

ANALISIS PERFORMA ARSITEKTUR VISION TRANSFORMER DALAM DETEKSI NOMINAL MATA UANG RUPIAH

Azzam Trisna Wibawa, Tri Aristi Saputri

Teknik Informatika, Universitas Dharma Wacana

Jl. Kenanga No.3, Mulyojati, Kec. Metro Barat, Kota Metro, Lampung

Alanisaurora0@gmail.com

ABSTRAK

Pengenalan nilai uang kertas Rupiah merupakan salah satu penggunaan *computer vision* yang memiliki potensi untuk dijadikan dasar sistem bantuan dalam mengenali uang. Tantangan dalam studi ini adalah bagaimana cara mengidentifikasi nilai uang kertas Rupiah secara otomatis dengan model yang mampu membedakan karakteristik visual dari setiap pecahan, khususnya untuk uang Rupiah emisi 2022 yang terdiri dari tujuh kategori nominal. Tujuan dari penelitian ini adalah untuk mengevaluasi kemampuan arsitektur *Vision Transformer* (ViT) dalam mengklasifikasikan nilai uang kertas Rupiah emisi 2022. Metode yang digunakan adalah klasifikasi citra menggunakan model *vit-base-patch16* dengan pendekatan *transfer learning*. Hasil dari penelitian ini menunjukkan bahwa model mencapai nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 1,0000 pada data uji internal. Hasil ini menunjukkan bahwa *Vision Transformer* sangat efektif dalam mengklasifikasikan nominal uang kertas Rupiah pada dataset internal yang digunakan, meskipun pada pengujian tambahan masih terdapat kesalahan pada citra dalam kondisi yang kurang ideal.

Kata kunci : *Vision Transformer, Deep Learning, Mata Uang Rupiah, Klasifikasi Citra.*

1. PENDAHULUAN

Uang Kertas Rupiah berfungsi sebagai alat pembayaran resmi yang masih banyak dipergunakan dalam beragam transaksi sehari-hari di masyarakat. Dalam pelaksanaannya, kemampuan mengenali nilai uang secara cepat dan tepat sangat krusial, terutama dalam situasi di mana transaksi dilakukan secara langsung dan dalam waktu singkat. Masalah ini menjadi semakin signifikan bagi mereka yang memiliki keterbatasan penglihatan, karena kurangnya kemampuan penglihatan membuat identifikasi nilai uang menjadi lebih sulit. Walaupun uang kertas Rupiah dilengkapi dengan *blind code*, fitur tersebut seringkali tidak efektif dikarenakan kondisi fisik uang yang beredar tidak selalu dalam keadaan baik, misalnya tercetak kusut, terlipat, atau sudah tua.

Permasalahan dalam penelitian ini adalah pengenalan nominal uang kertas Rupiah secara otomatis masih menghadapi kendala pada kondisi citra yang bervariasi, seperti perbedaan pencahayaan, sudut pengambilan gambar, dan kondisi fisik uang yang terlipat atau kusut. Beberapa penelitian sebelumnya menunjukkan bahwa metode OCR, pengolahan citra, dan *deep learning* telah mampu memberikan hasil yang baik, namun performanya masih dapat dipengaruhi oleh kondisi lingkungan dan kualitas citra yang digunakan. Bahar dkk. menunjukkan bahwa pendekatan OCR dapat mencapai akurasi tinggi pada kondisi pengujian tertentu [1], Antara dkk. menunjukkan bahwa metode pengolahan citra klasik masih cukup efektif [2], dan Prima dkk. menunjukkan bahwa pendekatan berbasis CNN dan SSD juga mampu mengenali nominal uang Rupiah, tetapi masih mengalami kendala pada kondisi pencahayaan tertentu [3].

Oleh karena itu, penelitian ini dilakukan untuk mengetahui bagaimana penerapan arsitektur *Vision Transformer* (ViT) dalam klasifikasi nominal uang kertas Rupiah emisi 2022 serta bagaimana kinerjanya berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score*.

Dalam konteks tersebut, penelitian ini bertujuan untuk menganalisis kinerja arsitektur *Vision Transformer* dalam klasifikasi nominal uang kertas Rupiah yang diterbitkan pada tahun 2022. Penelitian memanfaatkan dataset Citra Uang Rupiah Emisi 2022 yang mencakup tujuh kategori nominal, yaitu Rp1.000, Rp2.000, Rp5.000, Rp10.000, Rp20.000, Rp50.000, dan Rp100.000, dengan strategi *transfer learning* untuk memanfaatkan keahlian model yang telah dilatih sebelumnya. Penilaian kinerja dilakukan dengan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* agar performa model dapat dievaluasi secara komprehensif. Dengan cara ini, penelitian ini diharapkan dapat memberikan sumbangan untuk pengembangan metode klasifikasi citra uang kertas Rupiah sekaligus menjadi landasan bagi pengembangan sistem pendukung pengenalan nominal uang dalam penelitian-penelitian mendatang.

Untuk menyelesaikan masalah dalam penelitian ini dilakukan dengan menerapkan arsitektur *Vision Transformer* (ViT) pada proses klasifikasi nominal uang kertas Rupiah emisi 2022. Model *Vision Transformer* diterapkan dengan pendekatan *transfer learning* agar proses pembelajaran menjadi lebih efisien dan model dapat memanfaatkan pengetahuan awal dari pelatihan sebelumnya pada dataset citra yang lebih besar. Selanjutnya, performa model dianalisis menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui sejauh mana

Vision Transformer mampu mengklasifikasikan nominal uang kertas Rupiah emisi 2022 secara tepat.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian seputar pengenalan nilai uang kertas Rupiah telah dilaksanakan dengan beragam metode, termasuk *Optical Character Recognition* (OCR), pengolahan citra, *Convolutional Neural Network* (CNN), serta model deteksi objek seperti SSD dan YOLO. Bahar et al. menciptakan sebuah model pendeteksi nilai uang kertas Rupiah yang berbasis OCR dengan memakai algoritma *Regular Expression* pada perangkat *Android*, memanfaatkan kamera ponsel untuk menangkap gambar uang, mengkonversinya menjadi teks, dan kemudian memberikan respons berupa getaran dan suara untuk membantu pengguna tunanetra. Uji coba dalam studi tersebut menunjukkan tingkat keakuratan mencapai 100% pada 210 sampel data yang diambil dari tujuh jenis uang kertas Rupiah. Penelitian ini mengindikasikan bahwa OCR dapat dijadikan pendekatan awal dalam mengenali nilai uang, khususnya untuk sistem bantuan berbasis perangkat *mobile* [1].

