Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 32, Num. 1 (2025) 196-203

RESEARCH ARTICLE

Digital Fencing for Farms: Enhancing Object Detection through Multi-Dataset Integration

Cercas Digitais para Fazendas: Aprimorando a Detecção de Objetos por meio da Integração de Vários Conjuntos de Dados

Juliana Quintiliano Ferreira[1]*, Lucas Silva[1], Thiago L. Gomes[1], Michel M. Silva[1]

Abstract: Security in farm areas is a critical concern due to the risks posed by wild animal invasions, farm animals straying, and unauthorized human entry. Conventional object detection models, often trained on urban or context-specific datasets, frequently underperform in farm settings. To address this, our study investigates the integration of multi-datasets to enhance object detection for digital fences, thereby improving farm security. We propose a method to effectively utilize multi-datasets, even when they do not have the same classes, ensuring comprehensive coverage of all required categories. The proposed SmartClass methodology achieved more robust and adaptable detection approaches, suitable for agricultural environments, with considerable increases in recall, mAP50, and mAP50-95 metrics compared to models trained without the methodology. The code and data are available at github.com/MaVILab-UFV/Digital-Fencing-WVC-2024.
Keywords: object detection — farm security — digital fence — data integration

Resumo: A segurança em áreas agrícolas é uma preocupação crítica devido aos riscos representados por invasões de animais selvagens, animais de fazenda perdidos e entrada humana não autorizada. Modelos convencionais de detecção de objetos, geralmente treinados em conjuntos de dados urbanos ou específicos de contexto, frequentemente apresentam desempenho inferior em ambientes agrícolas. Para abordar isso, nosso estudo investiga a integração de vários conjuntos de dados para aprimorar a detecção de objetos para cercas digitais, melhorando assim a segurança da fazenda. Propomos um método para utilizar efetivamente vários conjuntos de dados, mesmo quando eles não têm as mesmas classes, garantindo uma cobertura abrangente de todas as categorias necessárias. A metodologia SmartClass proposta alcançou abordagens de detecção mais robustas e adaptáveis, adequadas para ambientes agrícolas, com aumentos consideráveis nas métricas de recall, mAP50 e mAP50-95 em comparação com modelos treinados sem a metodologia. Os códigos e dados estão disponíveis em github.com/MaVILab-UFV/Digital-Fencing-WVC-2024.
Palavras-Chave: Detecção de Objetos — Segurança de Fazendas — Cerca Digital — Integração de Dados

[1] *Department of Informatics, Universidade Federal de Viçosa (UFV), Brazil*
*__Juliana Quintiliano Ferreira__: juliana.q.ferreirar@ufv.br
DOI: http://dx.doi.org/10.22456/2175-2745.143486 • **Received:** 23/10/2024 • **Accepted:** 23/12/2024
CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The advancement of object detection models has significantly enhanced security across various settings, including businesses, commercial buildings, airports, and homes [1, 2, 3]. These models are increasingly deployed in systems that monitor city security cameras to detect stolen or suspicious vehicles, providing timely notifications to municipal authorities. Additionally, object detection is utilized in high-traffic areas for facial recognition, aiding in the capture of fugitives. For instance, B. Rohan Reddy et al. [4] applied object detection to improve transportation security by conducting baggage inspections and identifying prohibited items at airports. Similarly, Yaohui Xiao et al. [5] leveraged real-time object detection to enhance the security of substations, identifying unauthorized objects and vehicles in unusual positions and detecting early-stage fires.

We recognize that object detection can also play a crucial role in enhancing the security of farms and rural properties. Security in these areas is often challenged by various factors, including unauthorized human entry, livestock wandering off, and the intrusion of wild animals [6]. As Aishwarya D. Shetty et al. [7] note, preventing wild animal intrusions not only reduces risks to human safety and economic losses but also contributes to the conservation of wildlife.

