

RANCANG BANGUN SISTEM IDENTIFIKASI NOMINAL UANG KERTAS RUPIAH UNTUK LANSIA BERBASIS CNN DENGAN ARSITEKTUR VGG16 DAN FITUR TEXT TO SPEECH

Nur Hasanah Perangin Angin¹⁾, Arif Fadllullah^{2*)}

^{1,2} Program Studi Teknik Komputer, Universitas Borneo Tarakan

Email: arif.fadl@borneo.ac.id²⁾

Nomor Telp : +62 822 5366 1738

Asal Negara: Indonesia

ABSTRAK

Lansia merupakan individu yang telah berusia 60 tahun ke atas, yang umumnya mengalami penurunan fungsi tubuh, termasuk berkurangnya kemampuan penglihatan, termasuk dalam mengenal nominal uang kertas. Ketidakmampuan dalam mengenali nominal uang kertas dapat menimbulkan masalah, baik dalam kegiatan transaksi sehari-hari seperti berbelanja. Penelitian ini mengembangkan sistem identifikasi nilai nominal uang kertas rupiah berbasis CNN (*Convolutional Neural Network*) dengan arsitektur VGG16 dan fitur *text to speech* untuk membantu lansia. Dataset terdiri dari 4.200 citra tujuh jenis uang kertas yang dibagi dalam dua skema anotasi masing-masing 2.100 citra: skema pertama berdasarkan nominal dan skema kedua berdasarkan keseluruhan area uang. Dataset dibagi menjadi 70% train, 20% validasi, dan 10% uji. Skema pertama mencapai akurasi pelatihan 99,47% dan validasi 100%, sedangkan skema kedua memperoleh 74,37% dan 98,45%. Pada data uji, skema pertama menghasilkan *precision*, *recall*, dan *F1-score* 100%, sementara skema kedua mencapai rata-rata 94% dengan sedikit kesalahan pada Rp10.000 dan Rp20.000. Kemudian, pengujian secara *real-time* yang dilakukan terhadap sepuluh orang lansia menunjukkan bahwa skema pertama lebih akurat meskipun perlu perbaikan dikarenakan rendahnya performa deteksi. Selain itu, skor SUS (*System Usability Scale*) sebesar 70,25 menunjukkan kategori “baik”. Keterbatasan utama terletak pada ketergantungan pencahayaan dan posisi nominal, sehingga pengembangan selanjutnya disarankan menggunakan deteksi objek YOLO atau SSD.

Kata kunci: Lansia, CNN, VGG16, *Text to Speech*, Deteksi Nominal Uang Kertas

ABSTRACT

Elderly individuals, defined as those aged 60 years and above, generally experience a decline in physical functions, including reduced visual ability, which affects their capacity to recognize the denomination of banknotes. This inability to identify banknote values may cause difficulties in daily transactions such as shopping. This study develops a rupiah banknote denomination identification system based on a Convolutional Neural Network (CNN) using the VGG16 architecture and a text-to-speech feature to assist the elderly. The dataset consists of 4,200 images of seven types of banknotes, divided into two annotation schemes of 2,100 images each: the first scheme based on denomination and the second based on the entire banknote area. The data were divided into 70% for training, 20% for validation, and 10% for testing. The first scheme achieved a training accuracy of 99.47% and a validation accuracy of 100%, while the second scheme obtained 74.37% and 98.45%, respectively. On the test data, the first scheme achieved precision, recall, and F1-score values of 100%, whereas the second scheme reached an average of 94%, with minor misclassifications for Rp10,000 and Rp20,000 denominations. Furthermore, real-time testing involving ten elderly participants showed that the first scheme was more accurate, although performance improvements are still required due to lower detection capability. In addition, the System Usability Scale (SUS) score of 70.25 indicates a “good” category. The main limitation lies in the system’s dependence on lighting conditions and the position of banknotes; therefore, future work is recommended to implement object detection methods such as YOLO or SSD.

Keywords: Elderly, CNN, VGG16, *Text to Speech*, Banknote Denomination Detection

1. PENDAHULUAN

Uang kertas merupakan alat pembayaran yang sangat penting dalam aktivitas ekonomi sehari-hari karena memiliki nilai tukar yang telah ditentukan dan tercetak pada setiap lembarannya. Bank Indonesia pada tahun 2022 telah

meluncurkan tujuh denominasi uang rupiah tahun emisi 2022 dengan desain baru yang lebih aman dan modern (Gideon, 2022). Setiap pecahan uang rupiah memiliki ciri khas yang membedakannya, seperti ukuran, warna, desain visual, benang pengaman, gambar air (*watermark*), dan lambang

khusus (Rewina et al., 2024). Perbedaan visual ini menjadi kunci dalam proses identifikasi nilai uang oleh masyarakat dalam kehidupan sehari-hari.

Bagi sebagian kelompok masyarakat, khususnya lanjut usia (lansia), perbedaan visual tersebut justru menjadi tantangan dalam mengenali nilai nominal uang kertas. Lansia umumnya mengalami penurunan fungsi penglihatan seperti degenerasi makula, katarak, dan penurunan sensitivitas kontras yang menyebabkan berkurangnya kemampuan membedakan objek visual (Akbar et al., 2024). Kesulitan dalam mengenali nominal uang dapat menimbulkan kesalahan dalam transaksi, menurunkan tingkat kemandirian, serta meningkatkan risiko penipuan dalam aktivitas ekonomi sehari-hari. Kondisi ini menunjukkan perlunya pengembangan teknologi bantu yang dapat mendukung lansia dalam mengidentifikasi uang secara cepat dan akurat.

Berbagai penelitian sebelumnya telah mengembangkan sistem pengenalan uang berbasis pengolahan citra digital dan kecerdasan buatan, terutama menggunakan metode *Convolutional Neural Networks* (CNN) yang unggul dalam mengekstraksi fitur visual citra. Misalnya, penelitian oleh (Nur Hidayat et al., 2023) menggunakan struktur VGG-16 dengan alokasi data 80% untuk pelatihan dan 20% untuk pengujian, sistem ini memperoleh kinerja yang cukup memuaskan dalam mengidentifikasi nilai mata uang rupiah secara *real-time*, walaupun masih terdapat 5 kesalahan identifikasi akibat tingkat akurasi yang kurang akurat. Disisi lain, penelitian dari (Arrizqi et al., 2021) yang menggabungkan *Google Text to Speech* ke dalam aplikasi pendeteksi uang yang berbasis *Android* dengan penekanan pada uang kertas yang dikeluarkan tahun 2016, menerapkan sistem tingkat *threshold* sebesar 83.62% untuk mengoperasikan fitur suara panduan, yang berhasil memperoleh akurasi informasi hingga 98.6% dalam kondisi *frame* yang baik. Penelitian-penelitian tersebut membuktikan potensi penggabungan pengenalan citra dan *text-to-speech* dalam membantu pengguna mengenali objek visual tanpa bergantung pada penglihatan.

