Mask-Based IoU for Bounding Box Regression Using Medical Images

Serap Çakar Kaman[1,*] ID, Muhammed Kotan[2] ID, Cemil Öz[1] ID, Ahmet Furkan Sönmez[3] ID, Feyza Selamet[1] ID,
İbrahim Delibaşoğlu[4] ID

[1]*Department of Computer Engineering, Faculty of Computer and Information Sciences, Sakarya University, 54187, Sakarya, Türkiye.*
[2]*Department of Information Systems Engineering, Faculty of Computer and Information Sciences, Sakarya University, 54187, Sakarya, Türkiye.*
[3]*Department of Computer Engineering, Faculty of Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Türkiye.*
[4]*Department of Software Engineering, Faculty of Computer and Information Sciences, Sakarya University, 54187, Sakarya, Türkiye.*

ABSTRACT. Bounding box regression plays a pivotal role in the majority of object detection algorithms, significantly influencing the accuracy of object positioning and the regression speed of Convolutional Neural Networks (CNN). In object detection benchmarks, Intersection over Union (IoU) remains the widely adopted metric for evaluation. Traditional IoU-based loss functions often suffer from poor training outcomes and slow convergence, and they fail to account for situations where the predicted bounding box does not entirely capture the object's mask. This study introduces the Mask-based Intersection over Union (MbIoU) metric for improving bounding box regression in object detection using medical images. The proposed MbIoU metric incorporates the object mask into the bounding box regression process, offering a more precise evaluation of how well the predicted bounding box encapsulates the object. The developed MbIoU metric was tested on the MNIST: HAM10000 dermoscopic skin images dataset, COVID-19 CT dataset, and Brain Tumor dataset and compared to traditional IoU metrics. The results show that MbIoU enhances the prediction by better capturing the object's contained mask.

1. Introduction

Object detection continues to be a significant subject in the realm of computer vision, playing a crucial part in diverse applications such as autonomous vehicles, surveillance, and image comprehension. It has garnered extensive attention over the past few decades. The object detection task involves determining the location and classifying objects in an image. Region-based methods such as Region Based CNN (R-CNN) [20], Fast Region Based CNN (Fast R-CNN) [7], Faster Region-Based CNN (Faster R-CNN) [26], Single Shot Multibox Detector (SSD) [17], You Only Look Once (YOLO) [25] and RetinaNet [16] are used for object detection. Strategies can be developed to enhance the efficiency of applications using deep neural networks, such as using a better architectural backbone, developing a more

Corresponding Author

Email addresses: scakar@sakarya.edu.tr (S.Ç. Kaman), mkotan@sakarya.edu.tr (M. Kotan), coz@sakarya.edu.tr (C. Öz), afsonmez@beun.edu.tr (A.F. Sönmez), feyzacerezci@sakarya.edu.tr (F. Selamet), ibrahimdelibasoglu@sakarya.edu.tr (İ. Delibaşoğlu)

effective strategy for extracting reliable local features, or using a metric based on IoU rather than the $\ell 1$ and $\ell 2$ norms of regression losses [27].

Achieving high performance in object detection tasks demands a robust loss function that effectively captures both localization accuracy and classification precision. Mean Squared Error (MSE) is a traditional regression loss function commonly used in object detection. It penalizes the squared differences between ground truth and predicted bounding box coordinates. In recent years, most metrics used to evaluate segmentation, object detection, and tracking are based on the IoU metric [27]. The field of IoU-based bounding box regression has seen significant advancements, including the introduction of novel methods, loss functions, and refinement techniques to enhance the effectiveness and precision of object detection in computer vision. Nevertheless, the IoU-based loss functions that have been developed still encounter issues like suboptimal training outcomes and sluggish convergence speeds.

In addition, problems have been encountered, such as the IoU value not being zero when the predicted box does not contain the object mask that is in the ground truth box. Even though the predicted box covers very little of the object mask, the IoU value may be high, or even if the predicted box covers a large part of the object mask, the IoU value may be low. For these reasons, the MbIoU metric that considers the object mask was developed in this study. The developed MbIoU metric was tested on the MNIST:HAM10000 dermoscopic skin images dataset [31], COVID-19 CT lung and infection segmentation dataset [4], and Brain Tumor dataset [1].

2. Related Works

Object detection remains a central concern within the realm of computer vision, drawing extensive interest over the past few decades. The traditional approaches utilize manually designed feature descriptors, including Histogram of Oriented Gradients (HOG) [6], Local Binary Patterns (LBP) [21], and Scale-Invariant Feature Transform (SIFT) [19] for extracting image features. Subsequently, a machine learning-driven classifier is employed to systematically the complete image using a sliding window approach, aiming to identify and localize regions resembling objects.

