



## **Deteksi Objek Anjing dan Kucing Menggunakan Faster R-CNN**

**Muh Gunawan Hadi<sup>1</sup>, Nugroho Noto Susanto<sup>2</sup>, Reyhan Mahendra<sup>3</sup>, Perani Rosyani<sup>4</sup>**

<sup>1,2,3,4</sup> Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: <sup>1</sup>[hadi.gunawan1705@gmail.com](mailto:hadi.gunawan1705@gmail.com), <sup>2</sup>[reyhanmahendra840@gmail.com](mailto:reyhanmahendra840@gmail.com), <sup>3</sup>[ngrhnts@gmail.com](mailto:ngrhnts@gmail.com),

<sup>4</sup>[dosen00837@unpam.ac.id](mailto:dosen00837@unpam.ac.id)

(\* : coresponding author)

**Abstrak**—Penelitian ini bertujuan untuk mengatasi tantangan dalam deteksi objek hewan peliharaan, khususnya anjing dan kucing, yang memiliki variasi intra-kelas yang tinggi serta kemiripan visual yang signifikan antar ras. Metode yang diusulkan adalah arsitektur Deep Learning Faster R-CNN dengan backbone ResNet50-FPN. Untuk meningkatkan performa model pada dataset Oxford-IIIT Pet yang memiliki keterbatasan jumlah sampel dan ketidakseimbangan kelas, penelitian ini menerapkan strategi modifikasi berupa augmentasi data lanjutan menggunakan pustaka Albumentations dan teknik Gradient Accumulation. Augmentasi data diterapkan untuk memperkaya variasi pose dan kondisi pencahayaan, sedangkan Gradient Accumulation digunakan untuk menstabilkan proses pelatihan pada batch size kecil akibat keterbatasan memori GPU. Hasil eksperimen menunjukkan bahwa model yang dimodifikasi mampu mencapai performa optimal pada Epoch ke-11 dengan skor Mean Average Precision (mAP) sebesar 0.7519. Analisis kurva loss menunjukkan konvergensi yang stabil antara data latih dan validasi, mengindikasikan bahwa teknik augmentasi yang diterapkan efektif dalam mencegah overfitting. Penelitian ini membuktikan bahwa integrasi strategi augmentasi geometris dan akumulasi gradien dapat meningkatkan ketahanan model deteksi objek secara signifikan pada domain hewan peliharaan.

**Kata Kunci:** deteksi objek, Faster R-CNN, anjing dan kucing, albumentations, gradient accumulation, deep learning.

**Abstract**—This study aims to address the challenges in detecting pet objects, specifically dogs and cats, which exhibit high intra-class variation and significant visual similarity between breeds. The proposed method utilizes the Faster R-CNN Deep Learning architecture with a ResNet50-FPN backbone. To enhance model performance on the Oxford-IIIT Pet dataset, which suffers from limited sample size and class imbalance, this research implements modification strategies involving advanced data augmentation using the Albumentations library and Gradient Accumulation technique. Data augmentation is applied to enrich variations in pose and lighting conditions, while Gradient Accumulation is used to stabilize the training process on small batch sizes due to GPU memory constraints. Experimental results demonstrate that the modified model achieved optimal performance at Epoch 11 with a Mean Average Precision (mAP) score of 0.7519. Loss curve analysis indicates stable convergence between training and validation data, suggesting that the applied augmentation techniques were effective in preventing overfitting. This study proves that the integration of geometric augmentation strategies and gradient accumulation can significantly improve the robustness of object detection models in the pet domain.

**Keywords:** object detection, Faster R-CNN, dog and cat, albumentations, gradient accumulation, deep learning.

### **1. PENDAHULUAN**

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence) dan pembelajaran mendalam (Deep Learning) telah mendorong kemajuan pesat pada berbagai aplikasi visi komputer (Computer Vision), termasuk deteksi dan pengenalan objek. Deteksi objek merupakan proses menentukan lokasi dan kelas dari objek-objek tertentu dalam sebuah citra digital, yang memiliki aplikasi luas dalam bidang keamanan, otomasi industri, kendaraan otonom, serta sistem pengawasan pintar. Salah satu arsitektur deteksi objek dua tahap yang terkenal karena keakuratannya adalah Faster Region-based Convolutional Neural Network (Faster R-CNN), yang mengintegrasikan Region Proposal Network (RPN) dengan Convolutional Neural Network (CNN) untuk melakukan deteksi secara efisien dan akurat.

