

Д. К. МАРЧУК

старший викладач кафедри комп'ютерних наук
Державний університет «Житомирська політехніка»
ORCID: 0000-0001-8675-8047

М. С. ГРАФ

доктор філософії (Ph.D.) з комп'ютерних наук,
завідувач кафедри комп'ютерних наук
Державний університет «Житомирська політехніка»
ORCID: 0000-0003-4873-548X

МЕТОДИ ОЦІНКИ ЕФЕКТИВНОСТІ МОДЕЛЕЙ ВИЯВЛЕННЯ ОБ'ЄКТІВ У КОМП'ЮТЕРНОМУ ЗОРІ

Основними завданнями комп'ютерного зору є розпізнавання, виявлення об'єктів та сегментація. Розпізнавання зображень використовується в різних галузях від систем безпеки до діагностування у медицині. Виявлення об'єктів – це техніка визначення місцезнаходження з подальшим розпізнаванням у реальному часі. Сегментація – це процес розбиття зображення на багато сегментів. Процес побудови тієї чи іншої моделі може бути складним і щоб модель відповідала поставленому завданню в повному обсязі потрібно визначити її ефективність. Метою дослідження є огляд показників ефективності, точності, продуктивності моделей комп'ютерного зору. Було описано класична версія показника перетин через об'єднання (IoU). Представлені різні модифікації і покращення IoU такі як багато масштабний IoU (MSIoU), обмежувальний IoU (BalIoU) та BhIoU. Особу увагу приділено узагальненому перетину через об'єднання (GIoU), для усунення недоліків втрати IOU. Тобто втрата IOU завжди буде дорівнювати нулю, коли два блоки не взаємодіють – не перетинаються. Продуктивність алгоритмів комп'ютерного зору виявлення об'єктів і сегментації зазвичай перевіряється за допомогою середнього значення середнього (mAP). А так як mAP базується на різних субметриках, було розглянуто матриця плутанини, перетин через об'єднання, Recall та Precision. Для кращого розуміння показників було продемонстровано приклад з розрахунком продуктивності (Accuracy), точності (Precision), ефективності класифікації (Recall), гармонійного середнього значення точності та чутливості моделі (F1-Score). Насамкінець, очевидно, що це дослідження показало що за допомогою цих метрик можна перевірити, наскільки точна навчена модель.

Ключові слова: комп'ютерний зір, метрики оцінювання, IoU, Recall, F1-Score, Precision, Confusion matrix, mAP.

D. K. MARCHUK

Senior Lecturer at the Department of Computer Science
Zhytomyr Polytechnic State University
ORCID: 0000-0001-8675-8047

M. S. GRAF

Ph.D., Head of the Department of Computer Science
Zhytomyr Polytechnic State University
ORCID: 0000-0003-4873-548X

METHODS FOR EVALUATING THE EFFECTIVENESS OF OBJECT DETECTION MODELS IN COMPUTER VISION

The main tasks of computer vision are recognition, object detection, and segmentation. Image recognition is used in a variety of industries from security systems to medical diagnostics. Object detection is a technique for locating objects and then recognizing them in real time. Segmentation is the process of dividing an image into many segments. The process of building a model can be complex, and in order for the model to fully meet the task, it is necessary to determine its effectiveness. The purpose of the study is to review the performance, accuracy, and productivity of computer vision models. The classical version of the intersection of union (IoU) indicator is described. Various modifications and improvements of the IoU such as the multi-scale IoU (MSIoU), the bounding IoU (BalIoU), and the BhIoU are presented. Particular attention is paid to the generalized intersection through union (GIoU) to eliminate the disadvantages of IOU loss. That is, the IOU loss will always be zero when two blocks do not interact – do not intersect. The performance of computer vision algorithms for object detection and segmentation is usually tested using the mean of the mean (mAP). And since mAP is based on various submetrics, we considered the confusion matrix, Intersection through Merging, Recall, and Precision. For a better understanding of the metrics, an example was demonstrated with the calculation of accuracy,

precision, recall, harmonic mean of accuracy, and model sensitivity (F1-Score). Finally, it is obvious that this study has shown that these metrics can be used to check how accurate a trained model is.