The Deformable Part Model (DPM) [9] and its adaptations have been the primary techniques for object recognition from images over an extended period. Due to CNN's outstanding performance in large-scale object recognition [15], many CNN-based techniques have been developed [7, 8, 11], and [12], significantly improving object detection performance. Object identification techniques are built on R-CNN frameworks, which use bounding boxes to accurately propose the region that is considered to be the ground truth. Two-stage methods that combine a region proposal predictor with region-wise classifiers include Spatial Pyramid Pooling (SPP)-Net [11], Fast R-CNN [7], Faster R-CNN [12], R-FCN [5], FPN [10], Cascade R-CNN [5], and others. However, single-forward CNNs are used by one-stage methods like SSD [17], YOLO [13], and RetinaNet [16] to predict classes and localize the objects.

Recently, one-stage procedures gained popularity due to their computational efficiency, whereas two-stage approaches have become renowned for their superior performance. Proposed extensions seek to overcome the limitations that arise in certain complex cases. A rotation-invariant and Fisher discriminative CNN should be integrated, according to Cheng et al. [3], to enhance feature representation in R-CNN models. A search strategy based on Bayesian optimization was developed by Zhang et al. [36] with the goal of increasing the efficiency of region suggestions. R-CNN was trained using this method to improve object bounding box localization. Cai et al. [2] suggested training a set of three Faster R-CNNs with various IoU thresholds (0.5, 0.6, and 0.7) in order to improve performance in object identification and enrich the learning process.

In recent works, bounding box regression often relies on the utilization of IoU [24] and its improved versions. Generalized Intersection over Union (GIoU) loss [27] is a metric and loss function that improves the localization accuracy of predicted bounding boxes. IoU focuses on the overlap, but it doesn't take into account factors like the size, position, and aspect ratio of the bounding boxes. The GIoU loss was designed to solve these limitations with a more comprehensive measure of the accuracy of predicted bounding boxes. Distance Intersection over Union (DIoU) loss [38] also aims to be an improvement over the traditional IoU and GIoU metrics. DIoU aims to further refine the evaluation and training of object detection models by taking into account the distance between box centers and the aspect ratio of the bounding boxes. While GIoU and DIoU introduce improvements over traditional IoU, they can be computationally expensive due to their additional terms. Complete Intersection over Union (CIoU) [33] aims to strike a balance between effectiveness and computational efficiency. ICIoU proposes an improved penalty function over CIoU for better localization accuracy. N-IoU [30] loss incorporates the Dice coefficient into the computation of regression loss, presenting a novel metric that surpasses and can serve as a substitute for the IoU.

Also, Efficient-IOU (EIoU), Rotated-IoU (RIoU), Focal-RIoU (FRIoU), Gaussian-IoU (GsIoU), Manhattan-distance-IOU (MIoU), the combination of smooth $\ell1$-norm loss and IoU (LIoU), Valid-IoU (VIoU) metrics have also been improved for enhanced performance in object detection tasks [18, 23, 29, 32, 35, 37]. Although performance seems to increase with these methods, they are based on the proximity of the boxes to each other. These metrics do not take into account the object mask. Designing an effective bounding box regression loss function is crucial. The choice of loss function must balance the precision of object localization and the stability of the training process.

3. IoU-Based Bounding Box Regression

Object detection tasks typically contain Bounding Box Regression Loss and Classification Loss. The n-norm loss functions are commonly used in bounding box regression. However, they are sensitive to scale variations. IoU loss is also employed since Unitbox is scale-invariant [34]. Commonly used IoU-based bounding box regression loss functions are explained below. The area of union between the predicted bounding box B and the ground truth bounding box B_{gt} is measured by the IoU (also known as the Jaccard index) (Eq. (3.1)) and is shown in Figure 1.
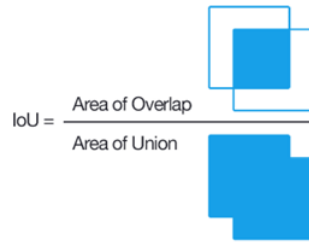

FIGURE 1. Intersection Over Union (IoU) [14]

$$J\left(B, B_{gt}\right) = IoU = \frac{area\left(B \cap B_{gt}\right)}{area\left(B \cup B_{gt}\right)}. \tag{3.1}$$

