International Journal of Applied Earth Observation and Geoinformation 126 (2024) 103613

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Identifying wind turbines from multiresolution and multibackground remote sensing imagery

Yichen Zhai [a], Xuehong Chen [a,b], Xin Cao [a,b,*], Xihong Cui [a,b]

[a] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

ARTICLE INFO

Keywords:
Wind turbine
Detection
Positioning
Multi-resolution
Multi-background
YOLOv5

ABSTRACT

The wind energy industry has expanded in recent years. Promotion of the Sustainable Development Goals (SDGs) is expected to further increase the scale of the wind energy industry. Determining the location and quantity of wind turbines is crucial for monitoring the development status of the wind energy industry and evaluating wind energy production. In this study, we propose a method for simultaneously detecting and positioning wind turbines in remote sensing images, namely, Wind Turbine YOLO (WT-YOLO), based on the You Only Look Once version 5 model (YOLOv5). The wind turbine hub, base, and shadow hub are treated as key points in the proposed method. Regression terms are incorporated into the head of the YOLOv5 basic framework to predict the location of these three key points. The base point is utilized to determine the exact position of the wind turbine. A multibackground and multiresolution wind-turbine image dataset is constructed by sampling high-resolution images from Google Earth. The WT-YOLO method outperforms existing methods in both wind turbine detection and positioning on different types of land cover backgrounds and multiresolution images of the constructed dataset. In the spatial resolution range of 0.6 m to 5.4 m, WT-YOLO exhibits enhanced wind-turbine detection, where the average precision (AP) is 5.92 % to 15.43 % higher than that of existing wind-turbine detection methods. Wind-turbine positioning by WT-YOLO has a mean distance error (MDE) that is 16.06 m to 21.59 m lower than that of existing wind-turbine positioning methods. A comparative analysis showed that the shadow and the three key points are effective features for wind-turbine detection. The proposed WT-YOLO model can support detection, positioning and counting for wind turbines worldwide.

1. Introduction

Wind is an environmentally friendly source of energy for modern society (Saidur et al., 2011) and an essential component of renewable energy (Sadorsky, 2021). The use of wind energy can help mitigate the greenhouse effect and facilitate progress toward realizing the SDGs (Olabi et al., 2023). According to the Global Wind Energy Report 2023 by the Global Wind Energy Council (GWEC), the past three years have witnessed the largest annual increase in wind capacity installations in history. On average, 88.8 GW of wind capacity is installed globally annually. Over the next five years, 680 GW of new wind capacity will be installed globally (GWEC, 2023). WindEurope estimates that the EU will need to reach 440 GW of installed wind capacity by 2030 to meet its renewable energy target (WindEurope, 2023). Several studies have projected broad prospects for development of the wind energy industry driven by technological advances, cost reductions, policy support and other favorable conditions (GWEC, 2023; European Commission et al., 2022; EIA, 2023; IEA, 2023). The growth rate of global installed wind capacity is expected to increase.

The rapid global expansion of the wind energy industry has led to a dramatic increase in both the quantity and range of spatial distribution of installed wind turbines. The location and quantity of wind turbines are crucial factors in assessing the effectiveness of wind farms (Parada et al., 2018; Song et al., 2016), predicting the annual energy production of wind farms (Grassi et al., 2014), optimizing the layout of wind farms (Wu et al., 2021), and evaluating the development potential of wind power projects (Sliz-Szkliniarz et al., 2019). However, a large number of wind turbines around the world are not precisely located. The complexity and diversity of the backgrounds of wind turbines installed in different areas make it difficult to obtain accurate wind turbine locations by remote sensing (Zhang et al., 2021a,b; Mandroux et al., 2022a; Mandroux et al., 2022b). Therefore, it would be useful to develop

* Corresponding author at: State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.
E-mail address: caoxin@bnu.edu.cn (X. Cao).

https://doi.org/10.1016/j.jag.2023.103613
Received 9 October 2023; Received in revised form 30 November 2023; Accepted 9 December 2023
Available online 22 December 2023

a fast and intelligent method for extracting the locations of wind turbines that is applicable to different land cover backgrounds around the globe to monitor the distribution and quantity of wind turbine installations worldwide.

With the development of sensor technology, remote sensing has become an indispensable tool for observing specific ground targets. Wind turbines consist of blades and cylindrical towers made of metallic materials. In an optical remote sensing image, a wind turbine appears as white highlights and a distinct spatial structure. In practice, a certain spacing must be maintained between wind turbines to reduce the impact of the wake of a wind turbine on the power generation efficiency (Wang et al., 2022), such that individual wind turbines appear as independent targets. However, the United States Wind Turbine Database compiled by Rand et al. (2020) shows that wind turbines come in a wide range of sizes, from 30 to 200 m. Insufficient spatial resolution can result in both blurring of blade-shape features and failure to detect small wind turbines. High-resolution optical remote sensing imagery is ideal for wind turbine identification because submeter spatial resolution can be achieved, which enhances the structural characteristics of wind turbines and increases the visibility of small wind turbines. In SAR imagery, offshore wind turbines exhibit features similar to manmade metallic targets in the oceanic background and therefore appear prominent on the sea surface (Ferrentino et al., 2019). However, onshore wind turbines are difficult to detect in SAR imagery because of land background clutter (Zhang and Hao, 2022). Thus, SAR imagery is unsatisfactory for observing wind turbines against complex land cover backgrounds.

The algorithms for extracting wind turbines have been continuously innovated in recent years. Scholars have proposed various methods that can be broadly categorized into three types based on the extraction algorithms and the output results: segmentation, positioning, and detection. Some scholars have adopted the **segmentation** technique to extract the outlines of wind turbines from remote sensing images. Chen et al. (2018) targeted the wind turbine body, used the traditional saliency object detection method to segment wind turbines in six Google Earth images in three counties in China, and produced binarization results. Han et al. (2018) developed the target U-net model to segment wind turbines in five Gaofen-2 images in the Shanxi and Shandong provinces in China, where the shadow feature was used to target wind turbines. The **positioning** method predicts the base position of wind turbines in an image. Mandroux et al., (2021, 2022a, 2022b) applied an *a contrario* approach in which the shadow feature was used to help locate wind turbines in several Sentinel-2 data covering the USA. Zhou et al. (2019) built five weakly supervised model structures based on the same backbone, namely, class activation mapping (CAM) (Zhou et al., 2016), improved gradient-weighted class activation mapping (Grad-CAM++) (Chattopadhyay et al., 2018), soft proposal networks (SPN) (Zhu et al., 2017), weakly supervised learning of deep convolutional neural networks (WILDCAT) (Durand et al., 2017), and peak response maps (PRM) (Zhou et al., 2018), and compared the performance of these five structures for high-resolution satellite imagery of the United States. WILDCAT was found to be the most effective weakly supervised structure for locating wind turbines. Manso-Callejo et al. (2020) located wind turbines by segmenting the base of a wind turbine in an image and calculating the location of the center of mass. Combinations of LinkNet or U-net segmentation structures with different backbone networks were tested on PNOA aerial images taken in Spain, and the combination of LinkNet and EfficientNet-b3 was found to be the most effective for wind turbine positioning. Manso-Callejo et al. (2021) subsequently proposed a framework for locating and categorizing wind turbines into three installed capacity intervals. The VGG model was used to filter out images containing wind turbines from PNOA aerial orthophotos from all over Spain, and the bases of the wind turbines were segmented to locate the turbines, which were then categorized into the three predetermined installed capacity intervals using the VGG model. Some research groups are working on the positioning of offshore wind turbines. Zhang et al. (2021a,b) applied adaptive thresholding, morphological operations, and center-of-mass computation to worldwide preprocessed Sentinel-1 SAR imagery to locate global offshore wind turbines. Hoeser et al. (2022) created a global dataset of offshore wind turbine locations that were determined by detecting offshore structures from Sentinel-1 SAR imagery, using two faster region-based convolutional network (Faster R-CNN) models, and pinpointing the exact turbine locations within a bounding box based on the peak of the backscatter coefficient. The object **detection** algorithm extracts wind turbine information and then creates a corresponding bounding box. Abedini et al. (2019) utilized Feature from Accelerated Segment Test (FAST) and Speeded Up Robust Features (SURF) as feature extractors along with Fast Library for Approximate Nearest Neighbors (FLANN) as a matcher to produce bounding boxes for wind turbines in ground-based photographs. Zhang et al. (2021a,b) used a Faster-R-CNN-based iterative detection framework to detect wind turbines and their accompanying shadows from 2-m resolution remote sensing imagery of regions throughout China.