Key words: computer vision, evaluation metrics, IoU, Recall, F1-Score, Precision, confusion matrix, mAP.

Постановка проблеми

Машинне навчання часто використовується для рішення задач які сприяють прийняттю того чи іншого рішення, для цього будується модель. Процес побудови такої моделі може бути складним і розроблена модель повинна відповідати задачі в повному обсязі. Визначення ефективності побудованої моделі є важливим кроком, тому що може впливати на прийняття рішення.

Для вирішення завдань, що пов'язані комп'ютерним зором, найчастіше використовуються різні моделі нейронних мереж на базі CNN. Наприклад, це можуть бути AlexNet, GoogleNet, VGGNet, тощо. Кожна із них має свої особливості та працюватиме по-різному залежно від різних факторів. Щоб визначитися з найкращою моделлю шляхом об'єктивного порівняння потрібно мати різні метрики оцінки та показники продуктивності.

Метою дослідження є аналіз різних показників ефективності і оцінки для моделей виявлення об'єктів у комп'ютерному зорі. Метрики оцінки виявлення об'єктів повинні враховувати як категорію, так і розташування об'єктів, і саме для цього використовується mAP. А щоб зрозуміти mAP, необхідно розуміти IoU, Precision, Recall та криву Precision-Recall.

Аналіз останніх досліджень та публікацій

Перетин через об'єднання (Intersection Over Union, IoU) є найпопулярнішим показником оцінки, який використовується в контрольних тестах виявлення об'єктів. Завдяки своїй привабливості і властивостям IoU використовується для оцінки виявлення об'єктів [1, 2, 3, 4] і відстеження [5, 6].

Автори статей [7, 8, 9] запропонували доповнення, покращення IoU. У роботі [7] автори підкреслили нечутливість популярних оціночних метрик до структури виявлених об'єктів і запропонували нову метрику для вирішення цієї проблеми. Автори пропонують Multiscale IoU (MIoU), яка є поєднанням популярної метрики оцінювання, а саме перетину через об'єднання (IoU) і геометричної концепції, яка називається фрактальним виміром. MIoU, як стверджують автори, чутлива до тонких граничних структур, які повністю не помічаються IoU та F1-Score. У документі [8] була розглянута проблема виявлення 2D/3D об'єктів, використовуючи IoU для двох повернутих Bbox. Автори запропонували уніфікований рівень втрат IoU, який можна застосувати для фреймворків виявлення 2D/3D об'єктів, вирівняних по осі або повернутих. Зокрема, запропоновані втрати IoU працюють набагато краще, коли поріг IoU встановлено на високе значення. У дослідженні [10] висвітлюється проблема великої похибки навчання та низької точності навчання на основі функції втрат IoU і запропоновано дві вдосконалені версії функцій втрат VaIoU та BhIoU, покращуючи форму IoU для покращення алгоритму IoU. Експеримент регресійного моделювання обмежувальної рамки доводить, що VaIoU та BhIoU можуть ефективно подолати проблеми повільної збіжності та великої похибки навчання функції втрат на основі IoU.

Викладення основного матеріалу дослідження

Виявлення об'єктів відноситься до області комп'ютерного зору, яка займається локалізацією та класифікацією об'єктів, що містяться на зображенні чи відео. В процесі виявлення об'єкту навколо нього малюється обмежувальна рамка і об'єкт в даному випадку вже є класифікованим.

Існує багато показників для оцінки моделей машинного навчання, а саме для виявлення об'єктів. Кожен показник має свої переваги та недоліки.

Наприклад, показник перетин через об'єднання (Intersection over Union, IoU), також відомий як коефіцієнт Жаккара (Jaccard), оцінює точність детектору об'єктів, показник, який кількісно визначає ступінь перекриття між двома регіонами.

IoU між двома обмежувальними рамками обчислюється як відношення площі перетину до площі об'єднання (рис. 1).

