Research on Rotational Object Recognition Based on HSV Color Space and Gamma Correction

Zhujun Nie[1], Teng Long[1,*] and Zhangbing Zhou[1,2]

[1] School of Information Engineering, China University of Geosciences (Beijing), Beijing 100083, China
[2] Telecom SudParis, Institut polytechnique de Paris, Paris, France

Abstract
In the current digital era, image processing and pattern recognition play a critical role in fields such as environmental monitoring in the Internet of Things and smart city planning. However, traditional target detection algorithms face performance challenges when dealing with rotated objects, such as low recognition accuracy. To address this issue, this study investigated a deep learning-based method for detecting rotated objects. Firstly, by applying hue, saturation, value (HSV) and gamma correction to the images, the image quality was optimized to enhance object recognition capability. Secondly, this research introduced the MMRotate framework dedicated to detecting rotated objects, which, compared to traditional target detection frameworks, better meets the detection needs for rotated objects, thereby improving detection accuracy and robustness. Finally, from the perspective of the Internet of Things, this study classified and experimentally validated relevant datasets in an IoT environment, showing the performance of different targets on different datasets. Overall, this study provides new ideas and methods for addressing the shortcomings of rotated object detection in image processing and pattern recognition in the IoT environment, offering valuable insights and guidance for the development of smart cities and environmental monitoring.

Keywords
HSV color space, Gamma correction, Internet of Things, image processing, pattern recognition

1. Introduction

The Internet of Things (IoT) technology represents a pivotal direction in modern information technology development, enabling various devices and objects to connect and exchange data through embedded sensors, software, and other technologies. The widespread adoption of this technology has profoundly transformed multiple sectors, including industry, agriculture, healthcare, urban management, and domestic life. With the rapid development of the IoT industry, a wide array of IoT devices has become ubiquitous in everyday life, particularly image acquisition devices, which play a crucial role in numerous application scenarios. However, the precision of these devices in recognizing rotated targets is often limited. Concurrently, as the IoT scales up rapidly, the influx of massive volumes of image data into IoT environments

The 6th International Symposium on Advanced Technologies and Applications in the Internet of Things (ATAIT 2024), August 19-22, 2024, Kusatsu, Japan
*Corresponding author.
†This work was supported by National Natural Science Foundation of China under Grants No. 62002332.
nzj@cugb.edu.cn (Z. Nie); longteng@cugb.edu.cn (T. Long); zbzhou@cugb.edu.cn (Z. Zhou)

poses a challenge due to slow recognition speeds, which is an urgent issue that needs to be addressed[1, 2, 3].

In recent years, as detection tasks have continuously evolved, traditional horizontal bounding boxes have become inadequate for meeting the demands of specific fields. Consequently, researchers have begun to reconsider the representation of objects. To address this challenge, methods such as increasing the degrees of freedom in regression have been adopted to achieve more flexible object detection. This novel detection approach is known as rotated object detection. Ensuring high precision while effectively conducting rotated object detection has become a current research hotspot. Numerous factors can influence the performance of deep learning-based detectors.

In today's digital age, image processing and pattern recognition technologies have become core components of Internet of Things (IoT) applications. With the widespread adoption of IoT technologies, the importance of high-resolution images in urban planning, environmental monitoring, and intelligent transportation is increasingly highlighted. However, as detection tasks continue to evolve, traditional horizontal bounding boxes have become inadequate for meeting the demands of specific fields. Accurately detecting and recognizing targets at various angles and postures within images remains challenging. To address this challenge, methods such as increasing the degrees of freedom in regression have been adopted to achieve more flexible object detection, known as rotated object detection. In this study, the dataset was enhanced using Gamma correction[4] in the HSV[5] color space and by conducting experiments with rotated bounding box definitions under the MMRotate[6] framework. These enhancement processes significantly improved image visual quality, increasing color contrast and saturation, and enhancing aesthetics and recognizability. Additionally, the impact of various conditions on rotated object recognition was clearly highlighted.

