunakan kamera merk Nikon D3400 lensa standart dan perlengkapannya. Selain itu proses pengambilan gambar (image) biji pada penelitian ini dilakukan dengan mendownload melalui Google Image Search. Ini dilakukan untuk mendapatkan varian biji yang bermacam-macam secara acak/random dengan jarak pengambilan gambar yang berbeda-beda karena biji- bijian yang diteliti mempunyai varian lebih dari satu.

B. Otsu Algorithm

Setelah data input didapatkan, selanjutnya dilakukan tahapan selanjutnya yaitu melakukan proses thresholding untuk memisahkan antara image foreground dan background dalam format gray image (Binarization) dimana hasil proses ini digunakan untuk melakukan ekstraksi fitur bentuk. Metode yang digunakan untuk proses thresholding ini adalah Otsu Algorithm. Metode Otsu thresholding diperkenalkan pertama kali oleh Nobuyuki Otsu, dalam jurnal ilmiahnya yang berjudul “A Threshold Selection Method from Grayscale Histogram” pada tahun 1979 Metode Otsu thresholding merupakan metode segmentasi yang cukup akurat dalam mendapatkan daerah yang merupakan obyek tersegmentasi dengan menggunakan histogram grayscale.[6]

Metode Otsu ini didasarkan pada histogramnya. Histogram menunjukkan sebarang nilai intensitas dari tiap piksel pada citra dalam 1 dimensi. Jadi sumbu x biasanya menyatakan level intensitas yang berbeda sedangkan sumbu y menyatakan jumlah piksel yang memiliki nilai intensitas tersebut. Dengan menggunakan histogram kita bisa melakukan pengelompokan terhadap piksel-piksel dalam citra. Pengelompokan ini didasarkan pada nilai ambang batas atau threshold.[7] Sebuah threshold yang mampu memisahkan kelas-kelas sehingga piksel-piksel antar kelas memiliki nilai intensitas yang berbeda maka threshold tersebut dikatakan sudah optimal.

Untuk sebuah gambar, metode segmentasi untuk foreground dan background dilambangkan dengan t, sedangkan untuk foreground dilambangkan dengan w₀, dan nilai mean nya dilambangkan dengan μ₀ sedangkan untuk background dilambangkan dengan w₁, dan nilai mean nya dilambangkan dengan μ₁. Nilai Mean keseluruhan image dihitung dengan rumus:

$$\mu = w_0 \times \mu_0 + w_1 \times \mu_1$$

Sedangkan nilai persamaan yang dibentuk (objective function) dengan menggunakan OTSU Algorithm :

$$g(t) = w_0 \times (\mu_0 - \mu)^2 + w_1 \times (\mu_1 - \mu)^2$$

Dimana g(t) adalah nilai on-class variance dan metode OTSU Algorithm dipakai untuk mendapatkan nilai maksimum dari g(t) dan metode ini dianggap yang terbaik.[8]

Sedangkan untuk ekstraksi fitur warna dan fitur tekstur tidak melalui tahapan ini.

C. Ekstraksi Warna RGB to HSV

Ekstraksi fitur warna cocok untuk menggambarkan dan mewakili gambar berwarna. Ekstraksi fitur warna adalah menganalisa warna dari sebuah citra (image) dimana sebuah citra tersusun dari piksel-piksel yang memiliki ukuran

intensitas warna masing-masing. Sebaran warna di tiap-tiap piksel ditunjukkan oleh histogram. Histogram menunjukkan distribusi piksel berdasarkan intensitas graylevel (derajat keabuan) yang dimiliki tiap-tiap piksel. Penggunaan histogram sebagai metode ekstraksi ciri didasarkan pada perbedaan sebaran atau distribusi piksel di masing-masing gambar. Histogram warna sangat leluasa untuk dibentuk dan dapat disusun dari gambar dengan berbagai rentang warna. Histogram warna dari sebuah gambar diproduksi dengan kuantisasi pada warna di gambar ke dalam bilangan biner, dan menghitung jumlah piksel pada tiap bilangan biner tersebut. [9]

Color Histogram adalah representasi dari distribusi warna dalam suatu gambar. Color Histogram suatu gambar dihasilkan oleh kuantisasi warna dalam gambar menjadi sejumlah piksel dalam satuan bin dan menghitung jumlah piksel gambar dalam setiap satuannya. Color Histogram warna dalam RGB dan HSV ruang warna dengan 8 bin di setiap saluran warna akan digunakan untuk mewakili data gambar.