Namun, sebagian besar sistem yang telah dikembangkan tersebut masih bersifat umum dan belum dirancang secara khusus untuk memenuhi kebutuhan pengguna lansia. Aspek antarmuka yang ramah lansia, kecepatan deteksi *real-time*, dan umpan balik suara langsung sering kali belum menjadi fokus utama. Celah inilah yang menjadi dasar pentingnya penelitian ini, karena kemampuan lansia dalam mengenali uang merupakan hal krusial untuk mempertahankan kemandirian dan keamanan mereka dalam bertransaksi.

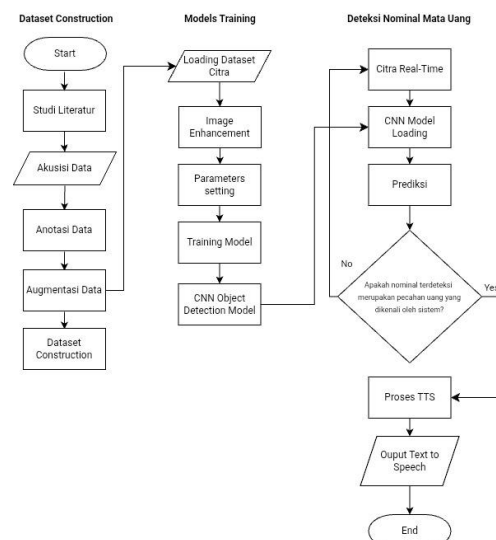
Oleh karena itu, penelitian ini bertujuan untuk mengembangkan sistem identifikasi nominal

uang kertas rupiah berbasis *convolutional neural networks* dengan arsitektur VGG16 dan fitur *text-to-speech* yang bekerja secara otomatis dalam membantu lansia mengenali uang secara cepat, akurat, dan mandiri. Sistem ini diharapkan dapat menjadi solusi inovatif sekaligus berkontribusi dalam pengembangan teknologi pendukung bagi kelompok rentan.

2. METODE PENELITIAN

Penelitian ini menggunakan jenis penelitian kuantitatif eksperimental, yang bertujuan untuk menguji dan mengukur efektivitas sistem identifikasi nominal uang kertas rupiah untuk lansia menggunakan metode *Convolutional Neural Network* dan *Text-to-Speech* dengan membandingkan dua skema anotasi data, yaitu 1) anotasi pada bagian nominal uang dan 2) anotasi pada seluruh area objek uang kertas.

2.1. Perancangan sistem



Gambar 1. Diagram alir perancangan sistem

Bagan alur pada Gambar 1 menunjukkan bagaimana perancangan sistem usulan yang dimulai dari tahapan *dataset construction*, *model training*, dan deteksi nominal mata uang.

2.1.1. Dataset Construction

Tahapan ini mencakup studi literatur, akuisisi data berupa pengambilan citra uang kertas, anotasi data untuk memberi label pada setiap nominal, serta augmentasi data guna meningkatkan variasi citra dan mengurangi *overfitting*. Data citra yang dikonstruksikan menjadi dataset adalah uang kertas Rupiah tahun emisi 2022 dengan nominal Rp1.000 hingga Rp100.000. Data dikumpulkan secara primer dengan memotret setiap pecahan menggunakan kamera *handphone*, masing-masing sebanyak 600 foto per kelas, sehingga diperoleh total 4.200 gambar dari dua posisi pengambilan berbeda. Sedangkan proses anotasi dan augmentasi

menggunakan platform *Roboflow*: <https://roboflow.com/>. Dua skema anotasi digunakan sebagai bahan perbandingan dalam penelitian ini. Skema pertama melakukan anotasi hanya pada bagian nominal uang, yaitu area yang memuat angka pecahan uang kertas, sehingga model difokuskan untuk mengenali informasi numerik secara spesifik. Skema kedua melakukan anotasi pada seluruh area objek uang kertas, sehingga model mempelajari keseluruhan visual uang kertas termasuk warna, motif, dan elemen desain lainnya. Perbandingan kedua skema ini bertujuan untuk mengetahui pendekatan anotasi mana yang menghasilkan performa klasifikasi lebih baik dalam mengenali nominal uang kertas rupiah.

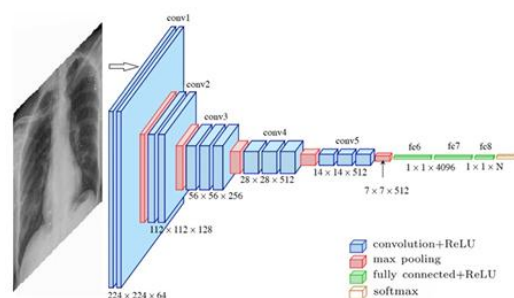
2.1.2. Model Training

Tahapan berikutnya adalah model training, di mana dataset citra dimuat dan dilakukan proses *image enhancement* untuk meningkatkan kualitas visual. Dataset tersebut dibagi ke dalam tiga bagian, yaitu data latih, validasi, dan uji dengan perbandingan 70%:20%:10%. Setelah itu dilakukan pengaturan parameter dan pelatihan model CNN dengan arsitektur VGG16 hingga diperoleh model deteksi objek uang kertas. Parameter tersebut diantaranya jumlah *epoch*, ukuran *batch*, *learning rate*, dan jenis augmentasi virtual generator yang digunakan. Proses pelatihan model dilakukan menggunakan *hardware* dan *software* dengan spesifikasi sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Alat dan Bahan

Nama Alat dan Bahan	Unit	Kegunaan
Laptop Intel(R) Core(TM) i3 N305 RAM 8GB	1	Digunakan untuk pelatihan model, membuat sistem, pengujian model secara <i>real time</i>
Software Roboflow	1	Digunakan untuk memberi label pada objek, mengubah ukuran gambar, dan augmentasi data
Google Collab	1	Digunakan untuk training data dan evaluasi data
Uang Kertas Rupiah Tahun Emisi 2022	7	Objek yang digunakan untuk <i>training</i> dan pengujian

Spesifikasi tersebut digunakan selama keseluruhan proses pelatihan model CNN berbasis VGG16, mulai dari pemuatan dataset, pengaturan parameter, hingga penyimpanan model hasil pelatihan.