Selain OCR, metode pengolahan citra konvensional juga sudah diterapkan untuk mengenali nilai uang Rupiah. Antara et al. menggunakan segmentasi K-Means, ekstraksi fitur warna HSV, dan klasifikasi *Random Forest* untuk memprediksi nilai uang kertas Rupiah dari tujuh denominasi, berdasarkan 100 citra yang dijadikan data penelitian [2].

Akhirnya, hasil menunjukkan bahwa model *Random Forest* yang dikembangkan mampu memprediksi nilai uang dengan akurasi mencapai 95%. Azis et al. juga melakukan penyelidikan mengenai klasifikasi nilai uang Rupiah menggunakan metode pengolahan citra melalui *Canny Segmentation*, *Hu Moments*, serta klasifikasi dengan *Decision Tree* dan *Random Forest* pada dataset 30 kelas yang masing-masing terdiri dari 100 gambar. Temuan ini menunjukkan bahwa teknik berbasis fitur visual masih sangat relevan untuk tugas klasifikasi nilai uang, meskipun umumnya sangat tergantung pada kualitas fitur dan konsistensi data citra [4].

Pada tahap selanjutnya, metode berbasis CNN semakin populer karena dianggap lebih efektif dalam mempelajari pola visual secara otomatis. Amin dan rekan-rekannya melakukan penelitian mengenai klasifikasi uang kertas Rupiah terbaru dengan CNN sebagai model utama untuk mengidentifikasi nilai nominal [5].

Mohammad Soeharto dan timnya juga memanfaatkan CNN untuk membedakan antara nominal Rp2.000 dan Rp5.000, serta menunjukkan bahwa model ini mampu mencapai tingkat akurasi validasi yang baik meski pada data pengujian yang terbatas [6].

Nur Hidayat dan rekan-rekannya menggunakan arsitektur VGG-16 untuk mengenali nilai nominal

uang Rupiah bagi individu tunanetra dan berhasil memperoleh akurasi sebesar 83% saat menguji aplikasi tersebut. Di sisi lain [7], Aprillia dan timnya melakukan perbandingan antara CNN dan *Feedforward Neural Network* (FNN) dalam mendeteksi nominal Rupiah yang dikeluarkan pada tahun 2022 di berbagai kondisi pencahayaan, dan hasilnya menunjukkan bahwa CNN memberikan kinerja yang lebih unggul dibandingkan FNN [8]. Umumnya, studi-studi tersebut menegaskan bahwa CNN cukup efektif untuk mengklasifikasikan nilai uang, tetapi kinerjanya masih dipengaruhi oleh faktor pencahayaan, jumlah data yang tersedia, dan kompleksitas kelas yang digunakan.

Di sisi lain, banyak penelitian yang berfokus pada deteksi objek menggunakan SSD dan YOLO karena keuntungannya dalam mendeteksi secara *real-time*. Prima et al. memanfaatkan arsitektur SSD MobileNetV3 untuk membantu orang tunanetra dalam mengenali nilai uang kertas Rupiah. Mereka menyatakan bahwa penggunaan fitur *blind code* sering kali tidak efektif disebabkan oleh kondisi uang yang tidak selalu terpelihara dengan baik [3].

Penelitian tersebut menunjukkan bahwa sistem dapat berfungsi dengan baik dalam pencahayaan yang memadai dengan akurasi antara 80% hingga 95%, tetapi masih mengalami masalah dalam situasi cahaya rendah atau malam. Catyaningga juga menerapkan metode *Single Shot Detector* untuk mengklasifikasikan nilai uang kertas Rupiah secara *real-time* menggunakan *webcam*, dan hasilnya menunjukkan bahwa model SSD cukup menjanjikan untuk pengenalan nilai berbasis gambar [9].

Temuan ini menunjukkan bahwa SSD sangat efisien dalam mendukung sistem bantu untuk pengenalan uang, tetapi tetap rentan terhadap kondisi lingkungan saat proses pengambilan gambar [3].

2.2. Computer Vision dan Deep Learning

Computer vision adalah cabang dari kecerdasan buatan yang memungkinkan perangkat komputer untuk menangkap, mengolah, dan memahami informasi yang berasal dari gambar atau video. Di dalam disiplin ini, komputer dilatih untuk dapat mengidentifikasi pola visual, membedakan antar objek, serta membuat keputusan berdasarkan gambar yang diberikan. Dalam studi mengenai uang kertas Rupiah, *computer vision* memiliki peranan yang krusial karena sistem perlu mampu membaca karakteristik visual pada uang tersebut, termasuk warna, bentuk, angka nominal, dan fitur lain yang terlihat pada gambar. Implementasi *computer vision* dalam deteksi nilai uang Rupiah juga telah diterapkan dalam berbagai penelitian sebelumnya, yang menunjukkan bahwa teknologi ini sangat berfungsi untuk mendukung proses pengenalan nilai secara otomatis [10], [11].

Salah satu metode yang sering diterapkan dalam *computer vision* adalah *deep learning*. *Deep learning* adalah bagian dari pembelajaran mesin yang

memanfaatkan jaringan saraf tiruan berlapis untuk mengidentifikasi pola dari data dalam volume besar. Melalui metode ini, sistem dapat secara otomatis mengenali fitur-fitur penting tanpa harus sepenuhnya ditentukan secara manual oleh para peneliti. Catyaningga menjelaskan bahwa *deep learning* menggunakan jaringan saraf tiruan untuk menyelesaikan masalah dengan memanfaatkan dataset yang besar, serta bisa mempelajari dan mengklasifikasikan data seperti gambar, teks, dan suara dengan tingkat *accuracy* tinggi [9].

2.3. Klasifikasi Citra

Klasifikasi citra adalah langkah untuk mengelompokkan suatu foto ke dalam kategori tertentu dengan mempertimbangan ciri-ciri atau pola visual yang ada. Dalam dunia *computer vision*, klasifikasi citra berfungsi agar sistem mampu mengidentifikasi konten gambar dan memberikan label yang tepat. Dalam penelitian ini, klasifikasi citra diterapkan untuk mengenali nilai uang kertas Rupiah berdasar gambar yang dimasukkan ke dalam sistem. Dengan cara ini, model akan memahami perbedaan ciri visual dari setiap nilai nominal, lalu menentukan kategori apa yang sesuai untuk gambar tersebut [6], [12].