Pada penelitian ini, ekstraksi fitur warna yang digunakan adalah RGB to HSV dimana hasil ekstraksinya didapatkan 18 fitur yang akan dijadikan parameter untuk pengklasifikasian. Ke-18 fitur itu antara lain R mean, R max, R min, G mean, G max, G min, B mean, B max, B min, H mean, H max, H min, S mean, S max, S min, V mean, V max, dan V min.

D. Ekstraksi Bentuk Morphology Threshold

Ekstraksi fitur bentuk merupakan konfigurasi oleh garis dan kontur. Fitur bentuk dikategorikan bergantung pada teknik yang digunakan. Kategori tersebut adalah berdasarkan batas (boundary-based) dan berdasarkan daerah (region-based). Teknik berdasarkan batas (boundary-based) menggambarkan bentuk daerah dengan menggunakan karakteristik eksternal.

Sebelum dilakukan ekstraksi fitur bentuk, terlebih dahulu melewati tahapan thresholding untuk memisahkan antara image foreground dan background dalam format gray image (Binarization).Metode yang digunakan untuk proses thresholding ini adalah Otsu Algorithm.

Setelah metode Threshold dilakukan, Morphology Extraction Method digunakan untuk mengekstraksi komponen citra yang berguna dalam merepresentasikan dan mendeskripsikan bentuk daerah seperti batas dan rangka.

Proses morfologi suatu image adalah merupakan kumpulan operasi non linear yang berkaitan dengan bentuk atau morphology dalam suatu image. Secara praktek, sistim biner sering digunakan dalam proses morfologi, yaitu proses bit 1 atau yang dikenal dengan foreground dan bit 0 atau background dengan cara merubah bagian tertentu dari foreground menjadi daerah background dan sebaliknya merubah sebagian background untuk menjadi daerah foreground. Manfaat penggunaan proses morfologi yaitu untuk menghilangkan noise yang ada.

Proses dilation dalam morfologi image identik dengan menambahkan pixels dalam lingkup image asal, dengan cara menempatkan satu demi satu pusat penataan elemen untuk masing-masing pixel background. Bila sembarang pixel neighbourhood bernilai pixel foreground (nilai 1) maka pixel background dirubah ke foreground.

Proses erosion merupakan proses menghilangkan pixel dalam lingkup objek image dengan cara meletakkan pusat penataan elemen satu demi satu dalam pixel foreground (nilai

1). Bila terdapat pixel neighbourhood bernilai pixel background (nilai 0), maka nilai foreground tersebut di rubah ke background.

Dengan proses erosi berakibat penyusutan ukuran obyek image sehingga dapat digunakan untuk memisahkan objek yang saling gandeng satu sama lain. Sedangkan dilation akan menaikkan ukurannya sehingga dapat menebalkan objek image dan menyambung object yang terputus ataupun meratakan tepi objek yang rusak.

Opening merupakan proses morfologi dengan menggunakan proses erosi dan dilanjutkan dengan proses dilation menggunakan penataan elemen (structuring element) yang sama. Opening banyak digunakan untuk proses menghilangkan objek kecil dalam suatu image tetapi tetap masih mempertahankan bentuk aslinya.

Sedangkan closing merupakan proses morfologi dengan cara melakukan operasi dilation yang diteruskan dengan operasi erosi dengan menggunakan penataan elemen yang sama. Metode closing digunakan bila ingin menutup lubang objek dengan tetap mempertahankan bentuk aslinya. [10]

Untuk ekstraksi fitur bentuk mencari nilai Rasio “L” (Area/Perimeter), Metric (roundness), Metric, Eccentricity, Equivalent Diameter, Rasio “M” (Major Axis Length/Minor Axis Length) untuk dijadikan parameternya dan didapatkan 5 parameter untuk masing-masing biji.

Untuk parameter Rasio L didapatkan dengan membagi antara Area dan perimeternya karena nilai rata-rata Area dan Perimeter dari image yang digunakan karena jarak pengambilan image yang berbeda-beda.

Begitu juga dengan Rasio M didapatkan dengan membagi antara Major Axis Length dan Minor Axis Length untuk mendapatkan panjang garis rata-rata dari setiap image biji-bijian yang diteliti.

E. Ekstraksi Tekstur Grey Level Co-occurrence Matrix (GLCM) dan Local Binary Pattern (LBP).

Ekstraksi fitur tekstur sebuah image merupakan informasi berupa susunan struktur permukaan suatu gambar. Fitur tekstur cocok untuk memvisualisasikan pola dan sifat permukaan. Ini menyediakan informasi tentang variasi intensitas permukaan gambar.