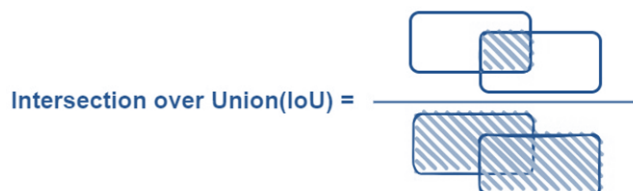


Рис. 1. Показник перетин через об'єднання

Також IoU можна представити у вигляді формули:

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Для обчислення IoU необхідно знайти координати області перетину і об'єднання, це мінімальні і максимальні значення координат двох обмежувальних рамок. Отримане значення буде коливатися від 0 до 1. За допомогою значення IoU можна визначити, чи є прогноз істинно-позитивним (True Positive, TP), істинно-негативним (True Negative, TN), хибно-позитивним (False Positive, FP) чи хибно-негативним (False Negative, FN) (рис. 2). Загалом поріг IoU, який можна використати для виявлення об'єктів, становить 0,5. Якщо IoU перевищує 0,5 прогноз можна вважати істинно-позитивним або хибно-позитивним, якщо менше 0,5. Найкраще значення буде дорівнювати 1, але це дуже мало ймовірно.

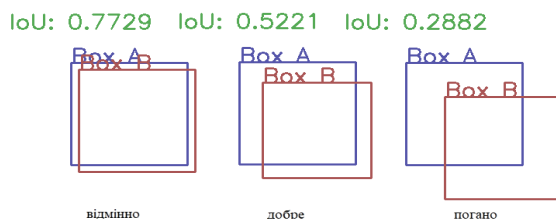


Рис. 2. Визначення IoU у трьох різних сценаріях

Якщо $IoU = 0$, це означає що обмежуючі рамки не перетинаються (рис.3), і фігури можуть бути рядом або знаходитися далеко одна від одної.

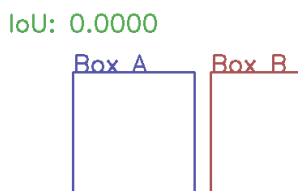


Рис. 3. $IoU=0$

В дослідження [9] було запропоновано узагальнену версію IoU яка вирішує проблему, де $|A \cap B| = 0$, $IoU(A, B) = 0$ під назвою узагальнений перетин через об'єднання (Generalized Intersection over Union, GIoU).

Мета методу полягає в тому щоб знайти для A і B найменший опуклий об'єкт C , де $C \subseteq S \in R$. Якщо $A, B \subseteq S \in R$. Потім обчислити співвідношення між об'ємом (площею), яку займає C , за винятком A і B , і поділити на загальний об'єм (площу), яку займає C . Це нормалізований показник, який зосереджується на порожньому об'ємі (площі) між A і B . Зрештою, GIoU визначається шляхом віднімання цього відношення від значення IoU, що можна представити наступною формулою:

$$GIoU = \frac{|A \cap B|}{|A \cup B|} - \frac{|C \setminus A \cup B|}{|C|}$$

GIoU завжди є нижньою межею для IoU, тобто $\forall A, B \subseteq S \text{ } GIoU(A, B) \leq IoU(A, B)$, і ця нижня межа стає більш жорсткою, коли A і B мають більшу подібність форми та близькість, тобто $\lim_{A \rightarrow B} GIoU(A, B) = IoU(A, B)$. $\forall A, B \subseteq S, 0 \leq IoU(A, B) \leq 1$, але GIoU має симетричний діапазон, тобто $\forall A, B \subseteq S, -1 \leq GIoU(A, B) \leq 1$. Подібно до IoU, значення 1 виникає лише тоді, коли два об'єкти накладаються ідеально, тобто якщо $|A \cup B| = |A \cap B|$, тоді $GIoU = IoU = 1$

Для вимірювання продуктивності моделі іноді використовують матрицю плутанини (Confusion matrix), яка підсумовує продуктивність моделі машинного навчання на тестовому наборі даних. Матриця відображає кількість показників: істинно-позитивних, істинно-негативних, хибно-позитивних і хибно-негативних.

Матриця відображається у вигляді таблиці. Для бінарної класифікації матриця буде представлена таблицею 2x2, де буде представлено чотири типи результатів (таблиця 1). Для класифікації з n класів матриця буде представлена таблицею nxn.

