

АЛГОРИТМ ОБНАРУЖЕНИЯ СЛАБЫХ ИНФРАКРАСНЫХ ЦЕЛЕЙ НА СЛОЖНОМ ФОНЕ ПОСРЕДСТВОМ НЕЙРОСЕТОВОЙ МОДЕЛИ YOLOV5

ALGORITHM FOR DETECTING WEAK INFRARED TARGETS ON A COMPLEX BACKGROUND USING THE YOLOV5 NEURAL NETWORK MODEL

Ch. Wang
G. Afanasyev
A. Afanasyev

Summary. The traditional algorithm for detecting infrared targets is based on the accurate selection and extraction of information from the surrounding background and does not always meet the requirements of detection in conditions of complex background and interference. The article discusses a modified algorithm for detecting weak infrared targets, which is based on the neural network model YOLOv5. This algorithm adds an attention mechanism to improve feature extraction ability and efficiency, and uses a modified loss function and a prediction frame filtering method to improve accuracy in detecting weak infrared targets.

Keywords: infrared weak targets, complex backgrounds, YOLOv5, attention mechanism, loss function.

Ван Чаочао

Магистрант, Московский государственный технический университет им. Н.Э. Баумана (национальный исследовательский университет)
wccbmstu@gmail.com

Афанасьев Геннадий Иванович

К.т.н., доцент, Московский государственный технический университет им. Н.Э. Баумана (национальный исследовательский университет)
gaipcs@bmstu.ru

Афанасьев Арсений Геннадьевич

Ассистент, Московский государственный технический университет им. Н.Э. Баумана (национальный исследовательский университет)
afanasievag@bmstu.ru

Аннотация. Традиционный алгоритм обнаружения инфракрасных целей основывается на точном выделении и извлечении информации из окружающего фона и не всегда удовлетворяет требованиям обнаружения в условиях сложного фона и помех. В статье рассматривается модифицированный алгоритм обнаружения слабых инфракрасных целей, который базируется на нейросетевой модели YOLOv5. Этот алгоритм добавляет механизм внимания для улучшения способности извлечения признаков и эффективности, а также использует модифицированную функцию потерь и метод фильтрации кадров предсказания для повышения точности в обнаружении слабых инфракрасных целей.

Ключевые слова: слабые инфракрасные цели, сложные фоны, YOLOv5, механизм внимания, функция потерь, нейронные сети.

Введение

Исследования в области обнаружения целей всегда придавали большое значение обнаружению слабых инфракрасных целей. Инфракрасные системы обладают целым рядом преимуществ: от хорошей скрытности до высокой чувствительности, большой дальности обнаружения, значительной проникающей способности в пространстве. Обнаружение слабых целей в инфракрасном диапазоне имеет важное значение в военном аспекте для раннего предупреждения, для инфракрасного наведения объектов на цель, в морском поиске и спасении, гражданской безопасности и других областях.

Существует три категории методов обнаружения слабых целей в инфракрасном диапазоне: методы в пространственной области, методы в области преобразования и методы нейронных сетей или глубокого обучения.

Методы в пространственной области и области преобразований в общем применяют к изображениям различные пространственные функции преобразования и фильтры, чтобы улучшить изображений. Такие методы на основе соответствующих преобразований лучше обнаруживают цели на слабом фоне, но плохо работают на сложном фоне с относительно большими волнистостями фона.

Методы глубокого обучения обладают мощными возможностями для решения рассматриваемой задачи. В 2020 году был предложен алгоритм YOLOv5, который значительно упростил модель глубокого обучения и улучшил скорость обнаружения требуемых целей с высокой точностью.