2.4. Vision Transformer

Vision Transformer (ViT) adalah salah satu model *deep learning* untuk pengolahan gambar yang diambil dari struktur *Transformer* yang sebelumnya banyak digunakan dalam *Natural Language Processing* (NLP). Berbeda dengan *Convolutional Neural Network* (CNN) yang mengambil fitur gambar dengan cara melakukan operasi konvolusi, ViT memproses gambar dengan membagi citra menjadi bagian-bagian kecil yang disebut *patch*. Setiap *patch* kemudian diubah menjadi bentuk vektor agar dapat diproses oleh model. Metode ini memperlakukan citra sebagai kumpulan bagian kecil yang saling terhubung, bukan sebagai satu gambar utuh [13], [14].

2.5. Transfer Learning

Transfer learning merupakan suatu metode dalam *deep learning* yang mengandalkan model yang telah dilatih sebelumnya dengan dataset besar, lalu disesuaikan untuk menyelesaikan tugas baru yang serupa. Dengan cara ini, model tidak perlu mempelajari seluruh pola dari awal, melainkan hanya perlu melakukan penyesuaian pada data yang relevan. Teknik ini sangat berguna saat jumlah dataset yang ada terbatas, karena dapat mempercepat proses pelatihan sekaligus meningkatkan kinerja model [13].

2.6. Uang Kertas Rupiah

Uang kertas Rupiah merupakan salah satu alat pembayaran yang sah di Indonesia dan digunakan dalam berbagai aktivitas transaksi sehari-hari. Sebagai

bagian dari mata uang nasional, uang kertas Rupiah memiliki nilai nominal, ciri visual, dan unsur keamanan yang membedakannya dari alat pembayaran lainnya. Dalam perkembangannya, uang kertas Rupiah diterbitkan dalam beberapa seri atau emisi yang menyesuaikan kebutuhan pembaruan desain, unsur pengaman, serta identitas visual yang ingin ditampilkan. Salah satu emisi yang digunakan dalam penelitian ini adalah uang kertas Rupiah emisi 2022, yaitu seri terbaru yang terdiri atas tujuh nominal dan memiliki karakteristik visual yang berbeda pada setiap pecahannya [5].

2.7. Confusion Matrix

Dalam kajian klasifikasi citra, penilaian model adalah langkah krusial untuk mengevaluasi seberapa efektif model dalam mengenali data yang telah disediakan. Salah satu metode yang lazim dipakai untuk menilai hasil klasifikasi adalah *confusion matrix*. *Confusion matrix* adalah sebuah tabel yang memperlihatkan perbandingan antara prediksi yang dihasilkan oleh model dan label yang sebenarnya. Melalui tabel ini, kita bisa mengetahui jumlah data yang terprediksi dengan benar maupun yang salah, sehingga kinerja model bisa dianalisis dengan lebih mendetail [14].

Berdasarkan *confusion matrix*, terdapat beberapa metrik evaluasi yang diaplikasikan dalam penelitian ini, antara lain *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* dipakai untuk mengukur seberapa besar proporsi prediksi yang benar di antara seluruh data. *Precision* menampilkan seberapa akurat model dalam memberikan prediksi positif. *Recall* memperlihatkan kemampuan model dalam menemukan semua data positif yang sebenarnya ada. Di sisi lain, *F1-score* merupakan rata-rata harmonis antara *precision* dan *recall*, sehingga metrik ini digunakan untuk menilai keseimbangan performa model. Rumus dari metrik tersebut dapat dinyatakan sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Dalam studi ini, *confusion matrix* dan ukuran evaluasi digunakan untuk menilai kinerja model *Vision Transformer* dalam mengklasifikasikan gambar uang kertas Rupiah edisi 2022 ke dalam tujuh kategori nominal. Dengan memanfaatkan ukuran tersebut, penilaian model tidak hanya berfokus pada tingkat *accuracy*, tetapi juga pada ketepatan dalam prediksi serta kemampuan model untuk mengidentifikasi setiap kategori dengan lebih komprehensif. Oleh karena itu, matriks kebingungan, *accuracy*, *precision*, *recall*, dan *F1-Score* menjadi elemen fundamental dalam proses analisis kinerja model dalam penelitian ini.

3. METODE PENELITIAN

3.1. Jenis dan Pendekatan Penelitian

Penelitian ini termasuk dalam kategori eksperimen dengan pendekatan kuantitatif. Pemilihan penelitian eksperimen dilakukan karena fokus utama studi ini adalah penerapan model *deep learning*, yaitu *Vision Transformer* (ViT), untuk mengklasifikasikan citra uang kertas Rupiah dari emisi 2022, serta menganalisis dan menilai kinerja model melalui proses pelatihan dan pengujian. Pendekatan kuantitatif diadopsi karena hasil dari penelitian dinyatakan dalam format data numerik yang diperoleh melalui pengukuran kinerja model, termasuk *accuracy*, *precision*, *recall*, dan *F1-score*. Pendekatan kuantitatif dalam eksperimen seperti ini juga terlihat pada penelitian Muhammad Fakhri Fadhlurrahman dkk. yang menerapkan model *Vision Transformer* dan mengevaluasi kinerja sistem melalui langkah-langkah pengumpulan data, pra-pemrosesan, pelatihan, pengujian, serta penilaian kinerja [15].

Pada studi ini, eksperimen dilaksanakan dengan memanfaatkan dataset Citra Uang Rupiah Emisi 2022 dari *Kaggle* sebagai sumber data utama, kemudian model *Vision Transformer* diaplikasikan untuk menganalisis pola visual dari tujuh kategori nominal uang kertas Rupiah. Proses penelitian tidak hanya terbatas pada penerapan model, melainkan juga mengevaluasi kemampuan model dalam menghasilkan klasifikasi yang tepat berdasarkan metrik evaluasi yang telah ditentukan. Dengan cara ini, penelitian ini lebih fokus pada analisis performa model dalam tugas klasifikasi citra dan bukan pada pengembangan aplikasi siap pakai. Karakteristik penelitian semacam ini sejalan dengan studi Figo dkk. yang menggunakan *Vision Transformer* untuk tugas klasifikasi citra dan menganalisis performa model melalui pengujian parameter, dampak *transfer learning*, dan hasil evaluasi numerik [13].

Selain itu, metode yang diterapkan dalam studi ini juga menggunakan *transfer learning*. Metode ini dipilih karena model yang telah dilatih sebelumnya pada kumpulan data yang lebih besar dapat digunakan kembali dan disesuaikan dengan kumpulan data yang lebih spesifik pada penelitian ini. Penerapan *transfer learning* dinilai dapat mempercepat tahap pelatihan dan meningkatkan kemampuan model dalam mengenali pola gambar, terutama saat jumlah data dalam penelitian ini tidak mencukupi. Oleh sebab itu, studi ini dapat dikategorikan sebagai penelitian kuantitatif eksperimental yang berbasis komputasi, dengan penekanan pada penerapan *Vision Transformer* dan evaluasi kinerjanya dalam mengklasifikasikan uang kertas Rupiah edisi 2022.

