

Dampak *Upsampling* dan *Downsampling* dalam Deteksi Keadaan Darurat dari Citra Udara Menggunakan RetinaNet

Avia Maharani Situmorang^{*1}, Fetty Tri Anggraeny², Achmad Junaidi³

^{1,2,3} Fakultas Ilmu Komputer, Universitas Pembangunan Nasional 'Veteran' Jawa Timur

Email: ¹19081010098@student.upnjatim.ac.id, ²fettyanggraeny.iff@upnjatim.ac.id, ³achmadjunaidi.iff@upnjatim.ac.id

^{*}Penulis Korespondensi

Abstrak

Ketidakeimbangan jumlah data pada dataset citra udara dapat mempengaruhi performa model deteksi citra udara, di mana hasil identifikasi terhadap keadaan darurat menjadi tidak optimal. Oleh karena itu, penelitian ini dilakukan untuk menganalisa teknik penyeimbang data, yaitu *upsampling* dan *downsampling* pada model RetinaNet yang menggunakan *backbone* ResNet50. Model ini dilatih menggunakan *Aerial Images Dataset for Emergency Response* (AIDER) untuk mengidentifikasi keadaan darurat, di mana *upsampling* menambahkan jumlah data pada kelas minoritas menggunakan augmentasi yaitu *gaussian blur*, *gaussian noise*, dan *contrast*, sedangkan *downsampling* mengurangi jumlah data pada kelas mayoritas. Model dilatih dengan parameter *optimizer* Adam dengan *learning rate* $2e-5$, *epoch* 5 dan 10, *batch size* 4, dan *GPU Tesla T4* yang tersedia pada *Kaggle*. Hasil latih dievaluasi menggunakan *mean Average Precision* (mAP), *class loss*, dan *regression loss*. Berdasarkan hasil penelitian, nilai mAP data *upsampling* memiliki hasil paling optimal yaitu 0.829917, sedangkan data *downsampling* memiliki hasil mAP paling rendah yaitu 0.801813. Selain itu, data *upsampling* memiliki nilai *class loss* terendah, sedangkan data *downsampling* memiliki *regression loss* yang lebih rendah.

Kata kunci: AIDER, Deteksi Objek, Downsampling dan Upsampling, Keadaan Darurat, RetinaNet.

Abstract

Class imbalance in the aerial images dataset can affect the performance of the aerial object detection model, leading to suboptimal emergency identification results. Therefore, this study was conducted to analyze data balancing techniques, namely upsampling and downsampling, in the RetinaNet model using the ResNet50 backbone. This model was trained using the Aerial Images Dataset for Emergency Response (AIDER) to identify emergencies, with upsampling increasing the amount of data in the minority class with extra augmentation of gaussian blur, gaussian noise, and contrast, while downsampling reducing it in the majority class. The model was trained using the Adam optimizer with a learning rate of $2e-5$ for 5 and 10 epochs, a batch size of 4, and a Tesla T4 GPU available on Kaggle. The training results were evaluated using mean Average Precision (mAP), class loss, and regression loss. Based on the study results, the upsampled data achieves the optimal mAP, 0.829917, while the downsampled data achieves the lowest mAP, 0.801813. In addition, upsampled data has lower class loss value, and downsampled data has lower regression loss value.

Keywords: AIDER, Downsampling and Upsampling, Emergency Response, Object Detection, RetinaNet.

I. PENDAHULUAN

Keadaan darurat memiliki berbagai macam bentuk, yakni bencana alam, seperti banjir dan kebakaran hutan, serta bencana yang disebabkan manusia, seperti bangunan roboh dan kecelakaan. Untuk meminimalisir efek dari bencana serta membantu yang terdampak, manajemen keadaan darurat yang efektif sangat diperlukan [1]. Salah satu langkah yang krusial adalah mengidentifikasi keadaan darurat itu sendiri. Untuk melakukan identifikasi, terdapat beberapa cara yang dapat dilakukan, salah satunya adalah penggunaan *unmanned aerial device* (UAV), atau yang lebih dikenal sebagai *drone*.