Gambar 2. Tahapan CNN berbasis arsitektur VGG16

Gambar 2 menunjukkan arsitektur VGG16 yang digunakan untuk mengekstraksi dan mengklasifikasikan fitur citra secara otomatis dan mendalam. Proses dimulai dari lapisan input yang menerima citra berukuran 224×224×3 piksel. Citra ini kemudian melewati serangkaian lapisan konvolusi (*convolutional layers*) yang berfungsi mengekstraksi fitur-fitur penting seperti tepi, sudut, tekstur, dan bentuk. Operasi konvolusi pada setiap lapisan dilakukan menggunakan filter berukuran 3×3 dengan *stride* 1 dan *padding* “same”, yang secara matematis dinyatakan dalam Persamaan (1).

$$y_{i,j}^{(k)} = \sum_m \sum_n x_{i+m,j+n} \cdot w_{m,n}^{(k)} + b^{(k)} \quad (1)$$

di mana x merupakan nilai piksel pada citra input, w adalah bobot kernel/filter ke- k , b adalah bias, dan $y_{i,j}^{(k)}$ adalah hasil konvolusi pada posisi (i, j) . Setiap hasil konvolusi kemudian dilewatkan melalui fungsi aktivasi ReLU (*Rectified Linear Unit*) untuk menambahkan sifat non-linear agar jaringan dapat mempelajari pola kompleks, dengan Persamaan (2).

$$f(x) = \max(0, x) \quad (2)$$

Setelah beberapa lapisan konvolusi dan aktivasi, dilakukan proses *max pooling* berukuran 2×2 yang bertujuan mengurangi dimensi spasial fitur sambil mempertahankan informasi paling penting. Operasi *pooling* ini dapat dinyatakan dengan Persamaan (3).

$$y_{i,j} = \max_{(m,n) \in R_{i,j}} x_{m,n} \quad (3)$$

Setiap blok konvolusi dan *pooling* menghasilkan representasi fitur dengan ukuran yang semakin kecil tetapi jumlah *channel* yang semakin banyak (64 → 128 → 256 → 512). Setelah tahap ekstraksi fitur selesai, hasilnya diratakan (*flattened*) dan diteruskan ke tiga lapisan fully connected (FC). Dua lapisan pertama masing-masing memiliki 4096 neuron dengan aktivasi

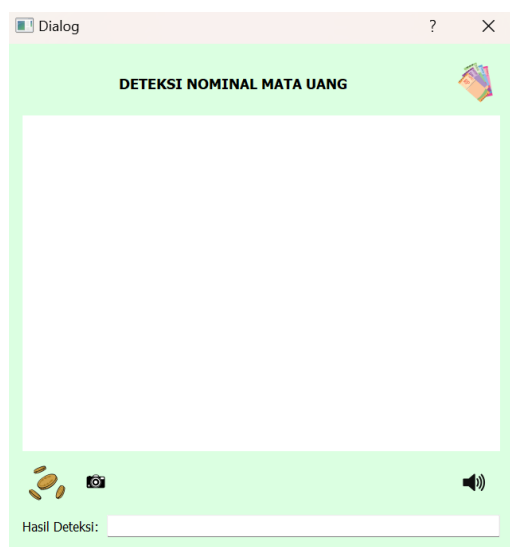
ReLU, sedangkan lapisan terakhir menggunakan fungsi Softmax untuk menghasilkan probabilitas tiap kelas, dengan Persamaan (4).

$$P(y = c | x) = \frac{e^{z_c}}{\sum_{k=1}^K e^{z_k}} \quad (4)$$

Dimana P adalah nilai probabilitas bahwa input x vektor input (fitur hasil ekstraksi dari lapisan sebelumnya) termasuk ke dalam kelas c yang dihitung dengan membandingkan nilai eksponensial dari output neuron ke- c (z_c) terhadap jumlah eksponensial dari semua neuron (z_k) di seluruh total kelas K . Melalui tahapan berurutan mulai dari konvolusi, aktivasi, *pooling*, hingga klasifikasi dengan *fully connected layer*, jaringan VGG16 mampu mengubah citra mentah menjadi representasi fitur yang bermakna dan akhirnya menghasilkan prediksi kelas secara akurat.

2.1.3. Deteksi Nominal Uang Kertas

Model yang telah dilatih dengan CNN berbasis arsitektur VGG16 kemudian digunakan pada tahap deteksi nominal mata uang, yang beroperasi secara *real-time*. Sistem memuat model CNN dan melakukan prediksi terhadap citra uang kertas yang diinput melalui kamera. Jika nominal yang terdeteksi sesuai dengan salah satu pecahan uang yang dikenal sistem, maka hasil prediksi akan diproses oleh modul *text to speech* (TTS) untuk menghasilkan output suara berupa penyebutan nominal uang dalam format audio. Tahapan ini diakhiri dengan keluaran suara yang membantu pengguna lansia mengenali nilai uang secara mandiri. Gambar 3 menunjukkan bagaimana *interface desktop* dari sistem usulan dibuat.



Gambar 3. Desain *interface* sistem usulan

Implementasi fitur *text-to-speech* pada sistem memanfaatkan beberapa *library* utama yang

saling terintegrasi. *Library Google Text-to-Speech* (gTTS) digunakan sebagai inti konversi teks menjadi audio dengan cara mengirimkan teks ke layanan *Google Translate* untuk kemudian menghasilkan keluaran berupa berkas audio dalam format MP3. *Library pygame* berperan dalam memuat dan memutar berkas audio hasil konversi tersebut secara langsung pada perangkat keras audio. Modul *datetime* dari *library* standar *Python* digunakan untuk menghasilkan *timestamp* unik pada setiap penamaan berkas audio, sedangkan modul *os* menangani operasi sistem berkas seperti pembuatan folder dan penghapusan berkas sementara. Selain itu, modul *QtCore.QTimer* dari *framework PyQt5* dimanfaatkan untuk menjadwalkan penghapusan berkas audio secara otomatis setelah pemutaran selesai tanpa mengganggu antarmuka pengguna yang sedang berjalan. Deklarasi impor seluruh *library* yang digunakan ditunjukkan pada Gambar 4.

```
import os
import pygame
from gtts import gTTS
from datetime import datetime
from PyQt5 import QtCore
```

Gambar 4. *Library* Implementasi *text-to-speech*

Modul *os* diimpor untuk menangani operasi sistem berkas, sedangkan *pygame* diimpor untuk keperluan pemutaran audio. Pustaka gTTS diimpor dari paket gtts sebagai komponen utama konversi teks menjadi suara. Kelas *datetime* diimpor dari modul *datetime* untuk keperluan pembangkitan *timestamp*, dan modul *QtCore* diimpor dari *framework PyQt5* untuk mengakses *QTimer* dalam penjadwalan penghapusan berkas audio.