Ekstraksi fitur tekstur sesuai untuk memvisualisasikan penampakan pola dan permukaan. Hal itu menyediakan informasi terkait variasi intensitas dari permukaan obyek pada gambar. Pada paper ini, metode yang dilakukan adalah dengan Gray Level Co-occurrence Level (GLCM)[11] dan Local Binary Pattern (LBP) [12].

Gray Level Co-occurrence Level (GLCM) adalah metode statistik yang menggunakan hubungan spasial dari piksel. Ini adalah perhitungan dari angka pertemuan dua piksel yang berdekatan dalam sebuah gambar, pada lokasi yang memiliki nilai abu-abu sebesar i dan j. ukuran dari GLCM tergantung pada tingkat keabu-abuan yang dapat mengatur ukuran gambar dengan mengatur nilai intensitas. Untuk membuat

GLCM, dua parameter menunjukkan hubungan piksel dari arah yang berbeda dan jaraknya didefinisikan dalam offset.

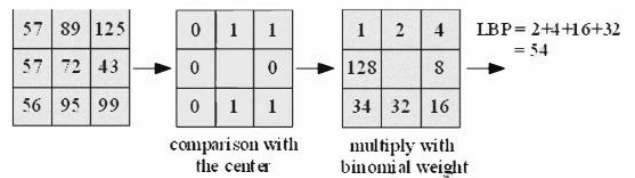
Untuk ekstraksi fitur tekstur dengan metode GLCM (Gray Level Co-occurrence Matrix) mencari nilai rata-rata Contrast, Correlation, Energy, Homogeneity untuk dijadikan parameternya dan didapatkan 16 parameter untuk masing-masing biji. [13] Parameter tekstur yang didapat diturunkan dari metode GLCM ini adalah Contrast, Correlation, Energy, Homogeneity untuk masing-masing sudut. Peneliti menggunakan 4 arah sudut pandang (degree = 0°, 45°, 90°, 135°), 8-jarak ketetanggaan (D = 1, 2,...,8) dan radius = 1.

Pola biner lokal (LBP) adalah jalur non-parameter yang menyederhanakan struktur spasial lokal dari suatu gambar. Ini adalah bentuk invariasi pada perubahan warna menjadi hitam putih. Untuk menghitung LBP, kotak berukuran 3x3 diterapkan pada gambar. Dengan diberi kotak, piksel yang akan lebih terlihat di antar piksel pusat dan sekitarnya dibandingkan dengan menggunakan persamaan. Hasil bilangan biner dikalikan dengan berat binomial yang diberikan untuk tiap piksel. Terakhir, hasil yang didapatkan dijumlahkan untuk memperoleh nilai LBP.

$$LBP(x_c, y_c) = \sum_{n=0}^7 s(i_n - i_c) 2^n$$

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Di mana i_n dan i_c adalah nilai kerapatan dari piksel pusat dan sekitarnya. [14]



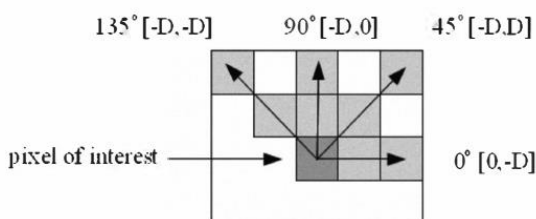
Untuk ekstraksi fitur tekstur dengan metode LBP (Local Binary Pattern) mencari nilai rata-rata dari histogram bin dan didapatkan 59 parameter yang digunakan untuk proses klasifikasi.

F. Random Forest

Setelah dilakukan ekstraksi fitur, dilakukan proses pengklasifikasian dengan berbagai metode untuk mendapatkan tingkat akurasi yang terbaik.

Pada paper ini, setelah proses ekstraksi fitur dilakukan, dan didapatkan beberapa parameter tiap ekstraksi fitur (warna, tekstur, bentuk) selanjutnya melakukan proses klasifikasi dengan menggunakan Random Forest dengan dataset dari nilai parameter yang didapat dari hasil ekstraksi yang tersebut diatas.

Breiman [15] mengusulkan Random Forest (RF), teknik klasifikasi yang disusun dengan menyusun sekumpulan pohon yang diinginkan. RF menggunakan variable respon ciri yang berbeda, yang akan di setiap pohon dan mengubah bagaimana klasifikasi atau pohon regresi terbentuk: setiap titik dipisah menggunakan yang terbaik di antara urutan prediksi yang dipilih secara acak dalam titik itu, dan pohon akan tumbuh meninggi tanpa mengalami pemendekan. Untuk prediksi data baru, RF mengumpulkan semua data hasil dari



pohon-pohon. Hal ini menjadi efektif dan cepat untuk data dengan jumlah yang banyak dan telah menunjukkan apabila dapat digunakan dengan baik dibandingkan program pengelompok yang lain.