We can identify a detection as accurate or inaccurate by comparing the IoU to a defined threshold, t. If IoU $\geq t$, the detection is deemed to be accurate. If IoU $< t$, the detection is deemed inaccurate [28, 34]. The formulation of IoU loss [22] is as follows:

$$L_{IoU} = 1 - IoU.$$

IoU loss, on the other hand, only functions in situations when the bounding boxes overlap; in non-overlapping scenarios, it would not produce a moving gradient. Next, by including a penalty word, GIoU loss [27] is suggested.

$$L_{GIoU} = 1 - IoU + \frac{\left|C - \left(B \cup B_{gt}\right)\right|}{|C|},$$

where the smallest box that covers B and B_{gt} is C. In non-overlapping circumstances, the prediction box will converge to the big truth box as a result of the penalty term's inclusion.

Since IoU loss does not introduce any motion gradient when the candidate frame and the actual frame do not overlap ($\mathcal{L}$IoU is always 1), GIoU loss introduces a penalty term. Due to the introduction of penalty time, the prediction box will move towards the ground truth box if there is no overlap. In the DIoU metric [22], the central point distance between two boxes is also taken into account, along with the overlap area in the suppression criterion. DIoU loss introduces a penalty term based on IoU loss, defined as follows:

$$L_{DIoU} = 1 - IoU + \frac{p^2\left(c_i, c_j\right)}{d_e^{\,2}},$$

where c_i and c_j are the center points of B and B_{gt}, respectively, $p(\cdot, \cdot)$ is the Euclidean distance between two locations, and d_e is the length of C's diagonal.

CIoU loss [22] introduces a penalty term taking into account the box aspect ratio and is defined as follows:

$$L_{CIoU} = 1 - IoU + \frac{p^2\left(c_i, c_j\right)}{d_e^{\,2}} + \mu\upsilon,$$

$$\upsilon = \frac{4}{\pi^2}\left(arctan\frac{w_i}{h_i} - arctan\frac{w_j}{h_j}\right)^2,$$

$$\mu = \frac{\upsilon}{(1 - IoU + \upsilon)},$$

where w_i and h_i (w_j and h_j) stand for the height and width of B (B_{gt}), respectively, and μ represents the weight of υ. Additionally, υ indicates the deviation of the aspect ratio.

4. Proposed Method

4.1. **Mask-Based IoU (MbIoU).** Despite being widely used, IoU has a specific limitation: It relies solely on the areas of the anchor box and the ground truth box, neglecting the actual shape of the object, such as that defined by a segmentation mask. This can lead to unfavorable assignments, as the IoU scores may paradoxically be lower or higher, disregarding the expected correlation with object overlap. For example, when the predicted bounding box does not contain the object mask, the IoU value does not include information in this context, as in some cases (Figure 3 (H), Figure 4 (L)). Even though the predicted box covers very little of the object mask, the IoU value may be high (Figure 3 (D), (E), Figure 4 (C), (G)), or even if the predicted box covers a large part of the object mask, the IoU value may be low (Figure 3 (A), (G), Figure 4 (B), (E), (F), (J)). For these reasons, the MbIoU metric was developed in this study as shown in Figure 2 and Equation 4.1. The MbIoU metric takes both bounding box areas and the object mask areas into account.

The MbIoU metric was tested on the MNIST:HAM10000 [31], COVID-19 CT [4], and Brain Tumor [1] datasets. MNIST:HAM10000 dataset contains 10015 dermoscopic skin images and seven different classes. These classes are Actinic Keratoses (akiec), Basal cell carcinoma (bcc), Benign keratosis (bkl), Dermatofibroma (df), Melanoma (mel), Melanocytic nevi (nv) and Vascular (vasc), respectively. COVID-19 CT dataset consists of three separate COVID-19 datasets and contains 2729 images. The Brain Tumor dataset contains 305 brain MRI images for brain tumor detection.