The farm environments present unique challenges for object detection algorithms, as models trained on urban imagery may not perform effectively in these settings, as noted by Zachary Pezzementi et al. [8]. Farm images often pose greater difficulty for algorithms due to the potential for people and animals to be camouflaged among vegetation or along farm roads. The distinct visual characteristics of farm areas—dominated by vegetation or open land—contrast sharply with the urban backdrop of buildings, houses, and asphalt, further complicating detection efforts [9].

Given the distinct characteristics of farm and urban environments and the pressing need to enhance farm security, this paper aims to develop a model capable of assisting in creating digital fences. These fences are intended to detect the presence of people, potentially dangerous wild animals that could pose a threat to humans and livestock, as well as other animals that may stray into inappropriate areas. Unlike most existing object detection models trained on the COCO dataset [10], our approach focuses on enhancing object detection specifically for farm environments, where objects may be camouflaged within the landscape. To achieve this, we include datasets in our training process that are more contextually aligned with farm settings, offering a closer match to the unique challenges of these environments.

Datasets are fundamental to the success of machine learning methods. However, there is currently no comprehensive dataset in the farm context that includes all relevant classes and provides a sufficient number of samples. Additionally, creating and labeling a new dataset was not feasible due to the significant time and labor costs involved. To overcome this challenge, we aggregated existing datasets for this study. By integrating these datasets, we trained an object detection model to enhance digital fences, ultimately improving security in farm environments. However, integrating different datasets is a complex task [11], as most datasets utilize different annotation formats and define varying classes. To address the issue of differing annotation formats, we adopted a preprocessing approach that converts all annotations into a standardized, widely accepted format. To handle the challenge of inconsistent class definitions across datasets, we proposed a set of methods collectively termed ”SmartClass” models. These methods are designed to train models in a way that prevents the absence of annotations for certain classes in one dataset from degrading the training of those classes.

To validate the effectiveness of our approach, we conducted experiments using four distinct datasets: COCO [10], NREC [8], Cows2021 [12], and APT-36 [13]. We focused on nine classes—person, cat, dog, horse, sheep, cow, elephant, zebra, and giraffe—that are represented in at least two or three of these datasets. Our results demonstrate that our strategy yields improved metric values and superior qualitative outcomes.

2. Related Work

Farm security can be significantly enhanced through object detection, which identifies people, animals, and vehicles that may unlawfully enter rural properties, thereby compromising the safety of individuals, livestock, and crops. By improving object detection in rural environments, it is possible to create a digital fence that enhances farm security. In this context, we review some studies related to this issue.

Several studies have focused on agricultural security, particularly on preventing animal intrusions that can harm farm animals or damage crops. For instance, Mamat et al. [14] proposed a model for detecting four categories of animals — elephant, monkey, tiger, and wild boar — that frequently invade agricultural areas. Their work demonstrated strong results using the YOLOv5 architecture, achieving high accuracy in animal detection. They also compared their results with the YOLOv4 architecture, which similarly yielded good accuracy. The study utilized a dataset with just 280 images, including 70 images per category, which restricts its applicability across diverse contexts. Naveenkumar M. et al. [15] also worked to prevent attacks on farms and people using DenseNet201. The work proposed the detection of nine classes of animals through video analysis. They adopted the classes bear, bison, elephant, fox, leopard, lion, tiger, pig, and horse. This work differs from ours as we explore and integrate multiple datasets to enhance object detection across a broader range of contexts and categories.

Akhil et al. [16] employed various techniques to develop a virtual fence aimed at preventing conflicts between humans and animals on properties with cultivated crops. Their approach involved using motion detection to identify the presence of animals. Once motion was detected, the system processed the image to confirm the presence of animals. If animals were detected, the system employed loud sounds and high-intensity strobe lights to deter them. For image detection, they utilized a convolutional neural network implemented in TensorFlow. The dataset comprised 100 images, including categories for cats, dogs, elephants, and wild boars. Although the study addressed the complete virtual fence process, the accuracy of animal detection was suboptimal. The authors noted the need for additional images and further model training to enhance detection accuracy.