Random Forest Classifier [16] merupakan pengembangan dari decision tree dengan menggunakan beberapa tree, dimana setiap tree telah dilakukan training menggunakan sampel individu dan setiap atribut dipecah pada tree yang dipilih antara subset yang bersifat acak. Dan pada proses klasifikasi, individunya didasarkan pada vote dari suara terbanyak pada kumpulan populasi tree.

Pada umumnya model klasifikasi ada dua proses yang akan dilakukan yaitu proses training dan proses testing. Data training didapatkan dari proses ekstraksi fitur (warna, tekstur dan bentuk) dengan berbagai parameternya untuk masing-masing biji yang diteliti seperti yang telah disebutkan diatas.

Sedangkan data testing, mengambil data secara random (acak) untuk masing-masing biji yang sudah dilakukan proses ekstraksi fitur.

Random forest yang dihasilkan memiliki banyak tree, dan setiap tree ditanam dengan cara yang sama. Tree dengan variabel x akan ditanam sejauh mungkin dengan tree dengan variabel y. dan dalam perkembangannya, sejalan dengan bertambahnya data set, maka tree pun ikut berkembang. Untuk prediksi data baru, RF mengumpulkan semua data hasil dari pohon-pohon. Hal ini menjadi efektif dan cepat untuk data dengan jumlah yang banyak. Beberapa fungsi learning yang dihasilkan Random Forest digunakan strategi ensemble “bagging” untuk mengatasi masalah overfitting apabila dihadapkan data set yang kecil.[17]

Hasil proses klasifikasi sebuah dataset secara random/acak dapat diidentifikasi secara tepat dari masing-masing biji yang diteliti. dan menunjukkan tingkat akurasi dari proses identifikasi 4 (empat) biji-bijian yaitu padi, jagung, kacang tanah dan kedelai.

III. PERANCANGAN

Pada pokok bahasan ini akan dibahas mengenai desain arsitektur dari system, alternative algoritma yang akan digunakan. Untuk blok diagram yang mempresentasikan proses penelitian digambarkan pada gambar 5.

A. Pre Processing

Proses pengambilan gambar (image) biji yang dilakukan peneliti dilakukan dengan menggunakan kamera merk Nikon D3400 lensa standart dan perlengkapannya. Selain itu proses pengambilan gambar (image) biji pada penelitian ini dilakukan dengan mendownload melalui Google Image Search. Ini dilakukan untuk mendapatkan varian biji yang bermacam-macam secara acak/random dengan jarak pengambilan gambar yang berbeda-beda karena biji- bijian yang diteliti mempunyai varian lebih dari satu..

B. Thresholding

Proses ini untuk memisahkan antara image foreground dan background dalam format gray image (Binarization) dimana hasil proses ini digunakan untuk melakukan ekstraksi fitur bentuk. Sedangkan untuk ekstraksi fitur warna dan fitur tekstur tidak melalui tahapan ini. Metode yang digunakan untuk proses thresholding ini adalah Otsu Algorithm.

C. Ekstraksi Fitur

Proses pengindeksan suatu database berupa citra (image) dengan isinya. Salah satu proses ekstraksi fitur adalah menganalisa berdasarkan isi visual seperti warna, bentuk dan tekstur.

• Ekstraksi Fitur Warna

Proses menganalisa warna dari sebuah citra (image) dimana sebuah citra tersusun dari piksel-piksel yang memiliki ukuran intensitas warna masing-masing. Sebaran warna di tiap-tiap piksel ditunjukkan oleh histogram. Histogram menunjukkan distribusi piksel berdasarkan intensitas graylevel (derajat keabuan) yang dimiliki tiap-tiap piksel. Penggunaan histogram sebagai metode ekstraksi ciri didasarkan pada perbedaan sebaran atau distribusi piksel di masing-masing gambar.

• Ekstraksi Fitur Tekstur

Merupakan informasi berupa susunan struktur permukaan suatu gambar. Fitur tekstur cocok untuk memvisualisasikan pola dan sifat permukaan. Ini menyediakan informasi tentang variasi intensitas permukaan gambar..

• Ekstraksi Fitur Bentuk

Merupakan konfigurasi oleh garis dan kontur. Fitur bentuk dikategorikan bergantung pada teknik yang digunakan. Kategori tersebut adalah berdasarkan batas (boundary-based) dan berdasarkan daerah (region-based). Teknik berdasarkan batas (boundary-based) menggambarkan bentuk daerah dengan menggunakan karakteristik eksternal.

