

Classification of Noised and Restored Deepfruit Image Using Random Forest Algorithm

Eva Fitriyaningsih
Departemen Ilmu Komputer
IPB University
Bogor, Indonesia
fitriaeva@apps.ipb.ac.id

Halida Fiadnin
Departemen Ilmu Komputer
IPB University
Bogor, Indonesia
halida.fiadnin@apps.ipb.ac.id

Illinia Malika Putri
Departemen Ilmu Komputer
IPB University
Bogor, Indonesia
ikaputri@apps.ipb.ac.id

Ismay Fana Fillah
Departemen Ilmu Komputer
IPB University
Bogor, Indonesia
ismyfanafillah@apps.ipb.ac.id

Novia Dwi Lestari
Departemen Ilmu Komputer
IPB University
Bogor, Indonesia
novialestari@apps.ipb.ac.id

Abstract—This study examines the impact of Gaussian and salt-and-pepper noise on the performance of image classification models using the DeepFruits dataset. The dataset contains annotated images of various fruits, enabling robust training and evaluation of object detection models. We simulated real-world conditions by adding Gaussian and salt-and-pepper noise to the images and applied various filtering techniques, including mean, median, Gaussian, and bilateral filters, to restore the images. Our findings show that Gaussian filtering is highly effective for Gaussian noise, while median filtering performs best for salt-and-pepper noise. However, adding noise significantly degrades classification performance. Using a Random Forest classifier, we observed a decrease in accuracy, precision, recall, and F1-score with noisy images. Post-restoration, the accuracy improved but remained suboptimal, indicating the need for more advanced denoising techniques. This research emphasizes the difficulties posed by noise in image processing and highlights the need for robust models and advanced filtering techniques to improve image classification accuracy.

Keywords—Image Processing, Gaussian Noise, Salt-and-Pepper Noise, Image Restoration, Image Classification, Random Forest

I. PENDAHULUAN

This template, modified in MS Word 2007 and saved as a “Word 97-2003 Document” for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-levelled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

II. TINJAUAN PUSTAKA

A. Image Restoration

Restorasi citra bertujuan untuk mengembalikan citra yang telah mengalami degradasi, memperbaiki ketidakakuratan geometris dan radiometrinya. Dalam era citra digital saat ini, restorasi citra memainkan peran krusial karena keterbatasan alat optik digital seperti kamera dapat menghasilkan citra yang kabur atau mengandung *noise*. *Noise* ini bisa disebabkan oleh sejumlah faktor, termasuk keterbatasan perangkat keras maupun intervensi manusia [1].

Restorasi citra melibatkan upaya untuk memperoleh citra yang lebih jelas dari citra yang telah terdegradasi, dengan memahami faktor-faktor degradasi yang mempengaruhinya. Meskipun sering disamakan dengan peningkatan kualitas citra, restorasi citra berbeda dari *image enhancement*. *Image enhancement* lebih menekankan pada penajaman fitur-fitur tertentu dalam citra, sementara restorasi citra memanfaatkan pemahaman tentang proses degradasi untuk mengembalikan citra aslinya [1].

B. Gaussian Noise

Gaussian Noise merupakan gangguan statistik yang memiliki *probability density function* yang mirip dengan distribusi normal [2]. *Probability density function* (PDF) dari sebuah variabel random gaussian adalah berikut:

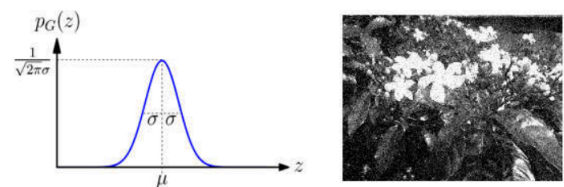


Fig. 1. (a) Probability density function Gaussian (b) Gambar dengan Gaussian noise

$$p_G(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/(2\sigma^2)} \quad (1)$$

Dalam persamaan (1), Z mewakili tingkat keabuan, μ adalah nilai rata-rata, dan σ adalah standar deviasi. Persamaan ini menunjukkan bagaimana probabilitas kejadian dari nilai keabuan tertentu dalam distribusi Gaussian noise, diukur oleh parameter μ dan σ . Gaussian noise dalam gambar digital utamanya muncul pada saat akuisisi, seperti ketika terjadi sensor noise akibat kondisi pencahayaan yang buruk atau suhu yang tinggi, serta pada saat transmisi data seperti noise yang dihasilkan oleh sirkuit elektronik [3]. Untuk mengurangi efek Gaussian noise dalam pengolahan gambar digital, filter spasial dapat digunakan. Namun saat meratakan gambar menggunakan filter ini ada kemungkinan efek yang tidak diinginkan seperti kabur pada tepi dan detail gambar. Hal ini disebabkan karena komponen-komponen tersebut termasuk dalam frekuensi tinggi yang akan difilter selama proses *smoothing*.

C. Salt and Pepper Noise

Salt and Pepper noise terjadi karena adanya gangguan tajam dan tiba-tiba dalam nilai abu-abu pada gambar. Gangguan ini bisa disebabkan oleh berbagai faktor, termasuk kesalahan selama proses akuisisi gambar, seperti kesalahan dalam transmisi data atau kerusakan pada sensor kamera. Misalnya, ketika gambar ditransmisikan melalui saluran yang tidak stabil atau terdapat gangguan saat gambar direkam, beberapa piksel dalam gambar dapat mengalami fluktuasi tiba-tiba yang menyebabkan mereka berubah menjadi nilai yang ekstrem, yaitu putih atau hitam yang tersebar secara acak di seluruh gambar [3].

$$\begin{aligned} PI(Z) &= P_a \text{ for } Z = a \\ PI(Z) &= P_b \text{ for } Z = b \\ PI(Z) &= 0 \text{ Otherwise} \end{aligned} \quad (2)$$

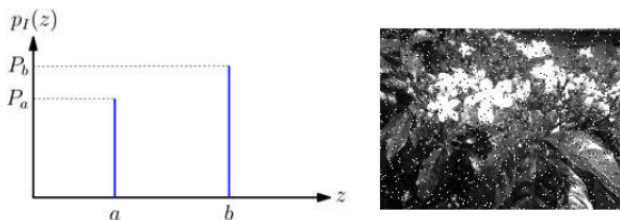


Fig. 2. (a) Probability density function Salt and Pepper (b) Gambar dengan Salt and Pepper noise

PDF dari model noise impuls bipolar diberikan oleh persamaan (2). Jika P_a lebih besar dari P_b , tingkat abu-abu b akan muncul sebagai titik putih (salt). Sebaliknya, a akan muncul sebagai titik gelap (lada). Jika salah satu dari P_a atau P_b adalah 0, maka PDF tersebut disebut unipolar [2].