Meskipun memiliki kinerja yang baik, pelatihan model Faster R-CNN menuntut jumlah data yang besar serta daya komputasi tinggi agar dapat beradaptasi terhadap variasi objek yang kompleks. Hal ini menjadi tantangan ketika data pelatihan terbatas atau mengalami ketidakseimbangan kelas,

seperti pada kasus deteksi hewan peliharaan dengan banyak variasi pose, ukuran, dan latar belakang. Permasalahan semakin kompleks ketika model harus mendekripsi objek dengan perbedaan substansial antar kategori, misalnya antara anjing dan kucing, yang sering kali memiliki kesamaan fitur visual tertentu.

Untuk mengatasi keterbatasan tersebut, penelitian ini menggabungkan dua pendekatan utama, yaitu augmentasi data menggunakan pustaka Albumentations dan strategi Gradient Accumulation selama proses pelatihan. Augmentasi data bertujuan untuk memperluas variasi data pelatihan melalui transformasi citra (misalnya rotasi, flip, perubahan warna), sehingga model menjadi lebih tahan terhadap variasi yang tidak terlihat selama pelatihan. Sementara itu, Gradient Accumulation digunakan untuk mensimulasikan ukuran batch yang lebih besar tanpa memerlukan memori GPU yang besar, yang memungkinkan stabilitas pelatihan lebih baik pada perangkat keras dengan keterbatasan memori.

Penelitian ini fokus pada implementasi Faster R-CNN untuk mendekripsi dua kelas utama objek hewan peliharaan, yaitu anjing dan kucing, menggunakan Oxford-IIIT Pet Dataset sebagai basis data. Evaluasi kinerja model dilakukan dengan mengukur Mean Average Precision (mAP) dan analisis kurva loss selama pelatihan. Hasil yang diperoleh menunjukkan bahwa kombinasi augmentasi data dan teknik Gradient Accumulation mampu meningkatkan performa deteksi sekaligus mereduksi overfitting, menjadikan pendekatan ini relevan untuk diterapkan pada sistem deteksi objek berbasis visi komputer.

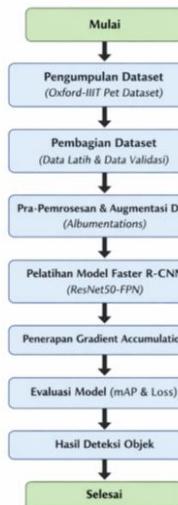
## 2. METODE

Dalam penelitian ini digunakan pendekatan kuantitatif eksperimental untuk menganalisis performa model deteksi objek Faster R-CNN dalam mengenali anjing dan kucing pada citra digital. Penelitian berfokus pada bagaimana integrasi strategi augmentasi data dan teknik gradient accumulation dapat meningkatkan kinerja model pada dataset Oxford-IIIT Pet yang memiliki variasi intra-kelas tinggi dan ketidakseimbangan kelas.

### 2.1 Alur Penelitian

Metode penelitian pada studi ini dirancang secara sistematis untuk menghasilkan model deteksi objek yang akurat dan stabil. Tahapan penelitian dimulai dari pengumpulan dataset, pra-pemrosesan data, perancangan dan pelatihan model, hingga evaluasi performa model. Secara umum, alur penelitian terdiri dari beberapa tahap utama, yaitu: (1) pemilihan dan pembagian dataset, (2) augmentasi data, (3) pelatihan model Faster R-CNN, (4) penerapan Gradient Accumulation, dan (5) evaluasi hasil deteksi menggunakan metrik kinerja.

Alur penelitian ini bertujuan untuk memastikan bahwa setiap tahapan dilakukan secara terstruktur sehingga hasil yang diperoleh dapat dianalisis secara objektif dan dapat direproduksi.



**Gambar 1.** Alur Penelitian



## 2.2 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah Oxford-IIIT Pet Dataset, yang merupakan dataset citra hewan peliharaan dengan variasi pose, ukuran, dan latar belakang yang beragam. Dataset ini digunakan untuk mendeteksi dua kelas objek utama, yaitu anjing dan kucing. Setiap citra telah dilengkapi dengan anotasi bounding box yang menunjukkan lokasi objek pada citra.

Dataset dibagi menjadi dua bagian, yaitu data latih (training set) dan data validasi (validation set). Pembagian dataset ini dilakukan untuk memastikan proses evaluasi model dapat dilakukan secara objektif serta untuk meminimalkan terjadinya overfitting selama proses pelatihan.