Several studies have utilized the Internet of Things (IoT) and sensors to detect human intrusion [17]. I. Oyelade et al. [18] implemented an intrusion detection system that combines IoT and computer vision to identify both humans and animals. The system performs facial recognition for people and alerts farm owners in the event of an intrusion. Their study focused on detecting cows, goats, and people, using Fast-RCNN for object detection. The dataset employed consisted of 3,752 images. Unlike their approach, which involves fewer animal classes, our work distinguishes itself by exploring and integrating multiple datasets to improve object detection across a wider range of contexts and categories. Research efforts have focused on improving the detection of people

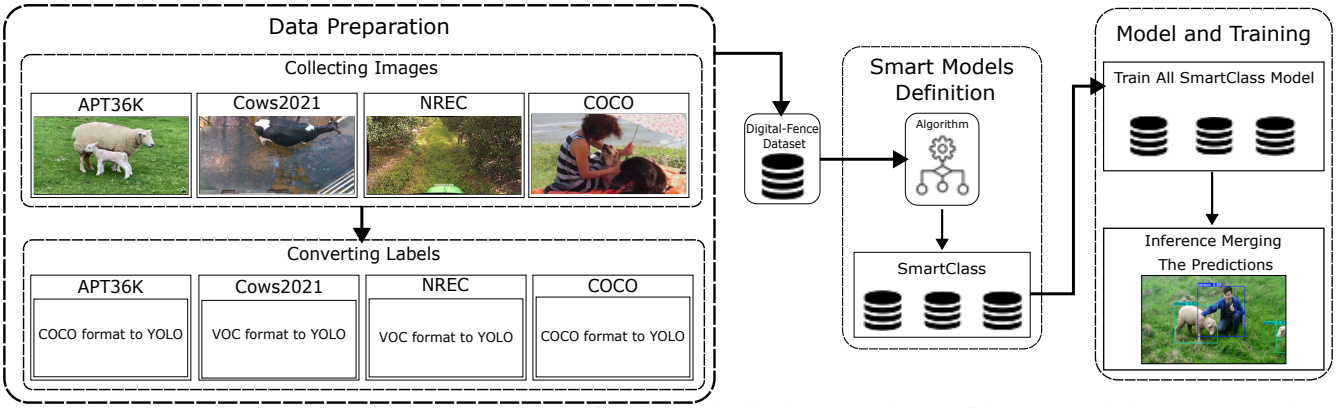


Figure 1. Images and annotations from datasets are unified into a single dataset. Subsets of datasets and classes (SmartClass) are defined to ensure consistent class annotations, followed by training the models using the YOLOv8 with transfer learning.

in farm environments to enhance the safety of autonomous vehicles in these areas [19]. Peter Neigel et al. [9] developed the OPEDD dataset (Off-Road Pedestrian Detection Dataset) to aid in this endeavor. Similarly, the NREC dataset [8] was introduced to advance autonomous driving research with an emphasis on human detection in agricultural settings. Unlike these studies, which focus on pedestrian detection for autonomous driving, our work includes animal detection and aims to enhance overall farm safety.

Xingyi Zhou et al. [11] explored training a model for multiple domains by integrating various datasets, highlighting challenges like semantic discrepancies and annotation inconsistencies. They proposed a unified object detection model with an automatic taxonomy to group similar classes. However, their approach is most effective when direct correspondences exist between classes across datasets. Still, it faces limitations with classes that lack annotations in certain datasets or do not have semantically similar counterparts annotated.

3. Methodology

To enhance object detection in farm contexts and ensure coverage of relevant classes across different datasets, we followed a structured approach. As shown in Figure 1, first, we collected images and annotations from various datasets and we unify this data in a new dataset. Second, we employed a strategy to define models (SmartClass) and select subsets of datasets where the chosen classes are consistently annotated across all datasets in the subset, thereby avoiding samples with unannotated present classes. Finally, we utilized the YOLOv8 architecture to train the selected models, leveraging transfer learning to develop the SmartClass models.