Dalam proses deteksi, gambar uang yang diambil dari area tertentu sesuai dengan *bounding box*, kemudian melalui langkah pra-pemrosesan yang meliputi perubahan ukuran menjadi 224×224 piksel dan penyesuaian nilai piksel sebelum dimasukkan ke dalam model CNN. Jumlah uang dengan nilai kemungkinan prediksi tertinggi kemudian diubah menjadi suara menggunakan gTTS dalam format MP3 dan diputar secara otomatis dalam waktu nyata. Dengan cara ini, pengguna yang lebih tua dapat mendapatkan informasi tentang jumlah uang tanpa harus melihat tampilan visual dari sistem. *file* audio yang dibuat disimpan sementara di folder *output_suara* dengan nama yang telah ditentukan berdasarkan waktu untuk menghindari bentrokan antara berkas. Setelah pemutaran selesai, berkas ini akan dihapus secara otomatis untuk menjaga efisiensi penyimpanan. Implementasi fungsi *text-to-speech* ditunjukkan pada gambar 5.

```

def speak(text):
    timestamp = datetime.now().strftime("%Y%m%d_%H%M%S")
    filename = os.path.join(OUTPUT_FOLDER, f"{text}_{timestamp}.mp3")
    tts = gTTS(text=text, lang="id")
    tts.save(filename)
    pygame.mixer.music.load(filename)
    pygame.mixer.music.play()

def remove_audio():
    pygame.mixer.music.stop()
    try:
        os.remove(filename)
    except PermissionError:
        pass

QtCore.QTimer.singleShot(3000, remove_audio)

```

Gambar 5. Code Implementasi fitur *text-to-speech*

Fungsi *speak()* menerima satu parameter yang berupa string teks yang merupakan hasil prediksi jumlah uang. Nama file audio dibuat secara khusus dengan menggabungkan teks dan timestamp dari *datetime.now()* menggunakan format *%Y%m%d_%H%M%S*. Ini dilakukan untuk mencegah terjadinya konflik nama file ketika fungsi ini dipanggil berulang kali. Teks itu diubah menjadi audio menggunakan *gTTS* dengan pengaturan bahasa Indonesia (*lang="id"*) supaya cara bacanya sesuai dengan konteks pemakaian. Kemudian, audio tersebut disimpan dalam folder *output_suara* dengan format MP3. Audio yang sudah disimpan kemudian dimuat dan diputar menggunakan modul *pygame.mixer.music*. Setelah pemutaran selesai, fungsi *remove_audio()* dipanggil secara terjadwal menggunakan *QtCore.QTimer.singleShot()* dengan jeda 3000 milidetik. Ini dilakukan untuk menghentikan pemutaran dan menghapus berkas audio sementara dari penyimpanan. Mekanisme *PermissionError* juga digunakan untuk mengantisipasi situasi di mana berkas masih sedang diakses oleh sistem saat proses penghapusan dilakukan.

```

img_resized = cv2.resize(cropped, (224, 224))
img_array = np.expand_dims(img_resized, axis=0) / 255.0
prediction = model.predict(img_array)
label_idx = np.argmax(prediction)
confidence = np.max(prediction)

detected_text = LABELS[label_idx] if label_idx < len(LABELS) else "Tidak diketahui"
self.last_detected_text = detected_text
self.dtk_hasil.setText(f"{detected_text} ({confidence * 100:.2f}%)")

speak(detected_text)

```

Gambar 6. Implementasi prediksi nominal uang

Citra hasil pemotongan berdasarkan area *bounding box* selanjutnya diubah ukurannya menjadi 224×224 piksel menggunakan fungsi *cv2.resize()* agar sesuai dengan ukuran masukan pada model VGG16. Setelah itu, citra dikonversi ke dalam bentuk *array NumPy* dengan menambahkan dimensi *batch* melalui fungsi *np.expand_dims()*. Proses normalisasi juga dilakukan dengan membagi nilai piksel sebesar 255.0 sehingga rentang nilainya berada antara 0 sampai 1. Data citra tersebut kemudian diproses oleh model untuk memperoleh nilai probabilitas dari setiap kelas nominal uang. Penentuan kelas prediksi dilakukan menggunakan fungsi

np.argmax() untuk memilih probabilitas tertinggi, sedangkan tingkat keyakinan prediksi dihitung menggunakan fungsi *np.max()*. Hasil klasifikasi kemudian dicocokkan dengan label nominal uang pada array "LABELS" disertai pemeriksaan indeks guna mencegah terjadinya kesalahan akses data. Nominal uang yang berhasil dikenali beserta nilai *confidence* ditampilkan pada antarmuka melalui komponen "dtk_hasil", kemudian fungsi *speak()* dijalankan untuk menghasilkan keluaran suara sebagai informasi nominal uang kepada pengguna.

2.2. Analisis Data

Dalam menganalisis kinerja model CNN berbasis arsitektur VGG16 dalam mengidentifikasi nilai mata uang kertas rupiah, pengujian dilakukan menggunakan *confusion matrix* konfusi untuk memperoleh nilai seperti *accuracy*, *precision*, *recall*, dan *f1-score* yang dapat dilihat rumusnya pada Persamaan (5) sampai dengan (8).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1_score = 2 \times \frac{Precision + Recall}{Precision \times Recall} \quad (8)$$

Keterangan:

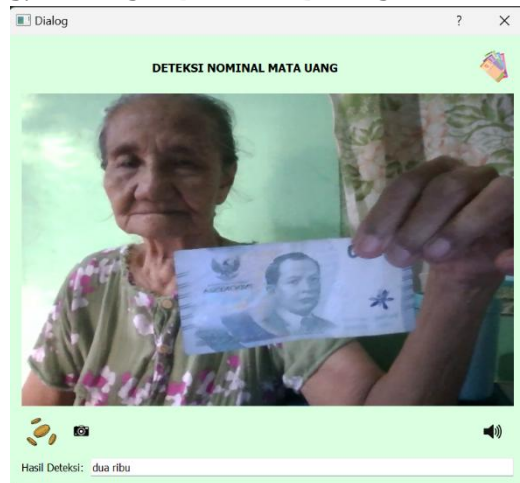
- True Positive* (TP): jumlah uang kertas yang terdeteksi dengan benar sesuai nominal aslinya, disimbolkan dengan (✓).
- True Negative* (TN): jumlah uang kertas yang benar dikenali bukan sebagai nominal tertentu.
- False Positive* (FP): jumlah uang kertas yang salah terdeteksi sebagai nominal lain padahal bukan, disimbolkan dengan (×).
- False Negative* (FN): jumlah uang kertas yang gagal terdeteksi sesuai nominal aslinya, disimbolkan dengan (–).