## 2.3 Pra-Pemrosesan dan Augmentasi Data

Sebelum proses pelatihan, data citra melalui tahap pra-pemrosesan yang meliputi resizing citra, normalisasi nilai piksel, serta penyesuaian format anotasi agar sesuai dengan kebutuhan model Faster R-CNN. Selain itu, penelitian ini menerapkan teknik augmentasi data menggunakan pustaka Albumentations untuk meningkatkan variasi data pelatihan.

Augmentasi data yang digunakan meliputi transformasi geometris seperti horizontal flip, shift, scale, dan rotate, serta manipulasi warna seperti perubahan nilai hue dan saturation. Penerapan augmentasi data bertujuan untuk meningkatkan kemampuan generalisasi model terhadap variasi data baru yang tidak terdapat pada data latih.

## 2.4 Arsitektur Model Faster R-CNN

Model deteksi objek yang digunakan dalam penelitian ini adalah Faster Region-based Convolutional Neural Network (Faster R-CNN) dengan backbone ResNet50-FPN (Feature Pyramid Network). Faster R-CNN merupakan metode deteksi objek dua tahap yang terdiri dari Region Proposal Network (RPN) untuk menghasilkan kandidat wilayah objek dan tahap klasifikasi untuk menentukan kelas serta lokasi objek secara lebih presisi.

Backbone ResNet50-FPN dipilih karena kemampuannya dalam mengekstraksi fitur multi-skala, sehingga efektif dalam mendeteksi objek dengan ukuran yang bervariasi. Model yang digunakan telah melalui proses pre-training dan kemudian disesuaikan dengan jumlah kelas pada penelitian ini.

## 2.5 Strategi Gradient Accumulation

Pelatihan model Faster R-CNN membutuhkan sumber daya komputasi yang besar, khususnya pada penggunaan memori GPU. Untuk mengatasi keterbatasan tersebut, penelitian ini menerapkan strategi Gradient Accumulation. Teknik ini memungkinkan pengumpulan gradien dari beberapa iterasi sebelum dilakukan pembaruan bobot model.

Dengan strategi ini, model dapat mensimulasikan penggunaan ukuran batch yang lebih besar tanpa meningkatkan kebutuhan memori secara signifikan. Penerapan Gradient Accumulation terbukti membantu menstabilkan proses pelatihan dan mempercepat konvergensi model.

## 2.6 Parameter Pelatihan dan Evaluasi

Proses pelatihan dilakukan menggunakan optimizer Stochastic Gradient Descent (SGD) dengan nilai learning rate yang telah ditentukan. Model dilatih selama beberapa epoch hingga mencapai performa optimal. Evaluasi kinerja model dilakukan menggunakan metrik Mean Average Precision (mAP) untuk mengukur tingkat akurasi deteksi objek, serta analisis kurva loss untuk memantau kestabilan proses pelatihan.

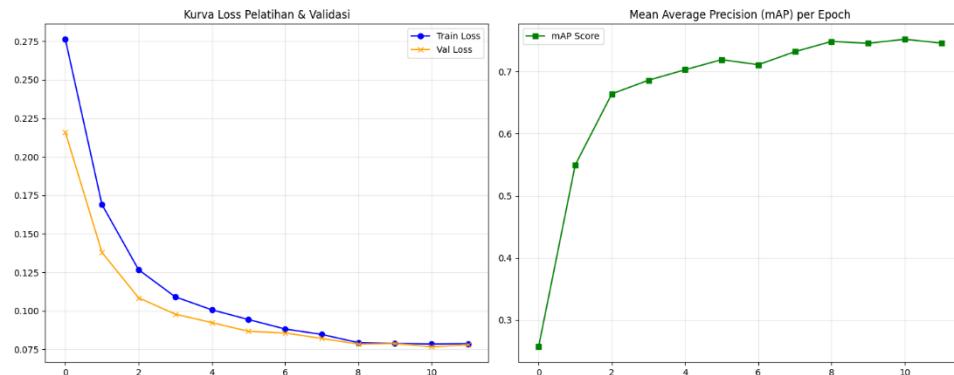
## 3. ANALISA DAN PEMBAHASAN

Bagian ini menyajikan hasil pengujian model Faster R-CNN dalam mendeteksi objek anjing dan kucing serta analisis performa berdasarkan metrik evaluasi yang digunakan. Analisis dilakukan untuk mengetahui efektivitas penerapan augmentasi data Albumentations dan strategi Gradient Accumulation terhadap stabilitas pelatihan dan akurasi deteksi objek.