Several critical issues have not been addressed in the aforementioned studies. First, the diversity of the data background has not been sufficiently considered. Images from a single region with a limited land cover background have primarily been investigated. However, landscapes vary considerably across regions worldwide and are substantially more complex than the limited backgrounds that have been used in these studies. Second, shadow features were incorporated into wind-turbine extraction methods in studies by Han et al. (2018), Zhang et al. (2021a,b), and Mandroux et al., (2021, 2022a, 2022b) but not in studies by Chen et al. (2018), Zhou et al. (2019), Abedini et al. (2019), and Manso-Callejo et al., (2020, 2021). However, the impact of shadows on wind-turbine extraction has not been explored. Third, the aforementioned methods are intended to be applied to single-resolution images and may not be suitable for images with multiple resolutions. Given the abundance of available remote sensing data at varying resolutions, it is crucial to develop algorithms that are capable of processing images with multiple resolutions. However, the effectiveness of applying these algorithms to multiresolution images has not been investigated. There is a lack of established algorithms for effectively identifying wind turbines at varying resolutions. In addition, wind-turbine detection algorithms rely on bounding boxes to identify the turbine but cannot precisely locate the turbine base. Mandroux et al., (2021, 2022a, 2022b) and Manso-Callejo et al., (2020, 2021) have developed a wind-turbine positioning model but it is still challenging to detect the wind-turbine base. Zhou et al. (2019) achieved 86.6 % accuracy using weakly supervised learning for positioning. However, weakly supervised learning may identify any part of the wind turbine and not specifically the base.

Therefore, a deep learning model, WT-YOLO, based on YOLOv5 was proposed in this study for identifying wind turbines. Within this model, a wind turbine and its shadow were treated as an objective and three key points of the wind turbine are identified: the hub, base, and shadow hub. Regression terms were added to the head to predict the location of these three key points using the YOLOv5 model as the base framework. High-resolution satellite images from Google Earth were used to create a dataset with various land cover background types, and the algorithm was used to detect and position wind turbines in different areas globally from the constructed dataset. The algorithm detection performance was studied in its correlation to the key-point and shadow features of the wind turbine. Assessments were carried out on the performance of the WT-YOLO model on images of varying resolutions and the influence of land cover backgrounds on the model performance. The proposed WT-YOLO algorithm can detect wind turbines in remote sensing images at a spatial resolution of meters. This approach could facilitate the recognition of global-scale wind turbines. The primary contributions of this study are summarized below.

- We identify key-point features in wind turbines that are integrated into wind-turbine detection to develop a deep learning model called Wind Turbine YOLO (WT-YOLO) that simultaneously performs detection and key-point positioning for wind turbines.

- Our model locates the exact position of a wind turbine according to the key points, achieving higher precision wind-turbine positioning across multiple spatial resolutions than existing models. We reduce the wind-turbine positioning bias to the meter level for the first time to achieve positioning results that far surpass existing methods.
- Our model outperforms existing methods for wind-turbine detection across multiple resolutions.
- We assess the impact of shadow and key-point features on wind-turbine detection.

2. Data, sampling and labeling

2.1. Data and sampling based on the background

This study was performed on true-color images with a spatial

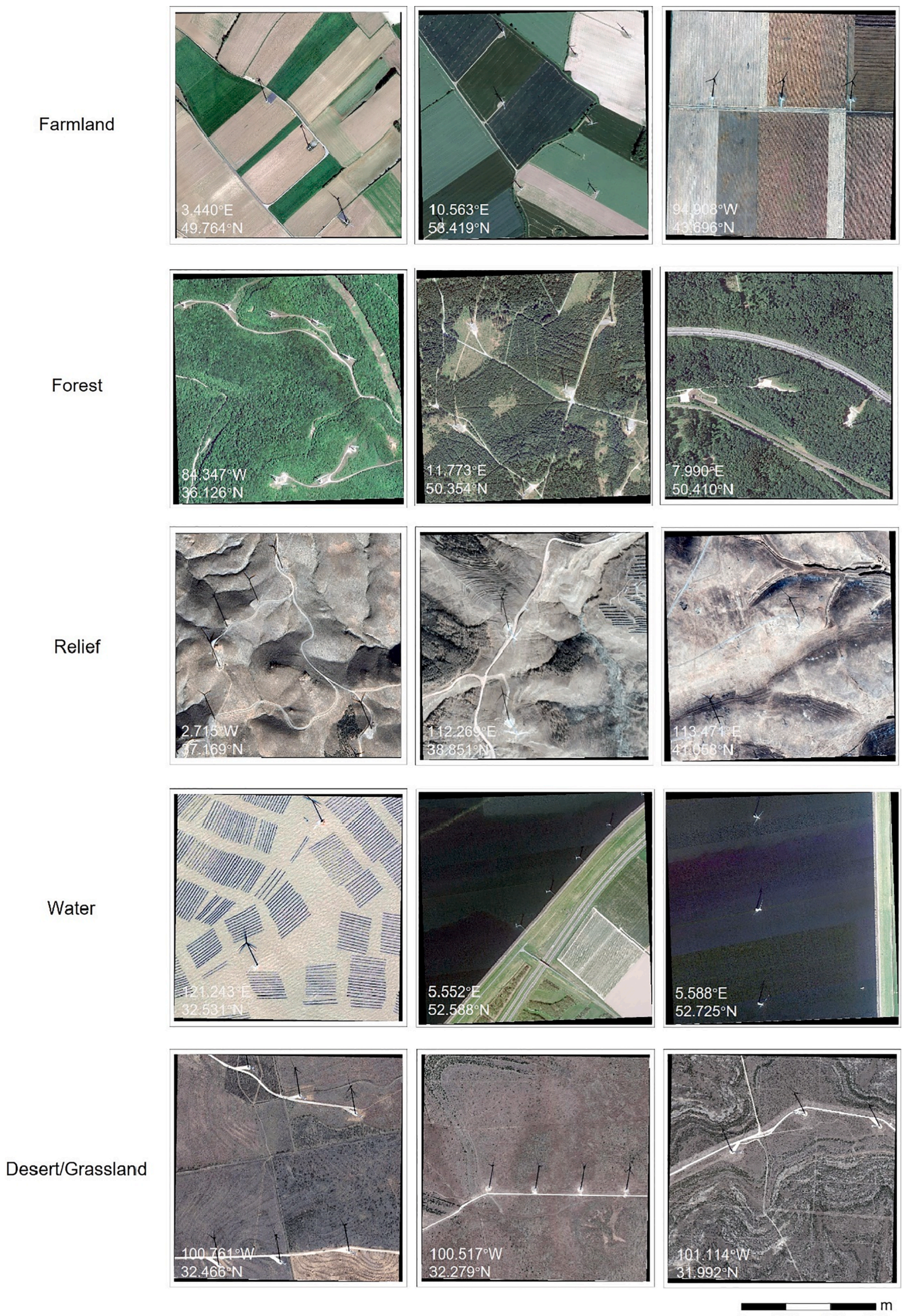

Fig. 1. The five classes of land cover backgrounds. Each image in the figure is projected into UTM projection.

resolution of approximately 0.6 m that were obtained from Google Earth (shown in Fig. 1). To acquire the images containing wind turbines, it is crucial to obtain the geographical coordinates of the corresponding area. Dunnett et al. (2020) created freely accessible worldwide datasets of locations of wind and solar farms. This dataset contains the center-of-mass coordinates of wind farms worldwide in 2020, which can be used to roughly identify areas with wind turbines. As the number of turbines on each wind farm is included in the dataset, we could capture images proficiently. We filtered out wind farms with fewer than 10 turbines to select wind farms located at the remaining coordinates as the image centers to obtain images covering a 1 km spatial extent around the centers. We meticulously inspected the acquired images and discarded those with missing wind turbines or turbines obstructed by clouds. To preserve the wind-turbine morphology accurately, we projected the images into the UTM coordinate system related to the UTM Zone centered on the image coordinates.

In several studies, land surfaces of narrow roads, relief, and agricultural fields have been found to easily confuse detectors (Zhang et al., 2021a,b; Mandroux et al., 2022a; Mandroux et al., 2022b). In particular, ridges within agricultural fields, mountain ridges, and roads highlighted in relief are ground targets that can be misdetected as wind turbines. Overlap between tall trees and the shadow of a wind turbine has been found to change the shape of the shadow in an image and water bodies have been found to obscure the shadow of a wind turbine in an image. Images are categorized into five classes based on the land cover background that detectors tend to misdetect: farmland, forest, relief, water bodies, and desert/grassland (Fig. 1). Both deserts and grasslands are homogeneous and do not affect the shadow or body morphology of wind turbines; thus, these two types of land cover combined into a single background category.

Twenty images were randomly selected for each background class that were evenly distributed across the globe (Fig. 2). Each image had an approximate width and height of 1700 pixels. These images were cropped into 416 × 416 pixel tiles with 75 % overlap. To ensure the feature integrity of each wind-turbine target, we meticulously screened all the tiles to select those containing at least one wind turbine with a complete body and shadow. The high overlap rate resulted in a single wind turbine target appearing in more than one tile. To minimize data redundancy, among tiles containing the same wind turbine, we only retained the tile in which the wind turbine was closest to the center. The filtered tiles were used to compile a dataset of wind-turbine images consisting of 353 tiles and 398 complete wind turbine instances. The classes of land cover background had similar numbers of tiles and instances (Table 1).