D. Spatial Filtering

Pengolahan spasial merupakan proses di mana kita dapat mengubah karakteristik dari sebuah gambar optik dengan memilih untuk menghapus sejumlah frekuensi spasial tertentu yang membentuk suatu objek. Misalnya, ini dapat dilakukan dengan menyaring data video yang diterima dari satelit dan pesawat ruang angkasa, atau menghapus garis-garis raster dari gambar televisi atau gambar yang

telah dipindai. Dalam konteks pengolahan citra digital, istilah "citra" mengacu pada fungsi intensitas cahaya dua dimensi, di mana tingkat amplitudo pada koordinat spasial (x, y) menentukan tingkat intensitas atau kecerahan pada titik tersebut dalam citra [4].

Filtering spasial merupakan teknik yang umum digunakan dalam interferometri digital [123,124]. Untuk menjaga *phase jump*, filtering dilakukan pada bagian sinus dan kosinus dari fasa. Beberapa jenis kernel dapat dipertimbangkan: *the moving average filter*, *filter median* (efektif untuk noise impuls), atau *filter Gaussian*. Ukuran kernel dapat bervariasi dari 3×3 hingga 21×21 untuk citra dengan ukuran 512×512 atau 1024×1024 piksel. Filter-filter ini umumnya mudah diimplementasikan dan menjadi metode referensi untuk membandingkan algoritma-algoritma yang lebih kompleks [4].

E. Order Statistic Filtering

Pengolahan citra dengan menggunakan *order statistic filter* adalah pendekatan yang efektif untuk mengurangi noise dalam citra. Filter ini bekerja dengan cara memilih nilai piksel tertentu dari sejumlah tetangga piksel, seperti yang dilakukan dalam median filtering, untuk menghasilkan nilai baru untuk piksel target. Median Filtering terkenal efektif dalam mengurangi jenis noise salt-and-pepper tanpa mengorbankan detail penting dalam citra. Namun, untuk menerapkan teknik ini, penting untuk memperhatikan ukuran kernel yang digunakan dan karakteristik noise yang ada dalam citra. Teknik ini sering digunakan dalam pengolahan citra untuk menjaga kejelasan citra sambil mengurangi gangguan noise [5].

Order Statistics Filtering menawarkan pendekatan yang sederhana namun efektif untuk mengatasi noise dalam citra. Keunggulan utamanya termasuk kemampuan untuk mempertahankan resolusi citra dan implementasi yang relatif mudah. Meskipun demikian, perlu diingat bahwa pemilihan ukuran kernel dan jenis filter yang sesuai sangat penting untuk mendapatkan hasil yang optimal. Selain itu, efektivitas teknik ini mungkin berkurang jika citra mengandung noise dengan distribusi spasial yang kompleks. Penggunaan Order Statistics Filtering harus disesuaikan dengan jenis dan tingkat keparahan noise dalam citra untuk hasil yang optimal [6].

F. Random Forest

Random Forest adalah salah satu algoritma pembelajaran mesin yang sering digunakan dalam klasifikasi dan regresi. Algoritma ini bekerja dengan menggabungkan beberapa pohon keputusan dalam suatu ensemble untuk meningkatkan akurasi prediksi. Setiap pohon keputusan dihasilkan secara acak dari sampel dataset yang diambil dengan penggantian, dan prediksi akhir didasarkan pada mayoritas suara dari semua pohon keputusan. Keunggulan utama dari Random Forest adalah kemampuannya untuk mengatasi overfitting dan mengatasi masalah data yang tidak seimbang [7].

Selain itu, Random Forest juga memiliki kemampuan untuk menangani dataset dengan fitur-fitur yang kompleks

dan korelasi yang tinggi antar fitur. Hal ini membuatnya menjadi pilihan yang populer dalam berbagai aplikasi, termasuk klasifikasi citra, pengenalan pola, dan analisis bioinformatika. Metode ensemble yang digunakan dalam Random Forest memungkinkan penggunaan beragam jenis fitur dan meminimalkan risiko overfitting. Namun, penting untuk memperhatikan bahwa Random Forest cenderung memiliki kompleksitas komputasi yang tinggi, terutama jika jumlah pohon dan fitur dalam ensemble besar [8].

Penerapan Random Forest telah berhasil dalam berbagai domain, termasuk di bidang kedokteran, keuangan, dan bioinformatika. Kemampuannya untuk mengatasi berbagai masalah dalam klasifikasi dan regresi membuatnya menjadi salah satu algoritma yang paling banyak digunakan dalam pembelajaran mesin. Meskipun demikian, seperti halnya dengan semua algoritma pembelajaran mesin, penting untuk memilih parameter yang tepat dan melakukan validasi yang cermat untuk memastikan performa yang optimal dari model Random Forest.

III. METODE

A. Data

Dalam penelitian ini, dataset DeepFruits yang tersedia di Roboflow Universe telah digunakan. Dataset ini dapat diakses pada tautan <https://universe.roboflow.com/research-sakib/deepfruits-9nwps>. Dataset DeepFruits terdiri dari sejumlah gambar berbagai jenis buah yang telah dianotasi untuk keperluan pelatihan model deteksi objek. Setiap gambar dilengkapi dengan anotasi yang menunjukkan lokasi dan jenis buah yang terdapat dalam gambar tersebut, yang sangat berguna untuk pengembangan dan evaluasi model berbasis visi komputer.

Pemilihan dataset ini didasarkan pada kualitas anotasi yang lengkap serta keragaman jenis buah yang disertakan, mencakup berbagai kondisi pencahayaan dan sudut pandang. Hal ini memungkinkan model yang diusulkan untuk belajar mengenali buah dalam berbagai kondisi, sehingga meningkatkan generalisasi dan akurasi model dalam aplikasi dunia nyata. Untuk memastikan verifikasi dan reproduksi hasil penelitian, dataset ini dapat diakses secara publik melalui tautan yang telah disebutkan.

B. Tahapan

1) Input Data

Pada tahapan ini, data citra diinputkan, data tersebut berasal dari dataset DeepFruits yang telah tersedia di Roboflow Universe yang dapat diakses melalui tautan berikut <https://universe.roboflow.com/research-sakib/deepfruits-9nwps>. DeepFruits terdiri dari gambar-gambar berbagai jenis buah yang telah dianotasi untuk keperluan pelatihan model deteksi objek.