D. Klasifikasi

Setelah proses ekstraksi fitur dilakukan, dan didapatkan beberapa parameter tiap ekstraksi fitur (warna, tekstur, bentuk) selanjutnya melakukan proses klasifikasi. Metode yang digunakan untuk proses klasifikasi ini adalah menggunakan Random Forest dengan dataset dari nilai parameter yang didapat dari hasil ekstraksi fitur.

IV. PENGUJIAN DAN ANALISIS

A. Dataset

Obyek penelitian adalah image (gambar) dari 4 varian biji-bijian (jagung, padi, kacang tanah, kedelai) dalam format gambar RGB dan JPEG dengan dimensi 300x500 pixel. Jumlah gambar yang diambil untuk 1 varian biji sebanyak 200 gambar. Jadi total gambar yang dijadikan obyek penelitian sejumlah 800 gambar. Untuk uji training sebanyak 75 % dari total gambar dan untuk testing sebanyak 25 % dari total gambar. Beberapa gambar yang menjadi data input:

a. Biji jagung



Gambar 1 Biji jagung

b. Biji kedelai



Gambar 2 Biji kedelai

c. Biji kacang



Gambar 3 Biji kacang

d. Biji padi



Gambar 4 Biji padi

B. Pengujian

Untuk pengerjaan tahap thresholding, tahap ekstraksi fitur (warna, bentuk dan tekstur) sampai tahap klasifikasi menggunakan library MATLAB R2015b.

Untuk uji training sebanyak 75 % dari total gambar dan untuk testing sebanyak 25 % dari total gambar. Untuk uji data training didapat sebanyak 150 gambar untuk masing-masing gambar biji yang sudah didapat.

Hasil Ekstraksi fitur warna dari data training sebanyak 75 % dari total gambar biji Jagung/Kacang Tanah/Kedelai/ Padi berdasarkan fitur warna RGB dan HSV menghasilkan 18 parameter. Ke-18 parameter itu antara lain R mean, R max, R min, G mean, G max, G min, B mean, B max, B min, H mean, H max, H min, S mean, S max, S min, V mean, V max, dan V min.

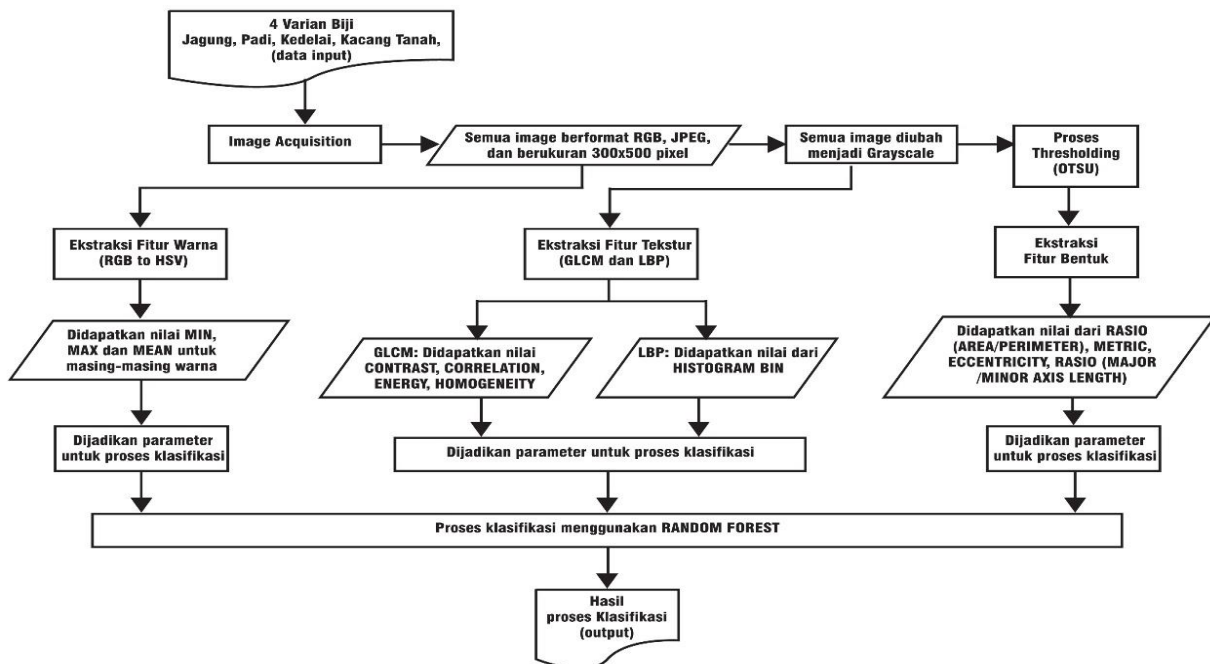
Untuk ekstraksi fitur tekstur dengan metode GLCM (Gray Level Co-occurrence Matrix) mencari nilai rata-rata Contrast, Correlation, Energy, Homogeneity untuk dijadikan parameter dan didapatkan 16 parameter untuk masing-masing biji.