3.1 Data Preparation

To train a model for creating digital fences on farms, obtaining a relevant dataset of images in these environments is essential. Due to the high cost of creating and labeling a new dataset, we leveraged existing ones. Since no dataset fully met our needs, we built the Digital-Fence-Dataset by combining images from NREC, Cows2021, APT-36k, and COCO.

The NREC (National Robotics Engineering Center) Person Detection Dataset [8] consists of off-road videos captured in orange and apple orchards, primarily focused on person detection. We selected it because the images scene is visually similar to the rural environment, which are the focus of interest of this project, and additionally, it contains the class People that plays a crucial role for creating digital fences to detect intruders. Annotated in the Pascal VOC format, it includes 76,000 images with a resolution of 720×480 pixels. The annotated classes are Person, Part-Person, and Background. For our work, we excluded images labeled as Background, as our focus is on images containing people. We converted the annotations from VOC to YOLO txt format. The Cows2021 dataset [12] features images of Holstein-Friesian cows, designed to identify individual cows based on their coat patterns. The dataset, annotated in Pascal VOC format, contains images with a resolution of 1280×720 pixels. We selected this dataset to improve the detection of cows viewed from above, considering that this class is relevant on farms context. We converted the annotations from VOC to YOLO format. The COCO (Common Objects in Context) dataset is a large-scale object detection, segmentation, and captioning dataset with over 200,000 labeled images containing annotations for 80 object categories. The dataset is widely used for training models in diverse object detection tasks, and we use it to help maintain detection of people and animals in various contexts, as some locations on farms and in agriculture may have urban characteristics, for *e.g.*, areas close to the farmstead. As with the other datasets, we migrated the annotations to the YOLO txt format to ensure consistency across the Digital-Fence-Dataset. APT-36k [13] is a dataset for animal pose estimation and tracking, comprising 36,000 images across various animal classes, annotated in COCO format. The annotated classes include a wide range of animals, from deer to elephants. For our purposes, we selected only the classes cat, dog, horse, sheep, cow, elephant, zebra, and giraffe, which overlap with classes in the COCO dataset. We converted the

annotations from COCO to YOLO txt format. We selected this dataset because it covers other classes of animals that may be found on farms or may invade farms, but we limited the animals to those also included in the COCO dataset to contain at least more than one dataset for each class.

3.2 Smart Models Definition

Integrating different datasets is a complex task [11]. A significant challenge that can undermine the accuracy of models trained on combined datasets is the inconsistency in class annotations. For example, in the APT-36k dataset, animals are annotated, but the images also contain unlabeled people. This inconsistency can decrease model performance for the people class. To address this issue, we proposed a suite of methods collectively termed SmartClass models. These methods are designed to train models in a way that minimizes the negative impact of missing annotations in one dataset, thereby preserving the accuracy of the trained classes.

We employed a strategy to define the SmartClass models by selecting subsets of datasets where the chosen classes are consistently annotated across all datasets within each subset, thereby avoiding samples with unannotated classes. Our goal is to maximize the number of classes represented in each model of the SmartClass model and ensure that each selected class is well-represented across the largest possible subset of datasets. The detailed process of this selection is in Algorithm 1.

Algorithm 1 Model and Class Selection

Input: Set of Classes, Datasets
Output: Selected Models, Selected Classes
$Classes \leftarrow \{person, \ldots\}$
$Models \leftarrow \emptyset$
$Selected_classes \leftarrow \emptyset$
for each subset S of $Classes$ ordered from largest to smallest **do**
  if there is a class $s \in S$ such that $s \in Selected_classes$ **then**
    Skip to the next subset
  else
    Identify the subset $T \subseteq S$ such that each element $t \in T$ is present in all datasets of the subset D of $Datasets$, where D is as large as possible.
    Add classes in T to $Selected_classes$
    Add model $M(T, D)$ to $Models$

3.3 Model and training

In this work, we utilize the Ultralytics framework [1], which includes the YOLOv8 architecture and pre-trained weights from the COCO dataset, which has 80 classes. We apply transfer learning using these pre-trained weights, focusing on nine specific classes: person, cat, dog, horse, sheep, cow, elephant, zebra, and giraffe.