Table 1
The numbers of tiles and instances in each of the five classes of land cover backgrounds.

Background	Tiles	Instances
Farmland	73	74
Forest	81	100
Relief	50	52
Water	76	76
Desert/Grassland	73	96

2.2. Labeling

Our model was designed to target wind turbines with shadows and their hubs, bases, and shadow hubs. In our image dataset, each turbine instance was manually annotated with a bounding box and three key points. To investigate the role of shadows and key points in wind-turbine detection, we defined two other types of targets: wind turbines with and without shadows. Each turbine was manually labeled with a bounding box for these two types of targets. Consequently, three distinct types of labels were used for each turbine (Fig. 3). For some wind turbines, cropping resulted in only the body or shadow of the wind turbine being visible at the edges of an image. These incomplete turbines (lacking essential characteristics) were not labeled in the training set. However, as the identification of incomplete targets was considered as a correct result during testing, incomplete turbines were labeled in the test set. The labeling process was performed using *Labelme* software.

3. Methods

3.1. WT-YOLO model

As the heights of wind turbines typically range from tens to hundreds of meters, the shadows of wind turbines are clearly visible in images. The bases of wind turbines are visible because of the non-nadir viewing angle of satellite images. Turbine towers and blades are slender and therefore, do not typically obscure shadows and bases. A wind turbine in an image has three visible points: the hub, base and shadow hub. We considered these three points as stable features and included them in the wind turbine targets. The base point of the wind turbine indicates its exact location. We predicted the location of the base point to determine the exact location of the wind turbine.

The You Only Look Once model (YOLO) series (Redmon et al., 2016; Redmon and Farhadi, 2017; Redmon and Farhadi, 2018; Bochkovskiy et al., 2020) is a classical one-stage model used in deep learning object

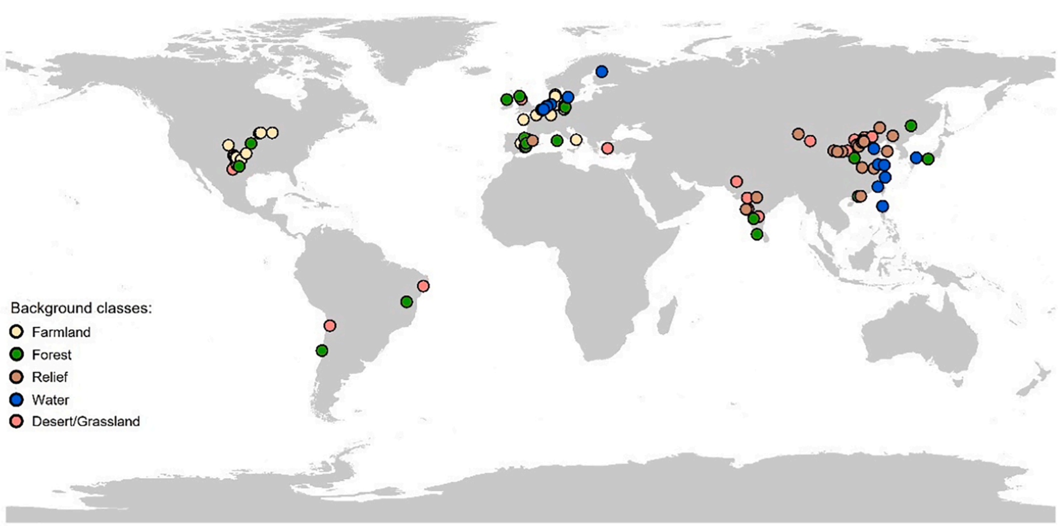

Fig. 2. Global distribution of sample images.

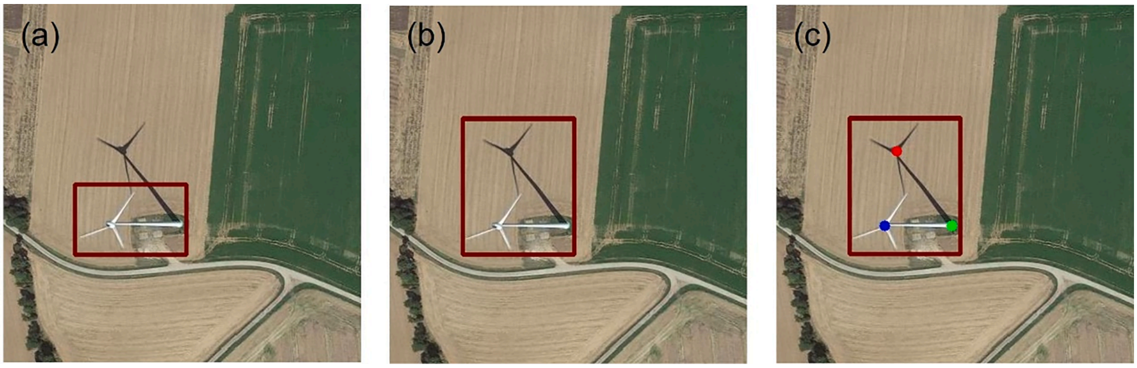

Fig. 3. The three types of targets used for wind turbine identification. (a) The wind turbine body alone. (b) The wind turbine and its shadow. (c) The wind turbine, its shadow and three key points.

detection that treats object detection as a regression problem to determine the location and size of a bounding box. The YOLO model has been applied to object detection in remote sensing images in many studies with excellent results (Xu and Wu, 2021; Zakria et al., 2022). Years of development of YOLOv5 have produced a fast and accurate object detection model. The proposed WT-YOLO model is an improvement on YOLOv5 (version 6.0). WT-YOLO fully inherits the data augmentation method as well as the backbone and neck structure of YOLOv5. The main improvement is the redesign of the head, which enables the model to incorporate the three key-point features. YOLOv5 regression in the head is used to predict the center position, width and height, confidence, and class of the bounding box. WT-YOLO extends the channel dimension

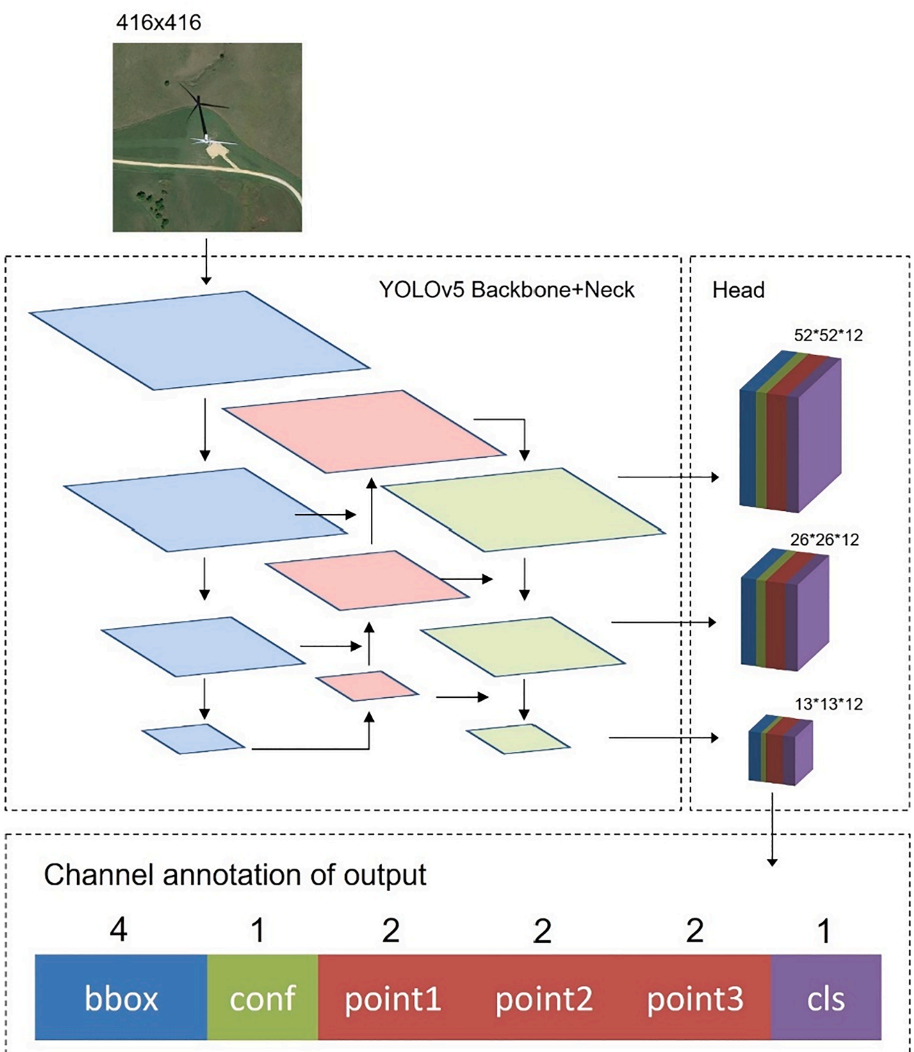

Fig. 4. Structure of the WT-YOLO model. The three cubes in the head are the model outputs for one anchor. Point1, point2, and point3 represent the regression terms that predict the hub, base, and shadow hub coordinates of the wind turbine, respectively. The terms bbox, conf and cls represent the regression terms for the bounding box, confidence and class, respectively.

length of the output in the head by incorporating regression terms for the three key-point coordinates, as shown in Fig. 4. For each anchor, the length of the output in the channel dimension is extended to 12. Considering that WT-YOLO retains 3 anchors, the total length of the output in the channel dimension is 36. WT-YOLO simultaneously regresses the locations of the target bounding box and the key points, achieving simultaneous wind-turbine detection and key-point positioning. As this simple structural addition only increases the computation slightly, WT-YOLO effectively inherits the advantage of the inference speed of YOLOv5.