Таблиця 1

Матриця плутанини 2x2

		Прогнозоване	
		Позитивне (P)	Негативне (N)
Фактичне	Позитивне (P)	Істино-позитивне (TP)	Істино-негативне (TN)
	Негативне (N)	Хибно-позитивне (FP)	Хибно-негативне (FN)

Для прикладу розглянемо модель бінарної класифікації, яка може передбачити результат розпізнавання машини на картинці.

Припустимо, що $TP = 4$, $FP = 1$, $FN = 3$, $TN = 2$, тоді матриця плутанини буде мати вид, представлений на рисунку 4.

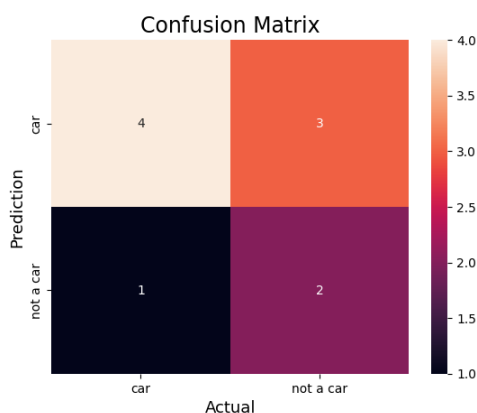


Рис. 4. Матриця плутанини

Отримавши матрицю плутанини можна знайти наступні метрики:

- продуктивність моделі (Accuracy);
- точність (Precision);
- ефективність моделі класифікації в ідентифікації всіх відповідних екземплярів із набору даних (Recall);
- оцінку загальної ефективності (F1-Score).

Продуктивність моделі – це відношення загальної кількості правильних випадків до загальної кількості випадків. Обчислюється за формулою:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Для наведеного вище випадку продуктивність моделі $= (4+2)/(4+2+1+2) = 6/9 = 0,667$.

Щоб визначити наскільки точними є позитивні прогнози моделі потрібно кількість істинно-позитивних прогнозів поділити на загальну кількість позитивних прогнозів. Обчислюється за формулою:

$$precision = \frac{TP}{TP + FP}$$

Точність $= 4/(4+1) = 4/5 = 0,8$.

Ефективність моделі класифікації в ідентифікації всіх відповідних екземплярів із набору даних – це відношення кількості істинно-позитивних випадків до суми істинно-позитивних і хибно-негативних випадків. Обчислюється за формулою:

$$recall = \frac{TP}{TP + FN}$$

Ефективність моделі класифікації в ідентифікації всіх відповідних екземплярів із набору даних $= 4/(4+3) = 4/7 = 0,57$.

Оцінки загальної ефективності – це гармонійне середнє значення точності та чутливості моделі. Обчислюється за формулою:

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Оцінки загальної ефективності $= (2 * 0,8 * 0,57) / (0,8 + 0,57) = 0,67$.

Чим показник F1-Score вище, тим краще точність моделі, якщо показник низький тим найгірше, є дисбаланс між точністю та продуктивністю моделі.

Продуктивність алгоритмів виявлення об'єктів і сегментації зазвичай перевіряється за допомогою метрики mAP (mean average precision). Багато алгоритмів використовують mAP для оцінки своєї продуктивності перед публікацією остаточних результатів, зокрема R-CNN, Faster R-CNN, YOLO, Mask R-CNN. Формула mAP базується на субметриках таких як Confusion matrix, IoU, Recall та Precision. mAP складається із середнього значення (AP) кожного класу. Але для оцінки виявлення об'єктів COCO AP і mAP інтерпретуються однаково.

mAP обчислюється шляхом визначення середньої точності (AP) для кожного класу, а потім усереднення для заданих класів. mAP можна обчислити за формулою:

$$mAP = \frac{\sum_{i=1}^N AP_i}{N},$$

де N – загальна кількість класів;

AP_i – середня точність для заданого класу i.