1. Нейросетевая модель YOLOv5

Нейросетевая модель YOLOv5 состоит из следующих четырех модулей: Input — на стороне входа,

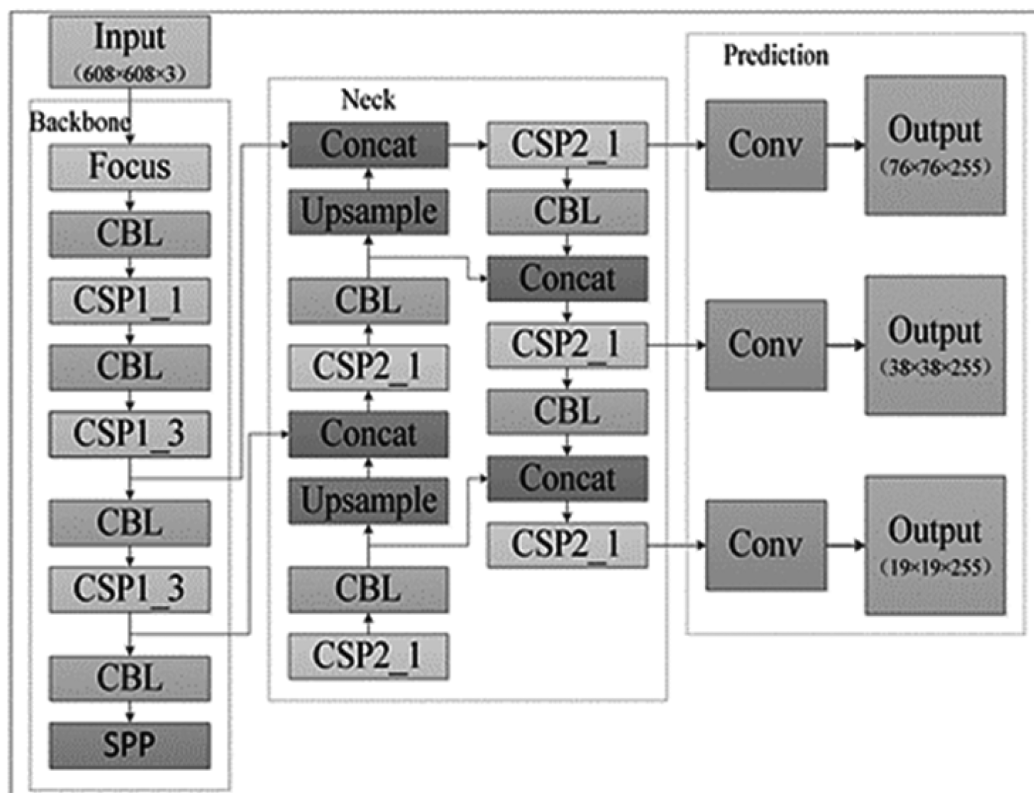


Рис. 1. Модель сети YOLOv5

Backbone — на стороне магистральной сети, Neck network и Prediction — на стороне предсказания [1]. Входные данные включают в себя улучшенные мозаичные данные, соответствующего масштабного размера изображений и адаптивные расчеты опорного кадра. Магистральная сеть включает в себя обработку Focus, обработку CBL, структуру CSP и пространственное пирамидальное объединение SPP. Модель сети изображена на рисунке 1.

Модуль прогнозирования

YOLOv5 использует метрику *GIOU_Loss* в качестве функции потерь при фильтрации целевых кадров [2]. *A* — целевой объект, *B* — предсказание целевого объекта. (рис. 2)

Диаграмма метрики IOU показана на рисунке 2.

$$IOU = \frac{A \cap B}{A \cup B} = \frac{M}{N} \quad (1)$$

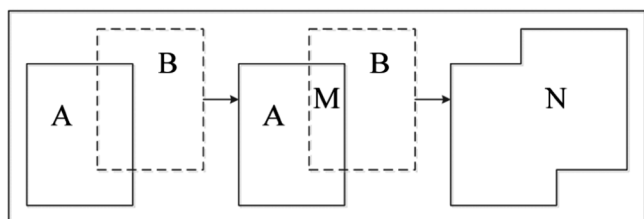


Рис. 2. IOU диаграмма

Диаграмма метрики GIOU показана на рисунке 3.