The total loss of the original YOLOv5 has three components: the bounding box, class, and confidence. WT-YOLO retains these three components and adds the key-point loss directly to the total loss. The key-point loss is based on a wing loss function (Feng et al., 2018). The wing loss function is sensitive to small errors, which facilitates learning subtle positional deviations of a key point. The wing loss function is given by

$$wing(x) = \begin{cases} w \cdot \ln\left(1 + \frac{|x|}{e}\right), & if\ |x| < w \\ |x| - C, & otherwise \end{cases} \tag{1}$$

where w and e are two set parameters. w defines the range of the nonlinear region and is set to 10 in this study. e regulates the curvature of the nonlinear region and is set to 2 in this study. C is a constant that connects the linear and nonlinear components and is given by $C = w - w \cdot ln(1 + w/e)$. Then, the key-point loss is given by

$$loss_p = \sum_{i=1}^{n} wing\left(p_x - p_x^{'}\right) + wing\left(p_y - p_y^{'}\right) \tag{2}$$

where p_x and p_y are the key-point coordinates output from the model and $p_x{'}$ and $p_y{'}$ are the ground truth key-point coordinates. n is the number of key-point coordinates, which is set to 3 in our model. The key-point loss is incorporated into the total loss as follows:

$$loss = \lambda_a \cdot loss_a + \lambda_b \cdot loss_b + \lambda_c \cdot loss_c + \lambda_p \cdot loss_p \tag{3}$$

where $loss_a$, $loss_b$ and $loss_c$ are the losses of the bounding box, class, and confidence, respectively. The same functions are used for these three losses are as in YOLOv5, where λ_a, λ_b, and λ_c are the weights of the three losses. In our model, λ_a, λ_b and λ_c are set to 0.05, 6.25E-3 and 4.225E-1. $loss_p$ is the key-point loss, and λ_p is the weight, which is set to 2E-3. As the key point only works in the correct bounding box, the bounding-box accuracy is prioritized over the key-point accuracy. Therefore, the key-point loss is weighted less than the bounding box loss in this study.

3.2. Experiment details

3.2.1. Division of training, validation and testing datasets

The dataset was divided into training, validation, and testing sets in a 7:2:1 ratio. We implemented a 10-fold cross-validation approach to mitigate the potential impact of the dataset division on the model performance and maximize the utilization of the dataset. The dataset was randomly separated into 10 subsets containing equal numbers of data. One subset was assigned as the test set. Two of the remaining subsets were randomly chosen to form the validation set, and the remaining seven subsets were used as the training set. Each subset needed to be used once as a test set, thereby creating ten group datasets derived from the original data with different partitions.

3.2.2. Simulation resolution

To assess the model performance at multiple resolutions, the acquired images were simulated to generate images at various resolutions. We obtained a Google Earth image with a spatial resolution of approximately 0.6 m. Mean resampling was applied to decrease the image resolution, followed by bilinear interpolation to restore the image size to simulate images at the meter level. As the smallest unit in the resampling process was one pixel, which was equivalent to 0.6 m, the simulated resolutions ranged from 0.6 m to 5.4 m in increments of 0.6 m. The coarsest resolution was restricted to 5.4 m because the slender turbine blades tended to become blurred at overly coarse resolutions. After the simulation, the resolution could be reduced without affecting the image size, which remained at 416 × 416 (Fig. 5). Identical labels were used for the original images and simulated images. The resolution simulation procedure was applied to each group dataset after data division. A total of 90 datasets were constructed, corresponding to 9 resolutions * 10 sets of samples.

3.2.3. Model training, validation, and testing

Transfer learning was used to train our model to achieve optimal performance in a shorter period of time, where all the layers were fine-tuned. The initial weights of WT-YOLO were the pretraining weights of YOLOv5 on the Common Objects in Context (COCO) dataset. The SGD optimizer was used with an initial learning rate of 1E-2, a final learning rate of 1E-3 after cosine learning rate decay, and a weight decay of 5E-4. The momentum was set to 0.8 during warmup and 0.937 after warmup. We set 3 warmup epochs, 300 total epochs, and a batch size of 16. WT-YOLO selects the best model in the validation based on the average precision at 0.5 (AP@0.5), corresponding to the average precision of detection when the intersection over union (IOU) threshold for a true positive is set to 0.5, without considering the accuracy of key-point positioning. Subsequently, the best model is evaluated on the test set to analyze the detection and positioning performance for wind turbines.

In this study, YOLOv5, the base model of WT-YOLO, was also trained to assess the impact of the shadow and key-point features on wind-turbine detection. To ensure comparable accuracy, we used the same values for the hyperparameters that are used in WT-YOLO. The YOLOv5 model was trained separately for the two types of targets: wind turbines with and without shadows. We trained the most effective wind-turbine detection and positioning models known to us to comparatively evaluate the capability of WT-YOLO. The current most effective wind-turbine detection technique was developed by Zhang et al. (2021a,b). The comparative analysis is only valid for their detection model used, which is the Faster R-CNN model with ResNet-18 as the backbone, because of the different overall frameworks. DeepWind proposed by Zhou et al. (2019) is currently the best wind-turbine positioning model available. To ensure an adequate number of negative samples for DeepWind training, we randomly selected an equal quantity of pure background tiles as the tiles used for our model. For the purpose of comparison, all the aforementioned models underwent 90 training iterations using the 10-fold multiresolution dataset.

The APs of all the models were measured on the test set. The APs were compared to assess the detection performance of the models for the wind turbines in the images. The MDEs of all the models, except for Faster R-CNN, were measured on the test set to compare the performance of wind-turbine positioning. WT-YOLO locates the turbine position using the predicted base point position, and YOLOv5 adopts the center point of the bounding box as the positioning outcome. The ground truth location of each wind turbine is determined by the base point location indicated in the labels of WT-YOLO. As 10-fold cross-validation was used, the mean and standard deviation of the AP and MDE were calculated at each resolution. The mean and standard deviation further reflect the comprehensive performance and robustness of the models.

The experiments were conducted on an Ubuntu server configured with an Intel Core i3-8350 K 4.00 GHz CPU, 32 GB RAM, and an NVIDIA GeForce RTX 1080ti GPU 12 GB environment. The models were developed using the PyTorch framework.