3.2. Objek dan Subjek Penelitian

Objek yang dicermati dalam studi ini adalah citra uang kertas Rupiah dengan Tahun Emisi 2022 yang terdistribusi dalam tujuh kategori nilai, yakni Rp1.000, Rp2.000, Rp5.000, Rp10.000, Rp20.000, Rp50.000, dan Rp100.000. Pemilihan objek ini didasarkan pada

fakta bahwa uang kertas Rupiah edisi 2022 merupakan emisi paling mutakhir yang masih berfungsi dalam transaksi sosial, serta memiliki ciri visual yang khas di setiap nilai, termasuk warna dominan, angka nominal, gambar utama, serta elemen desain lainnya. Variasi dari ciri visual inilah yang berfungsi sebagai landasan bagi model dalam memahami dan membedakan masing-masing kategori nominal.

Subjek yang terlibat dalam penelitian ini adalah kumpulan data citra uang kertas Rupiah Emisi 2022, yang berperan sebagai sumber data utama dalam tahapan pelatihan dan uji coba model. Kumpulan data ini diambil dari platform *Kaggle*, yaitu dataset Citra Uang Rupiah Emisi 2022, yang menyimpan gambar digital dari uang kertas Rupiah edisi 2022 dalam tujuh kategori nilai. Pemilihan dataset ini didasari oleh relevansinya dengan tujuan penelitian, yaitu mengkaji performa arsitektur *Vision Transformer* dalam mengelompokkan nilai uang kertas Rupiah. Dengan demikian, fokus penelitian tertuju pada citra uang kertas Rupiah emisi 2022, sedangkan subjek penelitian merujuk pada dataset citra yang dijadikan sebagai materi analisis dalam studi ini.

3.3. Teknik Pengumpulan Data

Teknik pengumpulan informasi dalam studi ini dilakukan melalui dokumentasi dan kajian literatur. Dokumentasi dilaksanakan dengan mengakses dataset gambar uang kertas Rupiah emisi 2022 yang tersedia secara terbuka di platform *Kaggle*, yaitu dataset Gambar Uang Rupiah Emisi 2022. Dataset ini dipilih sebagai sumber utama dalam penelitian karena menyimpan gambar digital uang kertas Rupiah yang diperlukan untuk mengklasifikasikan tujuh kategori nominal. Data yang diperoleh kemudian diorganisir dan dikelompokkan sesuai dengan kategori nominal agar memudahkan proses pengolahan pada tahap selanjutnya.

Di samping dokumentasi, penelitian ini juga mengandalkan kajian pustaka untuk mendapatkan dasar teori dan referensi yang berhubungan dengan pokok bahasan penelitian. Kajian pustaka dilakukan dengan menelaah buku, jurnal, artikel penelitian, dan sumber informasi lainnya yang membahas tentang *computer vision*, *deep learning*, *Vision Transformer*, *transfer learning*, klasifikasi citra, serta metrik evaluasi model. Teknik ini dimaksudkan agar penelitian memiliki landasan ilmiah yang solid dan dapat mendukung pengembangan metodologi serta analisis hasil penelitian.

Dengan begitu, teknik pengumpulan data dalam penelitian ini tidak dilakukan melalui observasi lapangan secara langsung, melainkan dengan mengumpulkan data sekunder berupa dataset gambar digital dan referensi ilmiah yang relevan. Data yang sudah terkumpul tersebut kemudian digunakan sebagai dasar pelatihan, pengujian, dan evaluasi model *Vision Transformer* dalam klasifikasi kategori nominal uang kertas Rupiah emisi 2022.

3.4. Teknik Analisis Data

Teknik analisis data di dalam studi ini dilakukan dengan pendekatan kuantitatif untuk hasil klasifikasi dari model *Vision Transformer* (ViT). Data citra uang kertas Rupiah emisi 2022 yang telah dikumpulkan sebelumnya dipersiapkan melalui tahap *preprocessing*, termasuk penyesuaian dimensi citra agar sesuai dengan ukuran input model. Proses ini sangat penting karena model *Vision Transformer* memproses citra dalam bentuk potongan dan memerlukan ukuran input yang konsisten [14].

Setelah tahap *preprocessing*, dataset kemudian dibagi antara data pelatihan dan data pengujian. Data pelatihan dimanfaatkan untuk melatih model agar dapat memahami pola visual dari setiap kategori nominal uang kertas Rupiah, sedangkan data pengujian digunakan untuk mengukur kemampuan model dalam mengkategorikan data yang belum pernah dilihat sebelumnya. Dalam penelitian ini, model *Vision Transformer* diterapkan menggunakan pendekatan *transfer learning* untuk membuat proses pelatihan lebih efisien dan memungkinkan model beradaptasi dengan lebih baik terhadap dataset penelitian. Pendekatan *transfer learning* pada model ViT juga diterapkan dalam studi klasifikasi citra oleh Figo dkk. dan Muhammad Fakhri Fadhlurrahman dkk. [13], [15].

Setelah penyelesaian sesi pelatihan, model diuji dengan menggunakan data pengujian guna menghasilkan prediksi kategori untuk setiap gambar uang kertas Rupiah. Hasil dari prediksi tersebut kemudian akan dianalisis dengan menggunakan *confusion matrix*. *Confusion matrix* berfungsi untuk membandingkan antara kategori hasil prediksi dari model dengan kategori yang sebenarnya, sehingga kita dapat mengetahui jumlah prediksi yang tepat atau keliru. Berdasarkan *confusion matrix* tersebut, dihitung nilai *accuracy*, *precision*, *recall*, dan *F1-score* yang berfungsi sebagai metrik utama dalam evaluasi pada studi ini. Penggunaan *confusion matrix* dan metrik evaluasi serupa juga diterapkan dalam penelitian oleh Firdaus dkk. dalam penilaian model *Vision Transformer* [14].