The content of this paper is shown as follows. Section 2 of this paper reviews recent related work on object detection in the Internet of Things. Section 3 introduces the Gamma image processing method based on HSV and discusses the application of the MMRotate framework built upon this method. Section 4 describes how we set up the experiments and present the results. Section 5 concludes with a brief summary and mentions of future work.

2. Related Work

The application of Internet of Things (IoT) object recognition refers to the use of IoT technology and image processing techniques to automatically detect and identify target objects in various scenarios. This technology demonstrates significant potential in enhancing security, optimizing resource utilization, improving efficiency, and enhancing the quality of life across different fields. Mohaimenul et al. [7] developed a system based on the ESP32-CAM platform and the YOLOv8 object detection model, which efficiently provides real-time alerts by recognizing endangered species and harmful animals in agricultural environments. Maithili et al. [8] proposed an IoT-based automated object recognition system that offers high-accuracy object detection and recognition for visually impaired individuals in both indoor and outdoor environments, simplifying their mobility challenges. Swapna et al. [1] introduced a method for implementing object detection in embedded IoT devices by integrating deep learning algorithms, achieving real-time object detection widely applicable in security, healthcare, and workplace environments.

Rotational object detection is used for detecting rotated, tilted, and deformed objects in images or videos, typically described by rotated boxes or polygons, and employing deep learning models with geometric transformations for precise detection. Cai et al. [9] introduced the RealSR dataset with low-resolution and high-resolution image pairs captured by digital zoom and post-processing. Feng et al. [10] developed RINet, a weakly-supervised, end-to-end rotation-invariant aerial object detection network utilizing multi-branch detectors to refine instances with varying rotational awareness, generating rotational consistency supervision and coupling predictions across branches to explore potential instances from different angles, achieving rotation-invariant learning and multi-instance mining. Concurrently, Wang et al. [11] proposed a method based on prediction-aware one-to-one label assignment and 3D Max Filtering to bridge the gap between fully convolutional networks and end-to-end object detection, which demonstrated superior performance on COCO and CrowdHuman datasets, especially with auxiliary losses.

Current rotating target detection faces challenges such as high model complexity, high computational demands, limited labeled datasets, and difficulty in handling various angles and shapes. This paper uses an HSV-based gamma correction to enhance image quality and mAP, compares different rotation bounding box definitions with the MMRotate framework, and integrates HRSC and DOTA datasets to explore image processing applications in the Internet of Things.

3. Methodology

3.1. HSV-based Gamma Correction

The HSV [5] (Hue, Saturation, Value) color model is a three-dimensional model used to describe colors by dividing their attributes into hue, saturation, and value. In the HSV model, hue corresponds to the type of color, saturation to the purity of the color, and value to the brightness. This model facilitates intuitive color adjustments, such as changing the hue to alter the color type, adjusting saturation to control the vividness, and modifying value to adjust brightness. HSV is particularly suitable for certain image processing algorithms due to its intuitive representation of color properties.

Due to the nonlinear nature of human visual perception of brightness, directly displayed images may exhibit distorted contrast in dark and bright areas. Gamma correction[4] is typically used to address this issue by adjusting the image's brightness and contrast. The primary purpose of gamma correction is to compensate for the nonlinear response of display devices, ensuring more accurate image rendering. In simple terms, gamma correction is a nonlinear process that modifies the image's grayscale values, making the output grayscale values follow an exponential relationship with the input values.In this Equation.(1), V_{out} represents the output luminance, V_{in} represents the input luminance, A is a constant, and γ (gamma) is the gamma value. When γ (gamma) < 1 , low increase in brightness and low grayscale details; γ (gamma) > 1, reduce brightness and highlight grayscale details. This relationship indicates that the output luminance is an exponential function of the input luminance, adjusted by the gamma value. Gamma correction is crucial in image processing to ensure that images are displayed correctly on different devices, compensating for the nonlinear way human eyes perceive brightness and the nonlinear response of display systems.

$$V_{out} = AV_{in}^{\gamma} \tag{1}$$

Figure.1 illustrates the relationship between input and output luminance values after Gamma correction: the horizontal axis represents input luminance, and the vertical axis represents output luminance. The blue curve shows mapping for Gamma values < 1, while the red curve shows mapping for Gamma values > 1. With Gamma < 1, image brightness increases, enhancing contrast in low luminance areas for better detail recognition in darker regions.

