




Article

An Approach to Large-Scale Cement Plant Detection Using Multisource Remote Sensing Imagery

Tianzhu Li ^{1,2}, Caihong Ma ^{1,*} , Yongze Lv ^{1,3}, Ruilin Liao ^{1,3}, Jin Yang ¹ and Jianbo Liu ¹¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China² University of Chinese Academy of Sciences, Beijing 100094, China³ College of information, Beijing Forestry University, Beijing 100083, China

* Correspondence: mach@aircas.ac.cn

Abstract: The cement industry, as one of the primary contributors to global greenhouse gas emissions, accounts for 7% of the world's carbon dioxide emissions. There is an urgent need to establish a rapid method for detecting cement plants to facilitate effective monitoring. In this study, a comprehensive method based on YOLOv5-IEG and the Thermal Signature Detection module using Google Earth optical imagery and SDGSAT-1 thermal infrared imagery was proposed to detect large-scale cement plant information, including geographic location and operational status. The improved algorithm demonstrated an increase of 4.8% in accuracy and a 7.7% improvement in MAP@.5:95. In a specific empirical investigation in China, we successfully detected 781 large-scale cement plants with an accuracy of 90.8%. Specifically, of the 55 cement plants in Shandong Province, we identified 46 as operational and nine as non-operational. The successful application of advanced models and remote sensing technology in efficiently and accurately tracking the operational status of cement plants provides crucial support for environmental protection and sustainable development.

Keywords: cement plant; SDGSAT-1; YOLOv5-IEG; remote sensing; Google Earth



Citation: Li, T.; Ma, C.; Lv, Y.; Liao, R.; Yang, J.; Liu, J. An Approach to Large-Scale Cement Plant Detection Using Multisource Remote Sensing Imagery. *Remote Sens.* **2024**, *16*, 729. <https://doi.org/10.3390/rs16040729>

Academic Editors: Gemine Vivone, Hossein M. Rizeei, Qi Zhao, Guangliang Cheng and Paolo Tripicchio

Received: 23 December 2023

Revised: 5 February 2024

Accepted: 7 February 2024

Published: 19 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The cement industry is a primary contributor to greenhouse gas emissions, particularly in terms of CO₂ emissions [1,2]. This is primarily due to the calcination of raw materials used in cement production and the fuel required to maintain high temperatures in the kiln [3]. Additionally, the cement industry plays a crucial role in the global economy, as it is utilized in the production of various infrastructure and construction projects [4,5]. In recent years, a primary objective of the global environmental agenda has been to curtail emissions to safeguard the Earth's climate patterns [6]. The escalating trend in atmospheric emissions necessitates the design and implementation of policies to address the challenges posed by climate change [7–10]. The National Development and Reform Commission (NDRC) of China has recently introduced several policies [11] aimed at mitigating environmental pollution and greenhouse gas emissions from the cement industry. Consequently, there is an urgent need to expedite the identification and monitoring of cement plants in China for effective management.

Traditional methods for monitoring cement plants heavily rely on manual on-site investigations and the statistical analysis of limited archival reports. However, these approaches suffer from narrow monitoring scope, low frequencies, and limited effectiveness in achieving comprehensive monitoring of extensive industrial zones [12]. With the evolution of remote sensing technology, applications based on remote sensing data have become increasingly prevalent. It is noteworthy that there have been some studies focusing on the detection of industrial heat sources, with researchers proposing methods that integrate the use of hotspot data to identify high-energy-consuming industrial sources. For example, Liu et al. [13] introduced a “time-space-temperature” approach, utilizing VIIRS

750-m nighttime fire (VNF) data to identify industrial heat sources. Additionally, this study presented temperature information resembling fingerprints for the classification of industrial heat source objects. Ma et al. [14] employed an improved adaptive K-means algorithm to identify industrial heat sources in China, successfully extracting 4771 such objects in the Chinese region. These approaches leverage hotspot clustering to detect industrial heat sources, mitigating the challenges associated with time-consuming on-site investigations. However, it is important to note that these methods struggle to distinguish between different types of industrial heat sources and face difficulties in obtaining precise location information.

Deep learning technologies, particularly Convolutional Neural Networks (CNNs), have advanced, and coupled with continuous progress in remote sensing image processing, their applications have yielded notable achievements in various domains, such as object detection [15,16], image segmentation [17], and speech processing [18]. Models designed for object detection tasks, such as RCNN [19], Faster-RCNN [20], YOLO series [21–24], and SSD [25], have been applied in a wide range of scenarios. Scholars have attempted to apply these models to object detection tasks in remote sensing imagery [16]. However, most research has primarily focused on typical objects, such as airplanes [26,27], ships [28], and vehicles [29], with relatively few studies addressing complex industrial sites like cement plants. Some researchers have employed deep learning techniques to detect industrial heat sources. Lu et al. [30], for instance, based their work on the deep learning object detection network, SSD. They constructed a deep learning object detection model suitable for GF-1 remote sensing imagery of steel plants. The model incorporated a maxout module, optimizing the negative sample pathway into a multi-branch structure, ultimately achieving a detection accuracy of over 80%. Another notable contribution was made by Xu et al. [31], who utilized the Faster RCNN deep learning framework and ResNet feature network to detect cement plant targets in satellite maps. They further enhanced the model's capabilities through image preprocessing, data augmentation, and the inclusion of negative sample training methods, achieving a 94% detection rate on an expanded test set. Additionally, researchers have employed large pretrained models to extract named entities, capturing global industrial heat sources [32]. While these studies have made progress in identifying different categories of industrial heat sources, they also demonstrate certain limitations. The primary challenges in this field currently revolve around two key issues: first, the construction of a dataset presents a significant obstacle due to the lack of publicly available datasets for training; second, the intricate and complex background of cement plants contributes to lower accuracy in these studies.

To the best of our knowledge, the YOLO family has found widespread application in the realm of target detection within remote sensing imagery [33,34]. Notably, YOLOv5 has gained considerable attention among researchers due to its robust functionality [35,36]. However, the literature reveals a scarcity of investigations focusing on the application of the YOLO family to the detection of industrial heat sources. This scarcity prompted our consideration of adopting YOLOv5 as the baseline model.

Furthermore, traditional optical remote sensing imagery typically provides position information about cement plants but falls short in intuitively reflecting their operational status. In the multifaceted process of cement production, the generation of clinker is the most energy-intensive and emission-intensive phase [37,38]. During this pivotal stage, the raw material mixture is conveyed into a kiln heated to over 900 °C, initiating the transformation of limestone (CaCO_3) into lime (CaO) and CO_2 [5]. This transformative process, known as calcination, subjects the raw materials to temperatures as high as 1450 °C, culminating in the formation of clinker. The extreme temperatures involved in this procedure yield distinctive thermal signatures, observable through thermal infrared imagery [32]. Hence, the integration of thermal infrared imagery provides a more comprehensive understanding of the thermal features of cement plants, facilitating the assessment of a plant's operational status.

In this study, high-resolution Google Earth optical imagery and SDGSAT-1 thermal infrared imagery were selected as experimental data. A comprehensive method based on YOLOv5-IEG and the Thermal Signature Detection (TSD) module using Google Earth optical imagery and SDGSAT-1 thermal infrared imagery was proposed to achieve large-scale cement plant detection. Through the combination of position information in optical imagery and thermal information in thermal infrared imagery, a more comprehensive understanding of the location and operational status of cement plants can be obtained. The methodology proposed in this study introduces a novel monitoring approach for cement plants, characterized by higher accuracy and broader detection coverage compared to traditional methods. Through an in-depth study of the unique environment of cement plants, this research seeks to provide specific and practical insights into the application of deep learning. The intention is to drive the widespread adoption of this technology in the field of industrial monitoring, offering the potential for more comprehensive, efficient, and accurate detection methods for factory surveillance tasks.

The main contributions of this paper can be summarized as follows:

1. We introduced a large-scale cement plant detection method that uses the YOLOv5-IEG model, achieving a detection and localization approach based on remote sensing imagery.
2. We established a monitoring model for the operational status of cement plants, leveraging SDGSAT-1 thermal infrared imagery.
3. A dataset of cement plants in China was created with higher accuracy than other available datasets.

2. Materials and Methods

2.1. Study Area

In this experiment, China (no data for Hong Kong, Macao, and Taiwan) was selected as the study area, as shown in Figure 1. Driven by the continuous growth in infrastructure and construction project demand, the cement industry has experienced significant expansion, particularly in China, India, and Southeast Asia. China is the world's largest producer and consumer of cement [11,39,40]. However, the industry's growth in China has been accompanied by environmental challenges such as air pollution and deforestation. Due to the difficulty in obtaining timely and comprehensive data on the construction status of cement plants, attributed to challenges in human resource utilization, China (no data for Hong Kong, Macao, and Taiwan) was chosen as the study area for this experiment.

2.2. Experimental Dataset

The LabelImg toolbox was employed in this study for manual annotation of samples in each cement plant image, generating files suitable for model training. In total, this study created a dataset comprising 1972 images (Figure 2; data available from the corresponding author), with a training-to-validation set ratio of 8:2. The positional offsets of the detection boxes generated by the network are depicted in Figure 3a, where x , y , width, and height represent the displacement values of the predicted box's center coordinates relative to the true coordinates. As evident from Figure 3b, the majority of detected boxes fall within the range of 0.20~0.4, indicating that the detected targets are predominantly of medium to small sizes. Notably, the prior boxes for the cement plant class generally exhibit larger sizes, and there is minimal variation in aspect ratios among cement plants. Furthermore, during the training phase, as our model automatically segments non-target regions within the input imagery as negative samples, there is no necessity to introduce negative samples into the training dataset.