[1] Available: ⟨https://docs.ultralytics.com/⟩

Training is conducted for each SmartClass model generated by the algorithm described in the previous section. Each model is tailored to a specific set of classes defined by the Smart Models Definition. After training, during the inference phase, we merge the predictions from the different SmartClass models to achieve comprehensive object detection across all classes.

4. Experiment

The dataset, Digital-Fence-Dataset, demonstrated in the previous section was used to perform these experiments. For all training, we trained for 100 epochs and the images were resized to 416x416 pixels. In addition, we leveraged the original YOLOv8 weights to start the training. The optimizer was SGD with initial lr = 0.01 and linear decay over the epochs. We trained Model 1, Model 2 and SmartClass, which are the models defined by the Algorithm 1, demonstrated in the Smart Models Definition methodology section.

The metrics used are Box P (Bounding Box Precision), Recall, mAP50 (Mean Average Precision) and mAP50-95 (Mean Average Precision at IoU=0.50 to 0.95). Box P refers to the proportion of positive predictions concerning the total predictions made by the model to check how correct a detection made by the model is. Recall indicates whether or not the model detects the majority of real objects present in the images. mAP50 returns the Average Precision (mAP) at an IoU threshold of 0.5 and mAP50-95 returns the Average Precision (mAP) at an IoU threshold of 0.5 to 0.95 in steps of 0.05.

The initial model trained was Model 1, which utilized all classes and images from the Digital-Fence-Dataset, except those derived from the COCO Dataset. Although COCO images were excluded from the training set, they were incorporated into the test set to assess model performance and obtain relevant metrics. This approach allowed us to evaluate whether the model's training generalized effectively to more generic problems or whether its performance was limited to the specific contexts in the Digital-Fence-Dataset.

The second model trained was Model 2, which utilized the full range of images and classes from the Digital-Fence-Dataset, including person, cat, dog, horse, sheep, cow, elephant, zebra, and giraffe. This model was designed to evaluate the effectiveness of directly integrating all available datasets, providing insights into how well the model performs with a comprehensive and unified data source.

The next models developed were the SmartClass models, as defined by the Algorithm 1. Three specialized datasets were created from the Digital-Fence-Dataset, leading to the training of three distinct SmartClass models. The first SmartClass model focused on detecting people, incorporating labeled images from two datasets. The second model was designed for detecting cows, using labeled images from three datasets. The third SmartClass model addressed the detection of cats, dogs, horses, sheep, elephants, zebras, and giraffes, and was trained on labeled images aggregated from two datasets.

Table 1. Comparison of object detection metrics across different models and object classes. Best in bold.

Class	Box Precision			Recall			mAP50			mAP50-95		
	Model 1	Model 2	SmartClass	Model 1	Model 2	SmartClass	Model 1	Model 2	SmartClass	Model 1	Model 2	SmartClass
person	***0.977***	0.901	0.918	0.543	0.758	***0.784***	0.669	0.851	***0.875***	0.439	0.56	***0.572***
cat	***0.971***	0.837	0.953	0.568	0.86	***0.925***	0.649	0.899	***0.958***	0.543	0.735	***0.812***
dog	***0.961***	0.874	0.917	0.616	0.757	***0.871***	0.721	0.861	***0.921***	0.586	0.685	***0.762***
horse	0.926	0.911	***0.945***	0.606	0.78	***0.825***	0.7	0.86	***0.905***	0.573	0.678	***0.741***
sheep	***0.941***	0.849	0.913	0.474	0.695	***0.746***	0.582	0.802	***0.859***	0.503	0.627	***0.679***
cow	***0.971***	0.935	0.959	0.768	0.854	***0.886***	0.864	0.927	***0.949***	0.756	0.779	***0.815***
elephant	***0.93***	0.897	0.922	0.637	0.871	***0.899***	0.748	0.929	***0.954***	0.619	0.734	***0.812***
zebra	0.944	0.922	***0.976***	0.778	0.873	***0.893***	0.848	0.94	***0.961***	0.691	0.756	***0.805***
giraffe	***0.97***	0.926	0.969	0.691	0.873	***0.905***	0.805	0.927	***0.957***	0.681	0.759	***0.822***