Penggunaan drone telah berkembang pesat dalam berbagai bidang pada beberapa tahun terakhir. Salah satunya dalam manajemen keadaan darurat dan bantuan kemanusiaan [2]. Perkembangan terkini, di mana drone diintegrasikan dengan sensor kamera, memberi peluang baru dalam manajemen keadaan darurat dengan lebih efisien. Namun, salah satu keterbatasan dalam penggunaan drone untuk keadaan darurat adalah menentukan algoritma yang dapat digunakan untuk menjalankan drone secara otomatis dan menghasilkan data *real-time* dengan efisien. Dalam hal ini, penggunaan dataset *Aerial Image Dataset for Emergency Response* (AIDER) menjadi alat bantu untuk menguji model terbaik dalam deteksi objek citra udara. Penggunaan algoritma *deep learning*, seperti *Convolutional Neural Networks* (CNN) dianggap memiliki probabilitas terbaik dalam pemrosesan data citra digital yang dihasilkan drone, sehingga dianggap

lebih cepat dalam perkiraan suasana dan meminimalisir risiko saat dilakukannya tanggap darurat. Salah satu bentuk CNN adalah RetinaNet.

RetinaNet merupakan salah satu model dari CNN. Pertama kali dikenalkan pada tahun 2017, model ini merupakan sebuah model tunggal dengan kumpulan jaringan yang berisikan sebuah jaringan *backbone* dan dua subnet yang memiliki tugas masing-masing, yaitu klasifikasi dan regresi. RetinaNet memiliki keunggulan dalam melakukan klasifikasi objek yang berskala kecil dan padat, juga mengatasi ketidakseimbangan data, hal inilah yang membuat RetinaNet lebih banyak digunakan sebagai model deteksi objek pada citra udara atau satelit [3]. Meskipun demikian, ketidakseimbangan data tetap saja terjadi, terutama jika distribusi jumlah objek yang digunakan tidak merata. Oleh karena itu, dibutuhkan sebuah teknik penyeimbangan data yang dapat membantu dalam meningkatkan kinerja dari model. Salah satu teknik yang banyak digunakan yaitu *upsampling* dan *downsampling* [4].

Upsampling adalah metode penambahan data pada kelas minoritas, sedangkan *downsampling* adalah metode pengurangan data pada kelas mayoritas. Kedua metode dilakukan untuk menyetarakan jumlah data antar kelas, sehingga pelatihan terhadap model dapat lebih seimbang. Terdapat beberapa penelitian awal yang menggunakan metode ini untuk memecahkan masalah ketidakseimbangan kelas [5]. Namun dalam RetinaNet, belum banyak ditemukan penggunaan teknik ini untuk mendukung kinerja dari model. Oleh karena itu, penelitian ini bertujuan untuk menganalisa dampak dari penggunaan teknik *upsampling* dan *downsampling* terhadap pelatihan RetinaNet pada dataset citra udara, yaitu AIDER. Performa dari model ini akan dievaluasi menggunakan metrik *mean Average Precision* (mAP), *class loss*, dan *regression loss* untuk menentukan mana yang menghasilkan performa terbaik.

II. METODE PENELITIAN

Dalam penelitian ini, dilakukan analisa terhadap pengaruh teknik *upsampling* dan *downsampling* terhadap kinerja model RetinaNet pada proses deteksi objek citra udara menggunakan dataset AIDER.