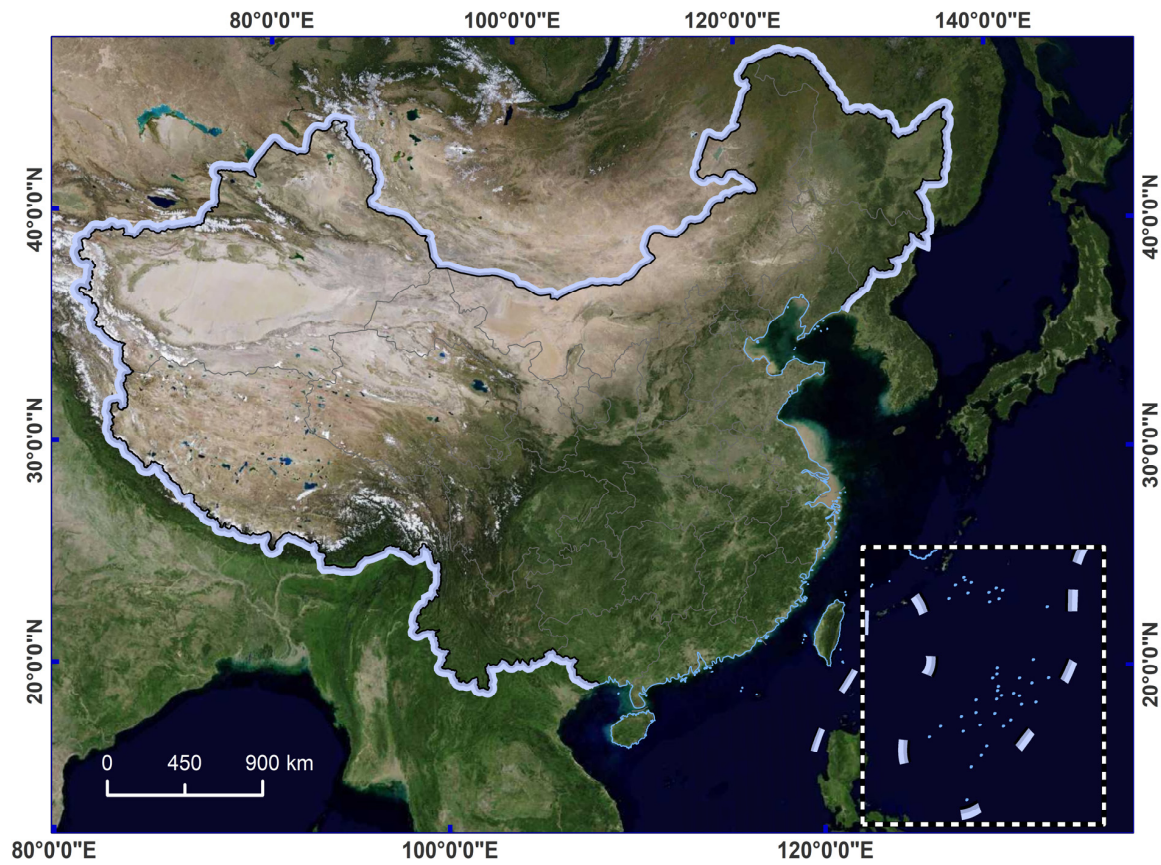


Figure 1. Geographical location of the study area.



Figure 2. Illustration of partial dataset.

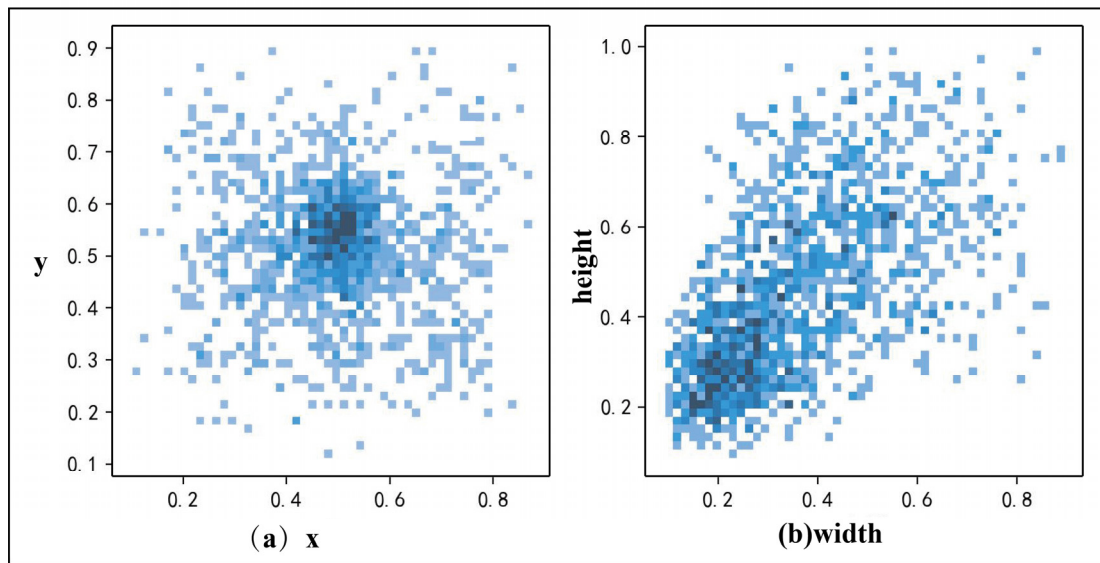


Figure 3. Distribution statistics of the training set.

2.3. Technical Route

To leverage the distinctive features of diverse remote sensing imagery, encompassing optical imagery for target detection and thermal infrared imagery for temperature analysis, we present a comprehensive method based on YOLOv5-IEG and the TSD module to detect large-scale cement plants. This approach aims to provide a more thorough understanding of the spatial distribution and operational status of cement plants. The methodology comprises two main components. In the first part, we employ the YOLOv5-IEG model, utilizing high-resolution Google Earth imagery with a spatial resolution of 2 m, to detect the spatial position information of cement plants. The second part introduces the TSD module, utilizing SDGSAT-1 thermal infrared imagery, proposed for the analysis of thermal signatures in cement plants. The technical framework of this method is illustrated in Figure 4, encompassing five key processes: data preprocessing, dataset construction, cement plant detection using Google Earth imagery, operational status monitoring of cement plants, and comparative analysis.

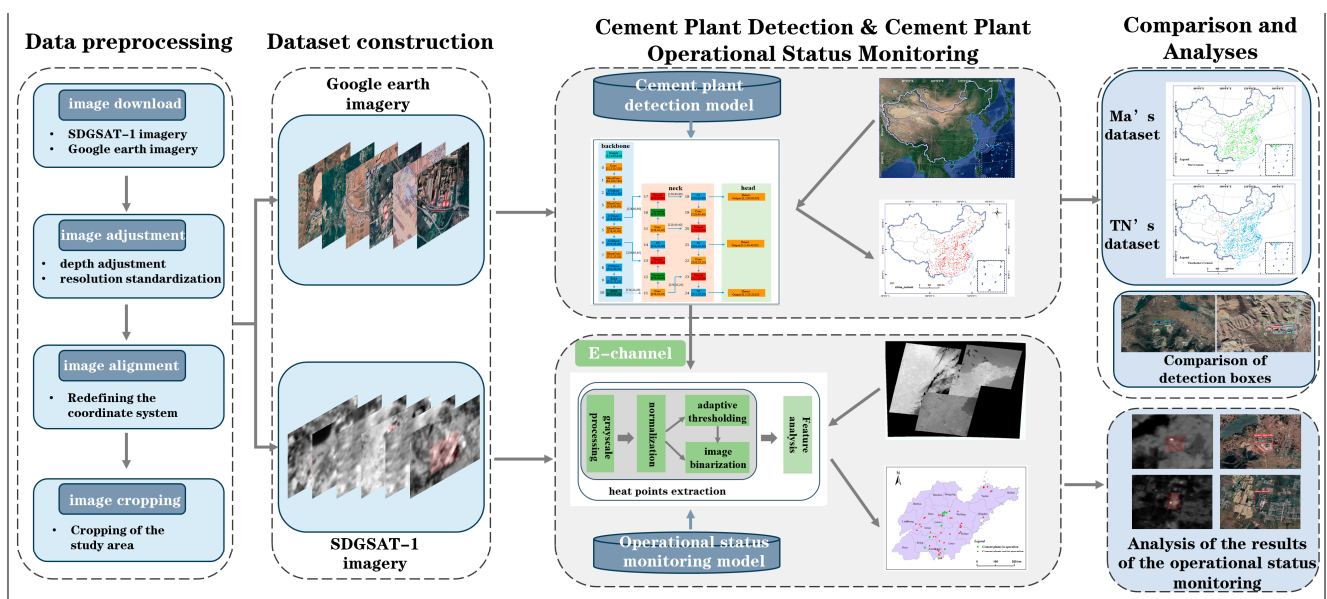


Figure 4. Technical framework of the proposed method.

1. Remote Sensing Image Preprocessing

Initially, high-resolution satellite imagery (levels 14–16) of the study area was obtained from Google Earth’s website (<https://earth.google.com/web/>, accessed on 1 August 2023). The imagery specific to cement plant areas was then extracted, forming the cement plant image dataset. Additionally, high-quality 2023 thermal infrared imagery with a resolution of 30 m was acquired from the SDGSAT-1 data website (<http://124.16.184.48:6008/home>, accessed on 1 August 2023). To facilitate matching of the two image types, preprocessing steps included depth adjustment and resolution standardization for Google Earth imagery, along with cropping procedures for the study area. Similar preprocessing steps, such as depth adjustment, resolution standardization, image mosaic, and region-specific cropping, were applied to the SDGSAT-1 thermal infrared imagery.

2. Dataset Construction

This study utilized two distinct datasets corresponding to the two types of remote sensing imagery. The first dataset, sourced from Google Earth imagery, served the purpose of spatial localization and detection of cement plants. The second dataset comprised SDGSAT-1 thermal infrared imagery, crucial for monitoring the operational status of cement plants. For the construction of the first dataset, existing cement plant information was gathered from various sources, and corresponding datasets, including geographical coordinates, were downloaded. The data were then cropped to a standardized size of 1200×800 pixels, resulting in a total of 1861 global cement plant images. Regarding the SDGSAT-1 thermal infrared dataset, imagery from multiple provinces within China with favorable data quality since 2023 was downloaded and cropped based on cement plant geographical coordinates, yielding the SDGSAT-1 thermal infrared dataset.

3. Cement Plants Detection Using Google Earth Imagery

To achieve rapid detection of cement plant geographical coordinates, YOLOv5 served as the foundational model in this study. Further enhancements were incorporated, including a fusion of the ghost module with the CSP module to improve detection speed. To enhance detection accuracy, an EMA module was inserted after the CSPGhost module, reinforcing meaningful channel information and suppressing background feature interference. These adaptations collectively enabled the swift and accurate detection of cement plants.

4. Cement Plant Operational Status Monitoring

Following the detection phase, thermal monitoring of the cement plants’ operational status was conducted by using the TSD module. Subsequent to the model’s prediction, a network linking layer was added to establish an “e-channel” between the detection results (bounding boxes) from the Google Earth imagery and the SDGSAT-1 thermal infrared imagery. This fusion allowed the reflection of the thermal status of cement plants through infrared imagery. The feedback loop culminated in the final output, encompassing both position information and operational status.

6. Comparative Analysis

To assess the detection results, we conducted comparative analyses by using several publicly available datasets in this study. Our evaluation focused on comparing the number of detected cement plants and the accuracy of the detection frames.

2.4. YOLOv5-IEG Algorithm

In this study, we propose a cement plant detection method (YOLOv5-IEG). The backbone network of YOLOv5 [24] consists of the CSP module, Focus structure, and SPP pyramid structure. The CSP module reduces redundant gradient calculations by segmenting gradients, the Focus structure effectively preserves crucial information, and the SPP pyramid structure enables the network to adapt to multi-scale inputs, strengthening the connection of contextual features. Inspired by the lightweight GhostNet [41] architecture, for the purpose of lightweighting the detection model, we optimized the CSP module in the backbone network to a CSPGhost module. This involved utilizing Ghost modules to

replace redundant feature maps, enhancing detection speed. To ensure detection accuracy, we inserted an Efficient Multi-scale Attention mechanism after the last CSPGhost module, establishing a structure with both feature grouping and multi-scale capabilities. This effectively establishes short-term and long-term dependencies, suppresses background feature interference, and optimizes network performance. During the prediction phase of the network, the Inner-IoU loss function was employed as the bounding box loss. This choice addresses the limitations of existing IoU loss functions [42] in detection tasks, enhancing the precision of detection with improved generalization capabilities and faster convergence. The specific architecture of the YOLOv5-IEG network is depicted in Figure 5.