2) Eksplorasi data

Tahapan eksplorasi data merupakan langkah awal untuk memahami karakteristik dan struktur dataset yang digunakan. Dataset dianalisis secara deskriptif untuk mendapatkan informasi dasar seperti jumlah gambar, dan distribusi label. Analisis ini mencakup pembuatan tabel ringkasan dan visualisasi data untuk memahami distribusi dan karakteristik utama dari dataset. Pada tahapan ini dilakukan visualisasi data untuk melihat pola-pola yang ada dalam dataset. Visualisasi ini berupa histogram, scatter plot, maupun heatmap yang sesuai untuk menggambarkan distribusi data dan hubungan antar variabel dalam dataset dari pola yang ada.

3) Penambahan Noise

Penambahan noise citra merupakan salah satu yang utama pemrosesan citra pada penelitian ini, bertujuan untuk mensimulasikan kondisi nyata yang mana citra dapat terpengaruh oleh berbagai jenis gangguan. Tahapan proses penambahan noise sebagai berikut.

a. Inisialisasi library dan pemuatan citra

Inisialisasi library dilakukan dengan mengimpor berbagai library yang diperlukan dalam mendefinisikan fungsi-fungsi untuk menambahkan berbagai jenis noise pada citra. Kemudian dilakukan pemuatan citra dengan mengambil citra dari path pada CSV dataset dan dibaca dalam mode grayscale untuk memudahkan pemrosesan lebih lanjut.

b. Penambahan Noise

Penambahan noise citra merupakan teknik dalam pemrosesan citra digital untuk mensimulasikan gangguan yang mungkin terjadi pada gambar dalam kondisi nyata. Tahapan ini dimulai dengan memastikan citra dalam format yang sesuai, seperti grayscale. Dilakukan beberapa percobaan untuk pemilihan jenis noise yang memungkinkan untuk digunakan pada penelitian ini seperti gaussian noise, salt and pepper noise, uniform noise, exponential noise, gamma noise, dan rayleigh noise. Pada penelitian ini, jenis noise yang akan ditambahkan adalah Gaussian Noise dengan distribusi normal untuk mensimulasikan noise sensor elektronik dan Salt-and-Pepper Noise yang menambahkan piksel putih-hitam acak untuk mensimulasikan gangguan impuls. Penambahan Gaussian Noise terhadap citra dilakukan dengan menentukan parameter mean dan standar deviasi, sedangkan Salt-and-Pepper Noise dengan menentukan proporsi piksel salt dan pepper. Citra hasil kemudian dinormalisasi untuk memastikan nilai piksel tetap dalam rentang yang sesuai, misalnya 0-255 untuk citra 8-bit.

c. Penyimpanan citra

Pada tahap ini, dibuat fungsi untuk melakukan iterasi pada seluruh citra input, termasuk dataset train, validation, dan test. Noise akan diterapkan pada setiap gambar, dan hasilnya akan disimpan dalam bentuk gambar yang telah diproses.

4) Restorasi

Setelah data ditambahkan noise, langkah selanjutnya adalah melakukan restorasi. Restorasi dilakukan kepada citra yang telah terkena noise, dengan tujuan mengembalikan citra ke kondisi yang lebih baik. Tahapan kegiatan restorasi mencakup:

a. Pengolahan gambar awal

Pengolahan gambar awal dilakukan dengan mengimport library yang diperlukan dan mendefinisikan fungsi-fungsi untuk menerapkan berbagai jenis filter pada gambar, seperti mean filter, median filter, gaussian filter, dan bilateral filter.

b. Pengujian filter

Pengujian dilakukan untuk beberapa jenis filter yaitu mean, median, gaussian, dan bilateral, dengan berbagai ukuran filter yaitu 3x3, 5x5, 7x7, 9x9. Mean filtering dan gaussian filtering merupakan contoh dari jenis Spatial Filtering, sedangkan median filtering merupakan contoh dari jenis Order Statistic Filtering, dan bilateral filtering merupakan contoh dari jenis Adaptive Filtering. Setelah pengujian, filter terbaik akan dipilih untuk digunakan dalam proses restorasi.

c. Proses Restorasi

Setelah pengujian, filter terbaik untuk setiap jenis noise akan dipilih. Filter kemudian diterapkan pada keseluruhan data yang mengandung noise tertentu. Hasil dari penerapan filter akan ditampilkan bersama dengan gambar awal yang mengandung noise. Penyajian gambar bertujuan untuk membandingkan hasil restorasi dan mengukur efektivitasnya.

d. Penyimpanan gambar dengan filter

Pada tahap ini, dibuat fungsi untuk melakukan iterasi pada seluruh gambar input, termasuk dataset train, validation, dan test. Filter akan diterapkan pada setiap gambar, dan hasilnya akan disimpan dalam bentuk gambar yang telah diproses.

5) Klasifikasi

Setelah data dilakukan restorasi, langkah selanjutnya adalah melakukan klasifikasi untuk menentukan label data pada tiap dataset. Tahapan klasifikasi mencakup:

a. Ekstraksi fitur

Tahap ini melibatkan pengambilan informasi fitur-fitur seperti tekstur, warna, dan bentuk yang dapat diekstrak dari gambar. Tujuan utamanya adalah untuk mengurangi kompleksitas data dan menyoroti aspek yang paling penting untuk analisis lebih lanjut.

b. Pemodelan dan evaluasi

Setelah fitur-fitur diekstrak, model pembelajaran mesin diterapkan pada data yang telah diproses. Model yang digunakan dalam penelitian ini adalah model Random Forest. Langkah pada tahap ini melibatkan pelatihan model menggunakan data latih, penyesuaian parameter, dan evaluasi kinerja model menggunakan data validasi atau uji. Metrik evaluasi seperti akurasi, presisi, recall, dan F1-score digunakan untuk mengevaluasi kinerja model.

c. Visualisasi

Visualisasi adalah komponen penting dalam proses analisis data. Visualisasi melibatkan representasi grafis dari informasi atau hasil dari model. Confusion matrix dapat digunakan untuk memvisualisasikan kinerja model secara lebih intuitif.

d. Prediksi dengan input data baru

Setelah model dilatih dan dievaluasi, model tersebut dapat digunakan untuk membuat prediksi pada data baru yang tidak pernah dilihat sebelumnya. Pada tahap ini, fitur-fitur dari data baru juga diekstrak, kemudian digunakan sebagai input untuk model yang telah dilatih untuk memprediksi kelas atau label yang sesuai.

C. Ruang Lingkup

- 1) Data yang digunakan merupakan data DeepFruits yang tersedia di Roboflow Universe.
- 2) Model noise yang digunakan merupakan model noise gaussian dan model noise salt and pepper.
- 3) Restorasi citra dilakukan menggunakan filter median 3x3 dan filter gaussian 5x5.
- 4) Klasifikasi dilakukan untuk menentukan label buah setelah citra ditambahkan noise dan menjalani proses restorasi. Metode Random Forest dari library

scikit-learn (sklearn) digunakan untuk melakukan klasifikasi.