Parameter tekstur yang didapat diturunkan dari metode GLCM ini adalah Contrast, Correlation, Energy, Homogeneity untuk masing-masing sudut. Peneliti menggunakan 4 arah sudut pandang ($degree = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}$), 8-jarak ketetanggaan ($D = 1, 2, \dots, 8$) dan radius = 1.

Untuk ekstraksi fitur tekstur dengan metode LBP mencari nilai rata-rata dari histogram bin dan didapatkan 59 parameter yang digunakan untuk proses klasifikasi.

Untuk ekstraksi fitur bentuk mencari nilai Rasio "L" (Area/Perimeter), Metric (roundness), Metric, Eccentricity, Equivalent Diameter, Rasio "M" (Major Axis Length/Minor Axis Length) untuk dijadikan parameter dan didapatkan 5 parameter untuk masing-masing biji.

Untuk pengerjaan uji testing yang diambil dari data training sebanyak 25 %, peneliti melakukan 5 kali uji testing untuk proses pengklasifikasian dengan pengambilan data gambar secara acak (random). Pada setiap uji testing diambil



Gambar 5 Alur Kerja

sebanyak 50 gambar untuk masing-masing gambar biji (jagung, kacang, kedelai, padi) sehingga didapat 200 gambar (image).

Pada uji testing ke-1, didapatkan tingkat akurasi sebesar 99 % dengan tabel confusion matrix dan tabel classification error untuk menampilkan tree dengan tingkat kesalahan yang paling kecil.

Pada uji testing ke-1 ini, tingkat classification error yang didapat terdapat pada number of grown trees ke 56 dengan nilai 0,0233. Dari nilai tersebut dapat dibangun sebuah decision tree untuk menentukan hasil klasifikasi dari sekumpulan data yang sudah dibuat.

TABEL 1
CONFUSION MATRIX TESTING 1

	Jagung	Kacang	Kedelai	Padi
Jagung	50	0	0	0
Kacang	0	50	0	0
Kedelai	0	0	48	2
Padi	0	0	0	50

Pada uji testing ke-2, didapatkan tingkat akurasi sebesar 100 % dengan tabel confusion matrix dan tabel classification error untuk menampilkan tree dengan tingkat kesalahan yang paling kecil.

Pada uji testing ke-2 ini, tingkat classification error yang didapat terdapat pada number of grown trees ke 56 dengan nilai 0,0233. Dari nilai tersebut dapat dibangun sebuah decision tree untuk menentukan hasil klasifikasi dari sekumpulan data yang sudah dibuat.

TABEL 2
CONFUSION MATRIX TESTING 2

	Jagung	Kacang	Kedelai	Padi
Jagung	50	0	0	0
Kacang	0	50	0	0
Kedelai	0	0	50	0
Padi	0	0	0	50

Pada uji testing ke-3, didapatkan tingkat akurasi sebesar 100 % dengan tabel confusion matrix dan tabel classification error untuk menampilkan tree dengan tingkat kesalahan yang paling kecil.

Pada uji testing ke-3 ini, tingkat classification error yang didapat terdapat pada number of grown trees ke 57 dengan nilai 0,0233. Dari nilai tersebut dapat dibangun sebuah decision tree untuk menentukan hasil klasifikasi dari sekumpulan data yang sudah dibuat.

TABEL 3
CONFUSION MATRIX TESTING 3

	Jagung	Kacang	Kedelai	Padi
Jagung	50	0	0	0
Kacang	0	50	0	0
Kedelai	0	0	50	0
Padi	0	0	0	50

Pada uji testing ke-4, didapatkan tingkat akurasi sebesar 100 % dengan tabel confusion matrix dan tabel classification error untuk menampilkan tree dengan tingkat kesalahan yang paling kecil.

Pada uji testing ke-4 ini, tingkat classification error yang didapat terdapat pada number of grown trees ke 57 dengan

nilai 0,0233. Dari nilai tersebut dapat dibangun sebuah decision tree untuk menentukan hasil klasifikasi dari sekumpulan data yang sudah dibuat.