3.3. Evaluation metrics

The AP and MDE were used as metrics of the wind-turbine detection

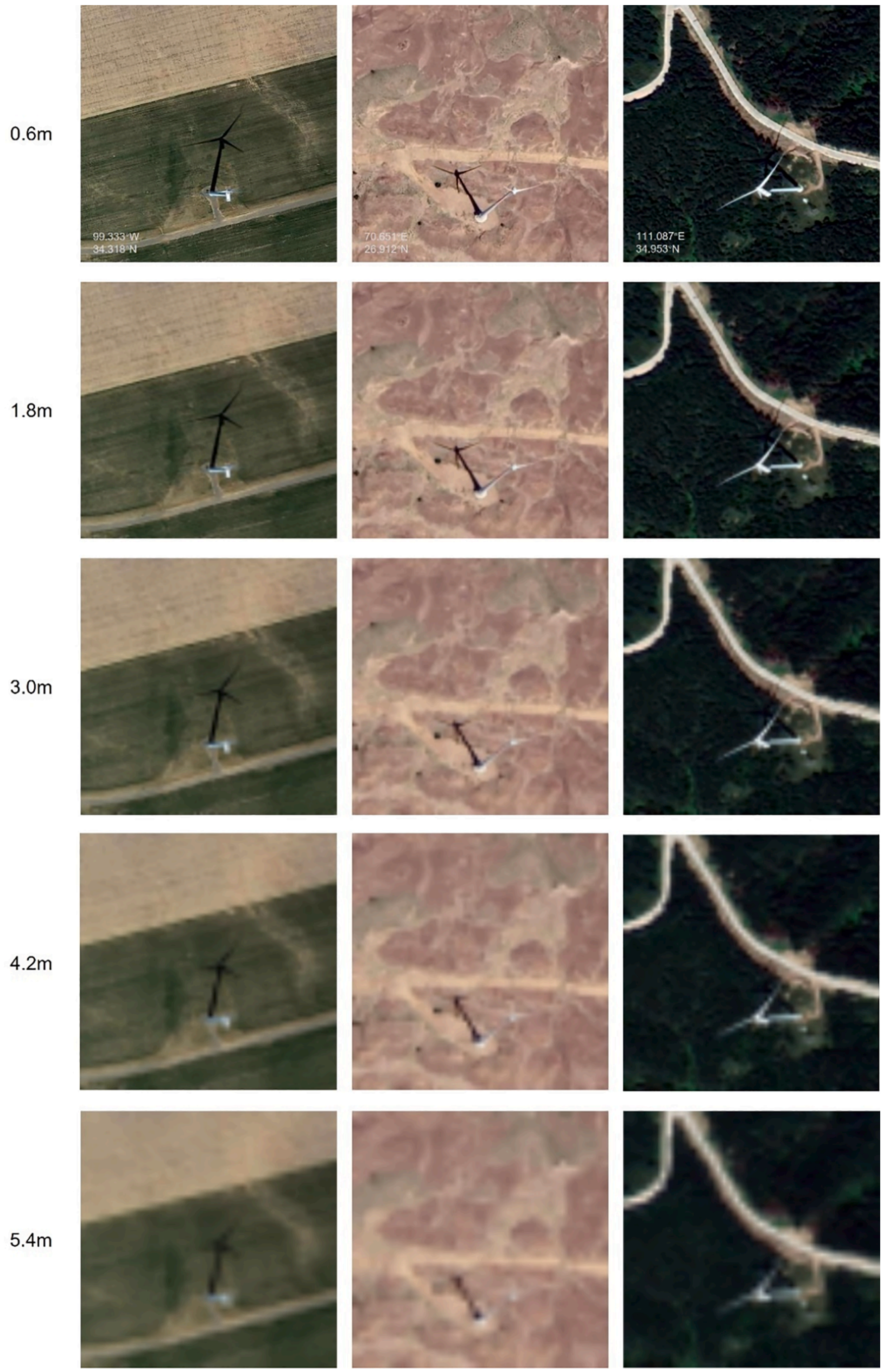

Fig. 5. Original tiles and the corresponding simulated tiles at partial resolution.

and positioning performance of the models, respectively. The AP is calculated as the area under the precision-recall curve (Padilla et al., 2020). The formulas for the AP and related metrics are

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$P_{interp}(R_{i+1}) = maxP\left(\widetilde{R}_{i+1}\right) \tag{6}$$

$$AP = \sum_{i=1}^{n}(R_{i+1} - R_i)\cdot P_{interp}(R_{i+1}) \tag{7}$$

where TP denotes true positives, FP denotes false positives, and FN denotes false negatives. n denotes the number of interpolation points. R_i denotes the recall of the i th interpolation point. $P(\widetilde{R}_{i+1})$ indicates a set of precisions corresponding to all recalls greater than R_{i+1}.

In this study, we established two criteria for defining True Positives (TP) to calculate the AP value. Except for DeepWind, other models employed the AP@0.5, which consider the predicted bounding box with an IOU greater than 0.5 with the ground truth bounding box and with the highest confidence as TP. For DeepWind, we consider predicted positioning points with the highest confidence within the ground truth bounding boxes containing shadow as TP.

The MDE is a measure of the difference between the predicted and actual locations of a key point. The smaller the MDE is, the higher the accuracy of point positioning is. The MDE is calculated as

$$MDE = \frac{\sum_i^N \sqrt{(x_i - x_i^{'})^2 + (y_i - y_i^{'})^2}}{N} \tag{8}$$

where N is the number of TP, x and y are the coordinates of the point in the image, and i denotes the i th TP.

4. Results

An evaluation was performed on the wind-turbine detection performance of the investigated models, the positioning results of WT-YOLO for three key points and the exact positioning results of the models for the wind turbines, except for Faster R-CNN, which was not used for positioning. All the results were analyzed at resolutions ranging from 0.6 m to 5.4 m. The mean cross-validation values were used to compare the accuracies of the models and analyze variations in the accuracy across multiple resolutions.

4.1. Wind-turbine detection performance

Table 2 shows the mean and standard deviation of the AP for each model on the test set. These data demonstrate that WT-YOLO performs highly accurate wind-turbine detection. The AP values of WT-YOLO are highly consistent, ranging from 92.53 % to 98.84 % under multi-resolution, with the standard deviation fluctuating between 0.73 % and 3.86 %. This result shows that WT-YOLO is highly accurate, as well as durable and flexible, when applied across multiple resolutions. The fluctuations in the standard deviation are comparable to or even smaller than those of YOLOv5 and Faster R-CNN. The revised head structure does not adversely affect the total stability of WT-YOLO. In fact, the revised head structure boosts the total stability of WT-YOLO in specific situations. Compared to the other models, the proposed model clearly exhibits exceptional accuracy at resolutions between 1.2 m and 5.4 m, underscoring its outstanding performance. WT-YOLO may not be optimal at a resolution of 0.6 m. However, WT-YOLO can predict the exact locations of three key points and achieves an AP that is only slightly lower than that of YOLOv5, which focuses on shadowed wind turbines. WT-YOLO outperforms DeepWind and Faster R-CNN by a large AP margin at all resolutions, demonstrating the superior performance of WT-YOLO over those of existing algorithms. The WT-YOLO output is presented in Fig. 6, indicating the excellent ability of WT-YOLO to detect wind turbines in various scenarios.

The results presented in Table 2 show that the same model detects wind turbines with shadows with a higher precision than wind turbines without shadows at every resolution, as has been found for YOLOv5. Thus, integrating the shadow feature into wind-turbine targets may enhance the efficacy of detection. Compared to YOLOv5, the detection performance of WT-YOLO for wind turbines with shadows is higher at all resolutions except 0.6 m and similar at the 0.6-m resolution. This result suggests that including key-point features can optimize detection at most resolutions.

The results presented in Table 2 also illustrate the multiresolution robustness of WT-YOLO and that WT-YOLO has a considerable advantage over the other models at coarse resolutions. From the finest to the coarsest resolution, WT-YOLO overcomes the negative impact of decreasing resolution on the detection performance more effectively than the other models: the AP reduction is only 6.31 % for WT-YOLO compared to higher values of 11.93 %, 15.82 %, 17.67 %, and 12.56 % for the other models. The APs of the other models decrease considerably as the resolution decreases, possibly because of gradual blurring of the blades, which makes the shape characteristics of the turbine less distinct. By contrast, the resolution has less of an effect on the three key points, which remain clearly visible even at low resolutions. The detection performance of WT-YOLO could be improved by relying on stable key-point features. At the 5.4-m resolution, the AP for WT-YOLO remains at an impressive 92.53 %, which is considerably higher than those of the other three models (of only 60.66 %, 77.10 %, 80.77 %, and 86.52 %), which exhibit a substantial drop in performance. The notably superior WT-YOLO AP indicates that WT-YOLO is more suitable for remote sensing images with coarse meter-level spatial resolution than the other models. Fig. 7 demonstrates that at submeter resolution, both WT-YOLO and YOLOv5 display comparable accuracy in detecting the target, whereas WT-YOLO also accurately identifies the three critical points. At roughly a meter-level resolution, WT-YOLO can still confidently detect the target and precisely mark its key points.

4.2. Key-point positioning performance

The positioning results of both WT-YOLO and YOLOv5 are correlated with the detection hyperparameters. We evaluated the positioning performance of these models using a threshold of 0.4 for the confidence level and a threshold of 0.45 for the nonmaximum suppression. The results presented in Table 3 indicate that WT-YOLO precisely locates the three key points at every resolution, with an MDE range of only approximately 4 m ~ 9 m. The MDE decreases gradually as the resolution decreases, but the decrease is not considerable. WT-YOLO exhibits the lowest MDE for base-point positioning for six out of nine resolutions and can therefore provide strong support for precise wind-turbine positioning. The hub point has the lowest MDE at the three finest

Table 2
Comparison of the AP of models on the test set for different resolutions. The numbers before and after "±" represent the mean and standard deviation of the cross-validation, respectively. The bolded values indicate the highest accuracy at one resolution.