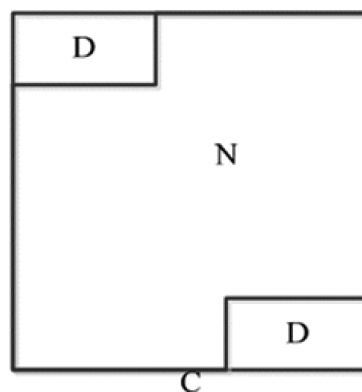


Рис. 3. Диаграмма GIOU

Метрика GIOU рассчитывается следующим образом

$$GIOU = IOU - \frac{|C - N|}{C} = IOU - \frac{D}{C} \quad (2)$$

Функция потерь имеет следующий вид:

$$GIOU_Loss = 1 - \left(\frac{M}{N} - \frac{D}{C} \right) \quad (3)$$

2. Модифицированный алгоритм обнаружения целей для YOLOv5

Как показано на рисунке 4, для обнаружения слабых инфракрасных объектов необходимы высокая точность и скорость обнаружения. В построенной модели ней-

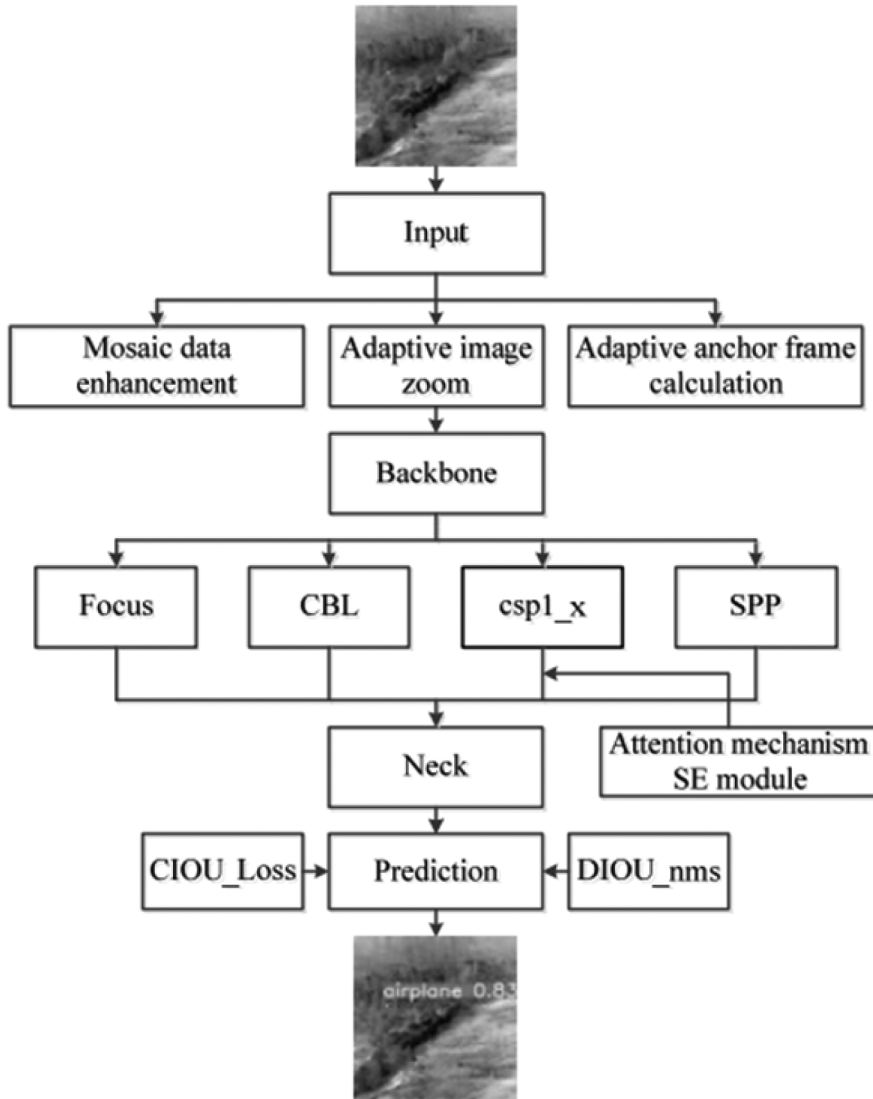


Рис. 4. Схема модифицированного алгоритма обнаружения целей YOLOv5

ронной сети YOLOv5 в основной части сети Backbone добавлен механизм внимания SE для повышения способности модели извлекать характеристики и повышения эффективности обнаружения. Модифицированы функции потерь и методы отбора предсказательных рамок в предсказательной части модели, что повышает точность обнаружения слабых инфракрасных объектов.