TABEL 4
CONFUSION MATRIX TESTING 4

	Jagung	Kacang	Kedelai	Padi
Jagung	50	0	0	0
Kacang	0	50	0	0
Kedelai	0	0	50	0
Padi	0	0	0	50

Pada uji testing ke-5, didapatkan tingkat akurasi sebesar 100 % dengan tabel confusion matrix dan tabel classification error untuk menampilkan tree dengan tingkat kesalahan yang paling kecil.

Pada uji testing ke-5 ini, tingkat classification error yang didapat terdapat pada number of grown trees ke 56 dengan nilai 0,0233. Dari nilai tersebut dapat dibangun sebuah decision tree untuk menentukan hasil klasifikasi dari sekumpulan data yang sudah dibuat.

TABEL 5
CONFUSION MATRIX TESTING 5

	Jagung	Kacang	Kedelai	Padi
Jagung	50	0	0	0
Kacang	0	50	0	0
Kedelai	0	0	50	0
Padi	0	0	0	50

Peneliti menggunakan metode klasifikasi Random Forest Classifier dengan membangun sebanyak 100 tree untuk mendapatkan keputusan yang terbaik berdasarkan tingkat classification error yang dibuat. Setelah didapatkan nilai dari tingkat classification error, dipilih nilai yang paling kecil pada number of grown trees. Untuk tampilan tree, peneliti memilih menggunakan RULE BASED DECISION TREE FOR CLASSIFICATION berupa text

Pada 5 uji testing yang dilakukan, uji testing ke-1, ke-2 dan ke-5 pada tree ke-56 didapatkan nilai classification error paling kecil yaitu sebesar 0,0233. Sedangkan pada uji testing ke-3 dan ke-4 pada tree ke-57 didapatkan nilai classification error paling kecil yaitu sebesar 0,0233.

Hasil dari penelitian menunjukkan persentase tingkat akurasi dari 5 uji testing adalah testing ke-1 sebesar 99%. Sedangkan tingkat akurasi uji testing ke-2 sampai ke-5 sebesar 100 %.

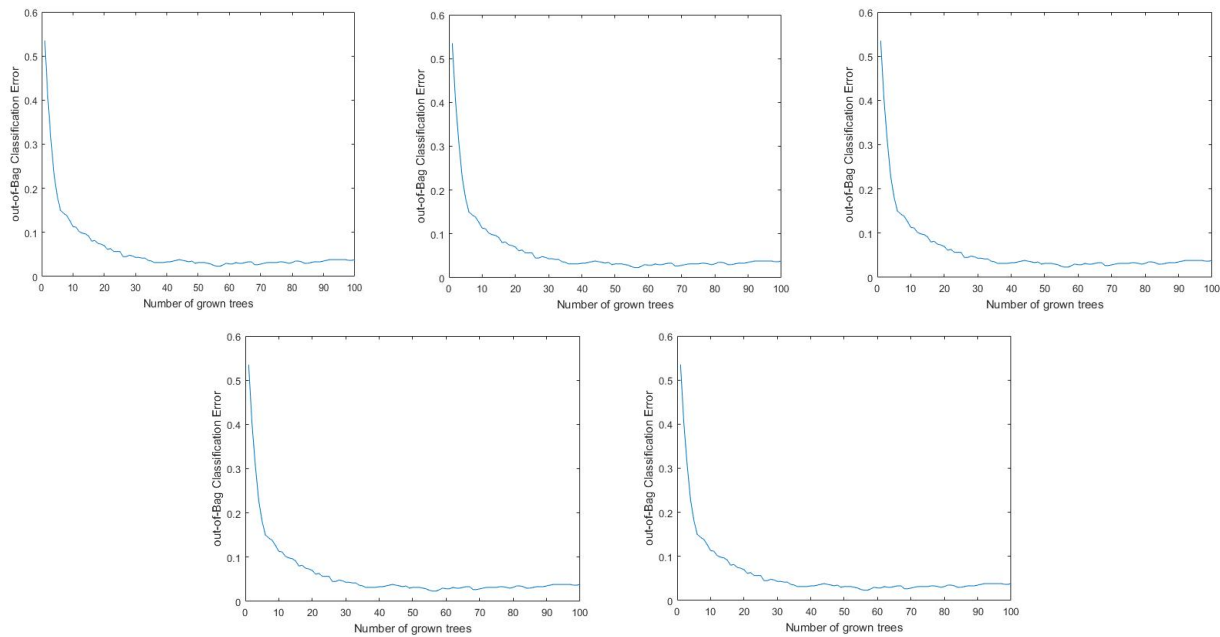
TABEL 6
TINGKAT AKURASI

Test 1	Test 2	Test 3	Test 4	Test 5	Rata-rata
99 %	100 %	100 %	100 %	100 %	99.8 %

Bisa disimpulkan persentase rata-rata dari uji testing untuk pengklasifikasian sebesar 99,8 % dan ini artinya bahwa identifikasi biji-bijian berdasarkan ekstraksi fitur warna, bentuk dan tekstur layak dilakukan.