2.1. RetinaNet

RetinaNet merupakan salah satu arsitektur dengan model one-stage object detector, model ini bekerja dengan cara melewati langkah pemilihan regional dari objek yang akan diklasifikasi dengan algoritma tersebut. *Model one-stage object detector* memiliki tingkat kecepatan lebih tinggi daripada two-stage object detector, namun tingkat akurasi lebih rendah. Hal ini dikarenakan model cenderung memiliki ketidakseimbangan dalam class-nya. Untuk mengatasi masalah tersebut, RetinaNet menggunakan *focal loss*, sehingga dapat meningkatkan tingkat akurasi. Berbeda dengan model one-stage lainnya, RetinaNet menggunakan sistem subnetwork, serta menggabungkan antara penggunaan dari Resnet dan *Feature Pyramid Network* (FPN), sehingga dapat meminimalisir waktu pengerjaan, sekaligus meningkatkan nilai akurasi dari hasil.

$$FL(pt) = -(1 - pt)^\gamma \log(pt) \quad [1]$$

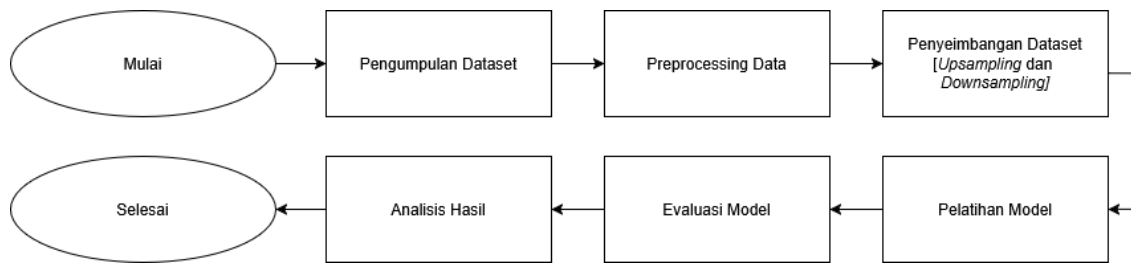
Nilai focal loss divisualisasikan sebagai sebuah nilai dari $\gamma \in [0,5]$. Saat sebuah objek mengalami misklasifikasi dan nilai pt (probabilitas estimasi dari kelas) kecil, faktor modulasi hampir 1, dan nilai loss tidak terafeksi, maka dengan $pt \rightarrow 1$, faktor akan berubah menjadi nilai nol dan nilai loss untuk objek yang diklasifikasikan dengan baik akan dikurangi bebannya. Saat $\gamma = 0$, focal loss akan sebanding dengan cross-entropy loss, dan saat γ ditingkatkan, maka efek dari faktor modulasi akan meningkat juga. Faktor modulasi nantinya akan mengurangi kontribusi dari nilai loss milik objek yang mudah diklasifikasi, dan meningkatkan jangkauan dari objek yang memiliki nilai loss rendah.

2.2. ResNet50

ResNet50 merupakan salah satu jenis *backbone* yang banyak digunakan pada model RetinaNet. ResNet adalah salah satu model CNN yang efektif dalam klasifikasi citra [6]. Dikembangkan pada tahun 2015 oleh *Microsoft Research*, ResNet50 melakukan deteksi objek dengan blok residual. ResNet50 merupakan salah satu bentuk dari ResNet, di mana ResNet50 memiliki 50 *neural networks layer*.

2.3. Metode Penelitian

Penelitian ini dilakukan melalui beberapa tahapan. Di mana alur proses penelitian ditunjukkan pada gambar 1 di bawah ini.



Gambar 1. Flowchart Penelitian

Berdasarkan bagan alur tersebut, langkah pertama yang harus dilakukan adalah pengumpulan dataset. Setelah itu dilakukan preprocessing data, di mana dilakukan augmentasi untuk meningkatkan jumlah data di setiap kelas. Langkah selanjutnya adalah penyeimbangan dataset, yaitu *upsampling* dan *downsampling*. Dataset asli, hasil *upsampling*, dan hasil *downsampling* kemudian dijadikan data latih pada model, yang kemudian dilanjutkan pada evaluasi model. Kedua hal ini dilakukan menggunakan *Kaggle* yang mendukung penggunaan CUDA dan GPU sehingga mempercepat pemrosesan.

1. Preprocessing Data

Sebelum diproses, dilakukan anotasi terhadap data yang akan digunakan. Anotasi merupakan pemberian *bouding box* pada setiap data untuk memberikan label pada target kelas di setiap objek. Anotasi dilakukan menggunakan *LabelImg*. Setelah itu, dilakukan augmentasi secara global untuk meningkatkan jumlah data asli. Pada tabel 1, tertulis jumlah data uji setelah dilakukan augmentasi.

