

OPTIMASI DETEKSI KELAS MINORITAS PADA DEEP LEARNING: EVALUASI KINERJA AUGMENTASI VISUAL MELAWAN TEKNIK SMOTE

Poetra Ebeline¹; Odi Nurdiawan²; Arif Rinaldi Dikananda³; Aris Pratama Putra⁴;

Program Studi Teknik Informatika¹
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi^{3,4}

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
poetraebelind@gmail.com

(*) Corresponding Author : destafebi26@gmail.com

Published : 30 Desember 2025

Abstract— *Class imbalance poses a substantial challenge in binary image classification, as the dominance of majority-class samples often reduces a model's ability to detect minority instances. This study compares two imbalance-handling strategies: image-level augmentation using visual transformations and feature-level balancing using the Synthetic Minority Oversampling Technique (SMOTE). The dataset comprises 1,001 cleaned and annotated images, divided using stratified sampling into 800 training and 201 testing samples. Two separate pipelines were developed. The first applies end-to-end training with the MobileNetV2 architecture and on-the-fly visual augmentation. The second extracts latent features using MobileNetV2, followed by SMOTE-based oversampling and classification with a Multi-Layer Perceptron (MLP). Model evaluation employed accuracy, precision, recall, and F1-score. Results indicate that the MLP + SMOTE model achieves superior accuracy (0.9552), precision (0.9697), and F1-score (0.9343), demonstrating more stable predictive performance. Meanwhile, the CNN + augmentation model attains the highest recall (0.9718), reflecting greater sensitivity to minority-class samples. These findings reveal a trade-off between sensitivity and predictive consistency, suggesting that the optimal technique depends on the performance priorities of the application. The study offers empirical insights to guide the selection of effective class imbalance mitigation strategies in binary image classification tasks.*

Keywords: *Class Imbalance, SMOTE, Image Augmentation, MobileNetV2, Binary Classification.*

Abstrak— *Ketidakeimbangan kelas merupakan salah satu tantangan paling signifikan dalam pengembangan model klasifikasi citra, terutama pada tugas klasifikasi biner keberadaan objek. Ketimpangan jumlah sampel antara kelas positif dan negatif menyebabkan model cenderung bias terhadap kelas mayoritas sehingga mengurangi kemampuan mendeteksi kelas minoritas. Penelitian ini bertujuan melakukan analisis komparatif terhadap dua pendekatan penyeimbangan data, yaitu augmentasi citra berbasis transformasi visual dan Synthetic Minority Oversampling Technique (SMOTE) berbasis ruang fitur. Dataset yang digunakan terdiri dari 1.001 citra hasil pembersihan dan pelabelan, yang kemudian dibagi secara stratifikasi menjadi 800 citra latih dan 201 citra uji. Dua pipeline pengolahan dibangun secara terpisah: (1) pelatihan end-to-end menggunakan arsitektur MobileNetV2 dengan augmentasi visual on-the-fly, dan (2) ekstraksi fitur menggunakan MobileNetV2 yang diikuti oversampling fitur melalui SMOTE serta klasifikasi menggunakan Multi-Layer Perceptron (MLP). Evaluasi performa dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score untuk memberikan gambaran menyeluruh mengenai kemampuan prediktif model. Hasil penelitian menunjukkan bahwa model MLP + SMOTE unggul dalam akurasi (0.9552), presisi (0.9697), dan F1-score (0.9343), menandakan kestabilan prediksi yang tinggi. Sebaliknya, model CNN + augmentasi memperoleh recall tertinggi (0.9718), menunjukkan kepekaan lebih baik terhadap kelas minoritas. Temuan ini memperlihatkan adanya trade-off antara sensitivitas dan konsistensi performa, sehingga pemilihan metode penyeimbangan data perlu disesuaikan dengan kebutuhan aplikasi. Penelitian ini memberikan kontribusi empiris penting sebagai dasar pengembangan strategi penanganan ketidakeimbangan data pada klasifikasi citra.*

Keyword: *Ketidakeimbangan kelas, SMOTE, Augmentasi Citra, MobileNetV2, Klasifikasi Biner.*