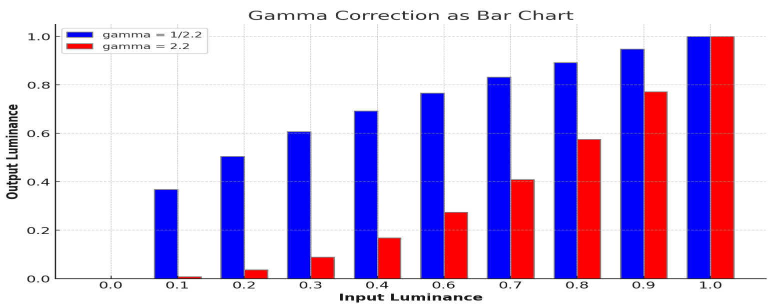

Figure 1: The relationship between input and output luminance values after Gamma correction. The blue curve shows the mapping for Gamma values less than 1, enhancing overall brightness and contrast in low luminance areas. The red curve represents Gamma values greater than 1.

In experimental data preprocessing, adjusting image saturation is crucial for color performance and visual quality. Existing methods have limitations: inconsistent results across different image types and inadequate handling of lighting variations. This study combines a novel HSV-based Gamma correction method to enhance saturation adjustment, improving visual quality and color performance. Compared to traditional methods, this approach offers enhanced robustness and applicability, achieving stable, accurate adjustments across scenarios.

The followings are images of the dataset processed in three ways: Figure.2 (a) shows the original image, Figure.2 (b) shows the image after gamma correction, and Figure.2 (c) shows the image after gamma correction based on HSV.

Figure 2: Comparison of three different processed images

3.2. Analysis of the MMRotate Framework

MMRotate[6] is a free, open-source toolkit based on PyTorch, focused on rotated bounding box detection, and is part of the OpenMMLab project. The main branch of the current version is compatible with PyTorch 1.6 and earlier versions. The toolkit features include providing multiple angle representation methods to accommodate different model configurations; it uses a modular design that decomposes the task of rotated bounding box detection into multiple modules, allowing users to easily build customized detectors; and it offers industry-leading performance with state-of-the-art algorithms and benchmark models, delivering robust support for rotated bounding box detection tasks.

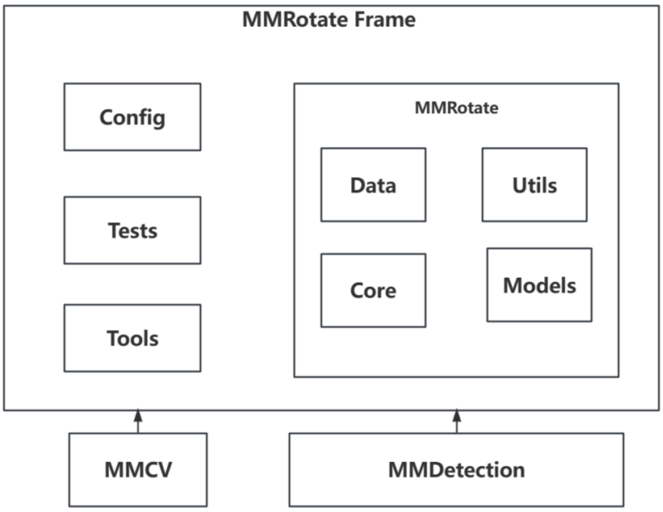

Figure 3: MMRotate framework diagram

The Figure.3 shows that MMRotate framework primarily consists of four components: datasets, models, core, and API. The dataset component handles data loading and preprocessing, including the datasets required for training, rotation frame data augmentation pipelines, and samplers for data loading. The model component is the core of the framework, encompassing rotation detection models and loss functions. The API component provides a user-friendly interface for model training, testing, and inference. Additionally, evaluation tools and custom hooks are integral parts of the model training core.