Selain itu, untuk mengukur tingkat kemudahan penggunaan sistem oleh lansia, dilakukan uji usability menggunakan System Usability Scale (SUS). Lansia diminta menggunakan sistem untuk mengenali pecahan uang kertas, kemudian mengisi kuesioner SUS yang terdiri dari 10 pernyataan, di mana nomor pernyataan ganjil berisi pernyataan positif, sedangkan nomor pernyataan genap memuat pernyataan negatif (Isnanto et al., 2023). Kuesioner ini menggunakan skala Likert 1-5 (sangat tidak setuju hingga sangat setuju) dan bertujuan mengevaluasi kemudahan, kenyamanan, serta kepercayaan pengguna terhadap sistem (Mu'min

Azis et al., 2024). Kuesioner SUS yang digunakan dalam penelitian ini telah melalui proses validasi dan penyesuaian berdasarkan acuan dari (Renaldo et al., 2025). Ketentuan yang perlu dipatuhi dalam proses perhitungan skor pada kuesioner:

- Pernyataan yang memiliki nomor tidak genap, skor dihitung dengan cara mengurangi nilai yang ditetapkan oleh pengguna sebanyak 1.
- Pernyataan bernomor genap, nilai dihitung dengan cara mengurangi hasil jawaban peserta dari angka 5.

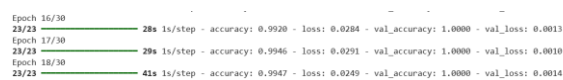
3. HASIL DAN PEMBAHASAN



Gambar 7. Hasil implementasi sistem usulan

Gambar 4 menunjukkan bagaimana keberhasilan sistem usulan yang mampu secara otomatis mendeteksi nilai mata uang sekaligus memutar suara nominal mata uang berdasarkan citra objek mata uang yang dipegang oleh lansia. Akan tetapi, untuk melihat sejauh mana validitas dan keakuratan sistem, berikut akan dijabarkan 1) hasil *training* model terhadap 70% data latih dan 20% data validasi; 2) evaluasi pengujian terhadap 10% data uji; 3) evaluasi kinerja sistem dan fitur *text to speech* secara *real-time* terhadap 10 lansia; 4) evaluasi *interface* sistem dengan SUS.

3.1. Hasil Training Model



Gambar 8. Hasil *training* skema 1



Gambar 9. Hasil *training* skema 2

Dalam penelitian ini, dilakukan pelatihan model VGG16 untuk mengidentifikasi nominal uang kertas menggunakan dua skema anotasi yang berbeda. Gambar 5 menunjukkan hasil *training* pada skema 1 yang menggunakan anotasi hanya

- Skor SUS diperoleh dengan mengakumulasi semua nilai dari setiap pernyataan, kemudian total tersebut dikalikan dengan 2,5 (Renaldo et al., 2025).

Untuk memperoleh total skor SUS dengan menggunakan Persamaan (9). Dimana

$$\bar{x} = \frac{\sum x}{n} \quad (9)$$

Keterangan:

$\bar{x}$: Skor rata-rata

$\sum x$: Jumlah skor SUS

n : Jumlah responden

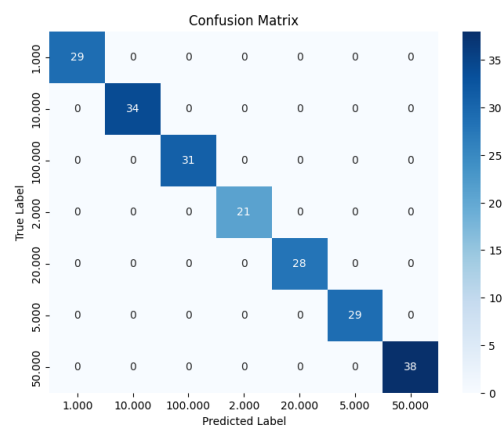
nominal mata uang. *Training* ini dirancang untuk 30 *epoch* namun dihentikan lebih awal pada *epoch* ke-18 menggunakan teknik *EarlyStopping* untuk mencegah *overfitting*. Pada *epoch* ini, model sudah mencapai akurasi pelatihan 99,47% dan akurasi validasi 100%, dengan *loss* pelatihan 0,0249 dan *validation loss* 0,0014, menunjukkan kemampuan model yang sangat baik dalam mengklasifikasikan nominal uang baik pada data latih maupun data validasi. Sementara itu, pada skema 2, model dilatih menggunakan anotasi seluruh objek uang hingga *epoch* ke-20 yang terhenti oleh *EarlyStopping*. Akurasi pelatihan tercatat 74,37% dan akurasi validasi 98,45%, dengan *loss* pelatihan 0,6312 dan *validation loss* 0,1749. Perbedaan akurasi antara data latih dan validasi menunjukkan kemungkinan *overfitting*, meski model tetap efektif mengenali nominal uang pada data baru. Tabel 1 menunjukkan perbandingan dua skema ini yang menyimpulkan bahwa skema pertama lebih stabil daripada skema yang kedua.

Tabel 1. Perbandingan hasil evaluasi *training* dua skema

Skema	Last Epoch	Acc	Loss	Val_Acc	Val_Loss
1	18	0,9947	0,0249	1,0000	0,0014
2	20	0,7437	0,6312	0,9845	0,1749

3.2. Evaluasi Sistem Terhadap Data Uji

3.2.1. Evaluasi Terhadap Data Uji Skema 1



Gambar 10. *Confusion matrix* terhadap data uji skema 1

Classification Report (Test):

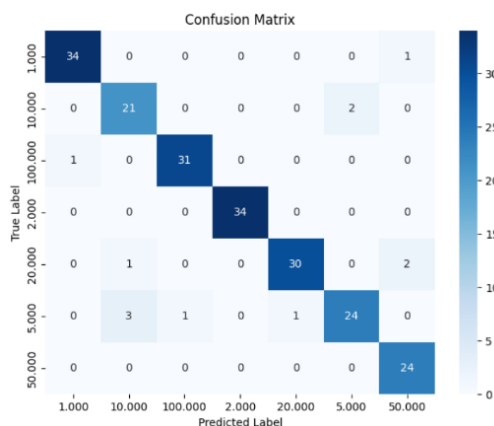
	precision	recall	f1-score	support
1.000	1.00	1.00	1.00	29
10.000	1.00	1.00	1.00	34
100.000	1.00	1.00	1.00	31
2.000	1.00	1.00	1.00	21
20.000	1.00	1.00	1.00	28
5.000	1.00	1.00	1.00	29
50.000	1.00	1.00	1.00	38
accuracy			1.00	210
macro avg	1.00	1.00	1.00	210
weighted avg	1.00	1.00	1.00	210

Gambar 11. Evaluasi model terhadap Data Uji Skema 1

Gambar 7 menampilkan *confusion matrix* hasil evaluasi model terhadap data uji skema 1. Dari matriks tersebut terlihat bahwa model berhasil mengidentifikasi semua kelas dengan sangat baik, ditandai oleh prediksi yang sepenuhnya berada pada diagonal utama. Distribusi prediksi yang benar per kelas adalah sebagai berikut: 1.000 (29 data), 2.000 (21 data), 5.000 (29 data), 10.000 (34 data), 20.000 (28 data), 50.000 (38 data), dan 100.000 (31 data), tanpa kesalahan identifikasi. Hasil ini menunjukkan bahwa model mencapai akurasi 100% dengan total 210 sampel data uji yang semuanya diklasifikasikan dengan benar. Berdasarkan Gambar 7 juga, seluruh kelas nominal uang kertas, mulai dari 1.000 hingga 100.000, berhasil diidentifikasi dengan sangat baik oleh sistem. Hal ini ditunjukkan oleh nilai TP yang tinggi (21–38) dan FP maupun FN bernilai 0, menandakan tidak ada kesalahan prediksi. Nilai TN juga tinggi, menunjukkan sampel dari kelas lain tidak salah diklasifikasikan.