IV. HASIL DAN PEMBAHASAN

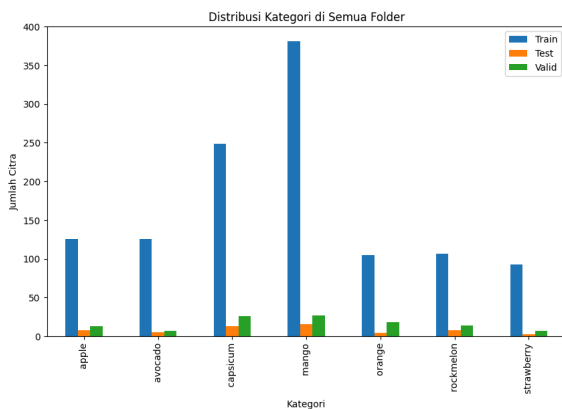
Proses pengolahan citra pada penelitian ini terdiri dari beberapa proses, diantaranya adalah eksplorasi data, penambahan noise pada citra, restorasi citra yang telah diberi noise, serta klasifikasi citra baik saat citra masih terkena noise maupun setelah proses restorasi.

A. Eksplorasi Data

Eksplorasi pada data dilakukan untuk mengetahui informasi pada dataset serta persebarannya. Terdapat 3 dataset yang dieksplorasi yaitu *classes* dari masing-masing folder yaitu, train, test, dan valid. Atribut (atau kolom) yang terdapat dari masing-masing file adalah 'filename' dan jenis buah pada citra. Terdapat 7 jenis buah yaitu 'apple', 'avocado', 'capsicum', 'mango', 'orange', 'rockmelon', dan 'strawberry', sehingga terdapat 8 atribut (nama file dan 7 jenis buah) untuk setiap file csv.

Eksplorasi dimulai dengan membaca masing-masing file, kemudian dilakukan pengecekan tipe data dan nilai null pada dataset. Hasilnya sama untuk ketiga folder yaitu 'filename' memiliki tipe data *object* dan untuk 7 jenis buah semua memiliki tipe data *integer*. Ketiga folder juga tidak ada yang memiliki nilai null, dengan jumlah data pada masing-masing folder yaitu 1188 baris untuk dataset train, 57 baris untuk dataset test, dan 112 baris untuk dataset valid.

Persebaran jumlah buah dalam dataset dicari menggunakan histogram dan table. Pada fig (?) terlihat bahwa jenis buah yang paling banyak adalah 'mango' pada folder train sebanyak 381 buah. Pada fig (?) juga dijelaskan lebih detail terkait jumlah persebaran masing-masing jenis buah.



fig

	Train	Test	Valid
apple	126	8	13
avocado	126	5	7
capsicum	249	13	26
mango	381	16	27
orange	105	4	18
rockmelon	107	8	14
strawberry	93	3	7

fig

B. Penambahan Noise

Terdapat 2 jenis noise yang akan ditambahkan pada citra, yaitu *gaussian noise* dan *salt and pepper noise*. Gaussian noise memiliki distribusi probabilitas berbentuk lonceng yang didefinisikan oleh dua parameter utama: mean dan standar deviasi. Mean diberi nilai ke 0 untuk memastikan bahwa noise tidak menambahkan bias pada citra. Standar deviasi menentukan seberapa besar variasi atau penyebaran nilai noise dari mean. Semakin besar nilai σ , semakin besar variasi noise. Pada penelitian ini nilai standar deviasi yang digunakan adalah 30.

Lalu salt-and-pepper noise mengubah nilai intensitas beberapa piksel dalam citra menjadi nilai maksimum (255 untuk putih) atau nilai minimum (0 untuk hitam) secara acak. Dua parameter utama yang digunakan adalah probabilitas salt (*salt_prob*) dan probabilitas pepper (*pepper_prob*). Nilai yang digunakan untuk *salt_prob* dan *pepper_prob* adalah 0.02.

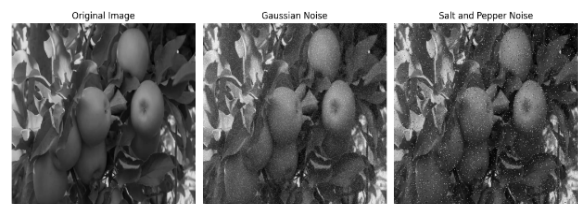


Fig. 3. Citra hasil penambahan noise

Selanjutnya, citra yang telah ditambahkan noise akan disimpan ke dalam folder baru.

```
Saved noisy image to C:\Users\nia\OneDrive\Documents\.\
\valid_salt\yum_png_jpg.rf.a5f72b744fa271db03a798f2bdf3:
Processed add noise using salt and pepper noise to image
Processed CSV file saved as: C:\Users\nia\OneDrive\Docu
mentation\valid_salt\_classes.csv
```

Fig. 4. Penyimpanan citra dan CSV file hasil

C. Restorasi Citra

Efektivitas teknik restorasi harus dievaluasi dengan menerapkan filter pada masing-masing citra yang telah diberi noise Gaussian dan noise Salt-and-pepper. Teknik yang digunakan adalah mean filtering, median filtering, gaussian filtering, dan bilateral filtering.

1) Restorasi Citra Gaussian Noise

Gaussian noise adalah jenis noise yang umum terjadi pada citra digital, yang disebabkan oleh fluktuasi acak dalam intensitas piksel yang mengikuti distribusi Gaussian. Noise ini sering diakibatkan oleh sensor kamera yang kurang optimal atau kondisi pencahayaan yang buruk. Untuk mengatasi noise ini, berbagai teknik filtering digunakan, yaitu mean filtering, median filtering, gaussian filtering, dan bilateral filtering. Setiap teknik memiliki kelebihan dan kelemahan tersendiri dalam menghilangkan noise dan mempertahankan detail citra.

Eksperimen dilakukan pada beberapa citra uji yang telah ditambahkan Gaussian noise. Parameter yang digunakan untuk setiap teknik filtering adalah mean filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9, median filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9, gaussian filtering dengan ukuran kernel 3x3, 5x5, 7x7, 9x9 dan sigma = 1, bilateral filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9 serta sigma warna = 75 dan sigma ruang = 75. Berikut adalah perbandingan antara keempat teknik tersebut berdasarkan hasil eksperimen yang telah dilakukan.