Figure 2 shows how the MbIoU metric is calculated on a sample image. The green area (*GA*) is the lesion mask, the blue area (*BA*) is the subtraction of the green area from the area of the ground truth bounding box, the red area (*RA*) is the lesion mask in the predicted bounding box, and the yellow area (*YA*) is the subtraction of the red area from the area of the predicted bounding box.

$$MbIoU = \frac{\frac{RA}{RA+YA+GA+BA+(GA-RA)+|YA-BA|}}{\frac{GA}{RA+YA+GA+BA}} \tag{4.1}$$
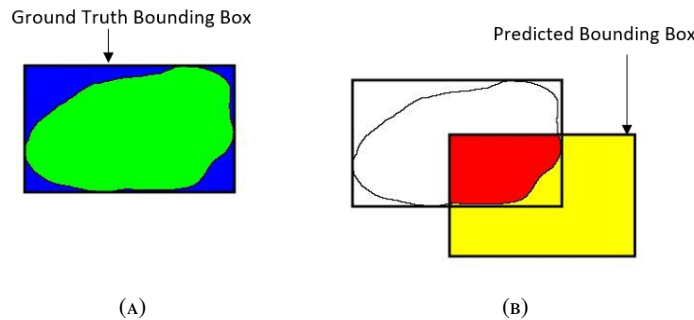

FIGURE 2. (A) The green area is the lesion mask, and the blue area is the subtraction of the green area from the area of the ground truth bounding box. (B) The red area is the lesion mask in the predicted bounding box, and the yellow area is the subtraction of the red area from the area of the predicted bounding box.

(A)
IoU:0.9159   GIoU:0.9159
DIoU:0.9157 CIoU:0.9157
MbIoU:0.9424

(B)
IoU:0.1595   GIoU:-0.031
DIoU:0.0566 CIoU:0.0566
MbIoU:0.1718

(C)
IoU:0.3238   GIoU:0.3238
DIoU:0.2571 CIoU:0.2570
MbIoU:0.2770

(D)
IoU:0.1945   GIoU:0.1945
DIoU:0.1233 CIoU:0.1233
MbIoU:0.1143

(E)
IoU:0.3985   GIoU:0.3985
DIoU:0.3553 CIoU:0.3504
MbIoU:0.1872

(F)
IoU:0.1395   GIoU:-0.091
DIoU:0.0259 CIoU:0.0249
MbIoU:0.0318

(G)
IoU:0.4686   GIoU:0.4686
DIoU:0.4557 CIoU:0.4557
MbIoU:0.5888

(H)
IoU:0.1546   GIoU:0.0886
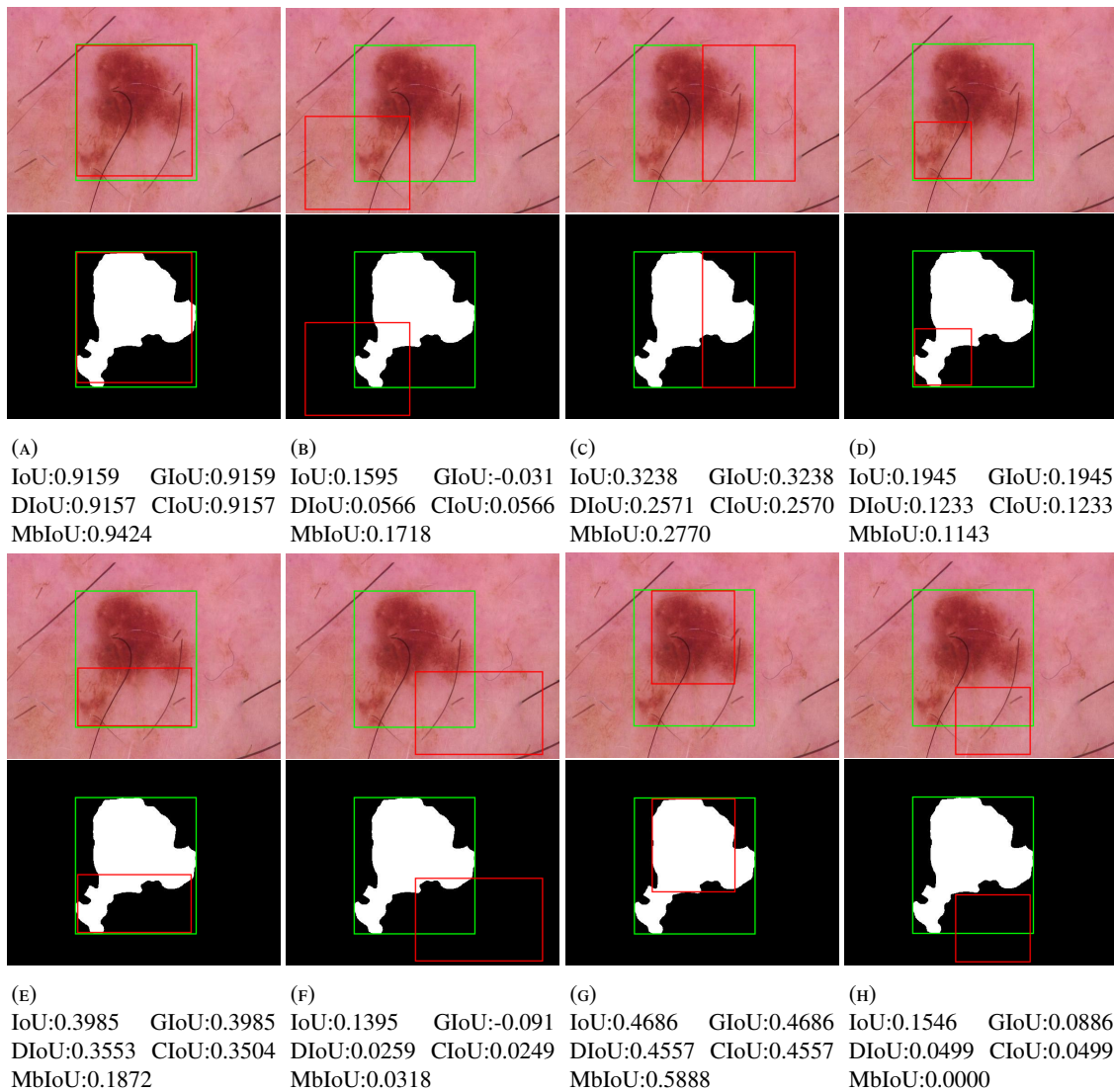DIoU:0.0499 CIoU:0.0499
MbIoU:0.0000