Resolution	AP (%)				
	DeepWind	Faster R-CNN(shadows)	YOLOv5(w/o shadows)	YOLOv5(shadows)	WT-YOLO (shadows & key points)
0.6 m	72.59 ± 11.45	92.92 ± 3.93	98.44 ± 1.21	**99.08** ± 0.61	98.84 ± 1.18
1.2 m	70.57 ± 13.59	92.18 ± 3.56	98.08 ± 1.81	98.42 ± 1.13	**98.53** ± 1.38
1.8 m	59.07 ± 19.52	92.14 ± 3.36	96.83 ± 2.59	98.03 ± 1.52	**98.74** ± 0.78
2.4 m	58.95 ± 22.03	91.83 ± 4.40	95.96 ± 2.80	97.46 ± 1.60	**98.90** ± 0.48
3.0 m	70.48 ± 11.46	88.52 ± 3.89	93.24 ± 4.35	97.44 ± 1.35	**98.09** ± 1.33
3.6 m	75.70 ± 17.19	85.66 ± 5.28	90.17 ± 4.76	95.87 ± 2.03	**96.18** ± 2.33
4.2 m	62.39 ± 18.19	82.93 ± 6.14	85.25 ± 5.12	90.41 ± 2.98	**94.88** ± 3.94
4.8 m	65.40 ± 18.15	82.44 ± 5.36	85.33 ± 3.53	90.18 ± 3.06	**94.89** ± 3.55
5.4 m	60.66 ± 17.61	77.10 ± 4.49	80.77 ± 6.37	86.52 ± 4.68	**92.53** ± 3.75

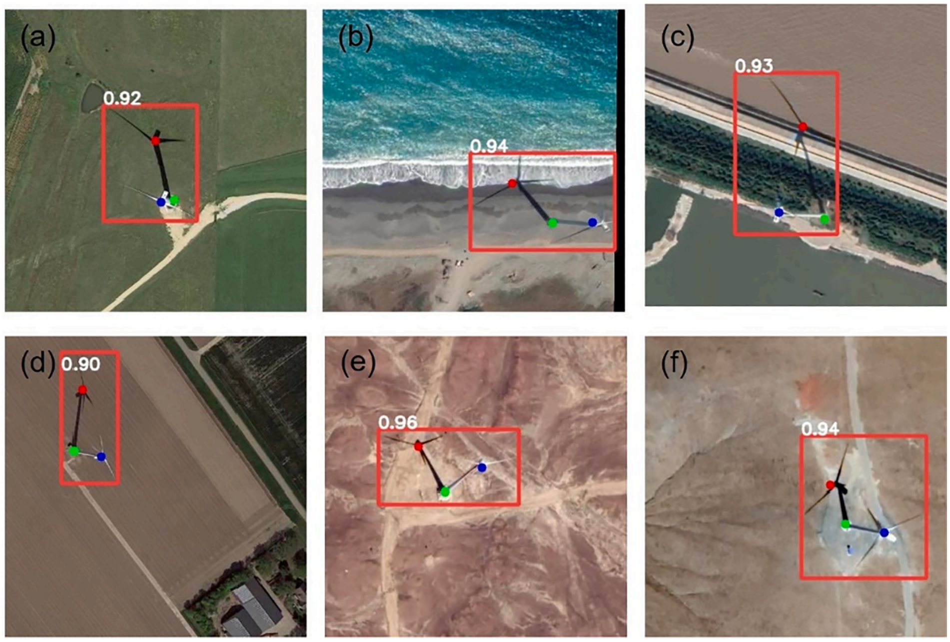

Fig. 6. Inference results of WT-YOLO. Each predicted target is generated with its bounding box, three key points and confidence level. The blue, green and red points represent the hub, base and shadow hub of the wind turbine, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

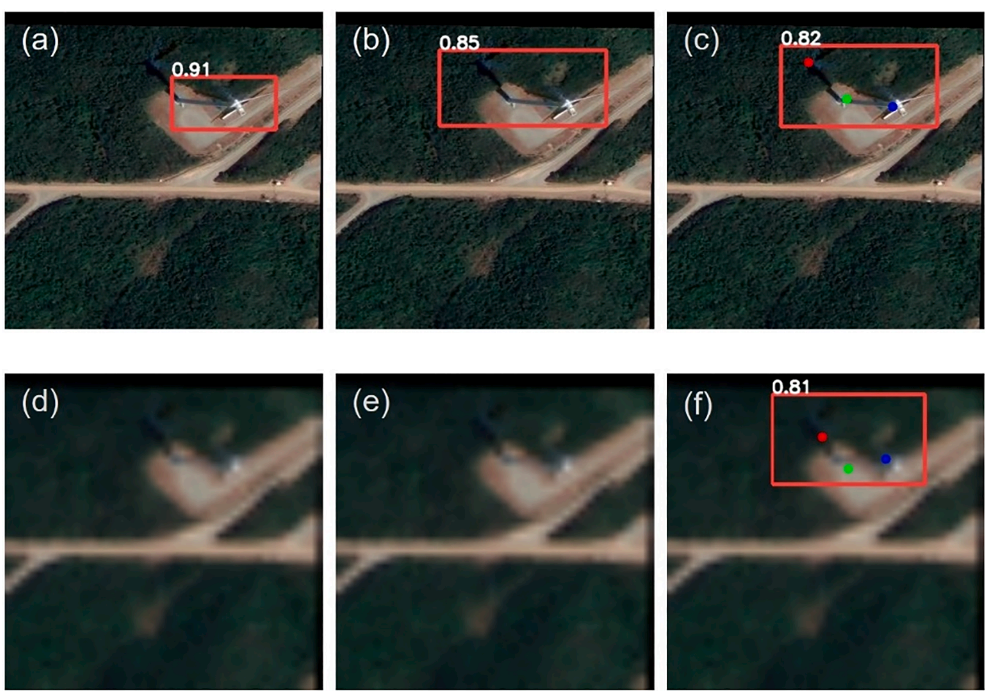

Fig. 7. Detection results of three wind turbine targets on the same tile at two resolutions: (a),(b) and (c) 0.6 m and (d),(e) and (f) 5.4 m. The detection results of YOLOv5 for a wind turbine (a) and (d) without a shadow and (b) and (e) with a shadow; (c)and (f) are the results obtained using WT-YOLO.

resolutions, possibly because the prominent wind-blade features facilitate hub positioning.

Table 4 is a comparison of the exact positioning performance of the wind turbine of the models. The ground-truth exact position of the wind turbine is the base-point position labeled on the images. WT-YOLO calculates the MDE by only considering the predicted position of the base point. YOLOv5 uses the center point of the predicted bounding boxes as the predicted location of the turbine, and DeepWind directly outputs the predicted turbine location. Compared to the other models, WT-YOLO provides considerably more accurate wind-turbine positioning results, especially at fine resolutions. At every resolution, the MDE is less than 7.4 m for WT-YOLO and over 21 m for the other models, indicating the suitability of using WT-YOLO for multiple resolutions. The comparative outcomes show that WT-YOLO is superior positioning to existing positioning methods and bounding box center-based methods for wind turbines. Compared to the other models, WT-YOLO displays superior positioning stability, with the lowest MDE standard deviation at every resolution. An analysis of the YOLOv5 results for the two types

Table 3
Summary of MDE results for positioning of three key points in WT-YOLO at multiple resolutions. The numbers before and after "±" represent the mean and standard deviation of the cross-validation, respectively. The bolded values indicate the lowest MDE at one resolution.

Resolution	MDE (m)			
	Point1(hub)	Point2(base)	Point3(shadow hub)	Average
0.6 m	**4.60** ± 0.80	4.62 ± 1.15	5.05 ± 1.28	4.76 ± 1.10
1.2 m	**4.57** ± 0.57	4.84 ± 0.75	4.95 ± 1.42	4.79 ± 1.00
1.8 m	**4.87** ± 0.87	5.04 ± 0.65	5.30 ± 1.28	5.07 ± 0.99
2.4 m	5.00 ± 0.44	**4.77** ± 1.17	5.46 ± 1.35	5.08 ± 1.10
3.0 m	5.34 ± 0.81	**5.32** ± 1.04	6.10 ± 0.94	5.59 ± 1.00
3.6 m	6.30 ± 0.97	**6.01** ± 1.33	6.77 ± 1.71	6.36 ± 1.40
4.2 m	6.67 ± 1.04	**6.35** ± 1.39	6.68 ± 0.93	6.56 ± 1.14
4.8 m	6.55 ± 1.04	**6.42** ± 1.50	7.80 ± 1.17	6.59 ± 1.26
5.4 m	8.28 ± 2.73	**7.40** ± 0.97	8.72 ± 2.21	8.13 ± 2.17

Table 4
Comparison of MDEs for wind-turbine positioning results of models at multiple resolutions. The numbers before and after "±" represent the mean and standard deviation of the cross-validation, respectively. The bolded values indicate the lowest MDE at one resolution.