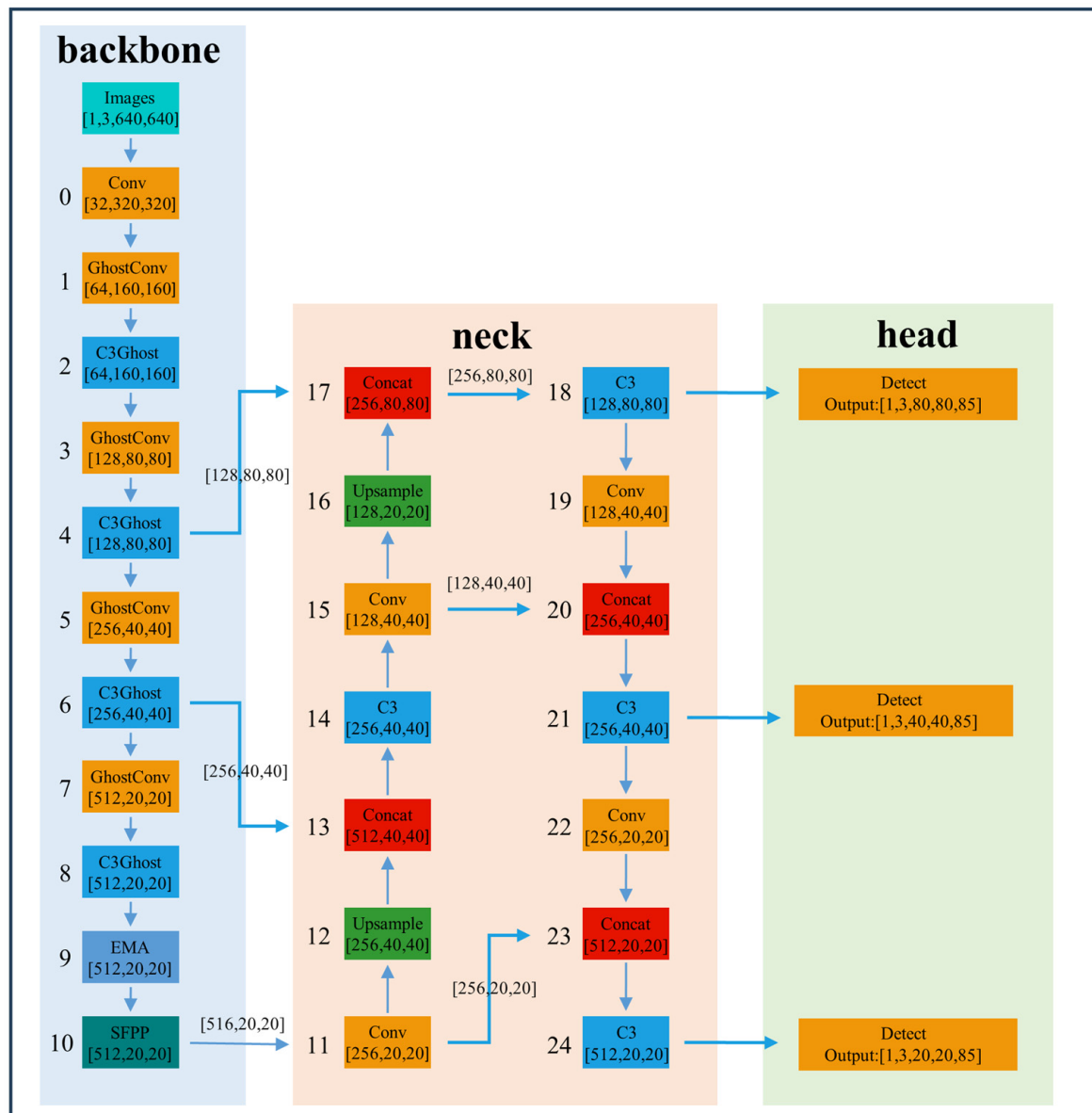


Figure 5. YOLOv5-IEG's architecture.

To improve visualization, the model has been streamlined into the flowchart depicted in Figure 6. YOLOv5-IEG comprises three main components: input, backbone network, and detection results. At the input stage, frames of the imagery to be detected are processed through operations such as feature extraction and feature fusion, ultimately yielding classifications and regression bounding box positions for the identified targets. Image preprocessing involves Mosaic data augmentation, adaptive scaling, and adaptive anchor box computation, strategically implemented to reduce computational overhead. At the

forefront of the backbone network, downsampling is achieved through the Focus structure, involving image segmentation and reassembly to diminish network parameters and augment local receptive fields.

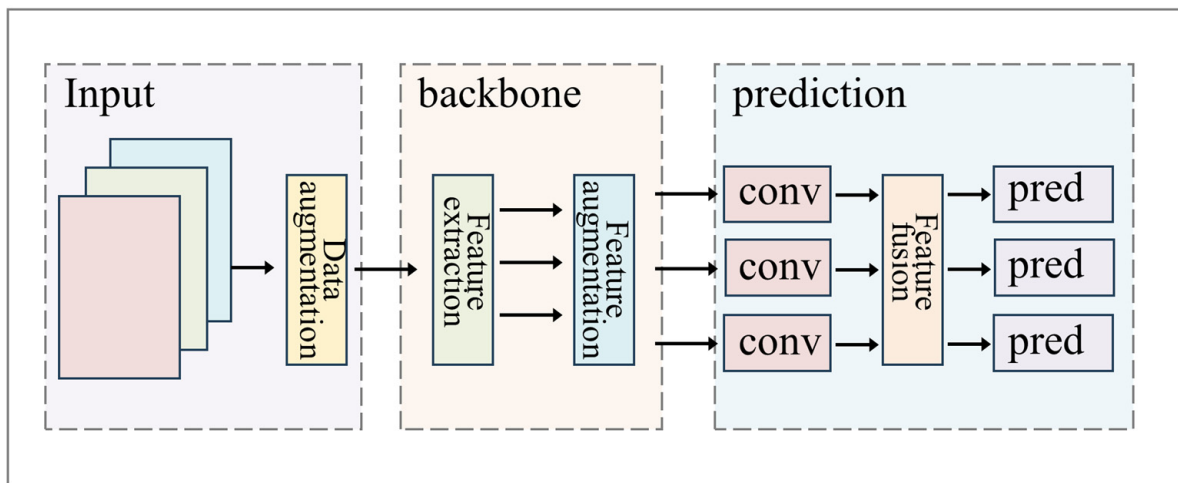


Figure 6. Simplified flowchart.

2.4.1. Efficient Multi-Scale Attention (EMA)

With the evolution of deep Convolutional Neural Networks (CNNs), attention-grabbing network topologies have been applied in the domains of image classification and object detection tasks. When expanding neural networks to multiple convolutional layers, they exhibit a remarkable ability to enhance learning feature representations. However, this entails stacking more deep convolutional methods, consuming significant memory and computational resources, which is a major drawback in constructing deep CNNs. As an alternative, attention mechanism methods, due to their flexible structural features, not only strengthen the learning of more discriminative feature representations but can also be easily inserted into the backbone architecture of neural networks. Consequently, attention mechanisms have garnered significant interest in the computer vision research community.

To better capture cross-spatial feature information of cement plants across different spatial dimensions and achieve more comprehensive feature aggregation, this study introduces Efficient Multi-scale Attention (EMA) [43] into the model. This module consists of only two convolutional kernels placed in parallel subnetworks. One parallel subnetwork is a 1×1 convolutional kernel, processed similarly to Coordinate Attention (CA), and the other is a 3×3 convolutional kernel. However, the CA (Coordinate Attention) module can be considered a similar method to the Squeeze-and-Excitation (SE) attention module [44], where global average pooling is used to model cross-channel information. Typically, channel statistics can be generated by using global average pooling, compressing global spatial position information into channel descriptors. Unlike SE, CA embeds spatial position information into the channel attention map to enhance feature aggregation. The parallel substructure helps the network avoid more sequential processing and large depths. The EMA module adopts a parallel substructure, as shown in Figure 7. The EMA module selects the shared components of the 1×1 convolutional kernel from the CA module and names it the 1×1 branch in EMA. To aggregate multi-scale spatial structural information, the 3×3 kernel is placed in parallel with the 1×1 branch for quick response and is named the 3×3 branch. This feature grouping and multi-scale structure effectively establish short-term and long-term dependencies, contributing to improved performance.

where B and B^{gt} represent the predicted box and the ground truth box, respectively. After defining IoU , the corresponding loss function can be expressed as follows:

$$L_{IoU} = 1 - IoU \quad (2)$$

To date, IoU-based loss functions have become mainstream and dominant, with many approaches extending IoU and adding new loss terms. For example, GIoU [46] addresses the gradient vanishing problem when the overlap between anchor boxes and GT boxes is zero; DIOU [47] introduces a new distance loss term; EIoU [42] minimizes the normalized differences in width, height, and center position between the target box and anchor box directly; and SloU [48] introduces an angle loss considering the impact of the angle between anchor boxes and GT boxes on bounding box regression. These bounding box regression loss functions can accelerate convergence and improve detection performance by adding new geometric constraints to the IoU loss function. However, they do not consider the inherent reasonableness of the IoU loss itself, which determines the quality of detection results.

To address the weaknesses of existing IoU loss functions in terms of weak generalization capabilities and slow convergence speed across different detection tasks, Zhang et al. [49] proposed Inner-IoU loss in 2023. This loss leverages auxiliary bounding boxes to calculate IoU loss and expedites the bounding box regression process. In Inner-IoU, a scale factor ratio is introduced to control the size of auxiliary bounding boxes. By using auxiliary bounding boxes of different scales for different datasets and detectors, the limitations in generalization capabilities of existing methods can be overcome. Inner-IoU loss is applied to existing IoU-based bounding box regression loss functions, such as $L_{Inner-IoU}$, $L_{Inner-GIoU}$, $L_{Inner-DIoU}$, $L_{Inner-CIoU}$, $L_{Inner-EIoU}$, and $L_{Inner-SIoU}$, as shown in Equations (3)–(8):

$$L_{Inner-IoU} = 1 - IOU^{inner} \quad (3)$$

$$L_{Inner-GIoU} = L_{GIoU} + IoU - IOU^{inner} \quad (4)$$

$$L_{Inner-DIoU} = L_{DIoU} + IoU - IOU^{inner} \quad (5)$$

$$L_{Inner-CIoU} = L_{CIoU} + IoU - IOU^{inner} \quad (6)$$

$$L_{Inner-EIoU} = L_{EIoU} + IoU - IOU^{inner} \quad (7)$$

$$L_{Inner-SIoU} = L_{SIoU} + IoU - IOU^{inner} \quad (8)$$

In this study, the *Inner-IoU* loss function will be incorporated into the YOLOv5 model, addressing the weaknesses of existing IoU loss functions in terms of weak generalization capabilities and slow convergence speed in detection tasks. This addition aims to enhance the generalization capabilities and improve the convergence speed of the model for the cement plant detection task in this study.