2.1. Совершенствование механизма внимания SE модуля

Механизм внимания был добавлен в нейросетевую модель обнаружения целей YOLOv5, чтобы улучшить способность извлечения признаков и повысить точность обнаружения модели. Основываясь на малом размере инфракрасных слабых целей, слой объединения признаков и масштаб признаков многомасштабного слоя обнаружения дополнительно оптимизированы, чтобы обученная модель обнаружения могла лучше адаптироваться к обнаружению малых целей и улучшить скорость обнаружения модели в то же время.

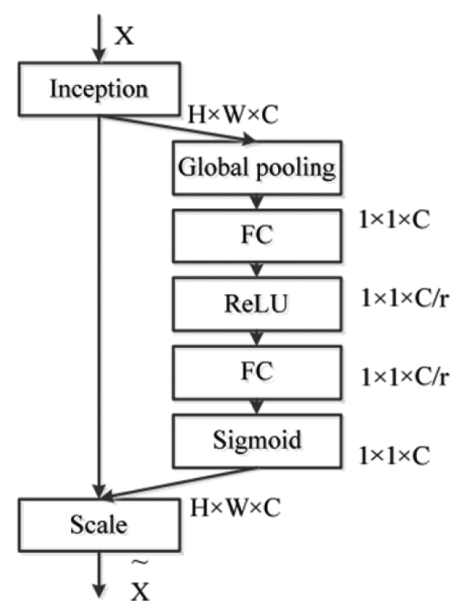


Рис. 5. Механизм привлечения внимания SE модуля

Этот модифицированный алгоритм имеет модуль SE механизма внимания, который работает в 3 этапа: сначала входная карта признаков (h,w,c) глобально объединяется с размером объединения (h,w) , и на выходе получается карта признаков размером $(1 \times 1 \times c)$; затем она проходит через два полностью связанных слоя с $c/12$ нейронами и с нейронами для увеличения нелинейного фактора; наконец, она проходит через сигмоидальную функцию активации. Функция активации Sigmoid имеет весовой коэффициент размером $(1 \times 1 \times c)$, который используется для умножения каждого канала исходной карты признаков (h,w,c) , чтобы в итоге получить карту признаков с различными весовыми каналами. Это показано на рисунке 5.

2.2. Улучшенная функция потерь CIOU_Loss

Как показано на рисунке 6 метрика GIOU_Loss не решает задачу возможности определения близости двух прямоугольных рамок, когда они не пересекаются и она не может определить положение предсказанной прямоугольной рамки, когда она находится внутри целевой реальной прямоугольной рамки [3].

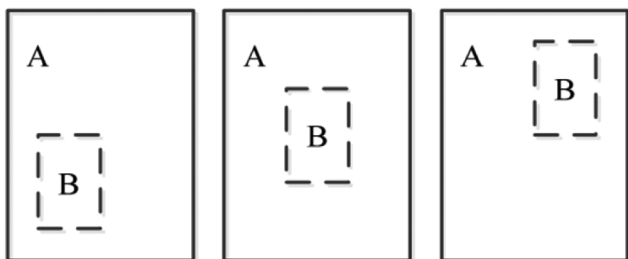


Рис. 6. Невозможность идентификации расположения прямоугольной рамки прогнозирования внутри целевой прямоугольной рамки

Поэтому в данном алгоритме вместо этого используется метрика CIOU_Loss (CompleteIntersectionOverUnion_Loss). Эта функция потерь учитывает не только расстояние между двумя прямоугольными рамками, но и информацию о соотношении сторон этих рамок, что может эффективно решить проблемы, существующие в алгоритме GIOU_Loss. Конкретный метод расчета выглядит следующим образом.