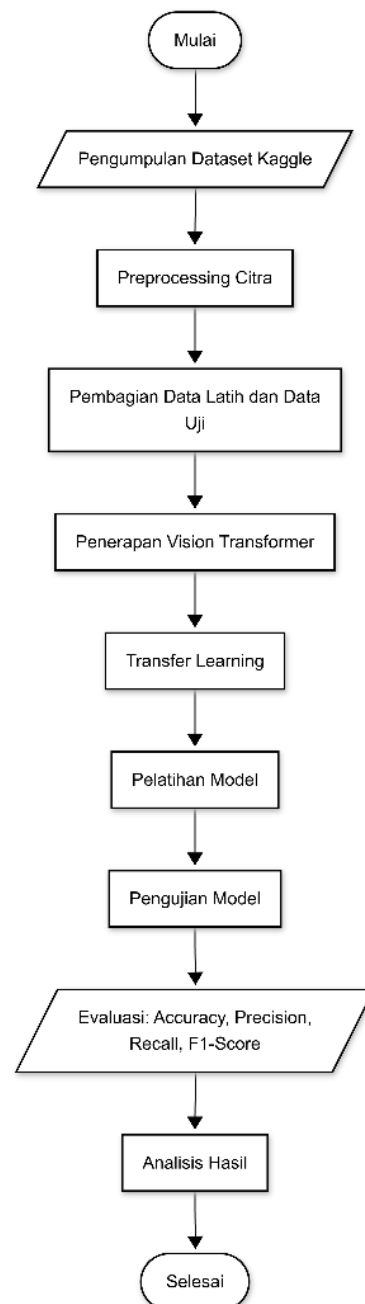
Berdasarkan hasil analisis ini, penelitian ini mengevaluasi kinerja struktur *Vision Transformer* dalam mengklasifikasikan nominal uang kertas Rupiah. Oleh karena itu, teknik analisis data dalam penelitian ini tidak hanya bertujuan untuk mengetahui tingkat *accuracy* model, tetapi juga untuk menilai ketepatan, konsistensi, serta kemampuan model dalam mengenali setiap kategori nominal dengan lebih mendalam.

3.5. Tahapan Penelitian

Tahapan penelitian ini disusun dengan rapi agar proses pengelompokan citra uang kertas Rupiah emisi 2022 melalui *Vision Transformer* (ViT) dapat dilakukan secara terstruktur. Penelitian dimulai dengan mengumpulkan dataset citra uang kertas Rupiah emisi 2022 dari media *Kaggle* sebagai sumber

data utama. Setelah mengumpulkan dataset, data ini kemudian dipersiapkan melalui tahap *preprocessing*, yang mencakup penyesuaian ukuran citra agar sesuai dengan spesifikasi input model. Tahap ini krusial karena *Vision Transformer* memproses citra dalam format *patch* dan memerlukan dimensi input yang konsisten.

Seluruh rangkaian proses penelitian, mulai dari pengumpulan data hingga evaluasi hasil akhir, secara sistematis disajikan dalam diagram alir pada Gambar 1 berikut ini.



Gambar 1. Diagram Tahapan Penelitian

Setelah menyelesaikan tahap *preprocessing*, dataset dibagi menjadi dua bagian: data latih dan data uji. Data latih berfungsi untuk melatih model dalam mengenali pola visual masing-masing nominal uang

kertas Rupiah, sedangkan data uji dipakai untuk menguji kemampuan model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya. Pada tahap pelatihan, model *Vision Transformer* diterapkan dengan pendekatan *transfer learning* untuk meningkatkan efisiensi proses pembelajaran dan agar model dapat beradaptasi dengan dataset yang digunakan dalam penelitian ini. Pendekatan ini juga diterapkan pada penelitian yang menggunakan ViT untuk tugas klasifikasi citra.

Tahap selanjutnya adalah pengujian model dengan menggunakan data uji untuk menghasilkan prediksi kelas dari setiap citra uang kertas Rupiah. Hasil prediksi tersebut kemudian dievaluasi dengan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. Dengan metrik tersebut, kinerja model dapat dianalisis secara komprehensif, baik dari segi keakuratan prediksi ataupun kemampuan model dalam mengidentifikasi setiap kelas nominal. Evaluasi yang didapat kemudian digunakan sebagai dasar untuk menarik kesimpulan mengenai kinerja *Vision Transformer* dalam klasifikasi nominal uang kertas Rupiah emisi 2022.

4. HASIL DAN PEMBAHASAN

4.1. Teknik Analisis Data

Penelitian ini dilakukan di platform *Google Colab* dengan dukungan GPU T4 yang menjadi perangkat komputasi utama. Pilihan menggunakan *Google Colab* berdasarkan pada kemampuan platform ini untuk menawarkan lingkungan pemrograman berbasis *cloud* yang memadai bagi pelatihan dan pengujian model *deep learning*. Model yang diterapkan dalam penelitian ini adalah *Vision Transformer* dengan arsitektur *google/vit-base-patch16-224-in21k*. Arsitektur ini dipilih karena relevan dengan tujuan penelitian, yaitu untuk menganalisis kinerja model yang berbasis *attention* dalam klasifikasi uang kertas nominal Rupiah.

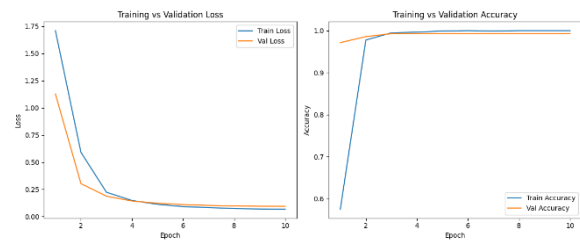
Dataset yang digunakan dalam penelitian berasal dari *Kaggle*, tepatnya dataset Citra Uang Rupiah Emisi 2022. Dataset ini terdiri dari 1.400 gambar yang dikelompokkan ke dalam 7 kelas nominal, yakni Rp1.000, Rp2.000, Rp5.000, Rp10.000, Rp20.000, Rp50.000, dan Rp100.000. Dalam proses eksperimen, dataset dibagi menjadi data pelatihan, data validasi, dan data pengujian dengan proporsi 80:10:10. Dari pembagian tersebut, dihasilkan 1.120 data pelatihan, 140 data validasi, dan 140 data pengujian yang digunakan untuk semua fase eksperimen.

Model dilatih selama 10 epoch. Selama tahap pelatihan, model belajar mengenali pola visual dari setiap kelas nominal berdasarkan gambar yang diberikan. Tujuan dari proses ini adalah untuk memastikan bahwa model dapat membedakan karakteristik dari setiap uang kertas Rupiah, termasuk warna, angka nominal, dan elemen desain utama. Melalui lingkungan eksperimen ini, penelitian dapat dilakukan secara sistematis untuk mengevaluasi

performa *Vision Transformer* dalam tugas klasifikasi uang kertas nominal Rupiah emisi 2022.