2.5. Cement Plant Operational Status Monitoring Model

This study aims to monitor the operational status of cement plants objectively. Upon detecting cement plants, the integration of SDGSAT-1 thermal infrared imagery is explored to monitor the operational status of the cement plants. After overlaying the thermal infrared image, the actively operating areas of the cement plant show clear thermal signatures in contrast to the non-operational areas, as shown in Figure 8.

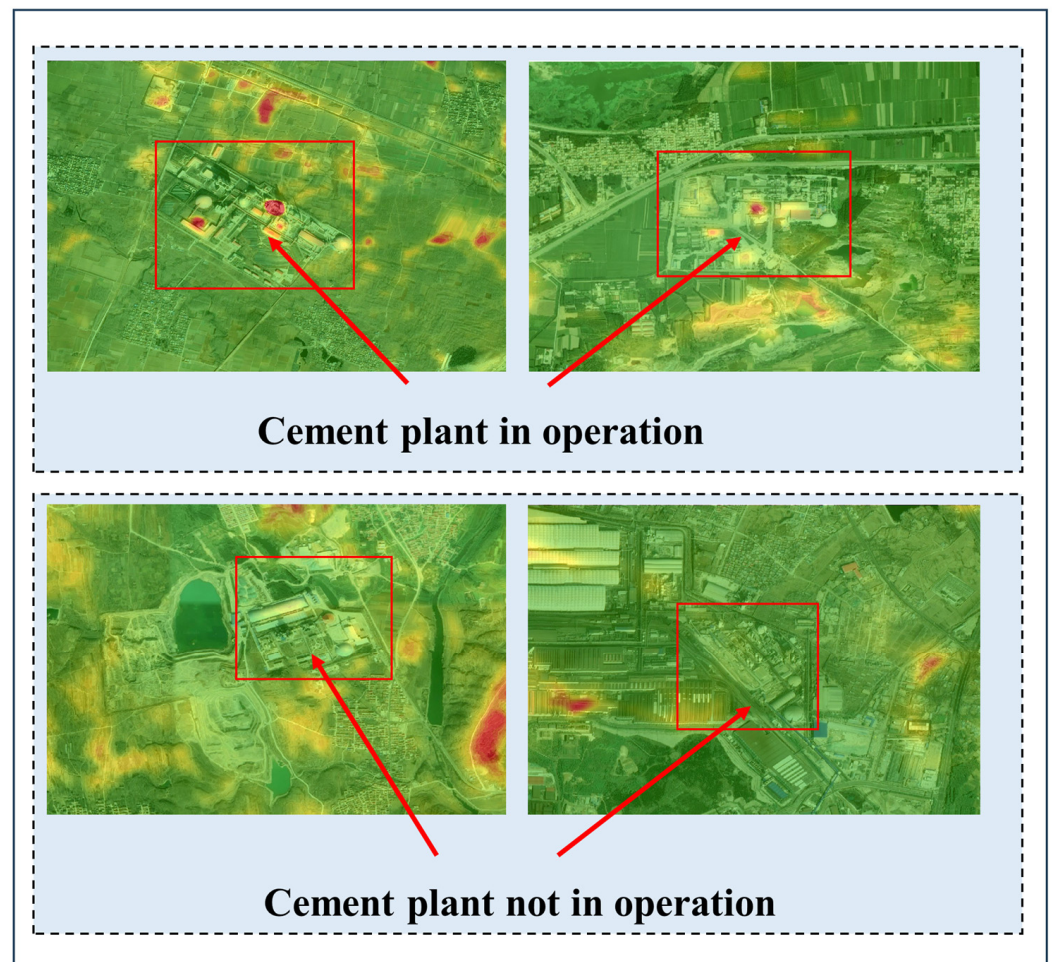


Figure 8. Diagram cement plant thermal signatures (red indicates high temperature, green indicates low temperature).

In this study, we have developed a Thermal Signature Detection (TSD) module, as illustrated in Figure 9, which is seamlessly integrated with the YOLOv5-IEG model through the E-channel. The YOLOv5-IEG model is responsible for discerning the precise location information of the cement plant, while the TSD module excels in evaluating the thermal signatures of the plant. This comprehensive approach enables the determination of the operational status of the cement plant. The detection process is elucidated in Figure 6, encompassing the following pivotal steps:

1. The YOLOv5-IEG model is employed to detect the precise location information of the cement plant.
2. The TSD module is utilized to identify thermal signature information within the SDGSAT-1 thermal infrared imagery.
3. Integration of the location information and thermal signature information is achieved through the E-channel, enabling a comprehensive assessment to determine the operational status of the cement plant.
4. As the TSD module relies on location information obtained from target detection models and requires no training, it can be seamlessly integrated with other target detection models without compromising their accuracy.

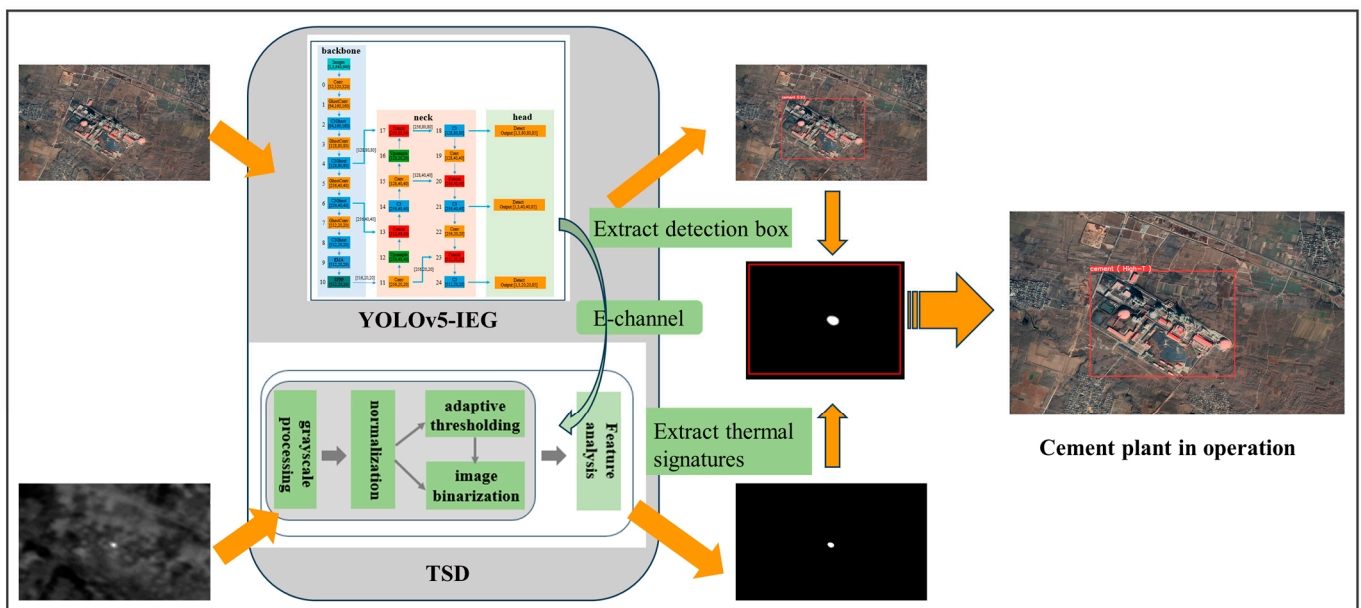


Figure 9. Flowchart of cement plant operational status monitoring.

3. Experiment and Results

3.1. Experimental Settings

The method proposed in this research, alongside comparative models, was executed within a Python 3.7 environment that employed the PyTorch-1.10.0 framework. The experiments themselves took place on a workstation equipped with two NVIDIA RTX 3090 GPUs. To maintain consistency throughout the training process, fixed hyperparameters were utilized, including the number of training epochs (set to 300), batch size (set to 48), and initial learning rate (set to 0.1). Furthermore, the PyTorch 1.10.0 framework's Stochastic Gradient Descent (SGD) optimizer was employed for optimizing the model parameters. To ensure rapid model convergence and prevent local minima, the momentum parameter was set to 0.8. Additionally, a weight decay of 0.0005 was implemented to mitigate model overfitting.

3.2. Accuracy Evaluation Method

In object detection tasks, the evaluation criteria for algorithm performance primarily involve classification accuracy and the precision of bounding box regression. The correctness of the assigned class is determined based on the confidence score, while the assessment of positional accuracy relies on whether the Intersection over Union (IoU) between the predicted and ground truth bounding boxes exceeds a predefined threshold. IoU quantifies the overlap between the predicted and annotated boxes and calculated as in Equation (1), where B represents the area of the predicted box, and B^{gt} is the area of the ground truth box.

The Confusion Matrix (CM) is employed to analyze the classification performance of a model, providing an intuitive representation of the extent to which each class is confused and the corresponding proportions. Each row of the matrix represents the true class of the samples, while each column represents the predicted class. Larger values on the diagonal indicate a higher probability of correct classification, reflecting better overall performance of the model.

Taking a binary classification confusion matrix as an example, as shown in Table 1, "Positive" and "Negative" denote sample labels, where "Positive" is the label for the target to be detected, and "Negative" is the label for the background. "True" and "False" represent the prediction results, where "True" signifies a correct prediction, and "False" denotes an incorrect prediction. Thus, TP (True Positive) signifies correct classification, FP (False Positive) represents the prediction of a negative sample as a positive instance, FN (False Negative) indicates the prediction of a positive sample as a negative instance, leading to a

missed detection, and TN (True Negative) denotes the prediction of a negative sample as negative, which is not considered in object detection tasks.

Table 1. Schematic diagram of confusion matrix.

Ground Truth \ Prediction	Positive	Negative
	Positive	True Positive (TP)
Negative	False Positive (FP)	True Negative (TN)

In object detection, commonly used metrics to assess algorithm performance are precision, recall, average precision (AP), and mean average precision (mAP). The specific meanings of these metrics are as follows:

Precision: Precision, also known as positive predictive value, represents the proportion of true positive predictions among all predictions of positive instances. It measures the accuracy of positive predictions. The calculation is given by the following formula:

$$P = \frac{TP}{TP + FP} \quad (9)$$

Recall: Recall, also known as sensitivity, indicates the proportion of true positive predictions among the total actual positive samples, reflecting the extent of missed detections. It is calculated as follows:

$$R = \frac{TP}{TP + FN} \quad (10)$$

Average Precision (AP): Precision and recall are complementary performance metrics, each having limitations as single-point values. To balance these two metrics, the mAP metric is introduced. The precision–recall (P–R) curve is constructed with precision as the y -axis and recall as the x -axis. The area under the P–R curve, denoted as AP, represents the performance for a specific class. The mAP is the mean of AP values across all classes in the dataset. The calculation is given by the following formula:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (11)$$

where N is the number of classes. For mAP at Different IoU Thresholds (e.g., mAP@0.5:0.95), the IoU (Intersection over Union) threshold is typically set to 0.5, indicating that predictions with IoU greater than 0.5 are considered valid. mAP can be calculated at different IoU thresholds, such as from 0.5 to 0.95 with an interval of 0.05. For example, mAP@0.5:0.95 represents the average mAP calculated over this IoU range.