Resolution	MDE (m)			
	DeepWind	YOLOv5 (w/o shadows)	YOLOv5 (shadows)	WT-YOLO (shadows & key points)
0.6 m	26.21 ± 3.64	25.89 ± 5.23	35.31 ± 4.06	**4.62** ± 1.15
1.2 m	25.66 ± 3.44	26.13 ± 5.39	35.33 ± 4.66	**4.84** ± 0.75
1.8 m	24.60 ± 4.78	26.41 ± 6.24	35.38 ± 3.95	**5.04** ± 0.65
2.4 m	24.91 ± 6.29	25.99 ± 8.17	35.13 ± 3.92	**4.77** ± 1.17
3.0 m	21.75 ± 4.79	27.74 ± 9.74	35.50 ± 4.51	**5.32** ± 1.04
3.6 m	22.07 ± 6.29	22.29 ± 9.04	36.26 ± 4.32	**6.01** ± 1.33
4.2 m	23.64 ± 5.42	26.93 ± 6.69	36.65 ± 4.04	**6.35** ± 1.39
4.8 m	25.04 ± 7.85	25.33 ± 6.98	36.70 ± 4.24	**6.42** ± 1.50
5.4 m	25.08 ± 4.10	28.04 ± 8.50	38.00 ± 5.89	**7.40** ± 0.97

of targets shows that the incorporation of shadow features amplifies the positioning bias because of the size of the bounding box is increased.

5. Discussion

5.1. Influence of background and sample

The ability of a model to generalize and mitigate background interference can be assessed from the model performance on images with backgrounds that were not previously seen. In this study, 0.6-m resolution images were used to test the capacity of WT-YOLO to detect wind turbines on previously unseen backgrounds. The test set comprised tiles with backgrounds belonging to the same class, and the training dataset consisted of tiles with the remaining four classes of backgrounds. Then, 30 % of the training set data were allocated for validation. In this study, images were used with five different classes of backgrounds, where each class was successively used as the test background. The results presented in Table 5 show the excellent performance of WT-YOLO on the detection and exact positioning of turbines against each test background, demonstrating the strong generalization capability of WT-YOLO. However, the MDE of the relief and the AP@0.5 of the water are relatively high and low. This result shows that shadow deformation may interfere with the exact turbine position determined by WT-YOLO because a shadow is prone to be distorted by rugged terrain. Poor shadow visibility may reduce the detection performance of WT-YOLO because the water background is likely to obscure the shadow.

Although the five background classes we categorized are applicable to most regions, backgrounds in different regions are intricate and diverse. Some backgrounds for wind-turbine images are not included in our dataset, such as the background of wind turbines installed near buildings, which should be classified as artificial surfaces. We are unable to assess the effectiveness of our proposed model against the backgrounds that are not included in our dataset. Therefore, it is essential to expand the background classes in the dataset by including backgrounds with fewer wind-turbine installations. Including a large quantity of data in the dataset can promote the ability of the model to extract potential features. Our constructed dataset comprises only 353 tiles, which is not a large sample size, despite the use of cross-validation to optimize data utilization. Unfortunately, as there are no publicly available high-resolution image datasets with wind-turbine bounding box labels, manual label creation is needed. It is imperative to increase the sample size when time and labor costs permit. The WT-YOLO labels also require three crucial points to be marked for each target, doubling the labeling effort required for bounding boxes alone.

Table 5
Overview of results for detection and positioning obtained by WT-YOLO for backgrounds not contained in the training set.

Background	AP@0.5 (%)	MDE (m)
Farmland	98.0	8.22
Forest	98.1	7.48
Relief	98.8	10.84
Water	96.9	8.16
Desert/Grassland	98.9	8.05

5.2. Evaluation of a single model trained on all-resolution data

A model that can be effectively applied to mixed-resolution data has very broad applicability. In this study, we assessed the mixed-resolution applicability of WT-YOLO by integrating data with multiple resolutions. We used the previously constructed dataset to integrate data with all resolutions for each cross-validation split to create a mixed-resolution dataset. Data from the training set at all resolutions were merged, and the same procedure was applied to the validation and testing sets. This procedure prevented the same image from being distributed at various resolutions across these three sets. The model performance on data at all resolutions was still evaluated using 10-fold cross-validation. Table 6 shows that WT-YOLO exhibits high applicability and robustness on mixed-resolution data, achieving an AP of over 95 % and an MDE of approximately 5 m. These results indicate that WT-YOLO can achieve high-precision detection for wind turbines and high-precision exact positioning for turbines even for mixed-resolution scenarios. WT-YOLO has a higher detection performance than YOLOv5 for both types of targets, and YOLOv5 with shadow targets has a higher detection performance than YOLOv5 without shadow targets. This result suggests that shadow and key-point features also have a noticeable impact on wind-turbine detection for mixed-resolution scenarios.

5.3. Challenges and future work

Although WT-YOLO achieves a high level of accuracy in key-point positioning, intricate backgrounds may impede further improvement of WT-YOLO performance. As the MDE results presented in Table 3 indicate that the bias in key-point positioning is nonnegligible, the current accuracy may not be sufficient for research and applications

Table 6
Overview of the results for detection and positioning of models on all-resolution mixed data. MDE denotes the MDE for the exact position of wind turbines. The numbers before and after "±" represent the mean and standard deviation of the cross-validation, respectively.

Models	AP@0.5 (%)	MDE (m)
WT-YOLO	95.81 ± 2.62	5.54 ± 1.57
YOLOv5 (shadows)	94.13 ± 1.89	36.39 ± 4.55
YOLOv5 (w/o shadows)	89.42 ± 2.67	22.08 ± 1.34

requiring precise key-point locations. A limitation of WT-YOLO is the high positioning bias of the shadow hub point. The shadow hub point consistently exhibits the highest MDE across all resolutions, whereas the other two points have lower and comparable MDEs. On average, at the same resolution, the MDE of the shadow hub point is 0.65 m higher than that of the point with the lowest MDE. The deformation of shadows caused by changes in terrain and solar altitude may pose a challenge to accurately predicting the shadow hub point.

Future work will concentrate on reducing the bias in key-point positioning, particularly for shadow hub points, and exploring application of key-point positioning. The potential of WT-YOLO to process images with meter-scale spatial resolution has been demonstrated. In future research, applying WT-YOLO to imagery with meter-scale spatial resolution and higher temporal resolution could be a promising approach for quickly monitoring the spatial dynamics of wind turbines.

6. Conclusion

We introduce a method for the simultaneous detection and positioning of wind turbines named WT-YOLO. The method consists of identifying a wind turbine, its shadow, and three key points (the hub, base, and shadow hub) as targets. The head structure of YOLOv5 is adjusted by adding regression terms for key-point localization. We used high-resolution images from Google Earth to create a dataset of images with multiple backgrounds and resolutions to conduct comprehensive comparative studies. The proposed approach outperforms existing methods in detecting and positioning wind turbines. The proposed method is highly stable at various resolutions and performs well against different backgrounds. The proposed method exhibits superior performance to existing methods at coarse resolutions, suggesting that the proposed method could be used to enhance the monitoring efficiency by being applied to satellites with 3–5 m resolutions and a shorter revisit period. The results of this study demonstrate that the shadow and distinctive key-point features of wind turbines considerably facilitate turbine detection.

The novel proposed method both enhances the efficiency of detecting wind turbines at multiple resolutions through the integration of key-point features and precisely locates the wind turbine using the base points. Next, we will primarily focus on refining the placement of key points to increase the precision of positioning of wind turbines. We will apply the model to other image-based data sources to monitor substantial changes in the spatial distribution of wind turbines on a large scale.

CRediT authorship contribution statement

Yichen Zhai: Methodology, Investigation, Validation, Writing – original draft, Visualization. **Xuehong Chen:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Xin Cao:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Xihong Cui:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (Grant No. 42371334).

References

Abedini, F., Bahaghighat, M., S'hoyan, M., 2019. Wind turbine tower detection using feature descriptors and deep learning. Facta Universitatis, Series: Electron. Energet. 33 (1), 133–153.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M., 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. https://doi.org/10.48550/arXiv.2004.10934.

Chattopadhyay, A., Sarkar, A., Howlader, P., Balasubramanian, V.N., 2018. Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. In: 2018 IEEE Winter Conference on Applications of Computer Vision, 839–847. https://doi.org/10.1109/WACV.2018.00097.

Chen, J., Yue, A., Wang, C., Huang, Q., Chen, J., Meng, Y., He, D., 2018. Wind turbine extraction from high spatial resolution remote sensing images based on saliency detection. J. Appl. Rem. Sens. 12 (01), 1. https://doi.org/10.1117/1.JRS.12.016041.

Dunnett, S., Sorichetta, A., Taylor, G., Eigenbrod, F., 2020. Harmonised global datasets of wind and solar farm locations and power. Sci. Data 7 (1), 130. https://doi.org/10.1038/s41597-020-0469-8.

Durand, T., Mordan, T., Thome, N., Cord, M., 2017. WILDCAT: Weakly Supervised Learning of Deep ConvNets for Image Classification, Pointwise Positioning and Segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5957–5966. https://doi.org/10.1109/CVPR.2017.631.

European Commission, Joint Research Centre, Telsnig, T., Georgakaki, A., Letout, S., et al., 2022. Clean Energy Technology Observatory, Wind energy in the European Union: status report on technology development, trends, value chains and markets: 2022. Publications Office of the European Union. https://data.europa.eu/doi/10.2760/855840.