4.2. Hasil Pelatihan Model

Hasil dari pelatihan model dapat dianalisis melalui grafik yang menunjukkan nilai *training loss*, *validation loss*, *training accuracy*, dan *validation accuracy* yang didapatkan selama 10 epoch. Dari grafik yang ditampilkan, terlihat bahwa *training loss* mengalami penurunan yang signifikan, mulai dari sekitar 1,70 di epoch pertama dan turun mendekati 0,07 pada epoch ke-10. Sementara itu, *validation loss* juga menunjukkan penurunan yang konsisten, yang dimulai dari sekitar 1,13 di epoch pertama dan mencapai sekitar 0,09 di epoch ke-10. Penurunan nilai *loss* ini mengindikasikan bahwa model semakin baik dalam menyesuaikan parameter pembelajarannya terhadap data yang ada.



Gambar 2. Hasil Pelatihan Model

Selain adanya penurunan pada nilai *loss*, grafik *accuracy* juga memperlihatkan peningkatan performa model yang sangat baik. Nilai *training accuracy* naik dari sekitar 0,58 di epoch pertama hingga hampir mencapai 1,00 mulai dari epoch berikutnya sampai akhir pelatihan. Nilai *validation accuracy* pun terjaga pada tingkat yang sangat tinggi, dimulai dari sekitar 0,96 dan terus stabil hampir mencapai 1,00 hingga epoch terakhir. Pola ini mengindikasikan bahwa model tidak hanya mampu mempelajari data latihan, tetapi juga menghasilkan performa yang sangat baik pada data validasi.

Dari grafik pelatihan tersebut, terlihat bahwa model telah menjalani proses pembelajaran yang cepat dan konvergen. Pada beberapa epoch awal, kemajuan performa model sudah terlihat secara jelas, menunjukkan bahwa model dengan mudah mempelajari ciri visual antara kelas dalam dataset yang digunakan. Setelah itu, nilai *accuracy* cenderung stabil dan nilai *loss* terus menurun, menandakan bahwa model sudah mencapai tingkat performa yang sangat baik pada data internal. Secara keseluruhan, hasil pelatihan menunjukkan bahwa *Vision Transformer* dapat berfungsi secara efisien dalam proses klasifikasi uang kertas Rupiah emisi 2022.

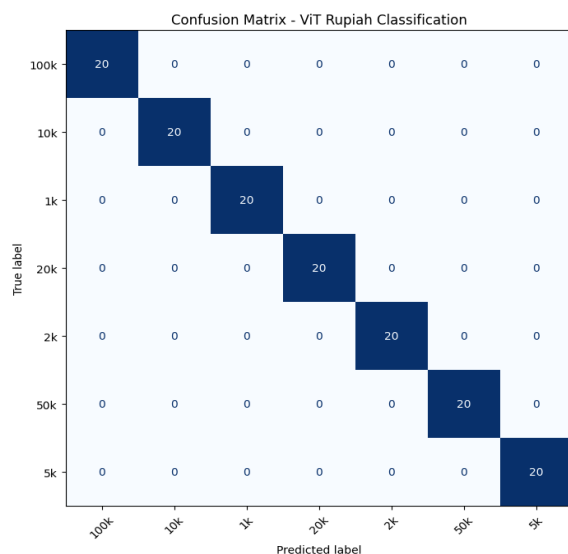
4.3. Hasil Pengujian Model

Setelah pelatihan selesai, model diuji menggunakan data pengujian untuk menilai kinerja akhir dalam klasifikasi nominal uang kertas Rupiah. Dari hasil pengujian, model ini mendapatkan nilai

evaluasi yang sangat memuaskan pada semua metrik yang digunakan. Angka *accuracy*, *precision*, *recall*, dan *F1-score* masing-masing mencapai 1,0000. Hasil ini menunjukkan bahwa pada data pengujian internal, model mampu mengklasifikasikan semua citra ke dalam kategori yang tepat.

Menurut laporan klasifikasi, setiap kategori nominal mendapatkan nilai *precision*, *recall*, dan *F1-score* masing-masing sebesar 1,0000. Semua kategori juga memiliki jumlah data pengujian yang seimbang, yaitu 20 citra per kategori, sehingga total data pengujian mencapai 140 citra. Kondisi ini menunjukkan bahwa kinerja model tidak hanya tinggi secara keseluruhan, tetapi juga konsisten di setiap kategori nominal yang diuji. Oleh karena itu, model *Vision Transformer* menunjukkan hasil yang sangat memuaskan dalam evaluasi internal dataset penelitian.

Di samping laporan klasifikasi, hasil pengujian juga didukung oleh *confusion matrix*. Matriks tersebut memperlihatkan bahwa semua prediksi terletak tepat pada diagonal utama, yang menunjukkan bahwa semua data pengujian berhasil diprediksi sesuai dengan label aslinya. Tidak ada nilai klasifikasi yang salah pada sel di luar diagonal, sehingga tidak ada kesalahan pertukaran antar kategori pada data pengujian internal. Hasil *confusion matrix* ini semakin memperkuat bahwa model mampu membedakan setiap nominal uang kertas Rupiah dengan sangat baik pada dataset yang digunakan.



Gambar 3. Confusion Matrix

Untuk memperjelas hasil evaluasi, ringkasan metrik pengujian model dapat ditunjukkan pada tabel berikut.

Tabel 1. Hasil Evaluasi Model

METRIK	NILAI
<i>Accuracy</i>	1,0000
<i>Precision</i>	1,0000
<i>Recall</i>	1,0000
<i>F1-Score</i>	1,0000

Selain itu, hasil klasifikasi untuk masing-masing kelas dapat ditunjukkan sebagai berikut.

Tabel 2. Hasil Klasifikasi per Class

Class	Precision	Recall	F1-score	Support
Rp100.000	1,0000	1,0000	1,0000	20
Rp10.000	1,0000	1,0000	1,0000	20
Rp1.000	1,0000	1,0000	1,0000	20
Rp20.000	1,0000	1,0000	1,0000	20
Rp2.000	1,0000	1,0000	1,0000	20
Rp50.000	1,0000	1,0000	1,0000	20
Rp5.000	1,0000	1,0000	1,0000	20

Hasil evaluasi yang sangat tinggi tersebut menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengenali dan membedakan pola visual antar nominal uang kertas Rupiah pada data uji internal. Setiap kelas dapat dikenali tanpa terjadi kesalahan klasifikasi, sehingga model terlihat sangat efektif pada dataset yang digunakan. Namun demikian, hasil ini tetap perlu dibahas lebih lanjut agar dapat diketahui apakah performa tersebut benar-benar merepresentasikan kemampuan model pada kondisi nyata. Oleh karena itu, selain pengujian internal, perlu juga diperhatikan hasil prediksi pada citra tambahan di luar data uji.