3.3. Experimental Results

Figure 10 and Table 2 present the qualitative and quantitative metrics, respectively, for the YOLOv5-IEG model on the cement plant dataset. The detection results on the cement plant dataset demonstrate the effectiveness of our proposed method in achieving accurate detection in complex background environments. YOLOv5-IEG achieves precision, recall, mAP@0.5 (%), and mAP@0.5:0.95 (%) scores of 96.8%, 93.7%, 96.9%, and 68.8%, respectively.

Table 2. Performance of the model on the test set.

P (%)	R (%)	mAP@.5 (%)	mAP@.5:.95 (%)
96.8	93.7	96.9	68.8

As illustrated in Figure 10, the red boxes represent the detection boxes, and the numbers indicate the confidence scores. In the experiments, a minimum confidence threshold of

0.80 was set, meaning confidence scores exceeding 0.8 are considered as cement plants. It is evident from the results that the proposed model accurately detects the positions of cement plants in diverse backgrounds, including mountainous regions, deserts, and industrial areas. These achievements can be attributed to the functionality of the EMA module, which effectively integrates cross-spatial feature information of cement plants in different spatial dimensions, enabling richer feature aggregation. Additionally, the Inner-IoU loss function enhances the model's convergence speed and generalization ability, enabling the rapid detection of cement plant targets of varying sizes.

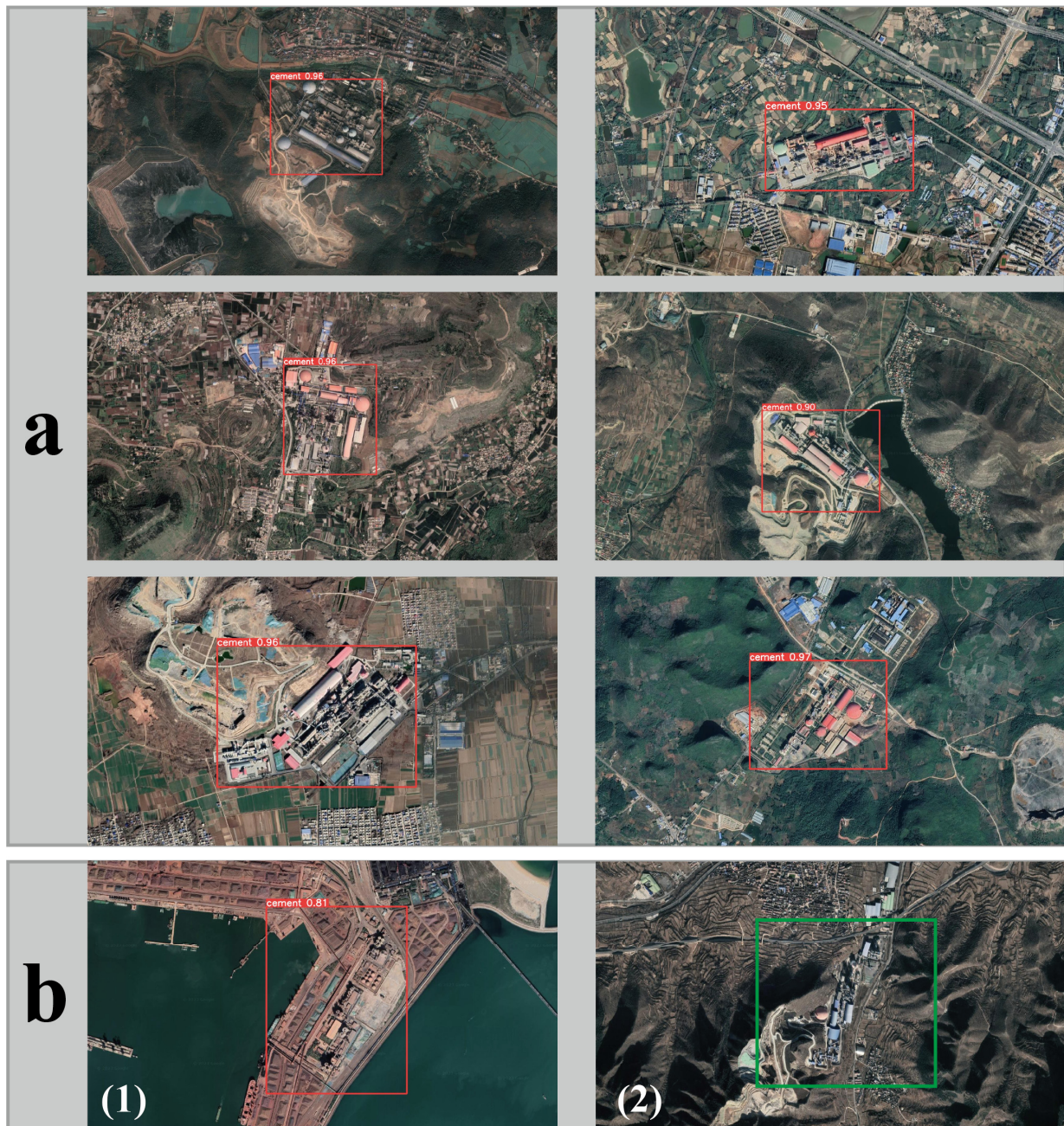


Figure 10. Schematic representation of detection results: (a) True Positive, (b): (1) False Positive, (2) False Negative.

3.4. Detection of Cement Plants in China Based on The YOLOv5-IEG Model

To detect spatial position information on cement plants in China, this study downloaded almost the entire 14–15 level Google Earth imagery for the Chinese region, totaling approximately 1.8 terabytes. In order to comprehensively detect all cement plants, an approach of overlapping and clipping was employed, with a resolution of 1200×800 pixels. In the end, a total of 781 large-scale cement plants were detected in the Chinese region, with a specific spatial distribution illustrated in Figure 11. Among them, Sichuan (62), Shandong (55), and Henan (48) have a higher quantity. Anhui, Hebei, Guangxi, Guizhou, and other provinces have a somewhat lower number of cement plants. The top ten provinces together contribute to 53.25% of the total, while the remaining 22 provinces with fewer cement plants account for 46.75%.

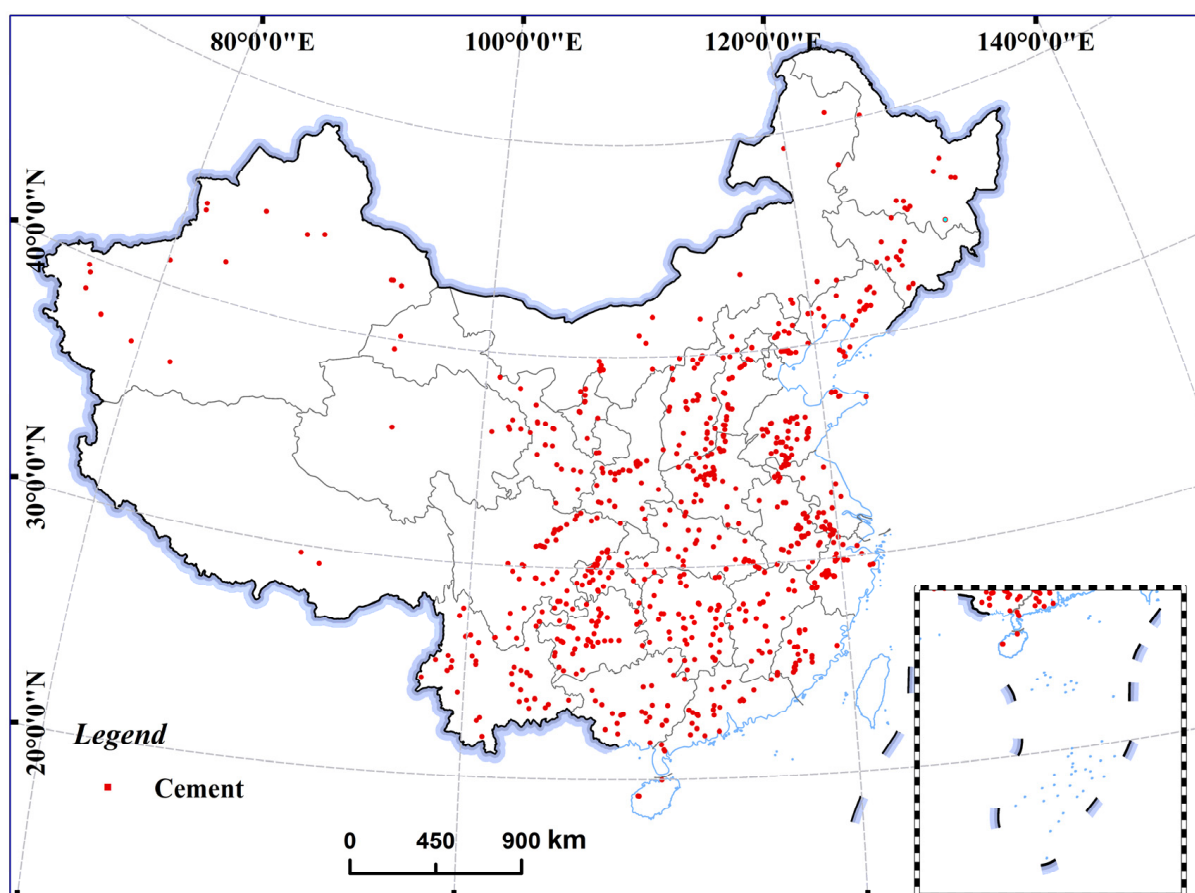


Figure 11. Distribution of detected cement plants in this study.

3.5. Monitoring the Operational Status of Cement Plants—Shandong Province

In this study, Shandong Province was selected as the research area to conduct monitoring of cement plant operations. In the preprocessing stage, infrared thermal images captured by SDGSAT-1 during the months of May to June in Shandong Province were acquired. A careful selection was made to choose a high-quality phase of remote sensing imagery for the experiment. Subsequently, cement plants in Shandong Province were detected and their thermal states were monitored. The results are illustrated in Figure 12. A total of 55 cement plants were detected, with 46 of them being in operation and nine not being in operation. These plants are primarily distributed in locations such as Zaozhuang, Zibo, and Yantai.

