

# Почему YOLOv8 быстрее и сильнее, чем YOLOv5 (на примере распознавания номеров)

## Оглавление

1.какие изменения в YOLOv8 и YOLOv5 .....	2
1.1.Структурная схема YOLOv8 и YOLOv5 .....	2
1.2.Основные изменения YOLOv8 .....	5
2. Показатели производительности .....	6
2.1. Матрица путаницы (confusion matrix) .....	7
2.2.Точность (Precision) .....	7
2.3 Точность(accuracy) .....	7
2.4.Скорость отзыва (Recall) .....	8
2.5.Скорость точного отзыва (PR_curve) .....	8
2.6.Оценка F1 (F1_curve) .....	8
2.7 .Индикаторы обнаружения .....	8
3. Экспериментально-сравнительный анализ (на примере распознавания номеров) .....	9
3.1. Набор данных .....	9
3.2 Настройки параметров обучающей модели .....	9
3.3.Сравнение результатов обучения .....	10
3.3.1. confusion_matrix .....	10
3.3.2 результаты .....	10
3.3.3. F1-curve .....	11
3.3.4.Сравнение результатов определения номерных знаков .....	12
4. Заключение .....	13
Литература .....	15

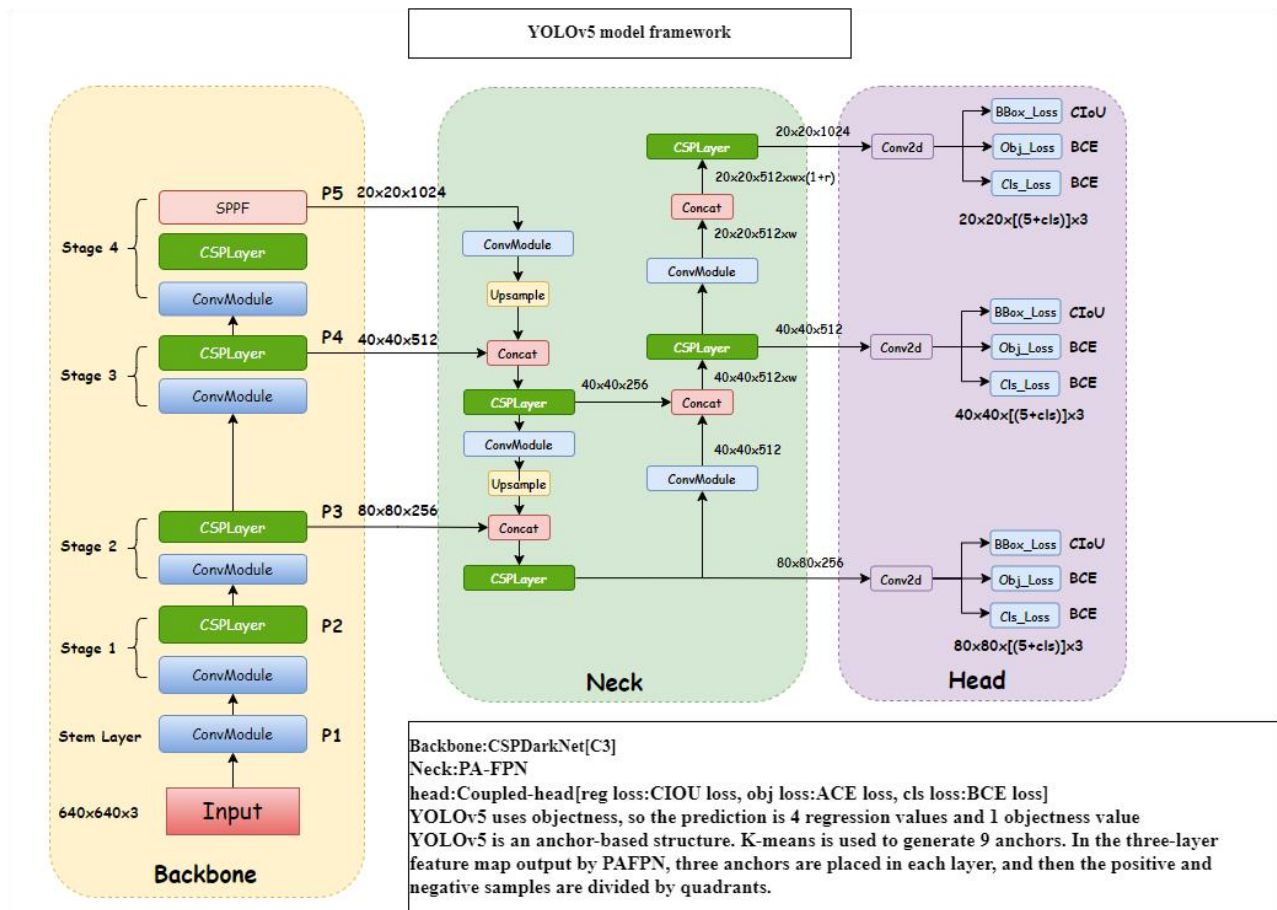
## 1.какие изменения в YOLOv8 и YOLOv5

### 1.1.Структурная схема YOLOv8 и YOLOv5

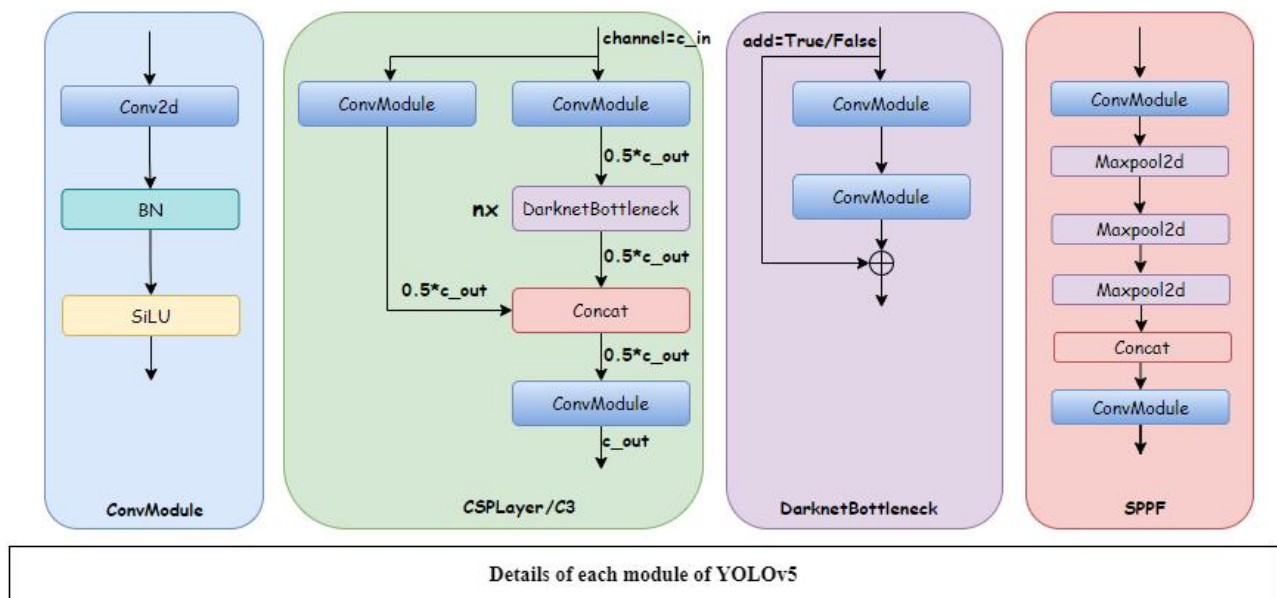
YOLOv8 и YOLOv5 созданы одной командой, поэтому YOLOv8 также можно рассматривать как оптимизацию YOLOv5, и структура также очень похожа Давайте сравним сходства и различия между YOLOv5 и YOLOv8.