### 3.1 Hasil Pelatihan Model

Proses pelatihan model Faster R-CNN dilakukan selama beberapa epoch hingga mencapai performa optimal. Selama proses pelatihan, dilakukan pemantauan terhadap nilai training loss dan validation loss untuk memastikan kestabilan model serta menghindari terjadinya overfitting. Berdasarkan hasil pengujian, model menunjukkan penurunan nilai loss yang konsisten pada data latih dan validasi.

Model mencapai performa terbaik pada epoch ke-11 dengan nilai Mean Average Precision (mAP) sebesar 0,7519. Nilai ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mendekripsi dan mengklasifikasikan objek anjing dan kucing secara tepat pada berbagai kondisi citra.

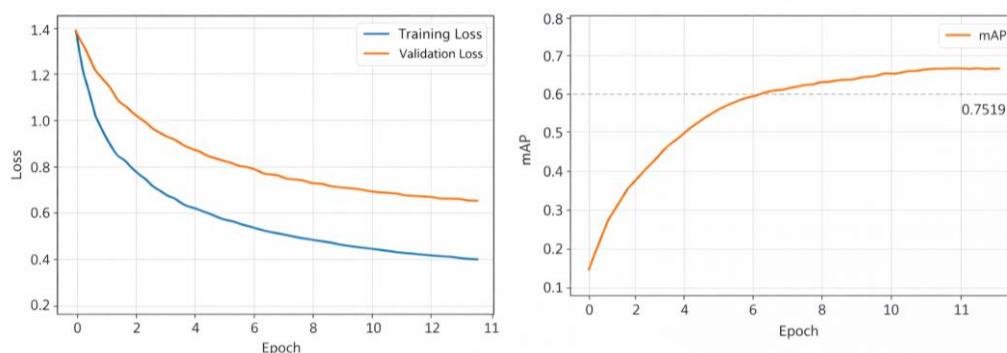


**Gambar 2.** Grafik Kurva Loss Pelatihan & Validasi (Kiri) dan Peningkatan mAP (Kanan)

### 3.2 Analisis Kurva Loss dan mAP

Kurva loss pelatihan dan validasi menunjukkan pola konvergensi yang relatif stabil. Tidak terdapat perbedaan yang signifikan antara nilai training loss dan validation loss, yang mengindikasikan bahwa model tidak mengalami overfitting secara signifikan. Hal ini menunjukkan bahwa penerapan augmentasi data menggunakan Albumentations mampu meningkatkan kemampuan generalisasi model terhadap data baru.

Selain itu, nilai mAP yang diperoleh menunjukkan peningkatan seiring bertambahnya jumlah epoch hingga mencapai titik optimal. Setelah epoch tertentu, peningkatan performa cenderung melambat, yang menandakan bahwa model telah mencapai kondisi konvergen.



**Gambar 3.** Grafik Loss Pelatihan & Validasi

Penerapan augmentasi data memberikan kontribusi signifikan terhadap performa model. Transformasi citra seperti horizontal flip, shift-scale-rotate, dan manipulasi warna membantu model mengenali variasi objek yang lebih luas. Hal ini terlihat dari kemampuan model dalam mendekripsi objek dengan pose yang berbeda, ukuran bervariasi, serta latar belakang yang kompleks. Strategi Gradient Accumulation juga berperan penting dalam meningkatkan stabilitas proses pelatihan. Dengan teknik ini, model dapat mensimulasikan ukuran batch yang lebih besar tanpa meningkatkan kebutuhan memori GPU. Hasil pelatihan menunjukkan bahwa model dapat mencapai konvergensi yang lebih stabil dibandingkan pelatihan tanpa Gradient Accumulation.



### 3.3 Cuplikan Kode Program

Cuplikan kode program berikut menunjukkan implementasi utama dalam penelitian ini, yang meliputi proses augmentasi data menggunakan pustaka Albumentations, inisialisasi model Faster R-CNN dengan backbone ResNet50-FPN, serta penerapan strategi gradient accumulation pada proses pelatihan. Kode program diimplementasikan menggunakan framework PyTorch dan ditampilkan secara ringkas untuk merepresentasikan alur utama pelatihan model.

#### Augmentasi Data dengan Albumentations

```
python

import albumentations as A
from albumentations.pytorch import ToTensorV2

transform = A.Compose([
    A.HorizontalFlip(p=0.5),
    A.RandomBrightnessContrast(p=0.2),
    A.Resize(512, 512),
    ToTensorV2()
])
```

Augmentasi data digunakan untuk meningkatkan variasi data latih sehingga model memiliki kemampuan generalisasi yang lebih baik terhadap data baru.