INTRODUCTION

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI) dan deep learning telah memberikan kontribusi signifikan terhadap peningkatan kemampuan sistem visi komputer dalam berbagai aplikasi, seperti deteksi objek, pengenalan pola, dan klasifikasi citra. Kemajuan ini didukung oleh perkembangan arsitektur jaringan saraf modern, ketersediaan dataset berskala besar, dan peningkatan kemampuan komputasi. Namun demikian, keberhasilan model-model tersebut sangat dipengaruhi oleh kualitas dan distribusi data yang digunakan dalam proses pelatihan. Dalam banyak domain, terutama visi komputer, data seringkali memiliki distribusi yang tidak seimbang antara kelas mayoritas dan minoritas. Kondisi ketidakseimbangan kelas ini dapat menurunkan performa model, terutama pada kelas yang jarang muncul, sehingga menimbulkan bias dan ketidakstabilan prediksi [1], [2].

Dalam konteks klasifikasi biner keberadaan objek, ketidakseimbangan kelas menjadi masalah kritis karena perbedaan jumlah sampel antara kelas "objek muncul" dan "objek tidak muncul" cenderung sangat besar. Model deep learning yang dilatih pada distribusi seperti ini memiliki kecenderungan memprioritaskan pola kelas mayoritas, sehingga performanya dalam mendeteksi kelas minoritas menurun drastis [3]. Kondisi ini mengakibatkan fenomena *accuracy paradox*, yaitu model menunjukkan akurasi tinggi secara keseluruhan namun gagal mengidentifikasi kelas minoritas secara akurat. Masalah tersebut menjadi semakin kompleks ketika objek minoritas memiliki variasi visual yang tinggi atau ukuran objek relatif kecil sehingga mempersulit proses ekstraksi fitur oleh model [4].

Berbagai penelitian telah menginvestigasi beragam pendekatan untuk menangani isu ketidakseimbangan kelas. Metode oversampling sintesis seperti Synthetic Minority Oversampling Technique (SMOTE) menjadi salah satu teknik yang paling banyak digunakan karena mampu meningkatkan representasi kelas minoritas tanpa mengurangi data mayoritas [5]. Namun, beberapa studi menyoroti bahwa SMOTE dapat menghasilkan sampel sintesis yang tidak sepenuhnya konsisten dengan distribusi kelas minoritas sehingga menyebabkan overlapping antar kelas dan meningkatkan risiko misclassification [6]. Merespons kelemahan ini, penelitian terbaru mengembangkan pendekatan oversampling berbasis fitur yang lebih adaptif dan stabil. Pendekatan tersebut bekerja pada ruang fitur hasil ekstraksi deep learning sehingga sampel sintesis dapat mencerminkan representasi abstrak objek minoritas yang lebih akurat [7], [8].

Di sisi lain, augmentasi citra berbasis transformasi visual telah menjadi strategi yang efektif untuk meningkatkan variasi data pelatihan. Teknik augmentasi sederhana seperti rotasi, flipping, zooming, dan cropping terbukti mampu memperkaya keragaman visual tanpa mengubah makna semantik citra [9], [10]. Augmentasi visual telah banyak digunakan dalam pipeline Convolutional Neural Network (CNN) karena kemampuannya membantu model mengenali pola objek dalam kondisi berbeda, termasuk pencahayaan, orientasi, atau latar belakang yang bervariasi. Namun, augmentasi citra tidak menyelesaikan ketidakseimbangan kuantitatif karena teknik ini tidak meningkatkan jumlah sampel minoritas secara proporsional. Kondisi ini membuat augmentasi kurang efektif pada skenario ketidakseimbangan ekstrem [4].

Selain isu distribusi data, penelitian lain menekankan bahwa performa model sangat dipengaruhi oleh kualitas anotasi dan representasi fitur [11], [12]. Anotasi yang tidak konsisten dapat menghasilkan fitur yang ambigu sehingga model mengalami kesulitan dalam membedakan kelas secara jelas. Lebih jauh lagi, representasi fitur hasil ekstraksi jaringan saraf juga dapat menunjukkan bias terhadap kelas mayoritas, terutama jika data minoritas menunjukkan variasi visual yang tinggi [13]. Oleh karena itu, pendekatan penanganan ketidakseimbangan kelas perlu mempertimbangkan baik aspek kuantitatif maupun kualitatif dari data.