Berikutnya berdasarkan Gambar 8, nilai rata-rata *precision*, *recall*, *f1-score*, dan *accuracy* seluruh kelas mencapai 1,00, yang menegaskan bahwa model bekerja konsisten dan optimal pada seluruh kelas.

3.2.2. Evaluasi Terhadap Data Uji Skema 2



Gambar 13. Confusion matrix data uji skema 2

Classification Report (Test):

	precision	recall	f1-score	support
1.000	0.97	0.97	0.97	35
10.000	0.84	0.91	0.88	23
100.000	0.97	0.97	0.97	32
2.000	1.00	1.00	1.00	34
20.000	0.97	0.91	0.94	33
5.000	0.92	0.83	0.87	29
50.000	0.89	1.00	0.94	24
accuracy			0.94	210
macro avg	0.94	0.94	0.94	210
weighted avg	0.94	0.94	0.94	210

Gambar 12. Evaluasi model terhadap Data Uji Skema 2

Gambar 9 menampilkan *confusion matrix* hasil evaluasi model terhadap data uji skema 2. Dari gambar terlihat model mampu mengidentifikasi sebagian besar citra uji dengan tepat, ditunjukkan oleh angka tinggi pada diagonal utama, misalnya nominal 1.000 dan 2.000 masing-masing 34 gambar, 100.000 sebanyak 31 gambar, dan 50.000 sebanyak 24 gambar. Namun, terdapat beberapa kesalahan, seperti nominal 10.000 dua gambar salah diprediksi sebagai 5.000, 20.000 satu gambar sebagai 10.000 dan dua gambar sebagai 50.000, serta 5.000 beberapa gambar salah diprediksi ke kelas lain, yang disebabkan oleh kemiripan visual antar nominal. Secara keseluruhan, model skema kedua menunjukkan identifikasi yang cukup baik, meski masih perlu peningkatan untuk membedakan nominal yang mirip.

Berikutnya berdasarkan Gambar 10, terlihat bahwa nilai *precision*, *recall*, dan *f1-score* pada setiap kelas berkisar antara 0,83 hingga 1,00, menandakan bahwa model memiliki performa yang cukup konsisten di seluruh kategori uang. Kinerja terbaik diperoleh pada nominal Rp 2.000, dengan nilai *precision*, *recall*, dan *f1-score* masing-masing mencapai 1. Sementara itu, performa terendah terdapat pada nominal Rp 5.000 dengan *f1-score* sebesar 0,87, yang menunjukkan bahwa masih terdapat beberapa kesalahan dalam mendeteksi kelas ini, kemungkinan disebabkan oleh kemiripan visual dengan nominal lain atau variasi kondisi citra seperti pencahayaan dan lipatan uang. Selain itu, nilai rata-rata *precision*, *recall*, *f1-score* dan *accuracy* seluruh kelas sebesar 0.94 menunjukkan bahwa model tidak hanya unggul pada kelas tertentu, tetapi juga mampu mempertahankan performa stabil di seluruh kategori uang.

Tabel 2 menunjukkan perbandingan evaluasi terhadap data uji untuk dua skema.

Tabel 2. Perbandingan hasil evaluasi data uji dua skema

Skema	Precision	Recall	F1-Score	Accuracy
1	1,00	1,00	1,00	1,00
2	0,94	0,94	0,94	0,94

3.3 Evaluasi Kinerja Sistem dan Fitur *Text to Speech* Terhadap Data *Real-Time*

Tabel 3. Hasil percobaan terhadap data *real-time* dua skema

Data	Skema	Nominal	Percobaan				
		Uang	1	2	3	4	5
1	Skema 1	Rp1.000	√	-	×	-	-
		Rp2.000	-	×	-	√	√
	Skema 2	Rp1.000	√	-	-	√	-
		Rp2.000	-	√	×	-	√
2	Skema 1	Rp10.000	√	-	-	×	-
		Rp5.000	-	×	√	-	√
	Skema 2	Rp5.000	√	-	×	-	-
		Rp10.000	-	√	-	×	√
3	Skema 1	Rp100.000	√	-	×	-	√
		Rp2.000	-	√	-	×	-
	Skema 2	Rp2.000	-	√	-	-	×
		Rp100.000	√	-	×	×	-
4	Skema 1	Rp2.000	√	-	-	√	-
		Rp1.000	-	√	×	-	√
	Skema 2	Rp1.000	√	-	-	×	-
		Rp2.000	-	√	√	-	√
5	Skema 1	Rp5.000	-	×	-	-	√
		Rp50.000	√	-	√	×	-
	Skema 2	Rp5.000	√	-	×	×	-
		Rp50.000	-	√	-	-	×
6	Skema 1	Rp50.000	√	-	-	-	×
		Rp5.000	-	×	√	√	-
	Skema 2	Rp50.000	√	×	-	√	-
		Rp5.000	-	-	×	-	√
7	Skema 1	Rp20.000	-	√	-	√	×
		Rp2.000	√	-	√	-	-
	Skema 2	Rp2.000	√	×	-	-	√
		Rp20.000	-	-	√	×	-
8	Skema 1	Rp10.000	-	√	√	-	×
		Rp100.000	√	-	-	×	-
	Skema 2	Rp10.000	√	-	×	-	-
		Rp100.000	-	×	-	×	×
9	Skema 1	Rp100.000	-	√	-	√	-
		Rp20.000	√	-	×	-	√
	Skema 2	Rp100.000	√	-	×	-	-
		Rp20.000	-	√	-	×	×
10	Skema 1	Rp5.000	√	-	√	-	√
	Skema 2	Rp1.000	-	×	-	×	-
	Skema 2	Rp1.000	-	×	√	-	-