Tabel 1. Distribusi data asli

Kelas	Jumlah Data Asli
Collapsed_Building	1183
Traffic_Incident	1117
Fire	1082
Flooded_Areas	1037

2. Penyeimbangan Dataset

Untuk melakukan metode *upsampling* dan *downsampling*, dilakukan dua metode yang berbeda. Pada *upsampling*, dilakukan peningkatan jumlah data pada kelas minoritas, dengan menambahkan augmentasi pada data hingga jumlah data menjadi sama dengan kelas mayoritas. Pada *downsampling*, dilakukan pengurangan jumlah data pada masing-masing kelas mayoritas sehingga semua kelas memiliki data seimbang. Pengurangan data di setiap kelas ini dilakukan secara acak.

3. Pelatihan Model

Pelatihan model dilakukan dengan mengimplementasikan RetinaNet dengan *backbone* ResNet-50. Menggunakan parameter berupa *optimizer* Adam dengan learning rate $2e-5$, epoch 5 dan 10, batch size 4, dan GPU Tesla T4 yang tersedia pada *Kaggle*.

4. Evaluasi Model

Evaluasi model dilakukan dengan menganalisa hasil berdasarkan nilai *class loss* yaitu *focal loss*, *regression loss* yaitu *Smooth L1 loss*, dan mAP.

III. HASIL DAN PEMBAHASAN

3.1. Distribusi Dataset

Berdasarkan tabel 2, dapat dilihat ketidakseimbangan pada jumlah data latih asli. Data uji untuk *collapsed_building* memiliki jumlah data yang paling banyak, sedangkan *flooded_areas* memiliki jumlah data paling sedikit. Perbedaan jumlah data ini dapat menimbulkan ketidakseimbangan kelas dan menghasilkan nilai yang kurang optimal.

Tabel 2. Distribusi data asli, *upsampling*, dan *downsampling*

Kelas	Jumlah Data Asli	Jumlah Data <i>Upsampling</i>	Jumlah Data <i>Downsampling</i>
Collapsed_Building	1183	1183	1037
Traffic_Incident	1117	1117	1037
Fire	1082	1082	1037
Flooded_Areas	1037	1183	1037

Untuk mengatasi hal tersebut, diaplikasikan teknik *upsampling* dan *downsampling*. Di mana pada *upsampling*, kelas dengan data minoritas, yaitu *flooded_areas* akan ditambahkan dengan secara acak melakukan augmentasi pada beberapa data. Augmentasi yang digunakan adalah *gaussian blur*, *gaussian noise*, dan *contrast*. Diterapkannya augmentasi ini guna menambahkan variasi data, sehingga model dapat lebih mengenali objek pada berbagai macam kondisi. Dengan dilakukannya *upsampling*, jumlah data *flooded_areas* akan setara dengan *collapsed_building*. Sedangkan pada *downsampling*, kelas dengan data mayoritas secara acak dilakukan pengurangan jumlah data. Hal ini untuk memastikan jumlah data setiap kelas setara.

3.1. Hasil *mean Average Precision* (mAP)

Mean Average Precision (mAP) merupakan salah satu metrik evaluasi yang kerap kali digunakan dalam deteksi objek [7]. Hal ini dikarenakan mAP memberikan evaluasi secara menyeluruh terhadap akurasi pada deteksi objek di setiap kelas. Perhitungan mAP didasarkan pada *precision* dan *recall*. Kedua variabel didasarkan pada nilai dari *Intersection over Union* (IoU). Nilai IoU ditentukan berdasarkan rasio area yang tumpang tindih antara *ground truth* dan *bounding box*. Untuk menentukan ini, ditentukan sebuah nilai *threshold*, yaitu ambang batas, di mana sebuah deteksi dianggap sesuai jika masih pada nilai ambang batas tersebut. Pada penelitian ini, nilai ambang batas yang digunakan adalah 0.45.