Terkait evaluasi model pada dataset tidak seimbang, sejumlah penelitian menunjukkan bahwa metrik akurasi tidak cukup mencerminkan performa yang sebenarnya [14], [15]. Dalam kondisi imbalance, sebuah model dapat mencapai akurasi tinggi meskipun tidak pernah memprediksi kelas minoritas. Oleh karena itu, metrik seperti recall, precision, dan F1-score dinilai lebih representatif dalam mengukur kemampuan model mengenali objek pada kelas minoritas.

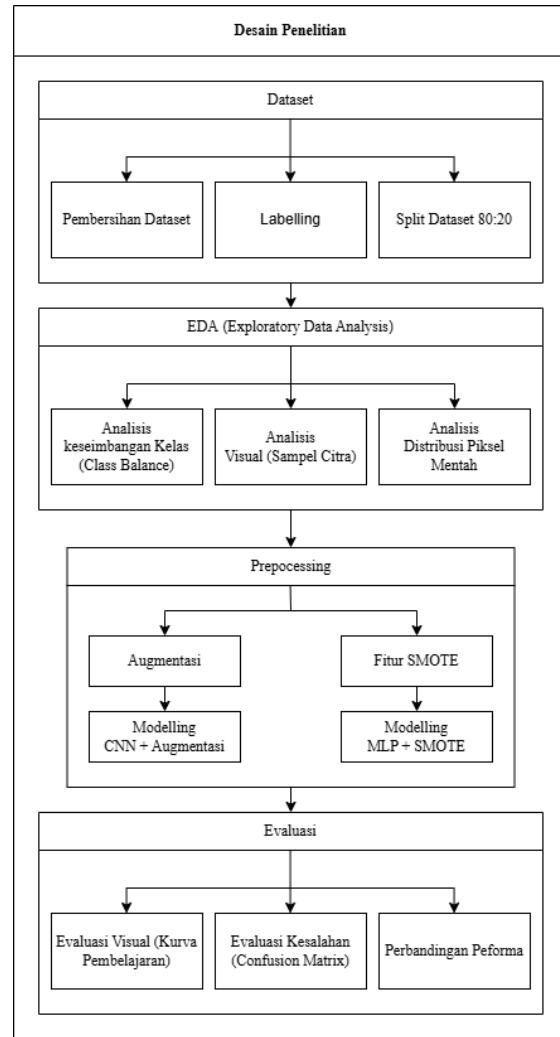
Di tengah banyaknya pendekatan penanganan imbalance, masih terdapat kesenjangan penting dalam literatur, yaitu minimnya studi yang secara langsung membandingkan performa SMOTE berbasis fitur dan augmentasi citra berbasis visual dalam satu kerangka eksperimen yang konsisten untuk tugas klasifikasi biner keberadaan objek. Sebagian besar penelitian hanya fokus pada satu pendekatan saja atau mengujinya pada dataset multi-kelas, sehingga belum memberikan pemahaman menyeluruh mengenai trade-off kedua teknik dalam konteks data biner yang sangat tidak seimbang [16].

Berdasarkan kesenjangan tersebut, penelitian ini menjadi penting untuk dilakukan

karena memberikan evaluasi komparatif yang objektif antara dua metode penanganan ketidakseimbangan kelas yang bekerja pada level representasi berbeda, yaitu augmentasi citra pada level visual dan SMOTE pada level fitur. Dengan menggunakan dua pipeline berbeda CNN dengan augmentasi citra dan MLP dengan SMOTE fitur penelitian ini memberikan gambaran empiris mengenai kelebihan dan keterbatasan masing-masing metode dalam menangani ketidakseimbangan data pada klasifikasi citra