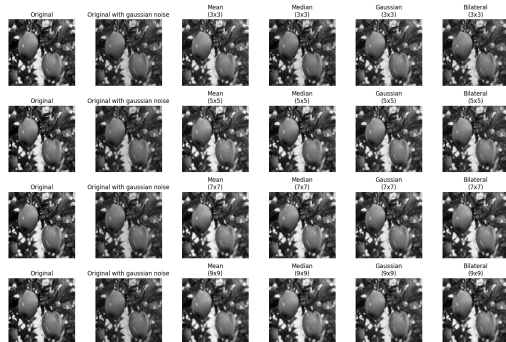


Fig. 5. Perbandingan 4 filtering terhadap citra yang diberi noise gaussian

Pada Fig. 5, hasil filtering menunjukkan bahwa mean filtering efektif dalam mengurangi Gaussian noise. Namun, efek samping dari mean filtering adalah membuat citra menjadi agak buram dan kehilangan beberapa detail. Ukuran kernel yang lebih besar, seperti 7x7 dan 9x9, lebih baik dalam menghilangkan noise, tetapi mengakibatkan penurunan yang lebih besar dalam ketajaman citra.

Median filtering tidak efektif untuk Gaussian noise. Teknik ini lebih cocok untuk mengatasi noise tipe impuls seperti salt-and-pepper, tetapi tidak efektif dalam menangani distribusi noise Gaussian. Hasil median filtering menunjukkan bahwa noise tetap ada dan detail citra juga terganggu.

Gaussian filtering sangat efektif dalam menghilangkan Gaussian noise. Dengan menggunakan ukuran kernel 5x5 dan sigma = 1, noise dapat dikurangi secara signifikan tanpa

mengorbankan banyak detail citra. Gaussian filtering mampu mempertahankan tepi citra dengan baik, menjadikannya pilihan yang ideal untuk restorasi citra dengan Gaussian noise.

Bilateral filtering juga menunjukkan hasil yang baik dalam menghilangkan Gaussian noise. Teknik ini tidak hanya mengurangi noise tetapi juga mempertahankan tepi dan detail citra dengan baik. Meskipun prosesnya lebih lambat dan memerlukan penyesuaian parameter yang lebih kompleks, bilateral filtering memberikan hasil yang sangat baik untuk citra dengan noise Gaussian.

2) Restorasi Citra Salt-and-Pepper Noise

Salt-and-pepper noise adalah jenis noise yang umum terjadi pada citra digital, di mana piksel acak berubah menjadi hitam atau putih. Noise ini sering diakibatkan oleh kesalahan dalam transmisi data atau kerusakan dalam memori. Untuk mengatasi noise ini, berbagai teknik filtering digunakan, yaitu mean filtering, median filtering, gaussian filtering, dan bilateral filtering. Setiap teknik memiliki kelebihan dan kelemahan tersendiri dalam menghilangkan noise dan mempertahankan detail citra.

Eksperimen dilakukan pada beberapa citra uji yang telah ditambahkan salt-and-pepper noise. Parameter yang digunakan untuk setiap teknik filtering adalah mean filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9, median filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9, gaussian filtering dengan ukuran kernel 3x3, 5x5, 7x7, 9x9 dan sigma = 1, bilateral filtering dengan ukuran kernel 3x3, 5x5, 7x7, dan 9x9 serta sigma warna = 75 dan sigma ruang = 75. Berikut adalah perbandingan antara keempat teknik tersebut berdasarkan hasil eksperimen yang telah dilakukan.

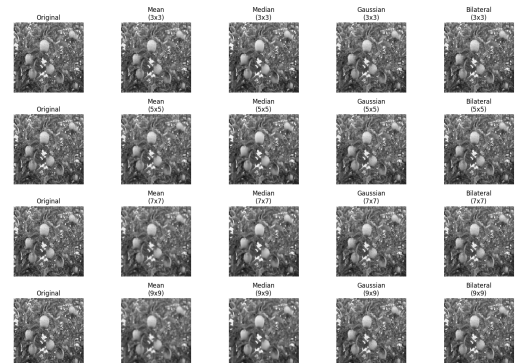


Fig. 6. Perbandingan 4 filtering terhadap citra yang diberi noise salt-and-pepper

Pada Fig. 4, dapat terlihat bahwa mean filtering kurang efektif. Mean filtering, meskipun merupakan salah satu teknik yang paling sederhana dalam mengatasi noise, namun kurang efektif dalam menghilangkan salt-and-pepper noise. Meskipun berhasil mengurangi sebagian kecil dari

noise, citra yang telah difilter cenderung memiliki tepi dan detail yang buram. Hasilnya, banyak noise yang tetap ada dalam citra yang difilter, mengakibatkan kualitas keseluruhan citra yang menurun.

Median filtering, dengan ukuran kernel 3x3, terbukti sebagai teknik yang paling efektif dalam menghilangkan salt-and-pepper noise. Dibandingkan dengan teknik lainnya, median filtering mampu mengurangi noise secara signifikan sambil mempertahankan tepi dan detail citra. Namun, ukuran kernel yang lebih besar, seperti 5x5, 7x7, dan 9x9, menghasilkan penurunan yang signifikan dalam ketajaman citra dan efektivitas dalam menghilangkan noise.

Gaussian filtering, meskipun sering digunakan untuk mengurangi Gaussian noise, ternyata tidak efektif dalam mengatasi salt-and-pepper noise. Teknik ini cenderung menghaluskan citra secara keseluruhan tanpa menghilangkan noise yang signifikan, sehingga noise masih terlihat jelas setelah proses filtering. Selain itu, gaussian filtering dapat mengakibatkan hilangnya detail citra.

Bilateral filtering menunjukkan hasil yang cukup baik dalam menghilangkan salt-and-pepper noise, tetapi tidak seefektif median filtering. Teknik ini mempertahankan tepi dan detail citra dengan lebih baik daripada mean dan gaussian filtering. Namun, prosesnya lebih lambat dan memerlukan penyesuaian parameter yang lebih kompleks untuk mencapai hasil yang optimal. Meskipun demikian, bilateral filtering tetap menjadi alternatif yang layak untuk mengurangi salt-and-pepper noise dalam citra.

D. Klasifikasi

1) Klasifikasi Noise

Penelitian ini dilakukan untuk menganalisis dan mengevaluasi kinerja model klasifikasi data awal dan membandingkannya dengan data yang diberikan noise. Data pertama merupakan citra asli tanpa modifikasi, sedangkan data kedua adalah citra yang telah diberikan noise Gaussian dan salt-and-pepper. Model klasifikasi yang digunakan adalah RandomForestClassifier, dan performanya dievaluasi berdasarkan akurasi, precision, recall, dan f1-score.