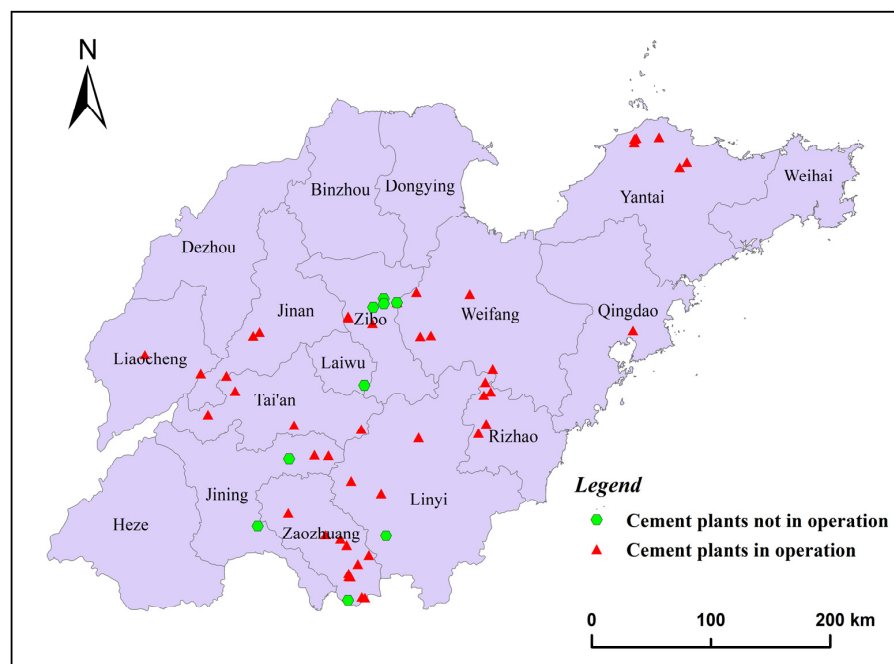


Figure 12. Distribution of detected cement plants in Shandong Province.

To enhance the clarity of the monitoring results, Figure 13 displays partial schematic results. While Figure 13a–c illustrate cement plants in operation, Figure 13d depicts a cement plant that is not operational. In this figure, the SDGSAT-1 thermal infrared image has undergone grayscale and normalization processing, with all values ranging from 0 to 255. Higher pixel values indicate greater thermal activity. The proposed Thermal Signature Detection module for cement plants in this study demonstrates effective detection of the thermal signatures of cement plants and monitoring of their operational status.

4. Discussion

4.1. Comparison with Other Target Detection Algorithms

In order to further validate the effectiveness of the improved YOLOv5-IEG algorithm, comparative experiments were conducted with other deep learning models in the same hardware environment. All seven network models were trained for 300 epochs. The final training results are presented in Table 3. The proposed algorithm demonstrated a 2.8% improvement in precision compared to Faster RCNN, with a significant increase of 17.41% in recall. Comparative experiments confirm that the improved algorithm in this study exhibits superior performance in the detection of large-scale cement plants, especially in complex background scenarios.

Table 3. Results of comparative experiments.

Model	P (%)	R (%)	mAP@.5 (%)	mAP@.5:.95 (%)
Faster-RCNN	94.0	76.3	94.1	68.2
Mask-RCNN	89.2	65.9	89.2	52.3
SSD	91.4	80.2	92.2	64.8
YOLOv6	91.5	88.2	90.1	61.8
YOLOv7	90.6	90.1	91.2	62.0
YOLOv8	92.2	90.2	88.2	61.2
Ours	96.8	93.7	96.9	68.8

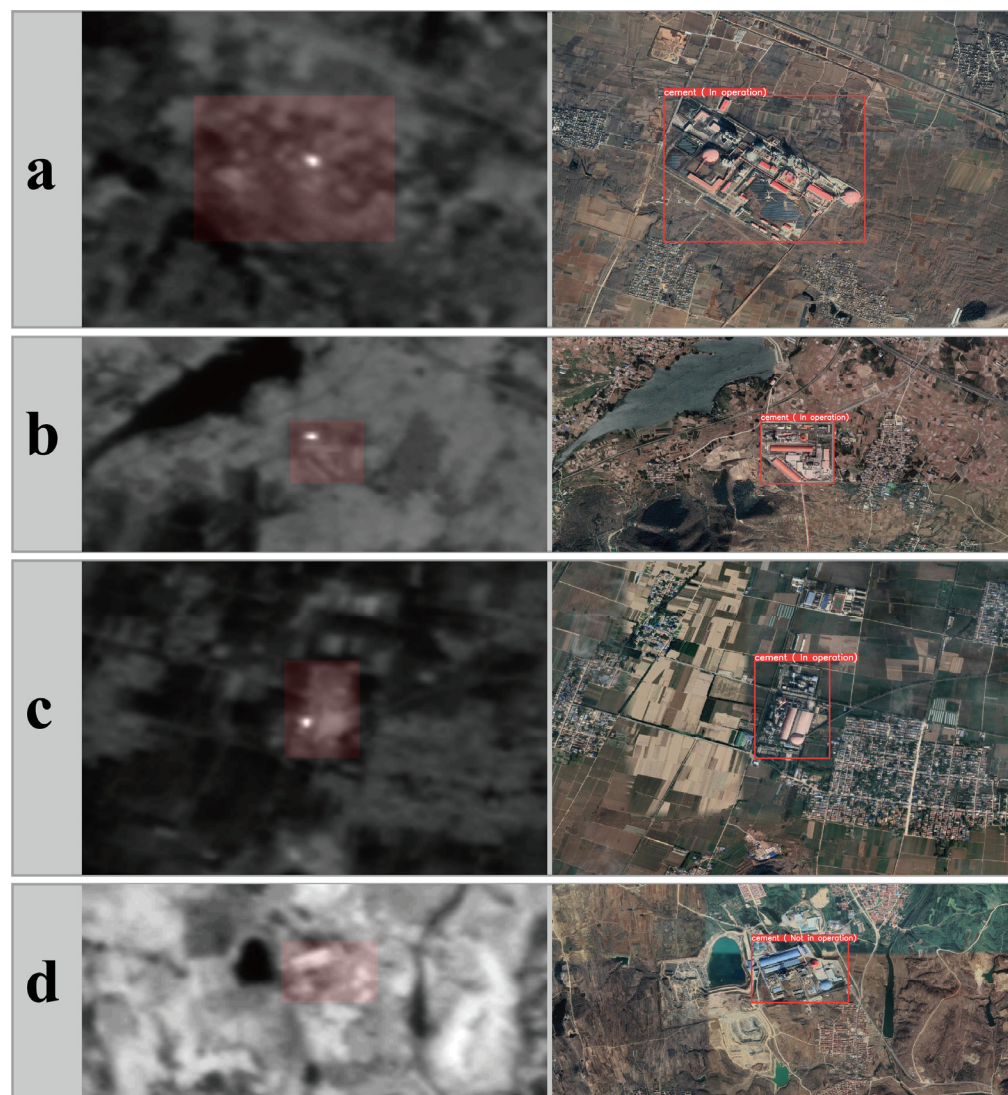


Figure 13. Schematic diagram for detecting the operational state of cement plants. (a–c) Operational cement plants, (d) Non-operational cement plants.

4.2. Ablation Experiments

In order to demonstrate the performance gains produced by the added CSPGghost module, EMA module, and Inner-IoU loss function in this study, we conducted ablation experiments. The experimental results are presented in Table 4. The improved YOLOv5 model showed positive optimizations with significant improvements in precision and mAP. It is noteworthy that the addition of the EMA mechanism resulted in improvements in all metrics. Ultimately, precision increased by 3.7%, recall increased by 1.61%, mAP@.5 increased by 1.6, and mAP@.5:95 increased by 7.7%. These indices collectively reflect the superiority of the improved YOLOv5-IEG model.

Table 4. Results of ablation experiments.

Method				P (%)	R (%)	mAP@.5(%)	mAP@.5:95(%)
YOLOv5s	Ghost	EMA	Inner-IoU				
✓				92.0	92.1	94.3	61.1
✓	✓			91.6	93.1	94.8	62.4
✓	✓	✓		94.6	94.8	96.1	66.2
✓	✓	✓	✓	96.8	93.7	96.9	68.8

4.3. Comparative Analysis between Our Results and Those of Others

To validate the quantity and accuracy of cement plants detected by the algorithm proposed in this study, two industrial heat source datasets were introduced. This includes a dataset from Ma et al. [14], which is based on long-term active thermal anomaly data and employs spatiotemporal density segmentation to construct a medium-to-low-resolution Chinese industrial heat source dataset spanning 2012 to 2021. This dataset comprises 4411 industrial heat sources, with a total of 719 cement plants, as shown in Figure 14a. Additionally, data from Tkachenko [34] were used, containing 1159 entries labeled as cement plants. However, due to the dataset creation process involving geographic spatial computer vision and large-scale language modeling techniques, many entries only have the word “cement” in their names or are cement dealerships, falling outside the scope of detection in this study. After manual interpretation, 865 entries were identified as actual cement plants, distributed as shown in Figure 14b. Comparing these two datasets in a contrastive experiment, the analysis revealed that the accuracy of cement plant detection in this study reached 90.8%, with a false-negative rate of 4.3% and a false-positive rate of 4.9%.

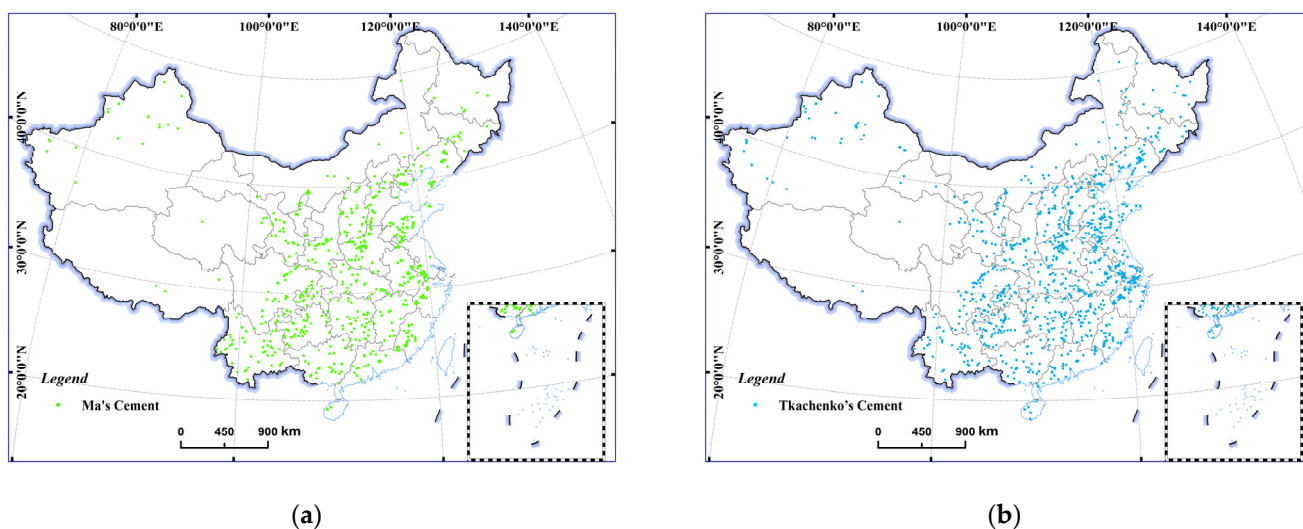


Figure 14. Spatial distribution of comparative datasets: (a) Ma et al.’s dataset, (b) Tkachenko’s dataset.