MATERIALS AND METHODS

Bagian ini menjelaskan secara sistematis bahan, prosedur, serta pendekatan metodologis yang digunakan dalam penelitian. Seluruh tahapan dirancang berdasarkan kebutuhan untuk melakukan evaluasi komparatif antara dua strategi penanganan ketidakseimbangan data, yaitu augmentasi citra dan SMOTE berbasis fitur. Alur penelitian mengacu pada skema yang terdiri dari empat rangkaian utama: persiapan dataset, exploratory data analysis, preprocessing dalam dua jalur (augmentasi dan SMOTE), serta evaluasi performa model. Pendekatan ini dipilih agar penelitian dapat mengamati pengaruh teknik penyeimbangan data tidak hanya pada performa akhir model, tetapi juga pada dinamika pembelajaran dan stabilitas prediksi. Dengan metodologi yang terstruktur, penelitian ini berupaya menghasilkan analisis empiris yang dapat menjadi dasar pemilihan teknik penanganan imbalance pada klasifikasi citra biner di berbagai konteks aplikasi.



Gambar 1. Alur Penelitian

Tahap berikutnya pada Gambar 1. Alur Penelitian adalah Penelitian diawali dengan tahap pengolahan dataset yang terdiri atas pembersihan data, pelabelan, dan pembagian data. Pada tahap pembersihan, seluruh citra diperiksa untuk memastikan tidak terdapat data rusak, duplikasi, atau label yang tidak konsisten. Konsistensi data menjadi faktor penting untuk menghindari bias pada tahap pelatihan model, sebagaimana disoroti dalam kajian [11]. Selanjutnya, proses pelabelan dilakukan dengan menetapkan setiap citra sebagai kelas positif atau negatif berdasarkan keberadaan objek, mengikuti prinsip representasi label biner yang banyak digunakan dalam studi klasifikasi objek [17]. Setelah seluruh data diberi label, dataset dibagi menggunakan *stratified split* dengan rasio 80:20, agar proporsi kelas pada data latih dan data uji tetap seimbang sehingga evaluasi performa model menjadi lebih representatif [15].

Tahap berikutnya adalah *Exploratory Data Analysis (EDA)*, yang berfungsi memahami karakteristik dasar dataset sebelum memasuki tahap pemodelan. Berdasarkan diagram, EDA

mencakup tiga analisis utama: distribusi kelas, inspeksi visual citra, dan distribusi piksel mentah. Analisis distribusi kelas dilakukan untuk mengidentifikasi tingkat ketidakseimbangan, mengingat ketimpangan jutaan sampel dapat memengaruhi sensitivitas model terhadap kelas minoritas [3], [18]. Inspeksi visual dilakukan untuk memahami variasi visual dalam citra misalnya perbedaan pencahayaan, ukuran objek, atau komposisi gambar yang dapat memengaruhi proses ekstraksi fitur CNN [9]. Selain itu, analisis distribusi intensitas piksel memberikan informasi mengenai kebutuhan normalisasi, karena nilai piksel yang terlalu bervariasi dapat mempengaruhi stabilitas pelatihan [13].

Tahap preprocessing pada diagram terbagi menjadi dua jalur utama, yaitu jalur augmentasi citra dan jalur SMOTE berbasis fitur. Pada jalur augmentasi, citra dimodifikasi melalui transformasi visual seperti rotasi, flipping, translasi, cropping, dan zooming. Teknik augmentasi seperti ini telah terbukti efektif meningkatkan generalisasi model dengan menciptakan variasi baru tanpa mengubah label kelas [4], [10]. Pada jalur kedua, representasi fitur diekstraksi menggunakan model CNN dan kemudian diseimbangkan menggunakan SMOTE. SMOTE dipilih karena mampu meningkatkan proporsi kelas minoritas melalui interpolasi fitur dalam ruang representasi laten, meskipun beberapa penelitian menunjukkan bahwa penggunaan SMOTE perlu dilakukan dengan hati-hati untuk menghindari overlapping antar kelas [5], [6]. Kedua jalur ini kemudian digunakan untuk melatih dua model berbeda: CNN pada data augmentasi dan MLP pada data fitur yang telah diseimbangkan.