Validation Accuracy: 0.79
Test Accuracy: 0.72

Validation Classification Report:				
	precision	recall	f1-score	support
apple	0.86	0.46	0.60	13
avocado	1.00	0.29	0.44	7
capsicum	0.96	0.88	0.92	26
mango	0.59	0.81	0.69	27
orange	0.79	0.83	0.81	18
rockmelon	0.81	0.93	0.87	14
strawberry	1.00	1.00	1.00	7
accuracy			0.79	112
macro avg	0.86	0.74	0.76	112
weighted avg	0.82	0.79	0.78	112

Test Classification Report:				
	precision	recall	f1-score	support
apple	1.00	0.38	0.55	8
avocado	1.00	0.00	0.00	5
capsicum	0.92	0.92	0.92	13
mango	0.54	0.88	0.67	16
orange	0.75	0.75	0.75	4
rockmelon	0.86	0.75	0.80	8
strawberry	0.75	1.00	0.86	3
accuracy			0.72	57
macro avg	0.83	0.67	0.65	57
weighted avg	0.80	0.72	0.68	57

Fig. 7. Evaluasi model klasifikasi data citra awal

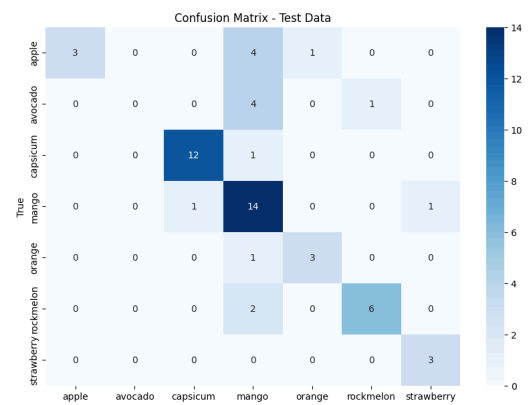


Fig. 8. Confusion matrix data citra awal

The predicted label for the image :
s: capsicum with confidence: 0.77

Fig. 9. Uji model klasifikasi data citra awal

Model pertama dilatih menggunakan data citra asli. Hasil evaluasi menunjukkan akurasi validasi sebesar 42% dan akurasi uji sebesar 30%. Pada validasi, kelas 0 menunjukkan performa terbaik dengan f1-score sebesar 0.54 dan recall sebesar 0.77, meskipun precision hanya 0.42. Kelas 1 dan 2 menunjukkan f1-score masing-masing sebesar 0.32 dan 0.33. Pada data uji, kelas 0 tetap memiliki performa terbaik dengan f1-score 0.35 dan recall 0.50, sementara precision sebesar 0.27. Kelas 2 memiliki recall tinggi sebesar 0.69 namun precision rendah sebesar 0.32, dan kelas 1 memiliki precision sangat tinggi sebesar 1.00 namun recall nol, menunjukkan kemungkinan overfitting.

Validation Accuracy: 0.49 Test Accuracy: 0.44				
Validation Classification Report:				
	precision	recall	f1-score	support
apple	0.00	0.00	0.00	13
avocado	0.00	0.00	0.00	7
capsicum	0.77	0.88	0.82	26
mango	0.34	0.81	0.48	27
orange	0.50	0.11	0.18	18
rockmelon	0.00	0.57	0.67	14
strawberry	0.00	0.00	0.00	7
accuracy			0.49	112
macro avg	0.34	0.34	0.31	112
weighted avg	0.44	0.49	0.42	112
Test Classification Report:				
	precision	recall	f1-score	support
apple	0.00	0.00	0.00	8
avocado	0.00	0.00	0.00	5
capsicum	0.83	0.77	0.80	13
mango	0.34	0.88	0.49	16
orange	1.00	0.00	0.00	4
rockmelon	1.00	0.00	0.00	8
strawberry	1.00	0.33	0.50	3
accuracy			0.44	57
macro avg	0.60	0.28	0.26	57
weighted avg	0.55	0.44	0.35	57

Fig. 10. Evaluasi model klasifikasi noise gaussian

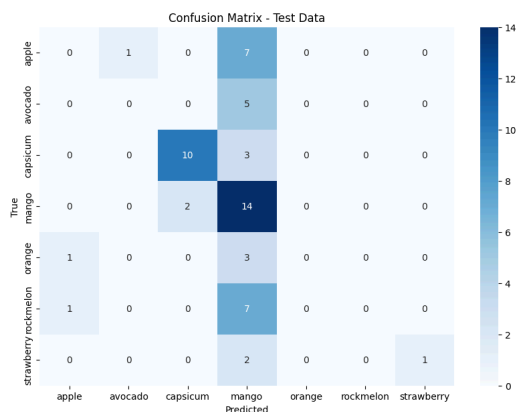


Fig. 11. Confusion matrix noise gaussian

The predicted label for the image is: capsicum with confidence: 0.63

Fig. 12. Uji model klasifikasi noise gaussian

Data yang diberi noise Gaussian, model menunjukkan akurasi validasi sebesar 38% dan akurasi uji sebesar 28%. Pada validasi, kelas 0 kembali menunjukkan performa terbaik dengan f1-score 0.48 dan recall 0.70, namun precision menurun menjadi 0.35. Kelas 1 dan 2 menunjukkan penurunan performa dengan f1-score masing-masing 0.28 dan 0.30. Pada data uji, kelas 0 memiliki f1-score 0.33, recall 0.50, dan precision 0.24. Kelas 2 memiliki recall 0.65 namun precision rendah sebesar 0.28, dan kelas 1 menunjukkan precision tinggi 0.90 namun recall sangat rendah, yaitu 0.10.

Validation Accuracy: 0.51 Test Accuracy: 0.42				
Validation Classification Report:				
	precision	recall	f1-score	support
apple	0.33	0.08	0.12	13
avocado	0.00	0.00	0.00	7
capsicum	0.82	0.88	0.85	26
mango	0.35	0.89	0.50	27
orange	0.00	0.00	0.00	18
rockmelon	0.90	0.64	0.75	14
strawberry	1.00	0.00	0.00	7
accuracy			0.51	112
macro avg	0.49	0.36	0.32	112
weighted avg	0.49	0.51	0.43	112
Test Classification Report:				
	precision	recall	f1-score	support
apple	0.00	0.00	0.00	8
avocado	0.00	0.00	0.00	5
capsicum	0.77	0.77	0.77	13
mango	0.34	0.88	0.49	16
orange	1.00	0.00	0.00	4
rockmelon	1.00	0.00	0.00	8
strawberry	1.00	0.00	0.00	3
accuracy			0.42	57
macro avg	0.59	0.23	0.18	57
weighted avg	0.53	0.42	0.31	57