Как показано на рисунке 7, пусть диагональ наименьшего внешней прямоугольной рамки C равна R , а расстояние между центральными точками реальной прямоугольной рамки цели A и прямоугольной рамки предсказания B равно r .

CIOU-Loss рассчитывается следующим образом:

$$CIOU = IOU - \frac{R^2}{r^2} - \frac{v^2}{(1 - IOU) + v} \quad (4)$$

$$CIOU_Loss = 1 - CIOU = 1 - IOU + \left(\frac{R^2}{r^2} + \frac{v^2}{(1 - IOU) + v} \right) \quad (5)$$

где: v — параметр, характеризующий постоянство соотношения сторон целевого кадра предсказания и рассчитываемый следующим образом:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right)^2 \quad (6)$$

где: w^{gt}, h^{gt} обозначают ширину и высоту целевого реального кадра; w^p, h^p обозначают ширину и высоту предсказанного кадра [4].

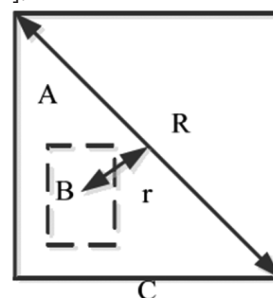


Рис. 7. Схематическая диаграмма CIOU

2.3. Модифицированный метод фильтрации прямоугольника предсказаний DIOU_NMS

В алгоритме YOLOv5 используется обычный метод NMS (Non-Maximum Suppression) для отбора предсказательных рамок, который использует коэффициент пересечения и объединения (IOU) для подавления избыточных предсказательных рамок. Этот метод основывается только на области перекрытия в качестве единственного критерия, что приводит к ошибочному подавлению в случае наличия перекрытия объектов. В данном алгоритме используется метод DIOU_NMS (Distance Intersection Over Union_Non Maximum Suppression) для отбора предсказательных рамок, в котором DIOU используется в качестве критерия подавления. Этот метод учитывает не только влияние области перекрытия, но и информацию о расстоянии между центрами двух рамок. Это позволяет повысить точность обнаружения перекрытых и затененных объектов, делая предсказания более логичными и точными.

3. Моделирование

3.1. Набор исходных данные

Для тестирования эффективности алгоритма использовался набор данных инфракрасных слабых целей, состоящий из изображений инфракрасных слабых целей с разрешением 256 x 256 пикселей. Этот набор данных содержал: изображения близкой одиночной цели на фоне открытого грунта; изображения близких пере-

крывающихся 2-х целей на фоне неба; изображения одиночной цели в группе инфракрасных объектов с дальнего и ближнего расстояния на фоне леса; изображения ближней и дальней одиночной цели на фоне лесистой местности; изображения одиночной цели на фоне леса; изображения маневрирующей одиночной цели на дальнем и близком расстоянии на фоне леса; изображения одиночной маневрирующей цели на фоне местности с холмами.

3.2. Конфигурация параметров среды моделирования

Моделирование было проведено на фреймворке глубокого обучения Pytorch, на операционной системе Ubuntu 22.04, видеокарта NVIDIA GeForce GTX 1660 SUPER, оперативная память 32ГБ, графический ускоритель CUDA11.1

3.3. Экспериментальные результаты

В данном моделировании оценка производится с использованием показателей точности (Precision), полноты (Recall) и средней площади под кривой точности-полноты (mAP). Формулы вычисления точности и полноты представлены в уравнениях (7) и (8).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

Precision (точность) представляет долю истинных слабых инфракрасных объектов TP относительно всех обнаруженных слабых инфракрасных объектов. FP — доля ошибочных обнаружений инфракрасных объектов.

Recall (полнота) представляет долю правильно обнаруженных слабых инфракрасных объектов TP относительно всех образцов слабых инфракрасных объектов. FN — доля необнаруженных слабых инфракрасных объектов.