Proses pemodelan dilakukan sesuai dengan jalur preprocessing. Pada jalur pertama, model CNN dengan arsitektur ringan seperti MobileNetV2 dilatih menggunakan citra hasil augmentasi. CNN dipilih karena mampu mengekstraksi pola visual secara hierarkis dan telah terbukti efektif dalam tugas klasifikasi citra biner [19]. Pada jalur kedua, model MLP dilatih menggunakan fitur yang telah diproses SMOTE. Pendekatan *two-stage learning* seperti ini sering digunakan ketika fitur abstrak dinilai memberikan representasi yang lebih stabil pada dataset tidak seimbang [7], [8].

Tahap akhir penelitian adalah evaluasi performa. Evaluasi dilakukan melalui analisis kurva pembelajaran, confusion matrix, serta perbandingan metrik performa. Kurva pembelajaran digunakan untuk menilai kestabilan pelatihan dan mendeteksi adanya *overfitting* atau *underfitting* [6]. Confusion matrix digunakan sebagai alat evaluasi utama karena mampu memberikan informasi detail terkait kesalahan prediksi model, terutama pada kelas minoritas [14].

Selanjutnya, model dibandingkan menggunakan akurasi, presisi, recall, dan F1-score empat metrik yang secara konsisten direkomendasikan dalam evaluasi dataset tidak seimbang [15]. Melalui pendekatan ini, penelitian dapat memberikan penilaian komprehensif terhadap kinerja kedua metode penanganan ketidakseimbangan kelas.

RESULTS AND DISCUSSION

Bagian ini menyajikan hasil penelitian yang telah diperoleh dari rangkaian proses dalam alur desain penelitian, kemudian diikuti dengan pembahasan yang mengintegrasikan temuan empiris dengan pemaknaan konseptual. Tahapan penelitian dimulai dari persiapan dataset, analisis eksploratif, pemrosesan melalui dua jalur teknik penanganan ketidakseimbangan data, pelatihan model, hingga evaluasi performa. Seluruh temuan dianalisis secara terpadu untuk menentukan pendekatan terbaik dalam menangani ketidakseimbangan kelas pada klasifikasi citra biner keberadaan objek.

Tahap awal menunjukkan bahwa dataset berhasil melalui proses pembersihan dan pelabelan dengan baik, sehingga seluruh citra dapat digunakan tanpa kendala struktural. Proses pelabelan memastikan setiap citra berada dalam salah satu dari dua kategori: kelas positif atau negatif. Setelah itu, dataset dibagi dengan rasio 80% data latih dan 20% data uji guna menjaga representativitas distribusi kelas. Ketidakseimbangan kelas yang teridentifikasi dalam analisis eksploratif menunjukkan bahwa kelas negatif jauh lebih dominan dibanding kelas positif, sehingga teknik penanganan ketidakseimbangan diperlukan untuk mencegah model hanya mempelajari pola kelas mayoritas.

```

TAHAP 1: DATASET (LOADING, CLEANING & MAPPING)
=====
[Proses] Membaca CSV Annotasi...
> Ditemukan 355 gambar dengan label 'Ada Objek' di CSV.
[Proses] Memuat dan Membersihkan 1001 file gambar...
100% ██████████ 1001/1001 [00:09<00:00, 209.01It/s]
> Selesai. 1001 Valid, 0 Skipped/Corrupt.

[BUKTI] Sampel Pemetaan Label (Filename -> Label)
  Nama File  Label (0/1)
vid_4_1000.jpg  1
vid_4_10500.jpg 1
vid_4_1020.jpg  1
vid_4_10720.jpg 0
vid_4_10040.jpg 1
vid_4_7760.jpg  0
vid_4_8720.jpg  1
vid_4_6520.jpg  1
vid_4_9020.jpg  1
vid_4_10480.jpg 1
    
```

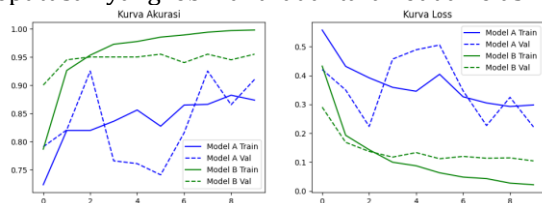
Gambar 2. Dataset

Selain analisis distribusi kelas, eksplorasi awal mencakup pemeriksaan variasi visual citra seperti pencahayaan, ukuran objek, serta latar belakang. Analisis distribusi piksel juga menunjukkan perlunya normalisasi agar proses pembelajaran lebih stabil. Hasil EDA ini memperkuat alasan

mengapa dua pendekatan berbeda augmentasi citra dan SMOTE dipilih untuk dibandingkan secara empiris.