5. Results

5.1 Quantitative Result

Table 1 presents the performance metrics for each class across the three trained models: Model 1, Model 2 and SmartClass.

Model 1 consistently underperformed compared to the other models, with the exception of the Box Precision (Box P) metric, which ranged from 0.926 (horse) to 0.97 (giraffe). This high Box P indicates that Model 1 is precise when it detects objects, but its low Recall scores reveal a significant limitation: the model fails to identify or predict all the objects present in the images. This suggests that while the model performs well in the original context it was trained, it struggles to generalize to new or varied contexts. Additionally, low mAP50 and mAP50-95 confirm its lack of robustness across scenarios.

Model 2 was trained on all classes in the Digital-Fence-Dataset (person, cat, dog, horse, sheep, cow, elephant, zebra, giraffe). It showed Box Precision ranging from 0.837 (cat) to 0.935 (cow), and Recall ranging from 0.695 (sheep) to 0.873 (zebra, giraffe). The mAP50 metrics ranged from 0.802 for sheep to 0.929 for elephants, while mAP50-95 values ranged from 0.56 for the person class to 0.779 for the cow class. Overall, Model 2 shows improvement over Model 1 across most metrics, particularly in recall and mAP scores. However, the metrics remain relatively low, except for Box P, indicating that while the model effectively localizes objects when it detects them, it still struggles to detect objects across various contexts. This suggests that while Model 2 performs better in recognizing and localizing objects, there is still room



Figure 2. The first line shows Model 1's prediction with unsatisfactory results in detecting people. The second line shows Model 2's prediction with failures in detecting people. The third line shows SmartClass's prediction with satisfactory results and good confidence in detecting people and animals.



Figure 3. Image shows people detection in a rural environment. The first line shows the prediction with the original YOLO, and the second line with SmartClass.

for improvement in ensuring the model's robustness across different scenarios.

As shown in Table 1, the SmartClass models consistently outperformed Models 1 and 2 across all classes, particularly in the Recall (R), mAP50, and mAP50-95 metrics. This indicates that the SmartClass approach significantly improves object detection performance across a variety of contexts. However, the SmartClass models demonstrated slightly lower performance in the Box P metric. This is likely because, by detecting more objects, the SmartClass models are also more likely to encounter cases where the bounding boxes are less accurate, leading to a slight decrease in precision.

5.2 Qualitative Result

Figure 2, Line 1, shows Model 1's poor performance on images outside its training context. In all three images shown in row 1, the model failed to detect any people, identifying only animals. This outcome confirms that a model trained in a specific context struggles to generalize effectively to different contexts, as evidenced by its inability to detect people in images that differ from the original training environment.

Row 2 of Figure 2 illustrates the performance of Model 2 on images that did not include the person label during training. In the first image of Line 2, no detections were made. In the second image, a detection occurred but with a low confidence score of 0.32. The third image showed a detection with a higher confidence score of 0.69. These results suggest that unlabeled objects in the dataset can negatively impact the model's performance, leading to variable detection quality and reduced overall effectiveness.