V. KESIMPULAN

Berdasarkan hasil pengujian yang dilakukan didapatkan tingkat akurasi sebesar 99.8 %. Dapat disimpulkan bahwa pengambilan dataset berupa gambar atau image biji-bijian yang diteliti dapat dijadikan patokan untuk pengidentifikasian dan dapat dikembangkan dalam proses selanjutnya.



Gambar 6 Classification Error Testing

VI. DAFTAR PUSTAKA

[1] S. Sinurat *et al.*, “Analisa tekstur citra biji kemiri menggunakan metode filter gabor,” vol. 13, pp. 50–54, 2018.

[2] M. M. Sobel, R. Canny, P. Teguh, K. Putra, N. Kadek, and A. Wirdiani, “Pengolahan Citra Digital Deteksi Tepi Untuk Membandingkan Metode Sobel, Robert dan Canny,” *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, vol. 2, no. 2, pp. 253–261, 2016.

[3] E. G. Dianta, “Deteksi Tepi Menggunakan Metode Canny Dengan Matlab Untuk Membedakan Uang Asli dan Uang Palsu,” *J. Jur. Tek. Inform. Fak. Teknol. Ind. Univ. Gunadarma*, vol. 1, no. 1, pp. 1–13, 2012.

[4] I. R. G. A. Sugiarta, M. Sudarma, and I. M. O. Widyantara, “Ekstraksi Fitur Warna, Tekstur dan Bentuk untuk Clustered-Based Retrieval of Images (CLUE),” *Teknol. Elektro*, vol. 16, no. 1, pp. 85–90, 2017.

[5] D. Li, Y. Liu, L. Gao, and T. Management, “Research of Maize Seeds Classification Recognition Based on the Image Processing,” vol. 9, no. 11, pp. 181–190, 2016.

[6] S. I. Syafi’i, R. T. Wahyuningrum, and A. Muntasa, *Segmentasi Obyek Pada Citra Digital Menggunakan Metode Otsu Thresholding*, vol. 13, no. 1, 2016.

[7] M. Otsu, “Optimum Global Thresholding Using Otsu’s Method,” vol. 1, pp. 1–6.

[8] K. Kiliç, I. H. Boyaci, H. Köksel, and I. Küsmenoglu, “A classification system for beans using computer vision system and artificial neural networks,” *J. Food Eng.*, vol. 78, no. 3, pp. 897–904, 2007.

[9] K. Kiratiratanapruk and W. Sinthupinyo, “Color and texture for corn seed classification by machine vision,” *2011 Int. Symp. Intell. Signal Process. Commun. Syst. “The Decad. Intell. Green Signal Process. Commun. ISPACS 2011*, pp. 7–11, 2011.

[10] R. Binary, “Morphological Image Processing,” pp. 1–7, 2017.

[11] B. Pathak, A. Bhuyan, and D. Barooah, “Gray-Level Co-occurrence Matrix Implementation based on Edge Detection Information for Surface Texture Analysis,” pp. 1–5, 2014.

[12] M. Pietikäinen *et al.*, *Computational Imaging and Vision 40*, vol. 40, no. Comput. Vis. Using Local Bin. Patterns. 2011.

[13] D. A. Muhammad, R. R. W. Ken, and E. Siswanto, “Pencarian Citra Digital Berbasis Konten dengan,” *J. Telemat.*, vol. 8, no. 2, pp. 8–13, 1858.

[14] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[15] Y. L. Pavlov, “Random forests,” *Random For.*, pp. 1–122, 2019.

[16] N. Horning, “Random Forests: An algorithm for image classification and generation of continuous fields data sets,” *Int. Conf. Geoinformatics Spat. Infrastruct. Dev. Earth Allied Sci. 2010*, pp. 1–6, 2010.

[17] P. T. T. Hong, T. T. T. Hai, L. T. Lan, V. T. Hoang, V. Hai, and T. T. Nguyen, “Comparative Study on Vision Based Rice Seed Varieties Identification,” *Proc. - 2015 IEEE Int. Conf. Knowl. Syst. Eng. KSE 2015*, pp. 377–382, 2016.