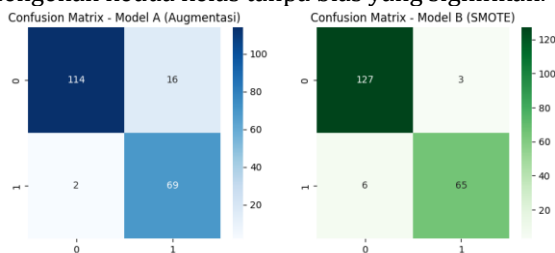
Pada proses pelatihan, Model A (CNN dengan augmentasi citra) menunjukkan peningkatan akurasi pelatihan dari epoch ke epoch. Namun, akurasi validasi terlihat fluktuatif dan *validation loss* cenderung tidak stabil. Hal ini mengindikasikan bahwa meskipun augmentasi berhasil menambah variasi visual, pendekatan ini tidak mampu sepenuhnya mengatasi bias akibat ketidakseimbangan kelas. Model cenderung mempelajari pola yang berlebihan dari kelas mayoritas, sehingga memiliki risiko *overfitting* terhadap data latih.

Sebaliknya, Model B (MLP dengan fitur hasil ekstraksi yang diseimbangkan SMOTE) menunjukkan kurva pelatihan yang jauh lebih stabil. Sejak awal pelatihan, akurasi validasi berada pada angka tinggi dan cenderung konsisten hingga epoch terakhir. Proses oversampling SMOTE yang bekerja pada ruang fitur memungkinkan representasi kelas minoritas diperluas secara seimbang, sehingga model dapat membentuk batas keputusan yang lebih akurat antara kedua kelas.



Gambar 3. Kurva Akurasi dan Loss

Ketika kedua model dibandingkan, terlihat bahwa Model A cenderung menghasilkan *false negative* yang tinggi pada kelas positif meskipun memiliki nilai recall tinggi. Ini menunjukkan bahwa model mendeteksi kelas positif secara agresif tetapi sering membuat prediksi positif yang keliru. Sementara itu, Model B menunjukkan profil kesalahan yang lebih seimbang, menekan *false negative* dan *false positive* dalam jumlah lebih proporsional. Perbandingan confusion matrix keduanya memperlihatkan bahwa Model B lebih andal dalam mengenali kedua kelas tanpa bias yang signifikan.



Gambar 4. Hasil Confusion Matrix

Perbandingan metrik performa lebih lanjut memperlihatkan bahwa Model B unggul pada

akurasi, presisi, dan F1-score, sementara Model A hanya unggul pada recall. Hasil ini menunjukkan bahwa augmentasi citra mampu meningkatkan sensitivitas model terhadap kelas positif, tetapi tidak cukup baik dalam menghasilkan prediksi yang tepat atau konsisten. Model B, dengan kombinasi fitur ekstraksi dan SMOTE, memberikan performa keseluruhan yang lebih stabil dan seimbang. Hal ini menjadikan pendekatan pada ruang fitur lebih efektif dibanding augmentasi pada tingkat citra mentah.

[TABEL PERBANDINGAN PERFORMA]

| Metrik | Model A (Augmentasi) | Model B (SMOTE) |
|----------|----------------------|-----------------|
| Akurasi | 0.910448 | 0.955224 |
| Presisi | 0.811765 | 0.955882 |
| Recall | 0.971831 | 0.915493 |
| F1-Score | 0.884615 | 0.935252 |

KESIMPULAN SISTEM: Berdasarkan F1-Score, metode terbaik adalah Model B (SMOTE).