Terdapat tiga nilai yang menjadi bagian dalam perhitungan IoU, yaitu *True Positive* (TP), *False Positive* (FP), dan *False Negative* (FN). TP adalah saat di mana nilai IoU tepat pada ambang batas yang telah ditentukan, FP adalah saat nilai IoU berada di bawah ambang batas, sedangkan FN adalah keadaan di mana model tidak dapat mendeteksi objek yang terdapat pada *ground truth* [8].

Ketiga nilai tersebut kemudian digunakan untuk menentukan nilai *precision* dan *recall*. *Precision* merupakan nilai yang merepresentasikan jumlah objek yang terprediksi dengan benar dari keseluruhan prediksi, sedangkan *recall* merepresentasikan proporsi objek yang terprediksi dengan benar dari semua objek yang ada pada dataset.

$$\text{Precision: } \frac{TP}{TP+FP} \quad [2]$$

$$\text{Recall: } \frac{TP}{TP+FN} \quad [3]$$

Nilai rata-rata dari tiap kelas yang didasarkan pada kurva *precision-recall* disebut sebagai *Average Precision* (AP). Menentukan nilai rata-rata dari keseluruhan AP tiap kelas merupakan hasil evaluasi akhir yang kita sebut sebagai mAP. Oleh karena itu, penggunaan mAP berguna sebagai satu metrik yang secara menyeluruh mengevaluasi performa dari sebuah model deteksi objek.

Tabel 3. Hasil mAP,

Metode	mAP (Epoch: 5)	mAP (Epoch: 10)
Data Asli	0.693382	0.824750
Upsampling	0.697165	0.829917
Downsampling	0.668942	0.801831

Berdasarkan hasil mAP pada tabel 3, dapat dipastikan penggunaan *upsampling* dan *downsampling* berpengaruh pada kinerja model. Pada tabel tersebut, nilai mAP tertinggi didapatkan pada epoch 10 dengan

nilai 0.829917 dengan data *upsampling*, sedangkan nilai terendah didapatkan pada epoch 5 dengan nilai 0.668942 dengan data *downsampling*. Secara keseluruhan, setiap mAP yang dihasilkan pada epoch 10 memiliki hasil yang lebih baik daripada mAP yang dihasilkan pada epoch 5. Hal ini menunjukkan penambahan epoch dapat membantu meningkatkan kinerja dari model dikarenakan semakin banyak pelatihan yang dilakukan, potensi akurasi hasil deteksi objek oleh model semakin tinggi. Hal ini dikarenakan, pada epoch 5, model belum mempelajari karakteristik objek dengan baik, sehingga deteksi belum stabil, sedangkan pada epoch 10, karakteristik objek lebih detail dan terperinci, sehingga nilai mAP lebih stabil dan meningkat.

Pada data *upsampling*, rata-rata hasil mAP lebih tinggi daripada data lainnya. Hal ini terjadi dikarenakan data *upsampling* meningkatkan jumlah data latih tanpa mengurangi data sebelumnya, sehingga lebih banyak data yang dilatih, model dapat lebih mengenali karakteristik objek terutama dengan tambahan augmentasi, yang mengenalkan objek dalam berbagai macam variasi. Sedangkan pada data *downsampling*, rata-rata hasil mAP lebih rendah daripada data lainnya. Hal ini dikarenakan pada *downsampling*, beberapa data justru dihapus dari data awal, hal ini mengurangi jumlah data latih, sehingga mengurangi kemampuan pengenalan karakteristik objek oleh model. Terutama jika perbedaan jumlah data pada kelas mayoritas dan minoritas tidak signifikan. Oleh karena itu, pada model RetinaNet dengan ResNet50 untuk deteksi citra udara ini, penggunaan *upsampling* memiliki hasil yang lebih signifikan daripada penggunaan *downsampling*.