Data	Skema	Nominal	Percobaan				
		Uang	1	2	3	4	5
		Rp5.000	√	-	-	×	√

Berdasarkan data dari Tabel 3 untuk skema 1, sistem berhasil mendeteksi dan mengaktifkan fitur *text-to-speech* nominal uang dengan benar sebanyak 33 kali (33%), mengalami kesalahan deteksi sebanyak 17 kali (17%), dan gagal mendeteksi nominal uang sebanyak 50 kali (50%). Sementara itu, pada Skema 2, jumlah deteksi benar tercatat sebanyak 27 kali (27%), kesalahan deteksi sebanyak 23 kali (23%), dan kegagalan deteksi sebesar 50 kali (50%). Jika dilihat dari *persentase* keberhasilan, Skema 1 menunjukkan performa yang lebih baik dengan tingkat keberhasilan sebesar 33% dan tingkat kesalahan 67%, dibandingkan Skema 2 yang memiliki keberhasilan 27% dan tingkat kesalahan 73%. Hasil ini menunjukkan bahwa konfigurasi anotasi pada Skema 1 yang hanya melingkupi area nominal angka lebih baik dibandingkan Skema 2 yang mencakup seluruh area uang kertas. Hal ini dapat dijelaskan bahwa dengan membatasi anotasi hanya pada area angka nominal, model lebih terfokus dalam mempelajari fitur diskriminatif yang paling relevan untuk membedakan antar kelas nominal, sehingga menghasilkan prediksi yang lebih tepat sasaran. Sebaliknya, pada Skema 2 model juga mempelajari elemen visual yang tidak relevan seperti latar belakang, tepi uang, dan motif dekoratif yang justru menambah noise dalam proses pembelajaran dan menurunkan kemampuan generalisasi model. Rendahnya nilai akurasi pada kedua skema, yaitu 33% untuk Skema 1 dan 27% untuk Skema 2, mengindikasikan adanya kesenjangan yang cukup signifikan antara kondisi data latih dan kondisi nyata saat pengujian *real-time*. Kesenjangan ini dapat disebabkan oleh beberapa faktor yang saling berkaitan. Pertama, variasi kondisi pencahayaan saat pengujian *real-time* tidak sepenuhnya terwakili dalam data latih, sehingga model kurang mampu beradaptasi terhadap perubahan intensitas cahaya yang memengaruhi kualitas citra secara langsung. Kedua, posisi dan orientasi uang kertas saat pengujian tidak selalu berada dalam kondisi ideal, seperti uang yang tidak sepenuhnya masuk ke dalam area *bounding box*, kemiringan sudut yang terlalu jauh dari posisi tegak lurus kamera, atau permukaan uang yang tidak rata akibat lipatan. Ketiga, kemiripan fitur visual antar beberapa pasangan nominal turut menjadi penyebab tingginya angka kesalahan klasifikasi. Berdasarkan data pada Tabel 3, kesalahan berulang terlihat pada pasangan nominal Rp1.000 dan Rp2.000 yang memiliki ukuran dan dominasi warna serupa, serta pasangan Rp50.000 dan Rp100.000 yang sama-

sama didominasi warna merah, sebagaimana terlihat pada data ke-1, ke-4, ke-5, dan ke-10. Tingginya nilai *false negative* sebesar 50% pada kedua skema juga menjadi perhatian serius. Kondisi ini menunjukkan bahwa sistem sangat sering gagal menghasilkan prediksi yang sesuai meskipun objek uang kertas telah berada di depan kamera. Hal ini mengindikasikan bahwa ambang batas kepercayaan model belum dikalibrasi secara baik untuk kondisi *real-time*, sehingga banyak prediksi yang tidak memenuhi syarat untuk ditampilkan dan disuarakan kepada pengguna. Selain itu, *false positive* yang lebih tinggi pada Skema 2 sebesar 23% dibandingkan Skema 1 sebesar 17% semakin memperkuat kesimpulan bahwa pendekatan anotasi pada seluruh area uang kertas membuat model lebih rentan menghasilkan prediksi yang salah.

Berdasarkan hasil perhitungan metrik evaluasi, diperoleh nilai *precision*, *recall*, dan *F1-score* untuk masing-masing skema sebagaimana ditunjukkan pada Tabel 5. Pada Skema 1, nilai *precision* mencapai 66,0%, *recall* sebesar 39,8%, dan *F1-score* sebesar 49,6%. Sementara itu, Skema 2 menunjukkan nilai *precision* sebesar 54,0%, *recall* sebesar 35,1%, dan *F1-score* sebesar 42,6%. Nilai *precision* yang lebih tinggi pada Skema 1 menunjukkan bahwa sistem pada skema tersebut memiliki ketepatan yang lebih baik dalam mengklasifikasikan nominal uang secara benar tanpa banyak menghasilkan deteksi yang salah. Selain itu, nilai *F1-score* yang juga lebih tinggi pada Skema 1 menandakan keseimbangan yang lebih baik antara ketepatan dan kemampuan mendeteksi nominal uang yang sebenarnya. Namun, nilai *recall* yang masih berada di bawah 50% pada kedua skema mengindikasikan bahwa sistem masih sering gagal mengenali sebagian besar nominal uang yang seharusnya terdeteksi, yang sejalan dengan tingginya angka *false negative* pada pengujian *real-time*. Secara keseluruhan, hasil pengujian ini menunjukkan bahwa meskipun model CNN berbasis VGG16 mampu mencapai performa yang baik pada data uji statis, kemampuan generalisasinya terhadap kondisi *real-time* masih perlu ditingkatkan. Diperlukan upaya perbaikan yang mencakup penambahan jumlah dan variasi data latih khususnya pada kondisi pencahayaan dan sudut pengambilan yang beragam, penyesuaian parameter jaringan, perbaikan tahap pra-pemrosesan citra, serta kalibrasi ambang batas kepercayaan prediksi agar sistem dapat bekerja secara lebih andal dalam skenario penggunaan nyata oleh pengguna lanjut usia.

Tabel 4. Perbandingan hasil *accuracy* data *real-time* dua skema

Skema	TN (✓)	FP (×)	FN (–)	Accuracy
1	33	17	50	0,33

Skema	TN (✓)	FP (×)	FN (–)	Accuracy
2	27	23	50	0,27

Tabel 5. Perbandingan hasil *precision*, *recall*, dan *f1-score* data *real-time* dua skema

Skema	Precision	Recall	F1-Score
1	0,660	0,398	0,496
2	0,540	0,351	0,426

Berdasarkan hasil perhitungan metrik evaluasi, diperoleh nilai *precision*, *recall*, dan *F1-score* untuk masing-masing skema sebagaimana ditunjukkan pada Tabel 5. Pada Skema 1, nilai *precision* mencapai 66,0%, *recall* sebesar 39,8%, dan *F1-score* sebesar 49,6%. Sementara itu, Skema 2 menunjukkan nilai *precision* sebesar 54,0%, *recall* sebesar 35,1%, dan *F1-score* sebesar 42,6%. Nilai *precision* yang lebih tinggi pada Skema 1 menunjukkan bahwa sistem pada skema tersebut memiliki ketepatan yang lebih baik dalam mengklasifikasikan nominal uang secara benar tanpa banyak menghasilkan deteksi yang salah (*false positive*). Selain itu, nilai *F1-score* yang juga lebih tinggi menandakan bahwa Skema 1 memiliki keseimbangan yang lebih baik antara ketepatan dan kemampuan mendeteksi nominal uang yang sebenarnya (*recall*). Namun, nilai *recall* yang masih berada di bawah 50% pada kedua skema mengindikasikan bahwa sistem masih sering gagal mengenali sebagian besar nominal uang yang seharusnya terdeteksi. Kondisi ini dapat disebabkan oleh variasi pencahayaan, kualitas citra yang kurang optimal, atau kemiripan fitur visual antar nominal uang. Dengan demikian, diperlukan upaya peningkatan performa model melalui perbaikan tahap pra-pemrosesan citra, penambahan jumlah data latih, serta penyesuaian parameter jaringan untuk memperoleh tingkat akurasi deteksi yang lebih baik pada penelitian selanjutnya.