Gambar 5. Hasil Tabel Perbandingan Performa

Pembahasan hasil ini mengarah pada pemahaman bahwa augmentasi citra memang bermanfaat dalam meningkatkan keragaman data, namun tidak mengatasi akar masalah ketidakseimbangan jumlah kelas. Ketidakseimbangan tetap memberi tekanan pada model sehingga performanya tidak stabil, terutama pada metrik yang sensitif terhadap distribusi kelas. Sebaliknya, SMOTE secara langsung memengaruhi distribusi kelas dengan menambah representasi kelas minoritas pada ruang fitur, sehingga pembelajaran model menjadi lebih seimbang dan generalisasi meningkat.

Secara keseluruhan, hasil menunjukkan bahwa pipeline Feature Extraction → SMOTE → MLP lebih unggul untuk dataset dengan ketidakseimbangan tinggi. Pendekatan ini memberikan performa yang stabil, meningkatkan sensitivitas terhadap kelas minoritas, sekaligus mempertahankan ketepatan prediksi. Kombinasi tersebut jauh lebih efektif dibanding pendekatan CNN + Augmentasi, yang meskipun mampu memperkaya variasi visual, tetap gagal mengatasi ketimpangan distribusi kelas yang mendasar.

CONCLUSION

Penelitian ini bertujuan mengevaluasi efektivitas dua pendekatan penanganan ketidakseimbangan data pada klasifikasi citra biner, yaitu augmentasi citra pada level visual dan SMOTE pada ruang fitur. Berdasarkan rangkaian proses mulai dari persiapan dataset, analisis eksploratif, pemrosesan dua jalur, pelatihan model, hingga evaluasi performa, diperoleh beberapa temuan penting.

Pertama, hasil analisis menunjukkan bahwa augmentasi citra mampu meningkatkan variasi visual pada data latih, tetapi tidak memberikan perubahan berarti terhadap distribusi jumlah sampel antar kelas. Hal ini menyebabkan model berbasis CNN tetap rentan terhadap bias kelas

mayoritas, sehingga performanya tampak fluktuatif pada data uji dan tidak mampu mempertahankan stabilitas prediksi. Meskipun model augmentasi menunjukkan nilai recall yang tinggi, presisi dan F1-score yang rendah menandakan bahwa model sering menghasilkan prediksi positif yang keliru dan tidak memiliki kemampuan generalisasi yang kuat.

Sebaliknya, pendekatan berbasis ekstraksi fitur yang diikuti oleh SMOTE menunjukkan performa yang lebih konsisten di seluruh tahapan evaluasi. Penyisipan sampel sintesis pada ruang fitur terbukti lebih efektif dalam memperbaiki representasi kelas minoritas tanpa menimbulkan noise visual. Model MLP yang dilatih pada fitur hasil SMOTE mampu mencapai akurasi, presisi, dan F1-score lebih tinggi dibanding model augmentasi, serta mempertahankan *learning curve* yang lebih stabil. Temuan ini mengindikasikan bahwa penanganan ketidakseimbangan pada tingkat representasi fitur lebih berdampak signifikan dibanding manipulasi pada level visual citra.

Secara keseluruhan, penelitian ini menyimpulkan bahwa pendekatan Feature Extraction → SMOTE → MLP merupakan strategi yang lebih optimal untuk menangani ketidakseimbangan data pada klasifikasi citra biner. Pendekatan ini tidak hanya menghasilkan performa prediksi yang lebih seimbang, tetapi juga memberikan stabilitas model yang lebih baik. Dengan demikian, penelitian ini memberikan kontribusi dalam menunjukkan bahwa penyeimbangan kelas pada ruang fitur merupakan solusi yang lebih efektif dibandingkan augmentasi visual semata, terutama pada dataset yang memiliki ketidakseimbangan kelas yang tinggi.