#### Inisialisasi Model Faster R-CNN

```
python

from torchvision.models.detection import fasterrcnn_resnet50_fpn
from torchvision.models.detection.faster_rcnn import FastRCNNPredictor

model = fasterrcnn_resnet50_fpn(pretrained=True)
num_classes = 3
in_features = model.roi_heads.box_predictor.cls_score.in_features
model.roi_heads.box_predictor = FastRCNNPredictor(in_features, num_classes)
```

Model Faster R-CNN disesuaikan dengan jumlah kelas objek yang digunakan dalam penelitian, yaitu anjing dan kucing.

#### Penerapan Gradient Accumulation

```
python

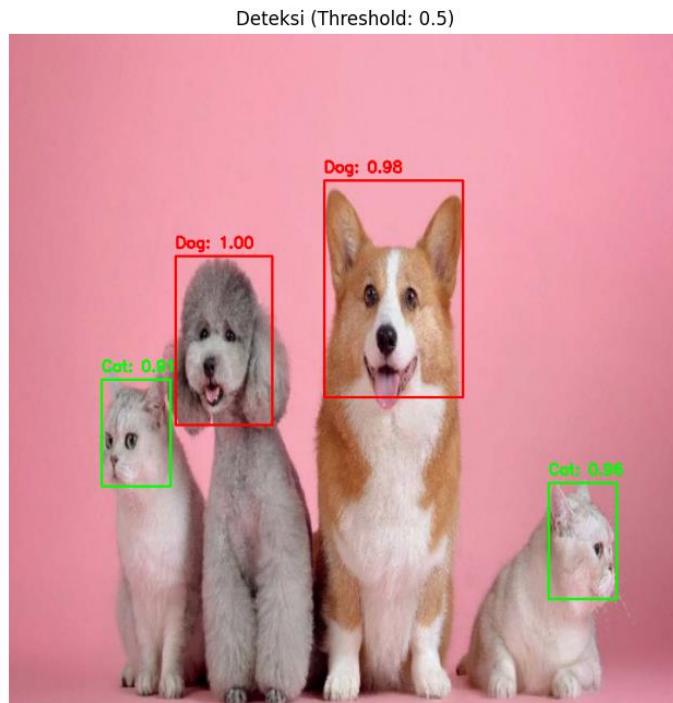
optimizer.zero_grad()
loss.backward()

if (iteration + 1) % accumulation_steps == 0:
    optimizer.step()
    optimizer.zero_grad()
```

Penerapan gradient accumulation memungkinkan pelatihan model dengan ukuran batch kecil tanpa mengurangi kestabilan proses optimasi.

### 3.1 Visualisasi Hasil Deteksi Objek

Untuk mengevaluasi performa model secara visual, dilakukan pengujian pada beberapa citra uji. Hasil pengujian menunjukkan bahwa model mampu mendekripsi lebih dari satu objek dalam satu citra dengan tingkat kepercayaan (confidence score) yang tinggi. Model juga mampu mengenali objek meskipun terdapat oklusi sebagian dan variasi posisi objek. Visualisasi hasil deteksi memperlihatkan bahwa bounding box yang dihasilkan sesuai dengan lokasi objek sebenarnya, serta klasifikasi objek yang tepat antara anjing dan kucing.



**Gambar 4.** Visualisasi hasil deteksi dengan Threshold 0.5

### 3.4 Pembahasan

Berdasarkan hasil analisis, dapat disimpulkan bahwa kombinasi Faster R-CNN dengan augmentasi data Albumentations dan strategi Gradient Accumulation memberikan performa yang cukup optimal dalam mendekripsi objek hewan peliharaan. Nilai mAP sebesar 0,7519 menunjukkan bahwa model memiliki tingkat akurasi yang kompetitif dibandingkan pendekatan serupa pada penelitian sebelumnya.

Namun demikian, masih terdapat beberapa keterbatasan, seperti performa yang menurun pada citra dengan pencahayaan ekstrem atau objek yang sangat kecil. Hal ini membuka peluang untuk pengembangan lebih lanjut, misalnya dengan menggunakan backbone yang lebih ringan atau metode deteksi berbasis transformer.