For SmartClass, we observed significant improvements in detecting people across various contexts, including challenging environments such as rural areas and crop fields where individuals may be camouflaged or in varying positions, such as crouching or lying down. Row 3 of Figure 2 highlights how the SmartClass model enhances people detection compared to Model 1 and Model 2. In the three images shown in this row, SmartClass achieves confidence scores exceeding 0.85 for each detection. SmartClass also delivered strong results in detecting cows and other animals. Figure 2 further shows the detection performance for sheep and cows across Model 1, Model 2, and SmartClass, underscoring the effectiveness of SmartClass in these categories as well.

Additionally, Figure 3 illustrates people detection in crop fields. The first row displays detections using the YOLO model originally trained on the COCO Dataset, while the second row presents results from SmartClass, demonstrating consistently satisfactory performance in detecting people in challenging scenarios.

6. Conclusion

Farm surveillance is challenging, especially given the unique conditions where traditional object detection often fails. To address this issue, we proposed an enhancement in object detection aimed at developing digital fences for farm environments. To tackle this problem, we introduced the Digital-Fence-Dataset, an aggregated dataset combining elements from NREC, Cows2021, APT-36k, and COCO.

Our findings highlight that while combining datasets is beneficial for addressing gaps in dataset coverage, careful management is required to avoid confusion during model training. Unlabeled images within the dataset can mislead the model, affecting performance. To mitigate this issue, we employed the SmartClass approach, which involves training models on fractional groups of classes to reduce confusion and improve detection accuracy.

Our experiments with COCO, NREC, Cows2021, and APT-36k showed that integrating these datasets with the SmartClass methodology led to enhanced metric values and better qualitative results. Consequently, the SmartClass model approach effectively supports the development of digital fences for farm security.

As a limitation, it is worth mentioning that this approach may require multiple models to cover all classes, instead of a single comprehensive model. For this reason, more processing is required and the computational cost increases. In addition, each model has its own accuracy and this may lead to interpretability issues regarding detection confidence.

Acknowledgements

The authors thank CAPES, CNPq and FAPEMIG for funding this work.

Author contributions

Juliana Quintiliano Ferreira implemented the proposed method, developed the GitHub repository, and wrote of the paper. Lucas Silva created all the images in the article and helped review the text and code on GitHub. Thiago L. Gomes and Michel M. Silva acted as consultants, offering critical feedback on the text and proposing essential ideas for implementing the main approach.

References

[1] YIONG, Y. T.; KHAIRUDIN, A. R. M.; REDZUWAN, R. M. Real-time substation detection and monitoring security alarm system. In: *Conf. on Systems, Process Control (ICSPC)*. [s.n.], 2023. Disponível em: ⟨https://ieeexplore.ieee.org/document/10419939⟩.

[2] BRINDHA, S. et al. Augmented security system for commercial buildings by manipulating object detection and admin panel. In: *Inter. Conf. on Communication, Computing and Internet of Things (IC3IoT)*. [s.n.], 2022. Disponível em: ⟨https://ieeexplore.ieee.org/document/9767997⟩.

[3] ALZAABI, A. et al. A systematic literature review on machine learning in object detection security. In: *Inter. Conf. on Computing Communication and Automation (ICCCA)*. [s.n.], 2020. Disponível em: ⟨https://ieeexplore.ieee.org/document/9250836⟩.

[4] REDDY, B. R.; CHOKKALINGAM, S. P. Enhancing transportation security: Automated prohibited object detection for baggage inspection leveraging yolo-nas. In: *Inter. Conf. on Image Information Processing (ICIIP)*. [s.n.], 2023. Disponível em: ⟨https://ieeexplore.ieee.org/document/10537641⟩.

[5] XIAO, Y. et al. Real-time object detection for substation security early-warning with deep neural network based on yolo-v5. In: *IAS Global Conf. on Emerging Technologies (GlobConET)*. [s.n.], 2022. Disponível em: ⟨https://ieeexplore.ieee.org/document/9872338⟩.