REFERENCE

- [1] N. Kaur and T. K. Gandhi, "A systematic review of class imbalance solutions in deep learning for image analysis," *Image Vis. Comput.*, vol. 138, p. 104752, 2023, doi: 10.1016/j.imavis.2023.104752.
- [2] W. Chen, K. Yang, Z. Yu, Y. Shi, and C. L. P. Chen, "A survey on imbalanced learning: latest research, applications and future directions," *Artif. Intell. Rev.*, 2024, doi: 10.1007/s10462-024-10759-6.
- [3] M. Hasan, T. Chowdhury, and G. Prince, "Challenges of object detection under severe class imbalance: A systematic analysis," *Image Vis. Comput.*, vol. 139, p. 104912, 2024, doi: 10.1016/j.imavis.2024.104912.
- [4] H. Wang, C. Sui, F. Jiang, S. Li, H. Liu, and A. Wang, "Value-guided adaptive data augmentation for imbalanced small object detection," *Electronics*, vol. 13, no. 10, p. 1849, 2024, doi: 10.3390/electronics13101849.
- [5] D. Elreedy and A. Atiya, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Mach. Learn.*, 2024, doi: 10.1007/s10994-022-06296-4.
- [6] S. Gholampour, "Impact of nature of medical data on machine and deep learning for imbalanced datasets: Clinical validity of SMOTE is questionable," *Mach. Learn. Knowl. Extr.*, vol. 6, no. 2, pp. 827–841, 2024, doi: 10.3390/make6020039.
- [7] S. Borah, S. Das, and S. Roy, "A novel data-level balancing strategy for minority object recognition in imbalanced datasets," *Expert Syst. Appl.*, vol. 233, p. 120846, 2023, doi: 10.1016/j.eswa.2023.120846.
- [8] E. Ersözlü, H. Taheri, and M. Koch, "Improved synthetic oversampling for deep feature-based classification on imbalanced datasets," *Sci. Rep.*, vol. 14, pp. 1124–1139, 2024.
- [9] Z. Wu and H. Zeng, "Image data augmentation techniques based on deep learning: A survey," *Math. Biosci. Eng.*, vol. 21, no. 6, pp. 6190–6224, 2024, doi: 10.3934/mbe.2024272.
- [10] H. Kaur, H. S. Pannu, and A. K. Malhi, "A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–36, 2022, doi: 10.1145/3343440.
- [11] R. Gupta and A. Singh, "Analyzing annotation quality and its impact on object detection and classification performance," *Inf. Sci. (Ny)*, vol. 660, p. 120097, 2024, doi: 10.1016/j.ins.2024.120097.
- [12] Y. Li, Q. Wang, and T. Zhang, "Revisiting object detection datasets for binary image classification through bounding-box guided supervision," *Pattern Recognit.*, vol. 140, p. 109573, 2023, doi: 10.1016/j.patcog.2023.109573.
- [13] J. Zhou, Y. He, and L. Huang, "A systematic evaluation of bounding-box transformations for downstream binary classification tasks," *Expert Syst. Appl.*, vol. 246, p. 123302, 2025, doi: 10.1016/j.eswa.2024.123302.
- [14] P. Thölke, G. Valente, B. Kießling, and T. Rohlfing, "Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain decoding with imbalanced data," *Neuroimage*, 2023, doi: 10.1016/j.neuroimage.2023.119844.

- 10.1016/j.neuroimage.2023.120253.
- [15] Y. Sun, Z. Zhang, and Z. Zhou, "A comprehensive evaluation framework for imbalanced data learning," *Pattern Recognit.*, vol. 141, p. 109537, 2023, doi: 10.1016/j.patcog.2023.109537.
- [16] M. Ahmad, F. Khan, and M. Alsaqer, "Explainable artificial intelligence for healthcare: A review," *Expert Syst. Appl.*, 2022, doi: 10.1016/j.eswa.2022.116952.
- [17] Z. Chen, H. Liu, and Y. Xu, "A unified representation of bounding box annotation for object-centric learning," *Neurocomputing*, vol. 509, pp. 150–162, 2022, doi: 10.1016/j.neucom.2022.07.089.
- [18] S. Bhattacharya, S. Gupta, and P. Mitra, "Addressing extreme class imbalance in image classification using adaptive hybrid sampling," *Pattern Recognit.*, vol. 154, p. 110723, 2024, doi: 10.1016/j.patcog.2024.110723.
- [19] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, "A survey of modern deep learning based object detection models," *Digit. Signal Process.*, vol. 126, p. 103514, 2022, doi: 10.1016/j.dsp.2022.103514.