FIGURE 3. Results of MbIoU and other IoU metrics on nv image in different situations. Green boxes are ground truth and red boxes are predicted bounding boxes.

MbIoU $\in [0, 1]$. According to Equation (4.1), as the RA/GA ratio increases, the MbIoU value increases, and as $(GA-RA)$ and $|YA - BA|$ values increase, the MbIoU value decreases. When the bounding boxes overlap, the MbIoU value becomes 1. Thus, the more the predicted bounding box covers the object mask, the higher the MbIoU. The penalty rate increases as the predicted bounding box cannot cover the object mask.

4.2. **Experimental Results.** The MbIoU metric has been tested on images selected from three databases for several different situations and compared with other commonly used IoU metrics, as shown in Figures 3 and 4. As seen in Figure 3 (A), (G), and Figure 4 (B), (E), (F), (J), the MbIoU metric result is higher than other metrics results when the predicted bounding box contains most of the object mask. In Figure 3 (D), (E), and Figure 4 (C), (G) the predicted box contained a very small part of the object mask, and thus the MbIoU metric result was lower than other metrics results. In Figure 3 (H), and Figure 4 (L), the predicted box does not contain the object mask at all. In this case, although the other metric results were non-zero, the MbIoU metric result was zero. Thus, more consistent loss values, eliminated unnecessary anchors, shortened training time, and faster convergence can be obtained with the MbIoU metric.
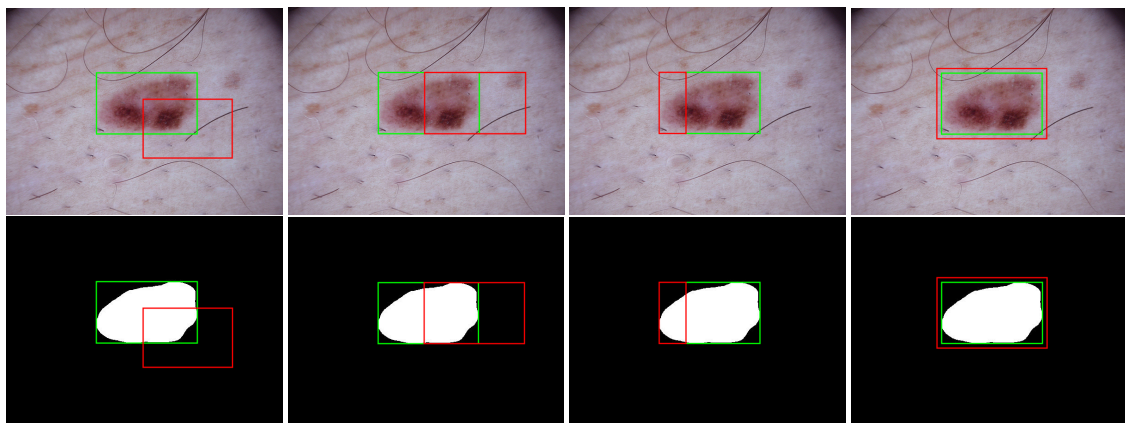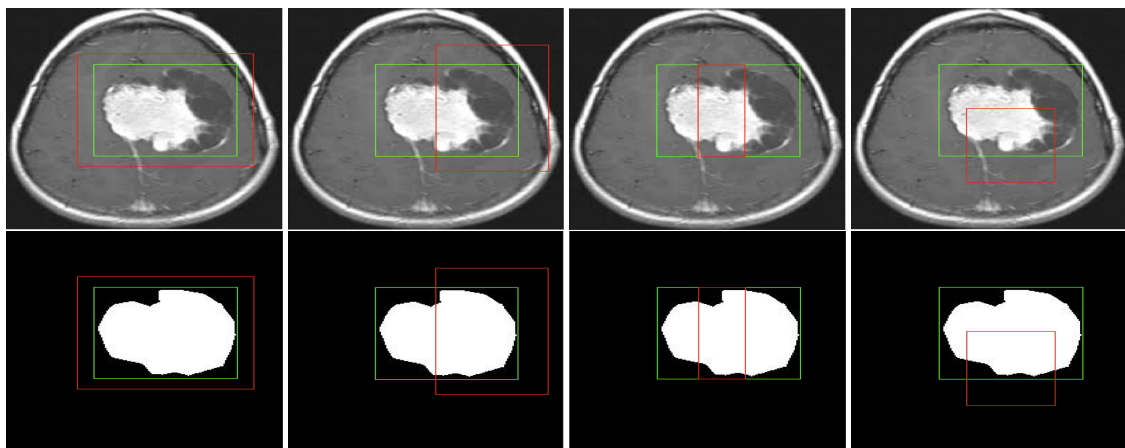
(A)
IoU:0.1971   GIoU:0.01997
DIoU:0.1073   CIoU:0.1073
MbIoU:0.2262