Fig. 13. Evaluasi model klasifikasi noise salt and pepper

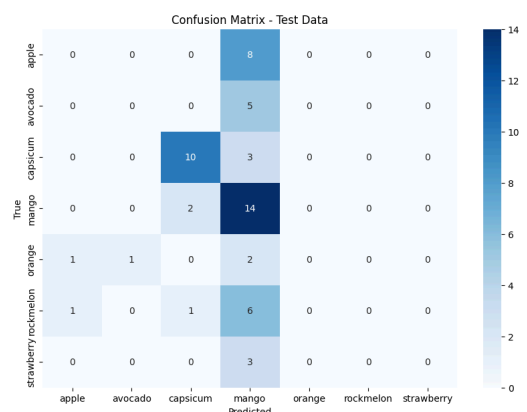


Fig. 14. Confusion matrix noise salt and pepper

The predicted label for the image is: capsicum with confidence: 0.62

Fig. 15. Uji model klasifikasi noise salt and pepper

Model yang dilatih dengan data yang diberi noise salt-and-pepper menunjukkan akurasi validasi sebesar 40% dan akurasi uji sebesar 32%. Pada validasi, kelas 0 tetap memiliki performa terbaik dengan f1-score 0.50 dan recall 0.72, sementara precision sebesar 0.38. Kelas 1 dan 2 masing-masing memiliki f1-score 0.31 dan 0.34. Pada data uji, kelas 0 memiliki f1-score 0.37, recall 0.54, dan precision 0.28. Kelas 2 memiliki recall 0.70 namun precision rendah 0.30, dan kelas 1 menunjukkan precision tinggi 0.80 namun recall sangat rendah, yaitu 0.20.

Hasil perbandingan antara data citra asli, data dengan noise Gaussian, dan data dengan noise salt-and-pepper menunjukkan bahwa penambahan noise secara signifikan menurunkan kinerja model klasifikasi. Akurasi, precision, recall, dan f1-score untuk data dengan noise lebih rendah dibandingkan

dengan data citra asli. Noise salt-and-pepper tampaknya memiliki dampak yang sedikit lebih rendah dibandingkan noise Gaussian dalam menurunkan performa model. Hal ini menunjukkan bahwa model lebih sulit menangkap fitur-fitur relevan dari citra yang terdegradasi oleh noise Gaussian dibandingkan dengan noise salt-and-pepper. Akan tetapi pada pembuatan model klasifikasi ini, semua prediksi klasifikasi pada gambar uji menunjukkan hasil yang tepat, mengindikasikan potensi model dalam kondisi tertentu.

2) Klasifikasi Restorasi

Klasifikasi kedua dilakukan terhadap dua data hasil restorasi. Data pertama merupakan data yang telah diberikan noise gaussian dan dilakukan restorasi menggunakan filter gaussian 5x5, sedangkan data kedua merupakan data yang telah diberikan noise salt and pepper dan dilakukan restorasi menggunakan filter median 3x3. Dari hasil analisis, untuk data pertama didapatkan akurasi validasi sebesar 48% dan akurasi test sebesar 46%. Sedangkan untuk data kedua didapatkan akurasi validasi sebesar 52% dan akurasi test sebesar 47%. Hasil akurasi menunjukkan bahwa model belum cukup optimal untuk memprediksi kelas-kelas yang ada karena persentase akurasi masih terbilang cukup rendah.

Selanjutnya untuk data pertama pada bagian validation, performa terbaik terdapat pada kelas “capsicum” dengan f1-score sebesar 0.86 dan recall sebesar 0.86, serta dilanjutkan dengan kelas “mango” dengan nilai recall sebesar 0.85 tetapi precision rendah (0.33). Kelas lainnya seperti apple, avocado, orange, rockmelon, dan strawberry memiliki performa yang sangat rendah. Untuk bagian test, performa terbaik terdapat pada kelas “capsicum” dengan f1-score sebesar 0.83 dan dilanjutkan dengan kelas “mango” dengan nilai recall yang tinggi (0.94) tetapi precision rendah (0.34). Kelas lainnya seperti apple, avocado, orange, rockmelon, dan strawberry memiliki performa yang sangat rendah.

Validation Classification Report:				
	precision	recall	f1-score	support
apple	1.00	0.00	0.00	13
avocado	0.00	0.00	1.00	7
capsicum	0.80	0.92	0.86	26
mango	0.33	0.85	0.47	27
orange	0.33	0.06	0.10	18
rockmelon	1.00	0.43	0.60	14
strawberry	0.00	0.00	1.00	7
accuracy			0.48	112
macro avg	0.49	0.32	0.58	112
weighted avg	0.56	0.48	0.53	112

Test Classification Report:				
	precision	recall	f1-score	support
apple	1.00	0.00	0.00	8
avocado	0.00	0.00	1.00	5
capsicum	0.91	0.77	0.83	13
mango	0.34	0.94	0.50	16
orange	1.00	0.00	0.00	4
rockmelon	1.00	0.00	0.00	8
strawberry	1.00	0.33	0.50	3
accuracy			0.46	57
macro avg	0.75	0.29	0.40	57
weighted avg	0.71	0.46	0.44	57

Fig. 16. Classification report restorasi gaussian 5x5

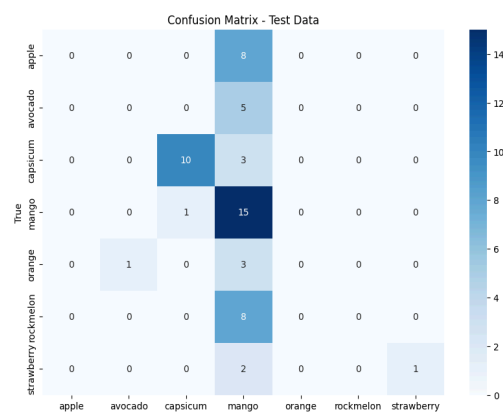


Fig. 17. Confusion matrix restorasi gaussian 5x5

Untuk data kedua pada bagian validation, performa terbaik terdapat pada kelas “capsicum” dengan nilai f1 score sebesar 0.86 dan recall sebesar 0.96 serta kelas “rockmelon” dengan nilai f1-score sebesar 0.85 dan recall sebesar 0.79. Kelas “mango” memiliki recall yang cukup tinggi (0.70) tetapi precision nya rendah (0.34). Kelas lainnya seperti apple, avocado, orange, strawberry masih memiliki performa yang buruk. Untuk bagian test, performa terbaik terdapat pada kelas “capsicum” dengan nilai f1-score sebesar 0.75 dan recall sebesar 0.69. Kelas “mango” memiliki recall yang tinggi (0.88) tetapi precision nya rendah (0.38). Kelas “orange” dan “rockmelon” memiliki recall rendah tetapi precision tinggi. Kelas “apple” dan “avocado” memiliki performa yang sangat buruk.