[6] PRIYANKA, S. et al. Smart farming: Intelligent animal detection framework in agriculture using iot sensors. In: *Inter. Conf. on Edge Computing and Applications (ICECAA)*. [s.n.], 2023. Disponível em: ⟨https://ieeexplore.ieee.org/document/10212297⟩.

[7] SHETTY, A. D.; ASHWATH, S. Animal detection and classification in image & video frames using yolov5 and yolov8. In: *Inter. Conf. on Electronics, Communication and Aerospace Technology (ICECA)*. [s.n.], 2023. Disponível em: ⟨https://ieeexplore.ieee.org/document/10394750⟩.

[8] PEZZEMENTI, Z. et al. Comparing apples and oranges: Off-road pedestrian detection on the nrec agricultural person-detection dataset. *Journal of Field Robotics*, 2017. Disponível em: ⟨https://doi.org/10.48550/arXiv.1707.07169⟩.

[9] NEIGEL, P. et al. Opedd: Offroad pedestrian detection dataset. *Journal of WSCG*, 2020. Disponível em: ⟨http://wscg.zcu.cz/wscg2020/journal/F13.pdf⟩.

[10] LIN, T.-Y. et al. Microsoft coco: Common objects in context. In: *ECCV*. [s.n.], 2014. Disponível em: ⟨https://doi.org/10.48550/arXiv.1405.0312⟩.

[11] ZHOU, X.; KOLTUN, V.; KRÄHENBÜHL, P. Simple multi-dataset detection. In: *CVPR*. [s.n.], 2022. Disponível em: ⟨https://doi.org/10.48550/arXiv.2102.13086⟩.

[12] GAO, J. et al. Towards self-supervision for video identification of individual holstein-friesian cattle: The cows2021 dataset. *arXiv preprint arXiv:2105.01938*, 2021. Disponível em: ⟨https://doi.org/10.48550/arXiv.2105.01938⟩.

[13] YANG, Y. et al. Apt-36k: A large-scale benchmark for animal pose estimation and tracking. In: *Advances in Neural Information Processing Systems*. [s.n.], 2022. Disponível em: ⟨https://doi.org/10.48550/arXiv.2206.05683⟩.

[14] MAMAT, N.; OTHMAN, M. F.; YAKUB, F. Animal intrusion detection in farming area using yolov5 approach. In: *Inter. Conf. on Control, Automation and Systems (ICCAS)*. [s.n.], 2022. Disponível em: ⟨https://ieeexplore.ieee.org/document/10003780⟩.

[15] NAVEENKUMAR, M. et al. Densenet201 for animal detection and repellent system. In: *Inter. Conf. on Electronic Systems and Intelligent Computing (ICESIC)*. [s.n.], 2022. Disponível em: ⟨https://ieeexplore.ieee.org/document/9783582⟩.

[16] AKHIL, V. et al. Virtual fencing to prevent human wild conflict. In: *Inter. Conf. on Innovations in Engineering and Technology (ICIET)*. [s.n.], 2023. Disponível em: ⟨https://ieeexplore.ieee.org/document/10220801⟩.

[17] HSU, C.-K. et al. Design and implementation of image electronic fence with 5g technology for smart farms. In: *VTS Asia Pacific Wireless Communications Symposium (APWCS)*. [s.n.], 2019. Disponível em: ⟨https://ieeexplore.ieee.org/document/8851659⟩.

[18] OYELADE, I. et al. Farmland intrusion detection using internet of things and computer vision techniques. *Inter. Journal of Information Technology and Computer Science*, 2024. Disponível em: ⟨https://www.mecs-press.org/ijitcs/ijitcs-v16-n2/v16n2-3.html⟩.

[19] TABOR, T. et al. People in the weeds: Pedestrian detection goes off-road. In: *Inter. Symposium on Safety, Security, and Rescue Robotics (SSRR)*. [s.n.], 2015. Disponível em: ⟨https://ieeexplore.ieee.org/document/7442951⟩.