(B)
IoU:0.3710   GIoU:0.3710
DIoU:0.2867   CIoU:0.2867
MbIoU:0.4557

(C)
IoU:0.2660   GIoU:0.2660
DIoU:0.1675   CIoU:0.1414
MbIoU:0.1226

(D)
IoU:0.7954   GIoU:0.7954
DIoU:0.7954   CIoU:0.7954
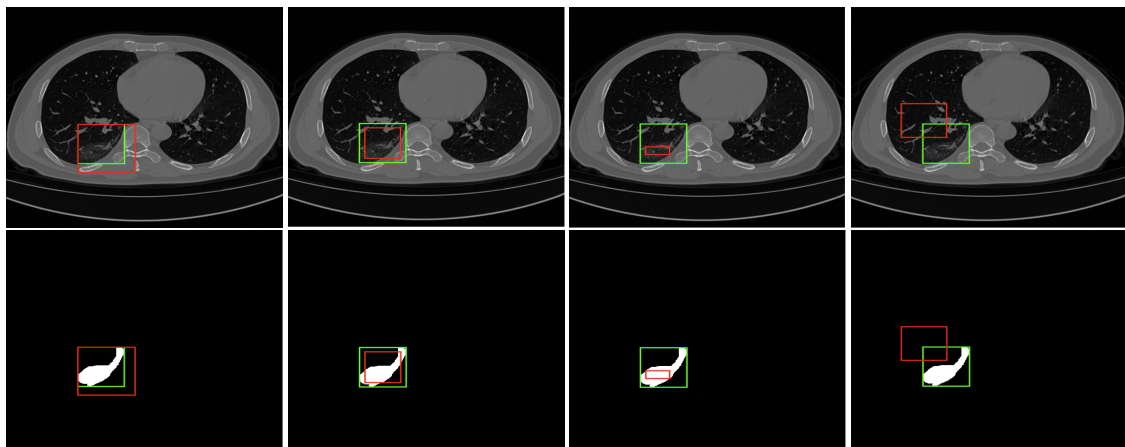MbIoU:0.8992