Ensuring high precision while effectively conducting rotated object detection has become a current research hotspot. This section will explore this topic from the perspectives of research approaches and definition methods.

OC[12] is a rotation angle representation method in a Cartesian coordinate system, typically using the coordinates of the upper-left and lower-right corners of a rectangular bounding box to represent the position and orientation of the target. OC is simple, intuitive, and easy to implement. It performs well for horizontal or vertical bounding boxes but may be less accurate for rotated bounding boxes as it struggles to describe rotation angles and aspect ratio changes. Suitable for simple horizontal or vertical bounding boxes or tasks with low rotation angle requirements.

LE135[13] is a length encoding method based on the direction of the target's principal axis, with the angle between the principal axis and the x-axis ranging from -45 degrees to 135 degrees. LE135 can more accurately represent large-angle rotated targets and performs better for such targets compared to OC. While effective for large-angle rotations, it may not perform well for certain angle ranges or occluded targets. Suitable for tasks requiring precise description of large-angle rotated targets, enhancing detection accuracy and robustness.

LE90[14] is a simplified form of LE135, using a length encoding range of -90 degrees to 90 degrees, making it a special case of LE135. LE90 is simpler and more intuitive in calculation and representation, suitable for scenarios with lower angle precision requirements. Due to its limited range, it may be less accurate for large-angle rotated targets compared to LE135. Suitable for scenarios with low angle precision requirements, simple calculation, and representation, especially in resource-constrained situations.

4. Experiment Setup and Results

4.1. Dataset Introduction

The DOTA dataset [15] is a comprehensive resource for remote sensing image object detection, featuring a large collection of high-resolution images from various sensors. It includes diverse object categories, scales, and complex occlusions, making it vital for advancing object detection algorithms in high-resolution imagery. This study analyzes targets like vehicles and ships across various scales in the DOTA dataset, important for military and civilian uses, to thoroughly validate detection model accuracy, Figure.4 show the details of the dataset.

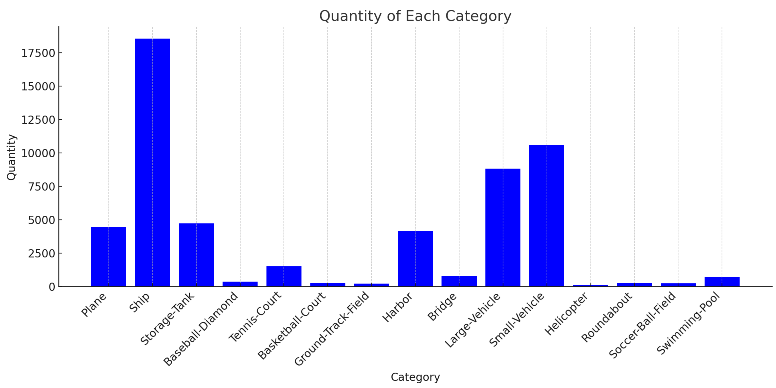

Figure 4: The number of various classes in the DOTA dataset

Additionally, this study also utilizes the HRSC dataset [16] as a supplementary resource. The HRSC dataset is a remote sensing image dataset used for ship detection and classification. The HRSC dataset comprises high-resolution aerial images from various angles and resolutions, along with detailed annotation information for ships associated with these images. The goal of this dataset is to provide a standard benchmark for the research and evaluation of algorithms for ship detection, classification, and recognition.

The dataset features aerial images of varying resolutions, taken under diverse temporal and weather conditions from platforms like satellites and drones. Additionally, the ships in the dataset are categorized into different types, including various kinds of vessels like cargo ships, fishing boats, and yachts. Each ship is annotated with detailed information, including its position, orientation, and dimensions. Every vessel is accurately marked with bounding boxes and potential directions. These annotations are crucial for training and evaluating ship detection algorithms. The dataset provides a wealth of aerial images and ship annotations, equipping researchers with ample data for studying ship detection, classification, and recognition tasks.

The HRSC dataset is a vital resource for ship detection and classification research, offering extensive image data and annotations to aid development in this field.