In addition, to compare the precision of detection boxes between our results and Ma et al.’s dataset, this study randomly selected 50 cement plants as samples, manually annotated standard detection boxes, and employed Intersection over Union (IoU) as the evaluation criterion. The box comparison chart, depicted in Figure 15, displays manually annotated ground truth boxes in blue, detection results from this study in red, and detection results from Ma et al.’s dataset in green. This study concludes that the average IoU (mIoU) for the detection boxes in this research is 78.2%, while Ma et al.’s dataset achieves an mIoU of 44.4%. This demonstrates that the detection boxes generated in this study exhibit higher accuracy and more precise localization.

4.4. Analysis of Thermal Signatures in Cement Plants

In this study, assessments were conducted on the operational status of 56 cement plants located in Shandong Province, China. Of these, nine cement plants were found to be non-operational, while 46 were identified as being in operation. As depicted in Figure 16, SDGSAT-1 thermal infrared imagery was overlaid onto Google Earth imagery with a transparency set at 50%. Observations revealed that the detected thermal signatures in the actively operating cement plants were predominantly located in the rotary kiln section, a crucial component used for the calcination of limestone to produce clinker. The observed phenomenon aligns with the objective facts. In contrast, the monitoring results for the nine non-operational cement plants, as illustrated in Figure 17, indicated a lack of discernible

thermal signatures within the plant boundaries. Consequently, it was determined that these cement plants were not in operation. It is essential to note that this assessment was based on high-quality SDGSAT-1 thermal infrared imagery acquired on a specific day in the months of May and June. Therefore, the findings of this experiment reflect the operational status of the cement plants at that moment. For dynamic monitoring of cement plant operations, continuous time series thermal infrared imagery from SDGSAT-1 can be utilized, providing insights into the sustained operational status of the cement plants over time.

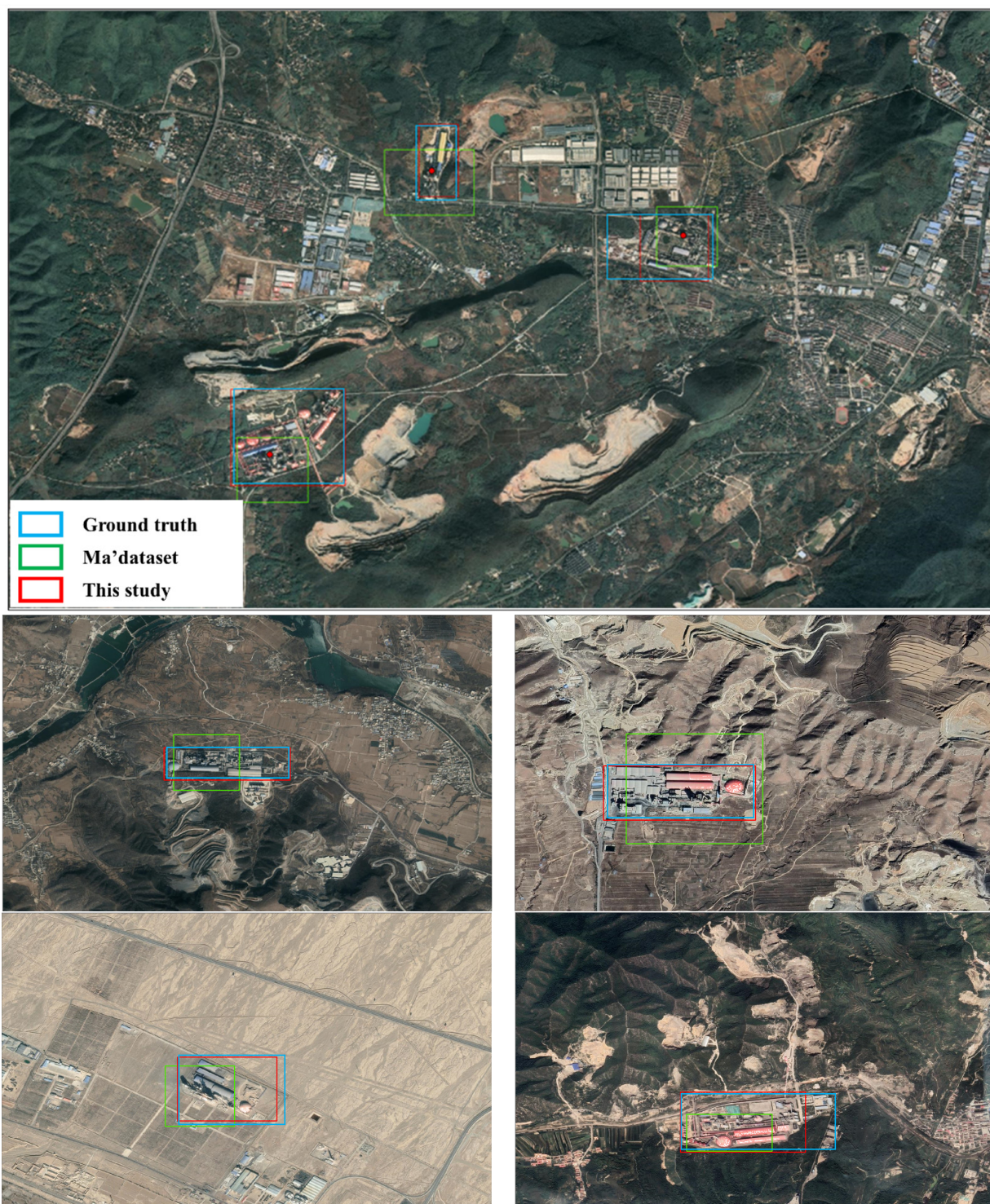


Figure 15. Comparison of detection boxes.

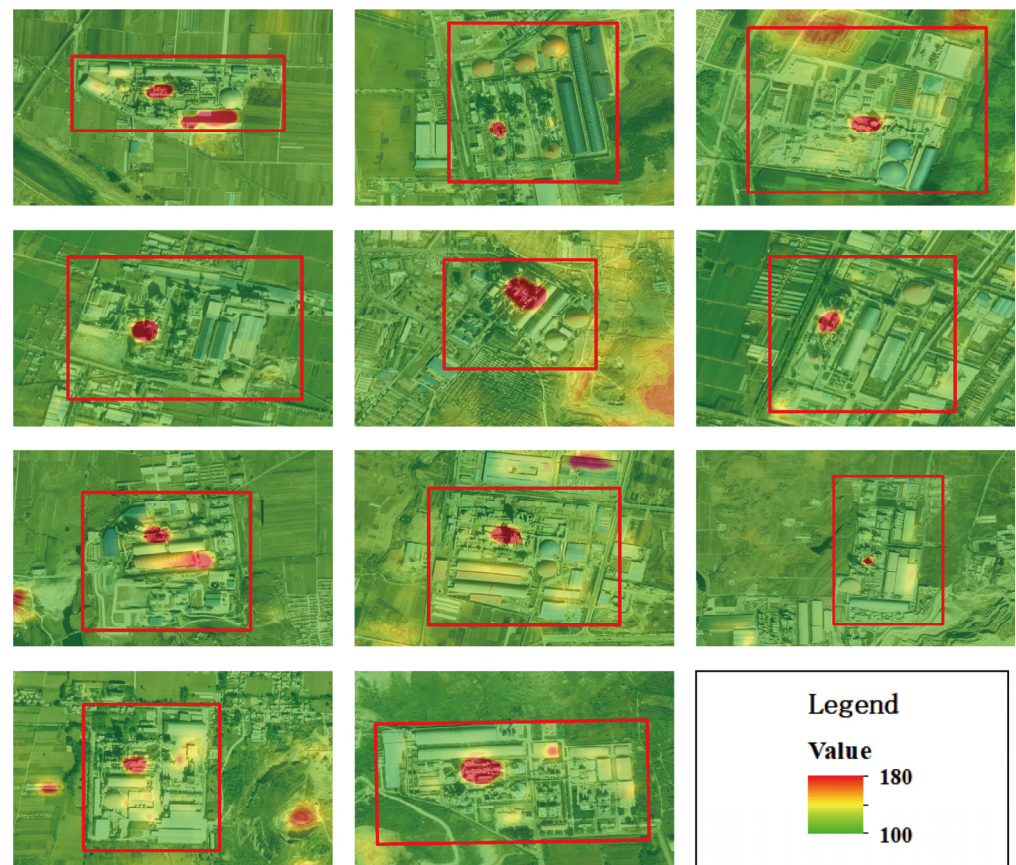


Figure 16. Schematic diagram of operational cement plants.



Figure 17. Schematic diagram of non-operational cement plants.

5. Conclusions

In this study, we introduced a deep learning model for detecting cement plant targets in complex background scenarios. To reduce the model's parameter count, the CSPGhost module was introduced, leading to a 29.2% reduction in model parameters. To address

the limitations of existing IoU loss functions in terms of weak generalization capability and slow convergence speed in detection tasks, the Inner-IoU loss function was introduced to enhance the generalization ability and improve the convergence speed of the model in the cement plant detection task. Additionally, to better capture cross-spatial feature information in different spatial dimensions for cement plants and achieve richer feature aggregation, a novel dimensionality-preserving EMA mechanism was incorporated into the model, mitigating detection issues such as false negatives to a certain extent. The final detection accuracy demonstrated a 4.8% improvement over the baseline model. Furthermore, we proposed a cement plant operational status monitoring model that combines Google Earth imagery and SDGSAT-1 thermal infrared imagery to detect cement plants and determine their operational status. In Shandong Province, 45 operational cement plants were monitored, with their predominant heat signatures observed primarily in the rotary kiln areas. Nine cement plants were identified as non-operational. This provides the capability to dynamically monitor the operational status of cement plants.

It is crucial to note that in object detection tasks, the quality of training samples significantly influences the model's detection performance. Therefore, obtaining a comprehensive dataset of cement plant samples is key to developing a reliable cement plant detection model. Inevitably, the model may struggle to detect novel or less distinctive cement plants. Addressing this challenge is one of our future research priorities, emphasizing the substantial enrichment of the training dataset to strengthen the model's reliability.

Author Contributions: Methodology, T.L. and C.M.; datasets, T.L., Y.L. and R.L.; experiments, T.L. and Y.L.; mapping the experimental results, T.L., C.M. and T.L.; results analysis, T.L. and C.M.; data curation, T.L. and Y.L.; writing—original draft preparation, T.L.; writing—review and editing, T.L., C.M., Y.L., J.Y., R.L. and J.L.; project administration, T.L., C.M., Y.L., J.Y., R.L. and J.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Guangxi Innovation-driven Development Special Project (GuiKe-AA20302022), the China-Pakistan Joint Research Center of Earth Sciences, CAS-HEC, Islamabad 45320, Pakistan, and the Youth Innovation Promotion Association of the Chinese Academy of Science under Grant 2021126.

Data Availability Statement: Data are contained within the article and available from the corresponding author.