(E)
IoU:0.6601   GIoU:0.6601
DIoU:0.6601   CIoU:0.6601
MbIoU:0.8312

(F)
IoU:0.3827   GIoU:0.2875
DIoU:0.3526   CIoU:0.3523
MbIoU:0.5274

(G)
IoU:0.3269   GIoU:0.3269
DIoU:0.3257   CIoU:0.3153
MbIoU:0.2724

(H)
IoU:0.2713   GIoU:0.1852
DIoU:0.2157   CIoU:0.2156
MbIoU:0.2985

(I)
IoU:0.6651   GIoU:0.6651
DIoU:0.6566   CIoU:0.6566
MbIoU:0.8339

(J)
IoU:0.5987   GIoU:0.5987
DIoU:0.5987   CIoU:0.5987
MbIoU:0.6237

(K)
IoU:0.1012   GIoU:0.1012
DIoU:0.0761   CIoU:0.0705
MbIoU:0.1654

(L)
IoU:0.1048   GIoU:-0.1467
DIoU:-0.0252 CIoU:-0.0252
MbIoU:0

FIGURE 4. Results of MbIoU and other IoU metrics on mel image (A, B, C, D), Brain Tumor image (E, F, G, H) and COVID-19 CT image (I, J, K, L) in different situations. Green boxes are ground truth and red boxes are predicted bounding boxes.

5. Conclusion

In this research, we introduced the MbIoU metric, which takes into account both the object mask and the proximity between the ground truth bounding box and the predicted bounding box. In this approach, when the predicted bounding box better encompasses the object mask, the MbIoU value increases. Additionally, the MbIoU value decreases when the diversity bounding box covers a small portion of the object mask. Thus, improved training outcomes and faster convergence can be achieved by eliminating the predicted boxes where the object mask is not covered or only slightly covered. The proposed metric underwent testing using the HAM10000, COVID-19 CT, and Brain Tumor datasets and was presented in the figures. As evident from the figures, in some cases, the MbIoU metric results differ from other common IoU metrics. In future studies, training and testing operations will be carried out with the loss function calculated using the MbIoU metric we created, and the results will be compared. Moreover, investigating the effectiveness of the MbIoU metric in real-time object detection systems could further validate its utility in practical applications.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Authors Contribution Statement

S. Ç. Kaman: Creation of the idea, organizing the execution of the study, literature research, contribution to article writing and study. M. Kotan: Data processing, organizing the execution of the study, contribution to article writing and study. C. Öz: Creation of the idea, organizing the execution of the study, contribution to article writing and study. A. F. Sönmez: Data collection, literature research, contribution to article writing and study. F. Selamet: Literature research, contribution to article writing and study. İ. Delibaşoğlu: Literature research, contribution to article writing and study.

References

[1] Brain Tumor Segmentation, Available at: https://github.com/rastislavkopal/brain-tumor-segmentation/tree/main/brain_tumor_data. Accessed on: 2024-09-20.

[2] Cai, Z., Vasconcelos, N., *Cascade R-CNN: Delving Into High Quality Object Detection*, Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2018.

[3] Cheng, G., Han, J., Zhou, P., Xu, D., *Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection*, IEEE Transactions On Image Processing, **28**(2018), 265–278.

[4] *COVID-19 CT Scan Lesion Segmentation Dataset*. Available at: https://www.kaggle.com/datasets/maedemaftouni/covid19-ct-scan-lesion-segmentation-dataset. Accessed on: 2024-09-20.

[5] Dai, J., Li, Y., He, K., Sun, J., *R-fcn: Object detection via region-based fully convolutional network*, Advances In Neural Information Processing Systems, **29**(2016).

[6] Dalal, N., Triggs, B., *Histograms of oriented gradients for human detection*, 2005 IEEE Computer Society Conference On Computer Vision And Pattern Recognition (CVPR'05), 2005.

[7] Girshick, R., *Fast r-cnn*,Proceedings of the ieee international conference on computer vision, **2**(2015), 1440–1448.

[8] Girshick, R., Donahue, J., Darrell, T., Malik, J., *Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation*, Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2014.

[9] Girshick, R., Iandola, F., Darrell, T., Malik, J., *Deformable Part Models are Convolutional Neural Networks*, Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2015.

[10] Gong, Y., Yu, X., Ding, Y., Peng, X., Zhao, J. et all., *Effective Fusion Factor in FPN for Tiny Object Detection*,Proceedings Of The IEEE/CVF Winter Conference On Applications Of Computer Vision (WACV), 2021.

[11] He, K., Zhang, X., Ren, S., Sun, J., *Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition*, IEEE Transactions On Pattern Analysis And Machine Intelligence, **37**(2014), 1904–1916.

[12] Jiang, H., Learned-Miller, E., *Face detection with the faster R-CNN*, 12th IEEE International Conference On Automatic Face & Gesture Recognition, 2017.

[13] Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B., *A Review of Yolo algorithm developments*, Procedia Computer Science, **199**(2022), 1066–1073.

[14] Junior, G., Ferreira, J., Millán-Arias, C., Daniel, R., Casado, A. et all. *Ceramic Cracks Segmentation with Deep Learning*, Applied Sciences, **11**(6)(2021), 6017.