4.2. Experimental Design and Result Analysis

To evaluate the effects of gamma correction with HSV on object recognition, a comparative experiment was conducted, analyzing color space characteristics on original images and applying gamma correction to enhance visual quality. Specifically, the HSV color space was chosen for its effectiveness in preserving color information. The results are presented as follows.

Table 1
Comparison of dataset performance under different processing methods

Processing method	Recall	Map
Original Image	0.64	0.74
Gamma Correction	0.85	0.77
Gamma Correction based on HSV	0.90	0.82

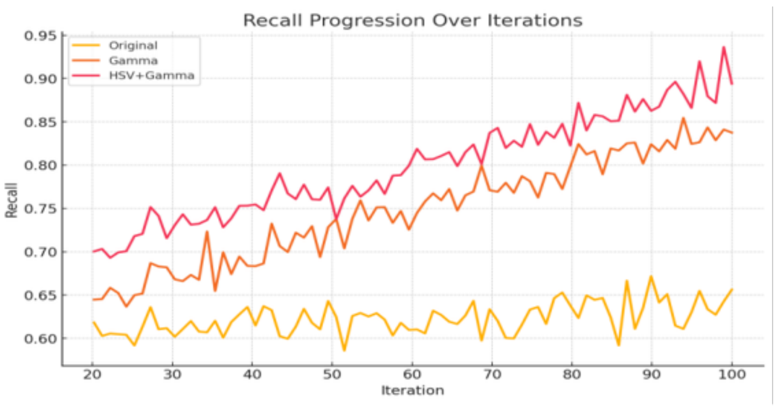

Figure 5: Experimental recall rate

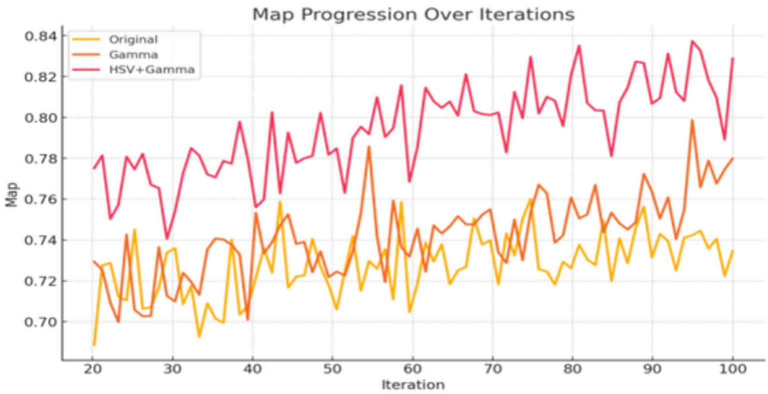

Figure 6: Experimental map rate

The analysis showed that images processed with Gamma correction had higher contrast and richer color saturation than the originals. Moreover, using Gamma correction in the HSV color space preserved original colors while improving visual quality. Table.1 presents the results of the relevant experiments, Figure.5 and Figure.6 visually demonstrate the regression rate andaccuracy of accuracy.

In object detection tasks, choosing an appropriate rotated bounding box definition is crucial for accurately identifying targets. Traditional rectangular bounding boxes may not effectively describe rotated or tilted objects, hence adopting more flexible rotated bounding box definitions could enhance recognition performance. This study utilized the MMRotate framework and replaced three different rotated bounding box definitions within the same model, including varying angle ranges, aspect ratios, and even non-rectangular shapes, to thoroughly investigate their impact on the accuracy of object recognition.

Table 2
Comparison of dataset performance under different processing methods

Rotation box definition method	Category	Map
OC	Plane	0.61
OC	Tennis-court	0.84
OC	Helicopter	0.60
LE90	Plane	0.60
LE90	Tennis-court	0.80
LE90	Helicopter	0.63
LE135	Plane	0.63
LE135	Tennis-court	0.77
LE135	Helicopter	0.68

Table.2 shows that the OC definition method cannot effectively describe the angle and shape changes of rotated objects, leading to poor detection performance for such targets. When detecting larger objects, such as sports fields and other large targets, the LE90 rotated bounding box definition method demonstrates certain advantages. This may be because the LE90 definition's rotated bounding box is better suited for capturing large, horizontal or nearly horizontal targets, thereby improving the recognition performance for these types of objects. In terms of average precision, LE90 might slightly lag behind LE135, but its advantages in specific scenarios are also noteworthy.