4.4. Hasil Prediksi

Selain melakukan pengujian dengan menggunakan data uji internal, model ini juga diuji dengan beberapa gambar tambahan untuk mengevaluasi kemampuan prediksinya terhadap nominal uang kertas Rupiah dalam kondisi yang lebih bervariasi. Pengujian tambahan ini sangat penting agar hasil penelitian tidak hanya bergantung pada metrik evaluasi numerik, tetapi juga dapat dianalisis secara visual melalui contoh prediksi yang nyata. Berdasarkan hasil yang didapatkan, model menunjukkan kemampuan untuk memberikan prediksi yang akurat pada beberapa gambar tambahan dalam berbagai kondisi, seperti posisi uang, sudut pengambilan gambar, latar belakang, dan keadaan fisik uang. Hasil ini menunjukkan bahwa model tidak hanya berfungsi dengan baik pada data uji internal tetapi juga menunjukkan performa yang memuaskan pada beberapa gambar tambahan yang diuji langsung.

Pada Gambar 4, terdapat beberapa contoh hasil prediksi yang tepat yang dihasilkan oleh model *Vision Transformer*. Contoh-contoh tersebut meliputi prediksi untuk nominal Rp20.000 dengan tingkat kepercayaan 0,8642, Rp10.000 dengan tingkat kepercayaan 0,8511, Rp1.000 dengan tingkat kepercayaan 0,8334, Rp100.000 dengan tingkat kepercayaan 0,8370, serta Rp50.000 dengan tingkat kepercayaan 0,8658. Selain itu, model ini juga berhasil meramalkan nominal Rp10.000 dengan tingkat kepercayaan 0,8474, nominal Rp10.000 lainnya dengan tingkat kepercayaan 0,8121, nominal Rp100.000 dengan tingkat kepercayaan 0,8720, dan nominal Rp2.000 dengan tingkat kepercayaan 0,7981. Dengan demikian, hasil prediksi tambahan tersebut

menunjukkan bahwa model masih mampu mengenali berbagai nominal uang kertas Rupiah dengan baik meskipun kondisi gambar tidak sepenuhnya konsisten.

Jika kita analisis lebih dalam, nilai kepercayaan dalam hasil prediksi tambahan tersebut tidak selalu mendekati 1,00 meskipun prediksinya akurat. Ini menunjukkan bahwa model tetap memiliki variasi dalam tingkat keyakinannya terhadap setiap gambar, tergantung pada kualitas visual dari objek yang dianalisis. Dalam beberapa kasus, uang tampak dalam posisi miring, sedikit terlipat, atau difoto dengan latar belakang yang cukup berbeda dibandingkan citra lainnya, tetapi model tetap dapat memberikan prediksi yang benar. Hasil ini menandakan bahwa model memiliki kemampuan klasifikasi yang baik, meskipun tingkat keyakinan model masih dipengaruhi oleh kondisi citra yang diujikan.



Gambar 4. Hasil Prediksi Benar



Gambar 5. Hasil Prediksi Salah

Sebaliknya, pengujian tambahan juga mengindikasikan bahwa model masih dapat menghasilkan prediksi yang keliru dalam keadaan tertentu. Pada Gambar 5, model memperkirakan nominal 2.000 dengan tingkat kepercayaan sebesar 0,7877, tetapi hasil itu tidak cocok dengan nominal yang sebenarnya. Gambar itu menunjukkan uang

dalam keadaan terlipat, tidak tegak lurus, dan sebagian area penting pada uang terhalang oleh lipatan, membuat informasi visual yang diterima model menjadi tidak lengkap. Temuan ini krusial karena menunjukkan bahwa meskipun performa model sangat baik dalam evaluasi internal, model tetap memiliki keterbatasan ketika menghadapi citra dengan kondisi visual yang kurang ideal.

Secara umum, hasil prediksi pada gambar tambahan menunjukkan gambaran yang lebih akurat tentang kapasitas model. Model ini terbukti dapat memprediksi sejumlah gambar tambahan dengan tepat, namun masih dapat melakukan kesalahan dalam situasi tertentu. Untuk itu, hasil pengujian tambahan ini sangat krusial untuk melengkapi hasil evaluasi internal yang sebelumnya memperlihatkan nilai metrik yang sangat tinggi. Dengan adanya contoh prediksi yang benar serta yang salah, analisis kinerja model menjadi lebih seimbang dan lebih mencerminkan kemungkinan penerapan model dalam keadaan nyata.

4.5. Ringkasan Hasil

Berdasarkan penelitian yang telah dilakukan, dapat ditarik kesimpulan bahwa model *Vision Transformer* berhasil diaplikasikan untuk mengelompokkan gambar uang kertas Rupiah emisi 2022 ke dalam tujuh kategori nominal. Selama proses pelatihan, terlihat adanya pola pembelajaran yang positif, yang ditunjukkan dengan penurunan nilai *loss* dan peningkatan *accuracy* secara konsisten hingga sesi pelatihan selesai. Hasil evaluasi internal juga menampilkan kinerja yang sangat mengesankan, dengan nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 1,0000. Hal ini menandakan bahwa model dapat mengenali semua kategori nominal dengan sangat baik pada data uji internal.

Di sisi lain, pengujian tambahan pada gambar di luar data uji internal mengindikasikan bahwa model masih bisa menghadapi kesalahan dalam prediksi. Kesalahan tersebut terjadi pada gambar yang berada dalam kondisi kurang ideal, seperti uang yang terlipat, posisi objek yang miring, dan latar belakang yang cukup mencolok. Temuan ini menunjukkan bahwa kinerja model yang sangat baik pada data internal tidak selalu mencerminkan kinerja dalam situasi nyata. Oleh karena itu, penelitian ini menegaskan bahwa *Vision Transformer* sangat efektif untuk klasifikasi nominal uang Rupiah dalam dataset internal, namun tetap perlu dilakukan pengujian lebih lanjut untuk menguji ketahanan model dalam kondisi penggunaan yang lebih bervariasi.

5. KESIMPULAN DAN SARAN

Berdasarkan evaluasi yang dilakukan, performa arsitektur *Vision Transformer* dalam studi ini menunjukkan hasil yang sangat memuaskan. Model ini memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 1,0000 pada data uji internal, dan matriks kebingungan menunjukkan bahwa semua data uji telah berhasil diklasifikasikan dengan tepat ke

dalam kelas masing-masing. Hasil ini mengindikasikan bahwa model memiliki kemampuan yang sangat tinggi dalam membedakan ciri visual dari uang kertas Rupiah dalam dataset yang digunakan. Oleh karena itu, tujuan penelitian yang bertujuan untuk mengukur kinerja *Vision Transformer* menggunakan parameter *accuracy*, *precision*, *recall*, dan *F1-score* telah tercapai.