[15] Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J., *A survey of convolutional neural networks: analysis, applications, and prospects*, IEEETransactions On Neural Networks And Learning Systems, 2021.

[16] Lin, T., Goyal, P., Girshick, R., He, K., Dollar, P., *Focal Loss for Dense Object Detection*, Proceedings Of The IEEE International Conference On Computer Vision (ICCV), 2017.

[17] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. et all., *Ssd: Single shot multibox detector*, Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, 21–37.

[18] Liu, X., Hu, J., Wang, H., Zhang, Z., Lu, X., et all., *Gaussian-IoU loss: Better learning for bounding box regression on PCB component detection*, Expert Systems With Applications, **190**(2022), 116178.

[19] Lowe, D., *Distinctive image features from scale-invariant keypoints*, International Journal Of Computer Vision, **60**(2004), 91–110.

[20] Malik, J., Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, ACM: New York, NY, USA, 2014.

[21] Ojala, T., Pietikainen, M., Maenpaa, T., *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*, IEEE Transactions On Pattern Analysis And Machine Intelligence, **24**(2002), 971–987.

[22] Padilla, R., Netto, S., Da Silva, E., *A survey on performance metrics for object-detection algorithms*, 2020 International Conference On Systems, Signals And Image Processing (IWSSIP), 237–242.

[23] Qian, X., Wu, B., Cheng, G., Yao, X., Wang, W. et all. *Building a bridge of bounding box regression between oriented and horizontal object detection in remote sensing images*, IEEE Transactions On Geoscience And Remote Sensing, **61**(2023), 1–9.

[24] Rahman, M., Wang, Y. *Optimizing intersection-over-union in deep neural networks for image segmentation*, International Symposium On Visual Computing, (2016), 234–244.

[25] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., *You Only Look Once: Unified, Real-Time Object Detection*, Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2016.

[26] Ren, S., He, K., Girshick, R., Sun, J., *Faster r-cnn: Towards real-time object detection with region proposal networks*, Advances In Neural Information Processing Systems, **28**(2015).

[27] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I. et all., *Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression*, Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR), 2019.

[28] Selamet, F., Cakar, S., Kotan, M., *Automatic detection and classification of defective areas on metal parts by using adaptive fusion of faster R-CNN and shape from shading*, IEEE Access, **10**(2022), 126030–126038.

[29] Shen, Y., Zhang, F., Liu, D., Pu, W., Zhang, Q., *Manhattan-distance IOU loss for fast and accurate bounding box regression and object detection*, Neurocomputing, **500**(2022), 99–114.

[30] Su, K., Cao, L., Zhao, B., Li, N., Wu, D. et all., *N-IoU: better IoU-based bounding box regression loss for object detection*, Neural Computing And Applications, (2023), 1–15.

[31] Tschandl, P., Rosendahl, C., Kittler, H., *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*, Scientific Data, **5**(1-9)(2018).

[32] Vakili, E., Karimian, G., Shoaran, M., Yadipour, R., Sobhi, J., Valid-IoU: An Improved IoU-based Loss Function and Its Application to Detection of Defects on Printed Circuit Boards, 2023.

[33] Wang, X.,Song, J., *ICIoU: Improved loss based on complete intersection over union for bounding box regression*, IEEE Access,**9**(2021), 105686–105695

[34] Yu, J., Jiang, Y., Wang, Z., Cao, Z., Huang, T., *Unitbox: An advanced object detection network*, Proceedings Of The 24th ACM International Conference On Multimedia, (2016), 516–520.

[35] Zhai, H., Cheng, J., Wang, M., *Rethink the IoU-based loss functions for bounding box regression*, 2020 IEEE 9th Joint International Information Technology And Artificial Intelligence Conference (ITAIC), **9**(2020), 1522–1528.

[36] Zhang, Y., Sohn, K., Villegas, R., Pan, G., Lee, H., *Improving Object Detection With Deep Convolutional Networks via Bayesian Optimization and Structured Prediction*, Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition (CVPR), 2015.

[37] Zhang, Y., Ren, W., Zhang, Z., Jia, Z., Wang, L. et all., *Focal and efficient IOU loss for accurate bounding box regression*, arXiv 2021, ArXiv Preprint ArXiv:2101.08158.

[38] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R. et all., *Distance-IoU loss: Faster and better learning for bounding box regression*, Proceedings Of The AAAI Conference On Artificial Intelligence, **34**(2020), 12993–13000.