Алгоритм розрахунку середнього значення (AP) кожного класу:

- 1) за допомогою моделі згенерувати оцінки прогнозу;
- 2) перетворити оцінки прогнозу на мітки класу;
- 3) обчислити матрицю плутанини (TP, FP, TN, FN).
- 4) обчислити показники Recall та Precision;
- 5) побудувати криву Precision-Recall, обчисліть площу під кривою;
- 6) обчислити середню точність.

Розраховуємо AP шляхом взяття площі під кривою Precision-Recall для приклада з розпізнавання машини на картинці (рис. 5).

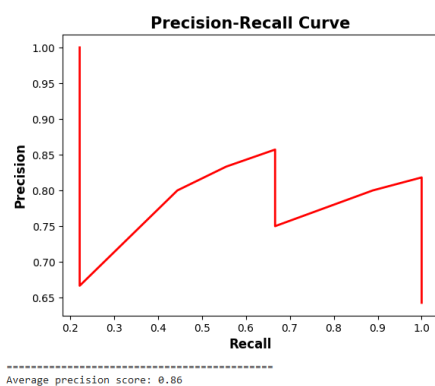


Рис. 5. Розрахунок AP для першого класу

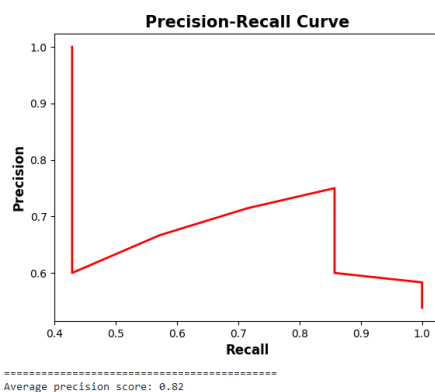


Рис. 6. Розрахунок AP для другого класу

Після обрахування AP для кожного класу в наборі даних обчислюється mAP.

Для наведеного прикладу $mAP = (0,86+0,82)/2 = 0,84$. Чим вищий показник, тим точніше виявлення моделі.

Висновки

Проведено дослідження різних показників ефективності та оцінки моделей виявлення об'єктів у комп'ютерному зорі. Були розглянуті різні метрики такі як IoU, Precision, Recall, крива Precision-Recall та mAP. За допомогою цих метрик можна перевірити, наскільки точна навчена модель із набором даних для перевірки. Як змінюються метрики, якщо змінювати порогові значення або параметри.

Список використаної літератури

1. C.-d. Yang, R.-c. Xie, S.-b. Shi and Z. Tang, "Multi-Target Tracking and Segmentation Method for Missile-Borne Image Based on IoU Association," 2019 12th International Conference on Intelligent Computation Technology and Automation (ICICTA), Xiangtan, China, 2019, pp. 207–211, doi: 10.1109/ICICTA49267.2019.00051.

2. C. Ma, L. Zhuo, J. Li, Y. Zhang and J. Zhang, "Prohibited Object Detection in X-ray Images with Dynamic Deformable Convolution and Adaptive IoU," 2022 *IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022, pp. 3001–3005, doi: 10.1109/ICIP46576.2022.9897684.
3. M. Sugang, L. Ningbo, P. Guansheng, C. Yanping, W. Ying and H. Zhiqiang, "Object detection algorithm based on cosine similarity IoU," 2022 *International Conference on Networking and Network Applications (NaNA)*, Urumqi, China, 2022, pp. 1–6, doi: 10.1109/NaNA56854.2022.00077.
4. V. Levkivskiy, D. Marchuk, N. Lobanchykova et al, . "Available parking places recognition system", 2022 *CEUR Workshop Proceedings 4th Workshop for Young Scientists in Computer Science & Software Engineering*, Vol. 3077, pp. 123–134, [Online], available at: <http://ceur-ws.org/Vol-3077/paper07.pdf>
5. L. Janos Lance, T. Edwin Sybingco and J. A. C. Jose, "Efficient Vehicle Counting Algorithm using Gaussian Mixing Models and IOU Based Tracker for LPWAN Based Intelligent Traffic Management Systems," 2022 *IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Boracay Island, Philippines, 2022, pp. 1–4, doi: 10.1109/HNICEM57413.2022.10109498.
6. Zhang, H., Wang, Y., Dayoub, F., & Sunderhauf, N. (2021). Varifocalnet: An iou-aware dense object detector. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8514–8523).
7. A. Ahmadzadeh, D.J. Kempton, Y. Chen and R.A. Angryk, "Multiscale IOU: A Metric for Evaluation of Salient Object Detection with Fine Structures," 2021 *IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA, 2021, pp. 684–688, doi: 10.1109/ICIP42928.2021.9506337.
8. D. Zhou, J. Fang, X. Song, C. Guan, J. Yin, Y. Dai, and R. Yang, "Iou loss for 2d/3d object detection," *In 2019 International Conference on 3D Vision (3DV)*, pp. 85–94. IEEE, 2019.
9. H. Rezatofighi, N.Tsoi, J. Gwak, A. Sadeghian, I. Reid and S. Savarese, "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*, pp. 658–666. 2019.
10. Z. Liu, J. Cheng, Q. Wang and L. Xian, "Improved Design Based on IoU Loss Functions for Bounding Box Regression," 2022 *IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Beijing, China, 2022, pp. 452–458, doi: 10.1109/IAEAC54830.2022.9929938.