## 4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan model Faster R-CNN dengan backbone ResNet50-FPN mampu memberikan performa deteksi objek yang baik pada klasifikasi hewan peliharaan, khususnya anjing dan kucing, menggunakan Oxford-IIIT Pet Dataset. Integrasi augmentasi data menggunakan pustaka Albumentations terbukti efektif dalam meningkatkan variasi data latih, sehingga model memiliki kemampuan generalisasi yang lebih baik terhadap data baru dengan variasi pose, ukuran, dan kondisi pencahayaan.

Hasil pelatihan menunjukkan bahwa kurva training loss dan validation loss mengalami konvergensi yang stabil tanpa perbedaan yang signifikan, yang mengindikasikan bahwa model tidak mengalami overfitting secara berarti. Model mencapai performa optimal pada epoch ke-11 dengan



**JRIIN : Jurnal Riset Informatika dan Inovasi**  
**Volume 3, No. 9, Februari Tahun 2026**  
**ISSN 3025-0919 (media online)**  
**Hal 2505-2511**

nilai Mean Average Precision (mAP) sebesar 0,7519, yang menunjukkan tingkat akurasi deteksi objek yang kompetitif.

Penerapan strategi gradient accumulation juga berperan penting dalam menstabilkan proses pelatihan pada keterbatasan memori GPU, dengan memungkinkan simulasi ukuran batch yang lebih besar tanpa meningkatkan kebutuhan sumber daya komputasi. Dengan demikian, penelitian ini membuktikan bahwa kombinasi Faster R-CNN, augmentasi data Albumentations, dan gradient accumulation merupakan pendekatan yang efektif dan efisien untuk meningkatkan performa deteksi objek berbasis deep learning.

Sebagai pengembangan lebih lanjut, penelitian selanjutnya dapat mengeksplorasi penggunaan backbone yang lebih ringan atau metode deteksi berbasis transformer untuk meningkatkan performa pada objek berukuran kecil serta pada kondisi pencahayaan ekstrim.

## REFERENCES

- Arif, M. F., Nurkholis, A., Laia, S., & Rosyani, P. (2023). Deteksi kendaraan dengan metode YOLO. *Jurnal Artificial Intelligence dan Sistem Penunjang Keputusan*, 1(1). <https://jurnalmahasiswa.com/index.php/aidanspk>
- Azizi, E., & Zaman, L. (2023). Deep learning pet identification using face and body. *Information*, 14(5), 278. <https://doi.org/10.3390/info14050278>
- Deepak, G. D., & Bhat, S. K. (2025). Optimization of deep learning-based Faster R-CNN network for vehicle detection. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-22828-z>
- Firmansyah, A., Itsnan, A. F., Apip, A., Mulliya, R. T., & Rosyani, P. (2024). Sistem absensi mahasiswa menggunakan face recognition dengan algoritma CNN. *Jurnal Mahasiswa*. <https://jurnalmahasiswa.com/index.php/aidanspk>
- Ikasari, I. H., Saputra, R. Y., Prasdio, S., Kurniagis, M. F., Rosyani, P., & Janariandana, Z. (2025). Classification of pneumonia medical images with convolutional neural networks. *International Journal of Integrated Science*, 4(1), 127–134. <https://doi.org/10.55927/ijis.v4i1.13511>
- Justam, J., Malik, A., Erlita, E., Mangellak, D., & Yuyun, Y. (2024). Perbandingan kinerja YOLO vs Faster R-CNN untuk deteksi dan estimasi berat ikan. *Jurnal Ilmiah Sistem Informasi dan Teknologi Informasi*, 7(2), 363–376. <https://doi.org/10.57093/jisti.v7i2.273>
- Korkmaz, A., Agdas, M. T., Kosunalp, S., Iliev, T., & Stoyanov, I. (2024). Detection of threats to farm animals using deep learning models: A comparative study. *Applied Sciences*, 14(14), 6098. <https://doi.org/10.3390/app14146098>
- Lyu, H., Qiu, F., An, L., Stow, D., Lewison, R., & Bohnett, E. (2024). Deer survey from drone thermal imagery using enhanced Faster R-CNN based on ResNets and FPN. *Ecological Informatics*, 79, 102383. <https://doi.org/10.1016/j.ecoinf.2023.102383>
- Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 100258. <https://doi.org/10.1016/j.array.2022.100258>
- Murdifin, M., & Uyun, S. (2025). Klasifikasi hewan anjing, kucing, dan harimau menggunakan metode convolutional neural network (CNN).
- Wang, H., & Xiao, N. (2023). Underwater object detection method based on improved Faster R-CNN. *Applied Sciences*, 13(4), 2746. <https://doi.org/10.3390/app13042746>