Namun, hasil dari pengujian tambahan pada gambar di luar data uji internal menunjukkan bahwa model masih dapat melakukan kesalahan prediksi dalam kondisi tertentu. Kesalahan ini muncul pada gambar dengan kondisi visual yang kurang ideal, seperti uang yang terlipat, sudut pengambilan gambar yang tidak sempurna, dan latar belakang yang sangat menonjol. Temuan ini menunjukkan bahwa performa model yang tinggi pada data internal tidak sepenuhnya merefleksikan performa yang sama dalam kondisi nyata yang lebih beragam.

Berdasarkan penelitian yang sudah dilakukan, ada beberapa rekomendasi yang bisa diberikan untuk pengembangan studi ke depan. Penelitian ini mengindikasikan bahwa model *Vision Transformer* dapat menghasilkan performa yang sangat memuaskan pada data uji internal, meskipun masih ada kesalahan dalam prediksi pada beberapa citra tambahan yang tidak dalam kondisi optimal. Oleh karena itu, penelitian mendatang disarankan agar memanfaatkan dataset dengan variasi kondisi citra yang lebih beragam, seperti perbedaan dalam pencahayaan, sudut pemotretan, latar belakang, serta kondisi fisik uang yang terlipat, kusut, atau sebagian terhalang. Dengan adanya data yang lebih bervariasi, diharapkan model dapat meningkatkan kemampuan generalisasinya dalam situasi nyata.

DAFTAR PUSTAKA

- [1] B. Bahar, R. D. Yc. Raban, and R. Arnje, "MODEL PENDETEKSI NOMINAL UANG KERTAS RUPIAH MENGGUNAKAN TEKNOLOGI OPTICAL CHARACTER RECOGNITION," *Ticom*, vol. 12, no. 1, pp. 8–13, Sep. 2023, doi: 10.70309/ticom.v12i1.101.
- [2] A. M. E. Antara, S. A. Sari, N. Riswanti, D. A. Amin, V. Verdila, and A. P. A. Masa, "Deteksi Nominal Rupiah Uang Kertas Berdasarkan Citra Warna Menggunakan Segmentasi K-Means Clustering dan Klasifikasi Random Forest," *kretisi*, vol. 1, no. 1, pp. 34–39, Jul. 2023, doi: 10.30872/kretisi.v1i1.776.
- [3] A. Prima, D. B. Santoso, and L. Nurpulaela, "DETEKSI OTOMATIS NOMINAL UANG KERTAS RUPIAH UNTUK TUNANETRA MENGGUNAKAN ALGORITMA ARSITEKTUR SSD MOBILENETV3," *teknokom*, vol. 6, no. 2, pp. 151–159, Aug. 2022, doi: 10.31943/teknokom.v6i2.166.
- [4] H. Azis, P. Purnawansyah, and N. Alfyyah, "Multiclass Classification on Nominal Value of Rupiah Banknotes Based on Image Processing," *Ilk. J. Ilm.*, vol. 16, no. 1, pp. 87–99, Apr. 2024, doi: 10.33096/ilkom.v16i1.1784.87-99.
- [5] A. S. Amin, M. A. R. Zahra, K. Fahreza, and A. P. Sari, "Klasifikasi Uang Kertas Rupiah Baru Menggunakan Metode CNN".
- [6] Mohammad Soeharto, Mohammad Jeky Hasan, Ahmad Rega Susanto, and Dimas Ahmad Fahrezi, "Mengklasifikasi Mata Uang Lima Ribu Rupiah dan Dua Ribu Rupiah dengan Menggunakan Algoritma CNN," *SABER*, vol. 2, no. 3, pp. 163–170, Jun. 2024, doi: 10.59841/saber.v2i3.1407.
- [7] A. M. Nur Hidayat, A. Antamil, and I. Zakiyah M, "Identifikasi Nominal Mata Uang Rupiah Bagi Penyandang Tunanetra Dengan Algoritma Convolutional Neural Network Berbasis Android," *shift*, vol. 3, no. 2, pp. 60–65, Jun. 2023, doi: 10.24252/shift.v3i2.102.
- [8] D. Aprillia, T. Rohana, T. A. Mudzakir, and D. Wahiddin, "Deteksi Nominal Mata Uang Rupiah Menggunakan Metode Convolutional Neural Network dan Feedforward Neural Network".
- [9] L. A. Catyaningga, "KLASIFIKASI UANG KERTAS RUPIAH BERDASARKAN ANGKA NOMINAL SECARA REALTIME MENGGUNAKAN METODE SINGLE SHOT DETECTOR".
- [10] A. D. E. Putra, M. F. Ardiyansyah, A. Harin, and R. M. Akbar, "Sistem Deteksi Nominal Uang Rupiah untuk Tunanetra Berbasis YOLOv8 dan Streamlit".
- [11] M. F. Mahfuzh, M. N. Abdillah, and B. Fatkhurrozi, "DETEKSI RUPIAH EMISI 2022 UNTUK DISABILITAS NETRA MENGGUNAKAN YOLOV5M DENGAN OUTPUT SUARA," *j. inti nm*, vol. 19, no. 1, pp. 01–09, Jun. 2024, doi: 10.33480/inti.v19i1.5295.
- [12] A. S. Romadhon, S. Patmanthara, and A. N. Handayani, "Development of a CNN-Based Knowledge System for Rupiah Currency Authenticity Detection and Nominal Classification," *generation*, vol. 10, no. 1, pp. 38–49, Jan. 2026, doi: 10.29407/gj.v10i1.27464.
- [13] J. A. Figo, N. Yudistira, and A. W. Widodo, "Deteksi Covid-19 dari Citra X-ray menggunakan Vision Transformer".
- [14] D. Firdaus, I. Sumardi, C. Chazar, and M. Z. Dafy, "Multiclass Klasifikasi Malware Berbasis Gambar Menggunakan Vision Transformer Architecture".
- [15] Muhammad Fakhri Fadhlurrahman, Munir, and Yaya Wihardi, "Pengenalan Ekspresi Wajah Peserta Didik di Ruang Kelas Menggunakan Vision Transformer (ViT)," *JUKTISI*, vol. 4, no. 2, pp. 1047–1058, Aug. 2025, doi: 10.62712/juktisi.v4i2.531.