Средняя точность (AP) — это площадь, заключенная между кривыми точности и полноты. Средняя точность (mAP) — это среднее значение точности по всем классам. mAP — среднее значение точности по всем классам обученной модели. mAP@0.5 означает вычисление AP для каждого класса в наборе данных изображений при пороге перекрытия IOU, установленном на 0.5, а затем нахождение среднего значения по всем классам.

В эксперименте, чтобы проверить производительность обученной модели для обнаружения целей, были сравнены алгоритмы SSD (Single Shot MultiBox Detector), Faster R-CNN, YOLOv3 и YOLOv5. В качестве показателя оценки была использована mAP, а также была учтена

скорость обнаружения модели. Результаты сравнения представлены в таблице.

Таблица 1.

Сравнение результатов моделирования

Параметры	mAP	Время (сек.)
SSD	76.67 %	7.90
Faster R-CNN	89.77 %	6.19
YOLOv3	94.35 %	3.77
YOLOv5	95.07 %	2.83
модифицированная YOLOv5	96.62 %	1.23

Согласно таблице, модифицированный алгоритм в рамках YOLOv5 в данном исследовании достигает значения mAP в 96.62 %. По сравнению с алгоритмами SSD, Faster R-CNN, YOLOv3 и YOLOv5, он улучшает показатель mAP на 19.95 %, 6.85 %, 2.27% и 1.55 % соответственно.

В рамках моделирования было выявлено, что алгоритм SSD имеет большой процент пропущенных тревог и способен обнаруживать слабые инфракрасные цели на относительно простых фонах, таких как граница раздела небо-земля и фон неба, но не на сложных фонах, таких как леса и холмы. Алгоритм Faster R-CNN не может эффективно извлекать признаки мелких целей, и у обученной модели есть ложные и пропущенные тревоги и уровень доверия к обнаруженным ложным целям довольно высок. Алгоритмы YOLOv3 и YOLOv5 улучшили производительность обученных моделей по сравнению с алгоритмами SSD и Faster R-CNN, но у них также есть ложные и пропущенные тревоги. Модифицированный алгоритм YOLOv5 смог эффективно идентифицировать слабые ИК-цели на семи сложных фонах с высокой точностью и низким коэффициентом ложных тревог, а уровень доверия обученной модели в обнаружении слабых ИК-целей составил около 0,83. Производительность значительно выше, чем у обученных моделей сравнения. Правильность и эффективность алгоритма и обученной модели подтверждены.

Заключение

Таким образом, для решения задачи по точному и быстрому выделению и извлечению информации об инфракрасных целях из изображений со сложными фонами рассмотренный в данной работе алгоритм в рамках нейросетевой модели YOLOv5 на тестовых данных показал свою эффективность относительно других рассмотренных нейросетевых моделей. Было подтверждено, что обученная нейросетевая модель типа YOLOv5 способна эффективно обнаруживать слабые инфракрасные цели. Средняя точность составила 96,62 %, что свидетельствует об эффективности рассмотренного алгоритма в рамках нейросетевой модели YOLOv5.

ЛИТЕРАТУРА

1. Yang X., Jiang W., Yuan H. Traffic sign recognition detection based on YOLOv5 // Information Technology and Informatization. 2021. 4. P.28–30.
2. Zhou F.Y., Jin L.P. and Dong J. A Review of Convolutional Neural Networks // Journal of Computers.2017. 40. P. 1229–1251
3. Zhou J. et al. Ship target detection algorithm based on improved YOLOv5 // Journal of Marine Science and Engineering. 2021. 9(8). P.908–912.
4. Zhang H., Ban Y., Guo L., et al. Detection method of remote sensing image ship based on YOLOv5 // Electronic Measurement Technology. 2021. 44(8). P.87–92.

© Ван Чаочао (wccbmstu@gmail.com); Афанасьев Геннадий Иванович (gaipcs@bmstu.ru); Афанасьев Арсений Геннадьевич (afanasievag@bmstu.ru)
Журнал «Современная наука: актуальные проблемы теории и практики»