3.2. Regression Loss dan Class Loss

Regression loss dan *class loss* menjadi dua variabel penentu tingkat keberhasilan sebuah model deteksi objek. Pada RetinaNet, *class loss* yang digunakan adalah hasil dari perhitungan *focal loss*, sedangkan *regression loss* yang digunakan adalah *Smooth L1 Loss*. *Smooth L1 Loss* adalah salah satu jenis *regression loss* yang menggabungkan kelebihan dari *L1 loss* dan *L2 loss* melalui sebuah parameter ambang batas. *Smooth L1 Loss* didefinisikan sebagaimana pada formula di bawah ini.

$$\text{SmoothL1}(\epsilon) = \begin{cases} \frac{\beta^2}{2} \epsilon^2, & \text{jika } \epsilon < \frac{1}{\beta^2} \\ |\epsilon| - \frac{1}{2\beta^2}, & \text{lainnya} \end{cases} \quad [4]$$

Di mana:

$$\epsilon = \|y - fw(x)\| \quad [5]$$

ϵ merepresentasikan prediksi yang error, yaitu perbedaan antara *bounding box* yang telah diprediksi, dengan *bounding box* pada *ground truth*. y merupakan kordinat lokasi dari *bounding box* pada *ground truth*. $fw(x)$ adalah *bounding box* yang berhasil diprediksi oleh RetinaNet, dan β adalah ambang batas yang menentukan bagaimana fungsi *loss* bekerja.

Formula ini bekerja pada dua kondisi. *Smooth L1 Loss* bekerja menyerupai *L2 loss* di saat prediksi yang error berukuran kecil dengan menggunakan fungsi kuadrat, sehingga model dapat melakukan optimasi yang lebih ringan dan penyesuaian ringan dalam menentukan lokasi objek. Kemudian, saat prediksi yang error berukuran besar, *Smooth L1 Loss* akan berubah dan bekerja menyerupai *L1 loss* di mana fungsi ini lebih tahan terhadap *outlier* karena memiliki gradien yang konstan, sehingga mencegah adanya *loss* besar pada bagian awal pelatihan.

Atas dasar sistem kerja tersebut, *Smooth L1 Loss* lebih tanggap dalam menangani prediksi error yang ekstrim, dibandingkan dengan *regression loss* lain seperti *Mean Squared Error* (MSE). Pada model deteksi objek, terutama dengan citra udara seperti dataset AIDER, kemungkinan terjadinya ketidakstabilan dalam memprediksi sebuah objek cukup besar, dikarenakan data citra udara umumnya memiliki sudut pandang atau latar yang kompleks. Dengan menggunakan *Smooth L1 Loss*, ketidakstabilan ini dapat dikurangi dengan meminimalisir efek dari error yang besar, yang menghasilkan gradien yang lebih stabil dan konvergen yang lebih baik selama pelatihan.

Tabel 4. *Regression loss* dan *class loss*

Metode	Reg Loss	Class Loss
Data Asli	0.22973	0.06147
<i>Upsampling</i>	0.24940	0.03475

Downsampling	0.22200	0.07281
--------------	---------	---------

Berdasarkan tabel 4, dapat dilihat bagaimana *class loss* pada *upsampling* memiliki tingkat *loss* paling rendah. Hal ini menunjukkan pada model RetinaNet dengan *backbone* ResNet50, penggunaan data *upsampling* mengurangi tingkat *loss* dalam memprediksi kelas selama deteksi objek. Hal ini juga menunjukkan bagaimana penambahan jumlah data meningkatkan kemampuan prediksi karakteristik objek oleh model. Sedangkan untuk *regression loss*, perbedaan nilai tidak terlalu signifikan, namun cenderung lebih rendah pada *downsampling*, hal ini dikarenakan dengan dikurangnya data mayoritas, distribusi data menjadi lebih merata, sehingga pelatihan terhadap setiap kelas juga merata. Oleh karena itu, penentuan lokasi dari *bounding box* menjadi lebih konsisten, terutama jika dataset yang digunakan. Oleh karena itu, *regression loss* pada *downsampling* menjadi lebih rendah.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, implementasi RetinaNet dengan *backbone* Resnet50 dianggap berhasil dalam mendeteksi dan mengklasifikasikan objek keadaan darurat menggunakan citra udara dengan dataset *Aerial Image Dataset for Emergency Response* (AIDER). Selain itu, penggunaan teknik penyeimbangan data, yaitu *upsampling* dan *downsampling* berpengaruh dalam performa dari RetinaNet dalam pelatihan model deteksi objek. Hal ini dapat dilihat dari pemaparan terhadap hasil pengujian menggunakan *mean Average Precision* (mAP), *class loss* yaitu *focal loss*, dan *regression loss* yaitu *Smooth L1 loss* pada epoch 5 dan epoch 10.