Acknowledgments: It is acknowledged that the SDGSAT-1 data are kindly provided by CBAS.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Korczak, K.; Kočański, M.; Skoczkowski, T. Mitigation Options for Decarbonization of the Non-Metallic Minerals Industry and Their Impacts on Costs, Energy Consumption and GHG Emissions in the EU—Systematic Literature Review. *J. Clean. Prod.* **2022**, *358*, 132006. [[CrossRef](#)]
2. Gao, T.; Shen, L.; Shen, M.; Chen, F.; Liu, L.; Gao, L. Analysis on Differences of Carbon Dioxide Emission from Cement Production and Their Major Determinants. *J. Clean. Prod.* **2015**, *103*, 160–170. [[CrossRef](#)]
3. Mikulčić, H.; Vujanović, M.; Duić, N. Reducing the CO₂ Emissions in Croatian Cement Industry. *Appl. Energy* **2013**, *101*, 41–48. [[CrossRef](#)]
4. Wang, Y.; Yi, H.; Tang, X.; Wang, Y.; An, H.; Liu, J. Historical Trend and Decarbonization Pathway of China's Cement Industry: A Literature Review. *Sci. Total Environ.* **2023**, *891*, 164580. [[CrossRef](#)] [[PubMed](#)]
5. Hendriks, C.A.; Worrell, E.; De Jager, D.; Blok, K.; Riemer, P. Emission Reduction of Greenhouse Gases from the Cement Industry. In Proceedings of the Fourth International Conference on Greenhouse Gas Control Technologies, Interlaken, Switzerland, 30 August–2 September 1998; IEA GHG R&D Programme: Interlaken, Austria, 1998; pp. 939–944.
6. Meinshausen, M.; Lewis, J.; McGlade, C.; Gütschow, J.; Nicholls, Z.; Burdon, R.; Cozzi, L.; Hackmann, B. Realization of Paris Agreement Pledges May Limit Warming Just below 2 °C. *Nature* **2022**, *604*, 304–309. [[CrossRef](#)] [[PubMed](#)]
7. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2021**, *23*, 18. [[CrossRef](#)] [[PubMed](#)]
8. Oggioni, G.; Riccardi, R.; Toninelli, R. Eco-Efficiency of the World Cement Industry: A Data Envelopment Analysis. *Energy Policy* **2011**, *39*, 2842–2854. [[CrossRef](#)]

9. Ali, M.B.; Saidur, R.; Hossain, M.S. A Review on Emission Analysis in Cement Industries. *Renew. Sustain. Energy Rev.* **2011**, *15*, 2252–2261. [[CrossRef](#)]
10. Schneider, M.; Romer, M.; Tschudin, M.; Bolio, H. Sustainable Cement Production—Present and Future. *Cem. Concr. Res.* **2011**, *41*, 642–650. [[CrossRef](#)]
11. Shen, W.; Cao, L.; Li, Q.; Zhang, W.; Wang, G.; Li, C. Quantifying CO₂ Emissions from China’s Cement Industry. *Renew. Sustain. Energy Rev.* **2015**, *50*, 1004–1012. [[CrossRef](#)]
12. Sawaya, K.E.; Olmanson, L.G.; Heinert, N.J.; Brezonik, P.L.; Bauer, M.E. Extending Satellite Remote Sensing to Local Scales: Land and Water Resource Monitoring Using High-Resolution Image. *Remote Sens. Environ.* **2003**, *88*, 144–156. [[CrossRef](#)]
13. Liu, Y.; Hu, C.; Zhan, W.; Sun, C.; Murch, B.; Ma, L. Identifying Industrial Heat Sources Using Time-Series of the VIIRS Nightfire Product with an Object-Oriented Approach. *Remote Sens. Environ.* **2018**, *204*, 347–365. [[CrossRef](#)]
14. Ma, C.; Yang, J.; Chen, F.; Ma, Y.; Liu, J.; Li, X.; Duan, J.; Guo, R. Assessing Heavy Industrial Heat Source Distribution in China Using Real-Time VIIRS Active Fire/Hotspot Data. *Sustainability* **2018**, *10*, 4419. [[CrossRef](#)]
15. Cheng, G.; Han, J. A Survey on Object Detection in Optical Remote Sensing Imagery. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 11–28. [[CrossRef](#)]
16. Li, K.; Wan, G.; Cheng, G.; Meng, L.; Han, J. Object Detection in Optical Remote Sensing Imagery: A Survey and a New Benchmark. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 296–307. [[CrossRef](#)]
17. Wen, D.; Huang, X.; Bovolo, F.; Li, J.; Ke, X.; Zhang, A.; Benediktsson, J.A. Change Detection from Very-High-Spatial-Resolution Optical Remote Sensing Imagery: Methods, Applications, and Future Directions. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 68–101. [[CrossRef](#)]
18. Purwins, H.; Li, B.; Virtanen, T.; Schlüter, J.; Chang, S.-Y.; Sainath, T. Deep Learning for Audio Signal Processing. *IEEE J. Sel. Top. Signal Process.* **2019**, *13*, 206–219. [[CrossRef](#)]
19. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [[CrossRef](#)] [[PubMed](#)]
21. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.
22. Bi, F.; Yang, J. Target Detection System Design and FPGA Implementation Based on YOLO v2 Algorithm. In Proceedings of the 2019 3rd International Conference on Imaging Signal Processing and Communication (ICISPC), Singapore, 27–29 July 2019; pp. 10–14.
23. Redmon, J.; Farhadi, A. Yolov3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767. [[CrossRef](#)]
24. Jocher, G.; Stoken, A.; Chaurasia, A.; Borovec, J.; Kwon, Y.; Michael, K.; Changyu, L.; Fang, J.; Skalski, P.; Hogan, A.; et al. *Ultralytics/Yolov5: V6. 0-YOLOv5n’Nano’models, Roboflow Integration, TensorFlow Export, OpenCV DNN Support*; Zenodo: Geneva, Switzerland, 2021. [[CrossRef](#)]
25. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In Proceedings of the Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 11–14 October 2016; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 21–37. [[CrossRef](#)]
26. Kharchenko, V.; Chyrka, I. Detection of Airplanes on the Ground Using YOLO Neural Network. In Proceedings of the 2018 IEEE 17th International Conference on Mathematical Methods in Electromagnetic Theory (MMET), Kyiv, Ukraine, 2–5 July 2018; pp. 294–297. [[CrossRef](#)]
27. Zhang, F.; Du, B.; Zhang, L.; Xu, M. Weakly Supervised Learning Based on Coupled Convolutional Neural Networks for Aircraft Detection. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5553–5563. [[CrossRef](#)]
28. Song, R.; Li, T.; Li, T. Ship Detection in Haze and Low-Light Remote Sensing Imagery via Colour Balance and DCNN. *Appl. Ocean. Res.* **2023**, *139*, 103702. [[CrossRef](#)]
29. Chen, X.; Xiang, S.; Liu, C.-L.; Pan, C.-H. Vehicle Detection in Satellite Imagery by Parallel Deep Convolutional Neural Networks. In Proceedings of the 2013 2nd IAPR Asian Conference on Pattern Recognition, Naha, Japan, 5–8 November 2013; pp. 181–185. [[CrossRef](#)]
30. Lu, K.; Li, G.; Chen, Z.; Zan, L.; Li, B.; Gao, J. Steel Plant Extraction Based on Multi-Channel Optimization of SSD Network with Negative Samples. *J. Univ. Chin. Acad. Sci.* **2020**, *37*, 352–359.
31. Xu, G.; Yue, J.; Dong, Y.; Lou, Q.; Xiong, W.; Nie, Y. Target Detection of Cement Plant in Satellite Imagery using Deep Convolutional Networks. *J. Image Graph.* **2019**, *24*, 550–561.
32. Tkachenko, N.; Tang, K.; McCarten, M.; Reece, S.; Kampmann, D.; Hickey, C.; Bayarara, M.; Foster, P.; Layman, C.; Rossi, C.; et al. Global Database of Cement Production Assets and Upstream Suppliers. *Sci. Data* **2023**, *10*, 696. [[CrossRef](#)]
33. Wang, Q.; Feng, W.; Yao, L.; Zhuang, C.; Liu, B.; Chen, L. TPH-YOLOv5-Air: Airport Confusing Object Detection via Adaptively Spatial Feature Fusion. *Remote Sens.* **2023**, *15*, 3883. [[CrossRef](#)]
34. Wang, Y.; Zou, H.; Yin, M.; Zhang, X. SMFF-YOLO: A Scale-Adaptive YOLO Algorithm with Multi-Level Feature Fusion for Object Detection in UAV Scenes. *Remote Sens.* **2023**, *15*, 4580. [[CrossRef](#)]

35. Wu, T.-H.; Wang, T.-W.; Liu, Y.-Q. Real-Time Vehicle and Distance Detection Based on Improved Yolo v5 Network. In Proceedings of the 2021 3rd World Symposium on Artificial Intelligence (WSAI), Guangzhou, China, 18–20 June 2021; pp. 24–28. [[CrossRef](#)]
36. Ting, L.; Baijun, Z.; Yongsheng, Z.; Shun, Y. Ship Detection Algorithm based on Improved YOLO V5. In Proceedings of the 2021 6th International Conference on Automation, Control and Robotics Engineering (CACRE), Dalian, China, 15–17 July 2021; pp. 483–487. [[CrossRef](#)]
37. Ishak, S.A.; Hashim, H. Low Carbon Measures for Cement Plant—A Review. *J. Clean. Prod.* **2015**, *103*, 260–274. [[CrossRef](#)]
38. Worrell, E.; Price, L.; Martin, N.; Hendriks, C.; Meida, L. Carbon Dioxide Emissions from the Global Cement Industry. *Annu. Rev. Energy Environ.* **2001**, *26*, 303–329. [[CrossRef](#)]
39. Xu, J.-H.; Fleiter, T.; Eichhammer, W.; Fan, Y. Energy consumption and co2 emissions in China’s cement industry: A perspective from LMDI decomposition analysis. *Energy Policy* **2012**, *50*, 821–832. [[CrossRef](#)]
40. Zhang, S.; Worrell, E.; Crijns-Graus, W. Evaluating co-benefits of energy efficiency and air pollution abatement in China’s cement industry. *Appl. Energy* **2015**, *147*, 192–213. [[CrossRef](#)]
41. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. GhostNet: More Features from Cheap Operations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020.
42. Zhang, Y.-F.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and Efficient IOU Loss for Accurate Bounding Box Regression. *Neurocomputing* **2022**, *506*, 146–157. [[CrossRef](#)]
43. Ouyang, D.; He, S.; Zhang, G.; Luo, M.; Guo, H.; Zhan, J.; Huang, Z. Efficient Multi-Scale Attention Module with Cross-Spatial Learning. In Proceedings of the ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–9 June 2023; pp. 1–5. [[CrossRef](#)]
44. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
45. Yu, J.; Jiang, Y.; Wang, Z.; Cao, Z.; Huang, T. UnitBox: An Advanced Object Detection Network. In Proceedings of the 24th ACM International Conference on Multimedia, MM ’16, Amsterdam, The Netherlands, 15–19 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 516–520. [[CrossRef](#)]
46. Rezatofghi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 658–666.
47. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12993–13000.
48. Gevorgyan, Z. SIoU Loss: More Powerful Learning for Bounding Box Regression. *arXiv* **2022**, arXiv:2205.12740. [[CrossRef](#)]
49. Zhang, H.; Xu, C.; Zhang, S. Inner-IoU: More Effective Intersection over Union Loss with Auxiliary Bounding Box. *arXiv* **2023**, arXiv:2311.02877. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.