Feng, Z.-H., Kittler, J., Awais, M., Huber, P., Wu, X-J., 2018. Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2235–2245. https://doi.org/10.1109/CVPR.2018.00238.

Ferrentino, E., Nunziata, F., Marino, A., Migliaccio, M., Li, X., 2019. Detection of Wind Turbines in Intertidal Areas Using SAR Polarimetry. IEEE Geosci. Remote Sensing Lett. 16 (10), 1516–1520. https://doi.org/10.1109/LGRS.2019.2905714.

Global Wind Energy Concil (GWEC), 2023. Global Wind Report 2023. https://gwec.net/globalwindreport2023/ (accessed 20 August 2023).

Grassi, S., Junghans, S., Raubal, M., 2014. Assessment of the wake effect on the energy production of onshore wind farms using GIS. Appl. Energy. 136, 827–837. https://doi.org/10.1016/j.apenergy.2014.05.066.

Han, M., Wang, H., Wang, G., Liu, Y., 2018. Targets mask U-Net for wind turbines detection in remote sensing images. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII–3, 475–480. https://doi.org/10.5194/isprs-archives-XLII-3-475-2018.

Hoeser, T., Feuerstein, S., Kuenzer, C., 2022. DeepOWT: A global offshore wind turbine data set derived with deep learning from Sentinel-1 data. Earth Syst. Sci. Data. 14 (9), 4251–4270. https://doi.org/10.5194/essd-14-4251-2022.

International Energy Agency (IEA), 2023. Renewable Energy Market Update - June 2023. https://www.iea.org/reports/renewable-energy-market-update-june-2023 (accessed 20 August 2023).

Mandroux, N., Dagobert, T., Drouyer, S., Grompone Von Gioi, R., 2022a. Single Date Wind Turbine Detection on Sentinel-2 Optical Images. Image Process. Line. 12, 198–217. https://doi.org/10.5201/ipol.2022.384.

Mandroux, N., Dagobert, T., Drouyer, S., Von Gioi, R.G., 2021. Wind Turbine Detection on Sentinel-2 Images. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 4888–4891. https://doi.org/10.1109/IGARSS47720.2021.9554578.

Mandroux, N., Drouyer, S., Grompone Von Gioi, R., 2022b. Multi-Date Wind Turbine Detection on Optical Satellite Images. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. V-2–2022, 383–390. https://doi.org/10.5194/isprs-annals-V-2-2022-383-2022.

Manso-Callejo, M.-Á., Cira, C.-I., Alcarria, R., Arranz-Justel, J.-J., 2020. Optimizing the Recognition and Feature Extraction of Wind Turbines through Hybrid Semantic Segmentation Architectures. Remote Sens. 12 (22), 3743. https://doi.org/10.3390/rs12223743.

Manso-Callejo, M.-A., Cira, C.-I., Garrido, R.P.A., Matesanz, F.J.G., 2021. First Dataset of Wind Turbine Data Created at National Level With Deep Learning Techniques From Aerial Orthophotographs With a Spatial Resolution of 0.5 M/Pixel. IEEE J. Sel. Top Appl. Earth Obs. Remote Sens. 14, 7968–7980. https://doi.org/10.1109/JSTARS.2021.3101934.

Olabi, A.G., Obaideen, K., Abdelkareem, M.A., AlMallahi, M.N., Shehata, N., Alami, A.H., Mdallal, A., Hassan, A.A.M., Sayed, E.T., 2023. Wind Energy Contribution to the Sustainable Development Goals: Case Study on London Array. Sustain.-Basel. 15 (5), 4641. https://doi.org/10.3390/su15054641.

Padilla, R., Netto, S. L., Da Silva, E. A. B., 2020. A Survey on Performance Metrics for Object-Detection Algorithms. In: 2020 International Conference on Systems, Signals and Image Processing, 237–242. https://doi.org/10.1109/IWSSIP48289.2020.9145130.

Parada, L., Herrera, C., Flores, P., Parada, V., 2018. Assessing the energy benefit of using a wind turbine micro-siting model. Renew. Energ. 118, 591–601. https://doi.org/10.1016/j.renene.2017.11.018.

Rand, J.T., Kramer, L.A., Garrity, C.P., Hoen, B.D., Diffendorfer, J.E., Hunt, H.E., Spears, M., 2020. A continuously updated, geospatially rectified database of utility-scale wind turbines in the United States. Sci. Data. 7 (1), 15. https://doi.org/10.1038/s41597-020-0353-6.

Redmon, J., Farhadi, A., 2017. YOLO9000: Better, Faster, Stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 6517–6525. https://doi.org/10.1109/CVPR.2017.690.

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. https://doi.org/10.48550/arXiv.1804.02767.
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 779–788. https://doi.org/10.1109/CVPR.2016.91.
Sadorsky, P., 2021. Wind energy for sustainable development: Driving factors and future outlook. J. Clean. Prod. 289, 125779 https://doi.org/10.1016/j.jclepro.2020.125779.
Saidur, R., Rahim, N.A., Islam, M.R., Solangi, K.H., 2011. Environmental impact of wind energy. Renew. Sust. Energ. Rev. 15 (5), 2423–2430. https://doi.org/10.1016/j.rser.2011.02.024.
Sliz-Szkliniarz, B., Eberbach, J., Hoffmann, B., Fortin, M., 2019. Assessing the cost of onshore wind development scenarios: Modelling of spatial and temporal distribution of wind power for the case of Poland. Renew. Sust. Energ. Rev. 109, 514–531. https://doi.org/10.1016/j.rser.2019.04.039.
Song, M., Chen, K., Zhang, X., Wang, J., 2016. Optimization of wind turbine micro-siting for reducing the sensitivity of power generation to wind direction. Renew. Energ. 85, 57–65. https://doi.org/10.1016/j.renene.2015.06.033.
U.S. Energy Information Administration (EIA), 2023. ANNUAL ENERGY OUTLOOK 2023. https://www.eia.gov/outlooks/aeo/ (accessed 20 August 2023).
Wang, L., Luo, W., Xu, J., Xie, J., Luo, Z., Tan, A.C.C., 2022. Comparative study of decentralized instantaneous and wind-interval-based controls for in-line two scale wind turbines. Renew. Energ. 189, 1218–1233. https://doi.org/10.1016/j.renene.2022.03.074.
WindEurope, 2023. Wind energy in Europe: 2022 Statistics and the outlook for 2023-2027. https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2022-statistics-and-the-outlook-for-2023-2027/ (accessed 20 August 2023).
Wu, C., Yang, X., Zhu, Y., 2021. On the design of potential turbine positions for physics-informed optimization of wind farm layout. Renew. Energ. 164, 1108–1120. https://doi.org/10.1016/j.renene.2020.10.060.
Xu, D., Wu, Y., 2021. FE-YOLO: A Feature Enhancement Network for Remote Sensing Target Detection. Remote Sens. 13 (7), 1311. https://doi.org/10.3390/rs13071311.
Zakria, Z., Deng, J., Kumar, R., Khokhar, M.S., Cai, J., Kumar, J., 2022. Multiscale and Direction Target Detecting in Remote Sensing Images via Modified YOLO-v4. IEEE J. Sel. Top Appl. Earth Obs Remote Sens. 15, 1039–1048. https://doi.org/10.1109/JSTARS.2022.3140776.
Zhang, W., Wang, G., Qi, J., Wang, G., Zhang, T., 2021. Research on the Extraction of Wind Turbine all over the China Based on Domestic Satellite Remote Sensing Data. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. 4167–4170. https://doi.org/10.1109/IGARSS47720.2021.9553559.
Zhang, Y., Hao, Y., 2022. A Survey of SAR Image Target Detection Based on Convolutional Neural Networks. Remote Sens. 14 (24), 6240. https://doi.org/10.3390/rs14246240.
Zhang, T., Tian, B., Sengupta, D., Zhang, L., Si, Y., 2021a. Global offshore wind turbine dataset. Sci. Data. 8 (1), 191. https://doi.org/10.1038/s41597-021-00982-z.
Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016. Learning Deep Features for Discriminative Positioning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2921–2929. https://doi.org/10.1109/CVPR.2016.319.
Zhou, Y., Zhu, Y., Ye, Q., Qiu, Q., Jiao, J., 2018. Weakly Supervised Instance Segmentation Using Class Peak Response. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3791–3800. https://doi.org/10.1109/CVPR.2018.00399.
Zhou, S., Irvin, J., Wang, Z., Zhang, E., Aljubran, J., Deadrick, W., Rajagopal, R., Ng, A., n.d., 2019. DeepWind: Weakly Supervised Positioning of Wind Turbines in Satellite Imagery. In: 33rd Conference on Neural Information Processing Systems.
Zhu, Y., Zhou, Y., Ye, Q., Qiu, Q., Jiao, J., 2017. Soft Proposal Networks for Weakly Supervised Object Positioning. In: Proceedings of the IEEE International Conference on Computer Vision. 1859–1868. https://doi.org/10.1109/ICCV.2017.204.