References

1. C.-d. Yang, R.-c. Xie, S.-b. Shi and Z. Tang, "Multi-Target Tracking and Segmentation Method for Missile-Borne Image Based on IoU Association," 2019 12th International Conference on Intelligent Computation Technology and Automation (ICICTA), Xiangtan, China, 2019, pp. 207–211, doi: 10.1109/ICICTA49267.2019.00051.
2. C. Ma, L. Zhuo, J. Li, Y. Zhang and J. Zhang, "Prohibited Object Detection in X-ray Images with Dynamic Deformable Convolution and Adaptive IoU," 2022 *IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022, pp. 3001–3005, doi: 10.1109/ICIP46576.2022.9897684.
3. M. Sugang, L. Ningbo, P. Guansheng, C. Yanping, W. Ying and H. Zhiqiang, "Object detection algorithm based on cosine similarity IoU," 2022 *International Conference on Networking and Network Applications (NaNA)*, Urumqi, China, 2022, pp. 1–6, doi: 10.1109/NaNA56854.2022.00077.
4. V. Levkivskiy, D. Marchuk, N. Lobanchykova et al, . "Available parking places recognition system", 2022 *CEUR Workshop Proceedings 4th Workshop for Young Scientists in Computer Science & Software Engineering*, Vol. 3077, pp. 123–134, [Online], available at: <http://ceur-ws.org/Vol-3077/paper07.pdf>
5. L. Janos Lance, T. Edwin Sybingco and J. A. C. Jose, "Efficient Vehicle Counting Algorithm using Gaussian Mixing Models and IOU Based Tracker for LPWAN Based Intelligent Traffic Management Systems," 2022 *IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Boracay Island, Philippines, 2022, pp. 1–4, doi: 10.1109/HNICEM57413.2022.10109498.
6. Zhang, H., Wang, Y., Dayoub, F., & Sunderhauf, N. (2021). Varifocalnet: An iou-aware dense object detector. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8514–8523).
7. A. Ahmadzadeh, D. J. Kempton, Y. Chen and R. A. Angryk, "Multiscale IOU: A Metric for Evaluation of Salient Object Detection with Fine Structures," 2021 *IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA, 2021, pp. 684–688, doi: 10.1109/ICIP42928.2021.9506337.
8. D. Zhou, J. Fang, X. Song, C. Guan, J. Yin, Y. Dai, and R. Yang, "Iou loss for 2d/3d object detection," *In 2019 International Conference on 3D Vision (3DV)*, pp. 85–94. IEEE, 2019.
9. H. Rezatofighi, N.Tsoi, J. Gwak, A. Sadeghian, I. Reid and S. Savarese, "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*, pp. 658–666. 2019.
10. Z. Liu, J. Cheng, Q. Wang and L. Xian, "Improved Design Based on IoU Loss Functions for Bounding Box Regression," 2022 *IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Beijing, China, 2022, pp. 452–458, doi: 10.1109/IAEAC54830.2022.9929938.