Berdasarkan hasil yang telah dipaparkan, metode *upsampling* memiliki nilai mAP tertinggi yaitu 0.829917 pada epoch 10. Selain itu, teknik *upsampling* juga memiliki *class loss* paling rendah dari metode lain. Hal ini menunjukkan penambahan data dengan augmentasi membantu model untuk bekerja lebih stabil dan akurat dikarenakan semakin banyak data yang digunakan, model dapat mempelajari karakteristik objek dengan lebih baik, sehingga model dapat bekerja lebih optimal. Pada teknik *downsampling*, nilai *regression loss* lebih rendah. Hal ini menunjukkan bagaimana prediksi *bounding box* menjadi lebih stabil dalam menentukan lokasi objek. Selain itu, hasil penelitian menunjukkan pelatihan model dengan epoch 10 lebih stabil dan optimal dibandingkan dengan epoch 5. Hal ini didasarkan pada penurunan nilai *class loss* dan *regression loss* yang konsisten, diikuti dengan peningkatan nilai mAP dalam pengujian.

Penelitian ini selanjutnya dapat dikembangkan dengan menggunakan jumlah dataset yang lebih besar dan variatif, diikuti dengan penggunaan augmentasi yang lebih kompleks. Selain itu, dapat dilakukan perbandingan antara RetinaNet dengan model lainnya, seperti YOLO. Penggunaan *backbone* berbeda, seperti ResNet101 juga dapat dilakukan untuk mengetahui pengaruh *backbone* terhadap performa RetinaNet dalam deteksi objek citra udara.

REFERENSI

- [1] C. Kyrkou, P. Kolios, T. Theodorides, and M. Polycarpou, "Machine Learning for Emergency Management: A Survey and Future Outlook," *Proceedings of the IEEE*, vol. 111, no. 1, 2023, doi: 10.1109/JPROC.2022.3223186.
- [2] C. Kyrkou and T. Theodorides, "EmergencyNet: Efficient Aerial Image Classification for Drone-Based Emergency Monitoring Using Atrous Convolutional Feature Fusion," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, 2020, doi: 10.1109/JSTARS.2020.2969809.
- [3] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, 2020, doi: 10.1109/TPAMI.2018.2858826.
- [4] D. Dumitrescu and C. A. Boiangiu, "A study of image upsampling and downsampling filters," *Computers*, vol. 8, no. 2, 2019, doi: 10.3390/computers8020030.
- [5] Y. Ye, M. Zhou, Z. Wang, and X. Shen, "Improved Upsampling Based Depth Image Super-Resolution Reconstruction," *IEEE Access*, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3274966.
- [6] M. Fauzan Novriandy, B. Rahmat, and A. Junaidi, "KLASIFIKASI CITRA PENYAKIT KANKER MULUT MENGGUNAKAN ARSITEKTUR RESNET50 OPTIMASI ADAM DAN SGD," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4732.

-
- [7] A. Gad, "Evaluating Object Detection Models Using Mean Average Precision," *KDnuggets*, 2021.
- [8] A. Hasby Bik, F. Tri Anggraeny, and E. Yulia Puspaningrum, "KLASIFIKASI PENYAKIT GINJAL MENGGUNAKAN ALGORITMA HIBRIDA CNN-ELM," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, 2024, doi: 10.36040/jati.v8i3.9807.