3.4 Evaluasi Interface Sistem dengan SUS

Tabel 6. Hasil evaluasi *interface* sistem dengan SUS

Data	Skor Hasil Hitung										Nilai (Jumlah × 2,5)
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	
1.	3	4	3	1	4	2	3	2	3	3	70
2.	4	2	3	2	3	2	3	3	4	2	70
3.	4	2	3	3	3	2	4	2	3	3	72,5
4.	2	2	2	3	4	3	3	2	3	2	65
5.	3	2	3	3	3	2	4	3	3	2	70
6.	4	3	4	2	3	3	4	2	3	2	75
7.	4	3	3	2	3	3	3	2	3	2	70
8.	3	2	4	2	4	3	3	2	2	2	70
9.	3	3	3	2	4	2	2	3	3	2	67,5
10.	4	3	3	2	3	2	4	3	3	2	72,5
Total skor $\sum x$											702,5
Skor rata-rata $\bar{x}$ (Hasil akhir)											70,25

Tabel 6 menunjukkan hasil evaluasi SUS (*System Usability Scale*) yang diambil setelah 10 orang responden lansia (dari rentang usia antara 60 hingga 82 tahun) melakukan uji coba sistem usulan secara langsung. Dari data dalam Tabel 6 diketahui bahwa skor total SUS yang diperoleh sebesar 702,5, dengan nilai rata-rata ($\bar{x}$) sebesar 70,25 yang menempatkan sistem dalam kategori baik dalam hal *usability*. Nilai ini mengindikasikan bahwa sebagian besar lansia merasa sistem ini cukup mudah digunakan serta mampu memenuhi kebutuhan mereka dalam mengenali nominal uang secara mandiri.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian yang telah dilakukan, sistem identifikasi nominal uang kertas rupiah berbasis CNN dengan arsitektur VGG16 dan fitur *text-to-speech* menunjukkan kinerja yang baik pada data uji statis. Skema pertama mencapai akurasi pelatihan 99,47% dan validasi 100%, dengan nilai *precision*, *recall*, dan *F1-score* masing-masing 100%, sedangkan skema kedua memperoleh akurasi pelatihan 74,37%, validasi 98,45%, dan rata-rata *F1-score* sebesar 94% dengan sedikit kesalahan pada nominal Rp10.000 dan Rp20.000. Meskipun demikian, pengujian secara *real-time* terhadap sepuluh responden lansia menunjukkan bahwa kedua skema masih memerlukan peningkatan performa, karena sistem belum sepenuhnya robust terhadap variasi pencahayaan dan posisi uang yang ditangkap secara *real-time*, meskipun skema pertama tetap menunjukkan ketepatan deteksi yang lebih tinggi dibandingkan skema kedua. Dari sisi usability, nilai SUS sebesar 70,25 mengindikasikan bahwa antarmuka sistem termasuk dalam kategori baik dan dapat digunakan dengan nyaman oleh pengguna lansia. Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan CNN berbasis VGG16 efektif dalam membantu lansia mengenali nominal uang kertas, dengan rekomendasi pengembangan ke depan berupa integrasi metode deteksi objek seperti YOLO atau SSD guna meningkatkan keandalan sistem pada kondisi *real-time*.

DAFTAR PUSTAKA

- Akbar, A. I., Dewi, S. R., & Suryaningsih, Y. (2024). Hubungan Dukungan Emosional Teman Sebaya Dengan Kualitas Hidup Lansia Di Upt Pelayanan Sosial Tresna Werdha Jember. *ASSYIFA: Jurnal Ilmu Kesehatan*, 2(2), 234–243. <https://doi.org/10.62085/ajk.v2i2.67>
- Arrizqi, N., Santoso, I., & Soetrisno, Y. A. A. (2021). Implementasi Google Text To Speech Pada Aplikasi Pendeteksi Uang Berbasis Android. *Transient: Jurnal Ilmiah*

Teknik Elektro, 10(3), 510–516. <https://doi.org/10.14710/transient.v10i3.510-516>

- Gideon, A. (2022). *Terbitkan Uang Baru 2022, Bank Indonesia Dapat Masukan dari Tunanetra*. Liputan6. <https://www.liputan6.com/bisnis/read/5045739/terbitkan-uang-baru-2022-bank-indonesia-dapat-masukan-dari-tunanetra?page=2>
- Isnanto, F., Muhammad, M. A., & Yulianti, T. (2023). Rancang Bangun Sistem Visualisasi Data Menggunakan Dashboard Pada Sistem Deteksi Hoaks Melalui Pendekatan Hcd (Human Centered Design). *Jurnal Informatika Dan Teknik Elektro Terapan*, 11(3s1), 806–814. <https://doi.org/10.23960/jitet.v11i3s1.3029>
- Mu'min Azis, M., Martanto, M., & Hayati, U. (2024). Analisis Usability Testing Menggunakan Metode System Usability Scale Pada Aplikasi Open Data Kabupaten Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3238–3243. <https://doi.org/10.36040/jati.v7i6.8154>
- Nur Hidayat, A. M., Antamil, A., & Zakiyah M, I. (2023). Identifikasi Nominal Mata Uang Rupiah Bagi Penyandang Tunanetra Dengan Algoritma Convolutional Neural Network Berbasis Android. *Journal Software, Hardware and Information Technology*, 3(2), 60–65. <https://doi.org/10.24252/shift.v3i2.102>
- Renaldo, F., Irsyad, H., & Hartati, E. (2025). *Penerapan Metode Design Thinking dalam Perancangan UI / UX Aplikasi Deteksi Mata Pengemudi*. 628–635.
- Rewina, A. E., Sulistyowati, Kurniawan, M., Dinarta, M. N., & Yunanda, S. F. (2024). Penerapan Metode CNN (Convolutional Neural Network) dalam Mengklasifikasi Uang Kertas dan Uang Logam. *Jurnal Terapan Informatika Nusantara*, 04(12), 778–785. <https://doi.org/10.47065/tin.v4i12.5128>