By experimentally comparing different rotated bounding box definition methods, we can gain a more comprehensive understanding of their performance in object detection tasks. This provides valuable reference for selecting the most suitable rotated bounding box definition for specific scenarios.

In the same model based on the MMRotate framework, studies on different datasets were conducted to explore the impact of various datasets on object recognition accuracy. Specifically, multiple datasets from different sources and with different characteristics were used, covering target images in various scenarios and environments. This diverse dataset selection aims to comprehensively evaluate the model's performance in different contexts, thereby better understanding its robustness and applicability. This experiment, in addition to the DOTA dataset, also included the HRSC dataset. The results are as follows.

Table 3
Supplementary dataset experiment

Dataset	Category	Recall	Map
DOTA	Tennis-court	0.85	0.80
	Basketball-court	0.636	0.65
	Swimming-pool	0.86	0.75
	Large-vehicle	0.74	0.71
	Plane	0.679	0.63
	Helicopter	0.909	0.62
	Roundabout	0.75	0.72
	Ship	0.354	0.52
HRSC	Ship	0.78	0.80

Table.3 shows the recognition of different categories in two datasets, Figure.7 is the experimental detection result of the DOTA dataset section , Figure.8 is the experimental detection result of the DOTA dataset section . Model performance on the DOTA dataset varies depending on the target type, with buildings and vehicles achieving higher recall and precision rates, while smaller or irregular targets may perform poorly. On the HRSC dataset, large commercial ships and warships generally exhibit better performance, whereas small vessels may result in decreased performance. The DOTA dataset encompasses a wider range of target types, thereby enhancing the model's generalization capabilities; in contrast, the HRSC dataset focuses on ship detection and performs better in high-resolution images.

(a) Ship inspection (b) Traffic vehicle detection

Figure 7: DOTA dataset correlation detection graph

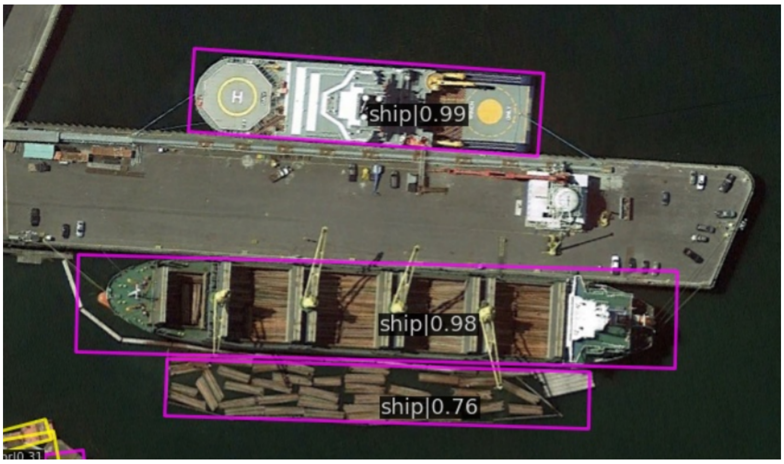

Figure 8: HRSC dataset correlation detection graph

5. Conclusion

In the paper, we used the HSV-based Gamma correction method to process images, which effectively improved visual quality and enhanced both color contrast and saturation. According to the experimental results, the mean average precision (mAP) of the original dataset is 0.74, while the mAP of the dataset processed with Gamma correction in the HSV color space is 0.82. Using the MMRotate framework, we conducted comparative experiments under three different rotated bounding box definitions to explore the effects of different conditions on rotated object recognition. Moreover, by integrating the HRSC dataset with the DOTA dataset, this study not only enriches the theoretical and practical aspects of the field but also explores the potential of these image processing techniques for object recognition and processing in the Internet of Things (IoT) environment. This provides new methods and insights for enhancing the performance of devices and services.