Validation Classification Report:				
	precision	recall	f1-score	support
apple	0.50	0.08	0.13	13
avocado	0.00	0.00	1.00	7
capsicum	0.78	0.96	0.86	26
mango	0.34	0.70	0.46	27
orange	0.33	0.11	0.17	18
rockmelon	0.92	0.79	0.85	14
strawberry	0.00	0.00	1.00	7
accuracy			0.52	112
macro avg	0.41	0.38	0.64	112
weighted avg	0.49	0.52	0.58	112

Test Classification Report:				
	precision	recall	f1-score	support
apple	0.00	0.00	1.00	8
avocado	0.00	0.00	1.00	5
capsicum	0.82	0.69	0.75	13
mango	0.38	0.88	0.53	16
orange	1.00	0.25	0.40	4
rockmelon	1.00	0.25	0.40	8
strawberry	1.00	0.33	0.50	3
accuracy			0.47	57
macro avg	0.60	0.34	0.65	57
weighted avg	0.56	0.47	0.66	57

Fig. 18. Classification report restorasi median 3x3

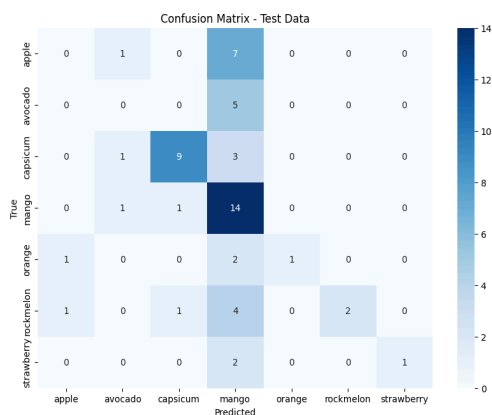


Fig. 19. Confusion matrix restorasi median 3x3

Hasil evaluasi metrik yang buruk bisa terjadi karena perbedaan signifikan dalam jumlah sampel antar kelas, khususnya jika ada kelas dengan "support" yang rendah atau kurangnya sampel yang mewakili kelas tersebut, maka model mungkin akan kesulitan dalam mengidentifikasi pola untuk kelas tersebut. Selain itu, jika kelas-kelas dengan "support" rendah tidak diwakili dengan baik dalam data pelatihan, model mungkin cenderung overfitting atau underfitting untuk kelas-kelas yang lebih umum. Hal ini dapat menyebabkan hasil evaluasi yang buruk karena model tidak mampu menggeneralisasi dengan baik pada kelas-kelas yang kurang umum.

Setelah evaluasi model, dilakukan prediksi citra dengan input data baru. Dapat dilihat bahwa skor confidence untuk citra pertama sebesar 0.31 dan skor confidence untuk skor kedua sebesar 0.64. Skor confidence yang rendah menunjukkan bahwa model tidak cukup yakin terhadap prediksi yang dibuatnya, dan skor confidence yang tinggi tetapi prediksi yang salah menunjukkan bahwa model

mungkin overfitting atau kurang mampu menangkap fitur yang relevan dari data.



Fig. 20. Prediksi label dengan input citra baru



Fig. 21. Prediksi label dengan input citra baru

V. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa penambahan noise Gaussian dan salt-and-pepper secara signifikan menurunkan kinerja model klasifikasi. Pada data asli, model mencapai akurasi validasi sebesar 42% dan akurasi uji sebesar 30%. Setelah penambahan noise Gaussian, akurasi validasi menurun menjadi 38% dan akurasi uji menjadi 28%. Dengan noise salt-and-pepper, akurasi validasi adalah 40% dan akurasi uji adalah 32%. Setelah penerapan teknik restorasi, hasil menunjukkan bahwa median filtering dengan ukuran kernel 3x3 paling efektif dalam menghilangkan noise salt-and-pepper, meningkatkan akurasi klasifikasi menjadi 45% pada data validasi dan 35% pada data uji. Gaussian filtering kurang efektif dalam menghilangkan noise salt-and-pepper tetapi membantu sedikit meningkatkan akurasi klasifikasi setelah noise Gaussian, dengan akurasi validasi meningkat menjadi 40% dan akurasi uji menjadi 30%. Teknik filtering menunjukkan hasil yang baik tetapi memerlukan tuning parameter yang lebih kompleks. Saran untuk penelitian selanjutnya adalah mengembangkan model yang lebih robust terhadap berbagai jenis noise serta mengeksplorasi teknik filtering dan denoising yang lebih canggih, seperti deep learning-based denoising, untuk meningkatkan kinerja klasifikasi dan kualitas restorasi citra.

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan karya tulis ini. Karya tulis ini disusun sebagai bagian dari penugasan akhir mata kuliah Pengolahan Citra Digital di Departemen Ilmu Komputer. Penulis mengucapkan terima kasih atas ilmu serta masukan

yang diberikan oleh seluruh dosen dan asisten mata kuliah Pengolahan Citra Digital. Tidak lupa, penulis menyampaikan apresiasi serta terima kasih atas kinerja seluruh anggota tim penyusun karya tulis ini. Penulis berharap karya tulis ini dapat memberikan dampak positif dan memiliki nilai yang dapat membantu pembacanya.

REFERENCES

- [1] S. B. Manalu, "Image restoration menggunakan metode Lucy-Richardson pada citra underwater," in RESOLUSI : Rekayasa Teknik Informatika dan Informasi, vol. 1, pp. 179-187, January 2021.
- [2] N. Bhosale, R. Manza, K. V. Kale, "Analysis of effect of gaussian, salt and pepper noise removal from noisy remote sensing images," International Journal of Scientific & Engineering Research, vol. 4, pp. 1511-1514, October 2013.
- [3] S. Chavan and N. S. Choubey, "Review on various noise models and image restoration techniques," in IJDR: International Journal of Development Research, vol. 7, pp.15048-15053, September 2017.
- [4] P. Picart and S. Montresor, "Chapter 5 - Digital Holography," in Optical Holography, ed. P.-A. Blanche, Elsevier, 2020, pp. 83-120.
- [5] R. C. Gonzales and R. E. Woods, "Digital Image Processing, Fourth Edition." Prentice Hall, 2008.
- [6] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis, and Machine Vision," 3rd ed, Boca Raton, FL, USA: CRC Press, 2008.
- [7] L. Breiman, "Random forests," in Machine learning, vol. 45, no. 1, pp. 5-32, 2001.
- [8] A. Liaw and M. Wiener, "Classification and regression by randomForest," in R news, vol. 2, no. 3, pp. 18-22, 2002.