In the future, due to the imbalance in the number of samples across different categories in remote sensing image datasets, we aim to enhance the samples for categories with fewer instances to achieve a more balanced distribution. Additionally due to the limitations of the experimental environment, we were unable to conduct experiments with more models. We will strive to incorporate additional models to enrich the experimental data.

References

[1] S. Borde, C. Patil, C. Sonawane, M. Singh, Object recognition based on deep learning algorithms using embedded iot with interactive interface, in: 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), 2023, pp. 1581–1586. doi:10.1109/ICICCS56967.2023.10142821.

[2] N. Shahid, S. Aneja, Internet of things: Vision, application areas and research challenges, in: 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2017, pp. 583–587. doi:10.1109/I-SMAC.2017.8058246.

[3] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, M. Mustaqim, Internet of things (iot) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5g-iot scenarios, IEEE Access 8 (2020) 23022–23040. doi:10.1109/ACCESS.2020.2970118.

[4] A. S. Tarawneh, A. Hassanat, I. Elkhadiri, D. Chetverikov, K. Almohammadi, Automatic gamma correction based on root-mean-square-error maximization, 2020 International Conference on Computing and Information Technology (ICCIT-1441) (2020) 1–5. URL: https://api.semanticscholar.org/CorpusID:227220297.

[5] C. Junhua, L. Jing, Research on color image classification based on hsv color space, in: 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control, 2012, pp. 944–947. doi:10.1109/IMCCC.2012.226.

[6] Y. Zhou, X. Yang, G. Zhang, et al., Mmrotate: A rotated object detection benchmark using pytorch, in: Proceedings of the 30th ACM International Conference on Multimedia, 2022, pp. 7331–7334.

[7] M. A. K. Raiaan, N. M. Fahad, S. Chowdhury, D. Sutradhar, S. S. Mihad, M. M. Islam, Iot-based object-detection system to safeguard endangered animals and bolster agricultural farm security, Future Internet 15 (2023). URL: https://doi.org/10.3390/fi15120372. doi:10.3390/fi15120372.

[8] M. K. Hebbar, P. K. Pullela, Object recognition system for the visually impaired: Leveraging iot and remote server integration, in: 2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM), 2023, pp. 1–7. doi:10.1109/IC-RVITM60032.2023.10434991.

[9] J. Cai, H. Zeng, H. Yong, Z. Cao, L. Zhang, Toward real-world single image super-resolution: A new benchmark and a new model, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 3086–3095. doi:10.1109/ICCV.2019.00318.

[10] X. Feng, X. Yao, G. Cheng, J. Han, Weakly supervised rotation-invariant aerial object detection network, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition

(CVPR), New Orleans, LA, USA, 2022, pp. 14126–14135. doi:10.1109/CVPR52688.2022.01375.

[11] J. Wang, L. Song, Z. Li, et al., End-to-end object detection with fully convolutional network, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 15849–15858.

[12] W. Li, R. Shang, Z. Ju, J. Feng, S. Xu, W. Zhang, Ellipse iou loss: Better learning for rotated bounding box regression, IEEE Geoscience and Remote Sensing Letters 21 (2024) 1–5. doi:10.1109/LGRS.2023.3345881.

[13] Y. Pu, Y. Wang, Z. Xia, Y. Han, Y. Wang, W. Gan, Z. Wang, S. Song, G. Huang, Adaptive rotated convolution for rotated object detection, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 6589–6600.

[14] X. Yang, X. Yang, J. Yang, Q. Ming, W. Wang, Q. Tian, J. Yan, Learning high-precision bounding box for rotated object detection via kullback-leibler divergence, Advances in Neural Information Processing Systems 34 (2021) 18381–18394.

[15] G.-S. Xia, et al., Dota: A large-scale dataset for object detection in aerial images, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3974–3983. doi:10.1109/CVPR.2018.00418.

[16] Z. Liu, L. Yuan, L. Weng, Y. Yang, A high resolution optical satellite image dataset for ship recognition and some new baselines, in: Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM), Porto, Portugal, 2017.