Module	YOLOv8	YOLOv5
Input size	640*640*3	640*640*3
Backbone	C2f(Added more information about gradient flow	C3
Neck	PA-FPN	PA-FPN
Head	Decoupling head	Coupling head
Anchor	Anchor-free	Anchor-based
Label	TAL	gt center point
Objectness	No	Yes
output	20*20*(4+cls)/40*40*(4+cls)/80*80*(4+cls)	20*20*(5+cls)/40*40*(5+cls)/80*80*(5+cls)

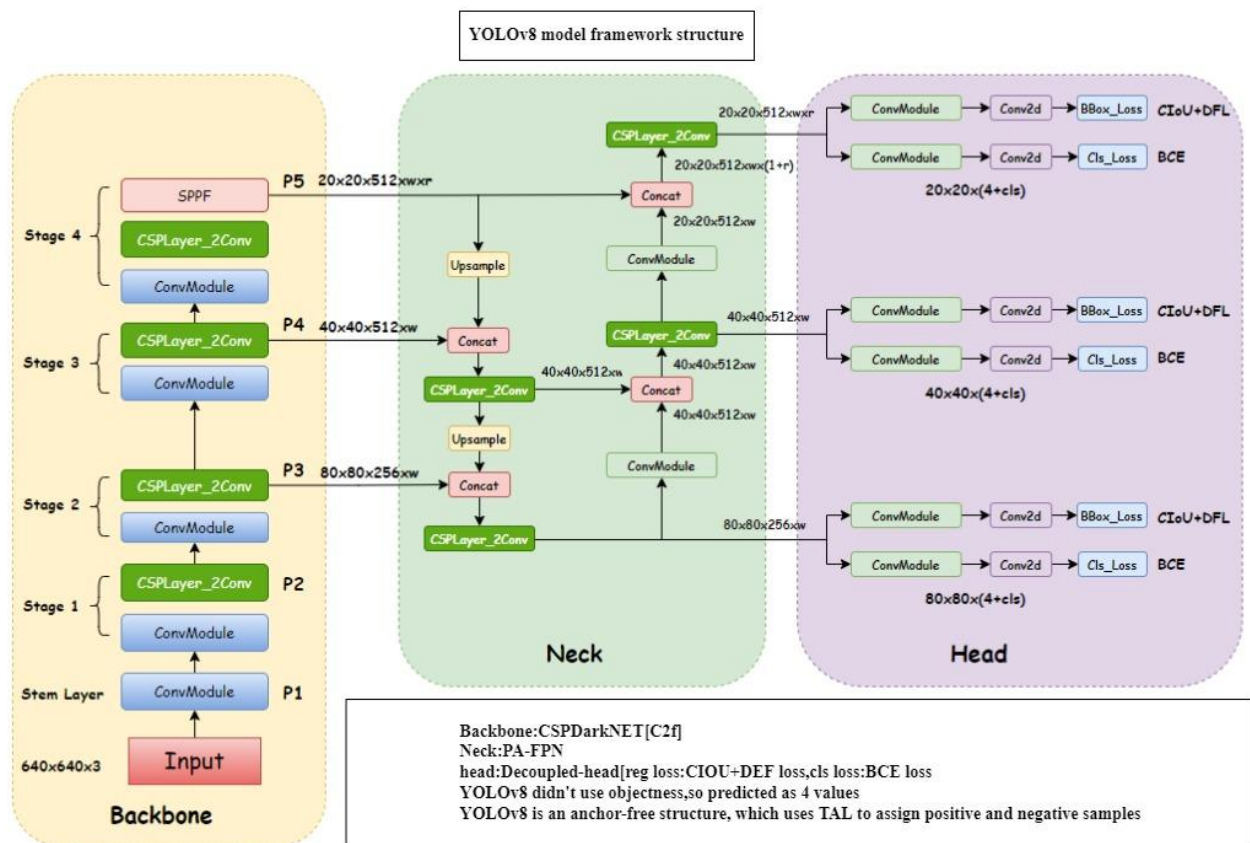
Структура модели YOLOv5 выглядит следующим образом:



Детали модуля YOLOv5 следующие:

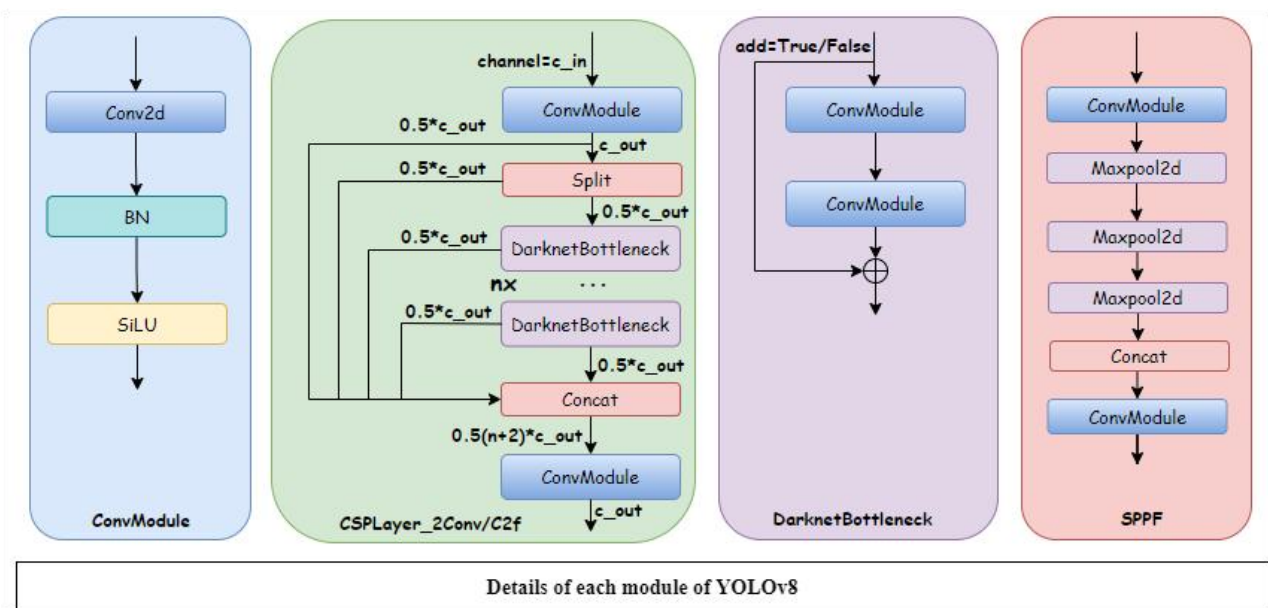


Каркасная структура модели YOLOv8 выглядит следующим образом:



Детали модуля YOLOv8 следующие:

1. Основное отличие состоит в том, что в YOLOv8 используется модуль C2f с более богатыми градиентами (производный от модуля E-ELAN в YOLOv7).
2. YOLOv5 использует модуль C3 только с одним градиентным шунтом.



## 1.2. Основные изменения YOLOv8

Основные функции и изменения можно резюмировать следующим образом:

1. Предоставляет совершенно новую модель SOTA, включая сеть обнаружения целей с разрешением P5 640 и P6 1280 и модель сегментации экземпляров на основе YOLACT. Как и в YOLOv5, также предоставляются модели различных размеров масштабов N/S/M/L/X в зависимости от коэффициента масштабирования для удовлетворения потребностей различных сценариев.
2. Магистральная сеть и шейная часть могут относиться к идее дизайна YOLOv7 ELAN Структура C3 YOLOv5 заменена структурой C2f с более богатым градиентным потоком, а количество каналов настроено для разных масштабных моделей, которые принадлежат для точной настройки структуры модели. Это больше не безмозглый набор параметров, которые можно применять ко всем моделям, что значительно повышает производительность модели. Однако такие операции, как Split в этом модуле C2f, не так удобны для конкретного развертывания оборудования, как раньше.
3. По сравнению с YOLOv5, головная часть сильно изменилась, она была заменена на текущую основную структуру развязывающей головки, которая разделяет головки классификации и обнаружения, а также изменилась с Anchor-Based на Anchor-Free.
4. Что касается расчета потерь, принята стратегия распределения положительной выборки TaskAlignedAssigner и введена функция распределения фокусных потерь.
5. Часть обучения по улучшению данных представляет последние 10 эпох в YOLOX, чтобы отключить операцию улучшения Mosaic, что может эффективно повысить точность.

Из вышеизложенного видно, что YOLOv8 в основном относится к родственному дизайну алгоритмов, таких как недавно предложенные YOLOX, YOLOv6, YOLOv7 и PPYOLOE, Сам по себе нововведений не так много, и он смещен в сторону инженерной практики.

Далее будут подробно представлены различные улучшения обнаружения целей YOLOv8 в соответствии с четырьмя частями расчета потерь, улучшения обучающих данных, стратегии обучения и процесса рассуждения модели, а часть сегментации экземпляра не будет временно описана.

### 1. расчет потерь

Процесс расчета потерь включает две части: стратегию распределения положительных и отрицательных выборок и расчет потерь. Алгоритм YOLOv8 напрямую ссылается на TaskAlignedAssigner TOOD. Его

стратегия сопоставления просто резюмируется следующим образом: отбирать положительные образцы на основе взвешенных оценок классификации и оценок регрессии.

Формула:

$$t = s^{\alpha} + u^{\beta}$$

$s$  — оценка предсказания, соответствующая категории метки,  $u$  — значение  $iou$  поля предсказания и поля  $gt$ , а степень выравнивания может быть измерена путем их умножения.

Расчет потерь включает 2 ветки: ветки классификации и регрессии, без предыдущей ветки объектнос ти.

Ветвь классификации по-прежнему использует BCE Loss.

Ветвь регрессии должна быть привязана к представлению в интегральной форме, предложенному в разделе «Распределение фокусных потерь», поэтому используется распределение «Фокусные потери», а также потери CIOU. Три потери могут быть взвешены с определенным весовым коэффициентом.

## 2. Улучшение обучающих данных

С точки зрения улучшения данных особых отличий от YOLOv5 нет, но введена операция закрытия Mosaic в последних 10 эпохах, предложенная в YOLOX. Учитывая, что разные модели должны использовать разные степени улучшения данных, некоторые гиперпараметры будут изменены для моделей разных размеров. Как правило, MixUp и CopyPaste будут включены для больших моделей.

## 3. Стратегия обучения

Стратегия обучения YOLOv8 ничем не отличается от стратегии YOLOv5. Самое большое отличие состоит в том, что общее количество эпох обучения модели увеличилось с 300 до 500, что также приводит к резкому увеличению времени обучения.

## 4. Процесс рассуждения модели

Процесс рассуждения YOLOv8 почти такой же, как и YOLOv5. Единственное отличие состоит в том, что форма  $bbox$  интегрального представления в Distribution Focal Loss должна быть декодирована, чтобы стать обычным 4-мерным  $bbox$ . Последующий процесс вычисления такой же как у YOLOv5.

## 2. Показатели производительности

Оценка производительности модели машинного обучения (то есть способности к обобщению), такая

как результаты прогнозирования и классификации тестового набора, хороша или нет. Но для количественной оценки критериев оценки у нас есть ряд метрик производительности. С индикаторами мы можем количественно оценивать и сравнивать разные модели при их сравнении. При этом мы продолжим оптимизировать нашу модель за счет дальнейшей настройки параметров индикаторов.

## 2.1. Матрица путаницы (confusion matrix)

Если у нас есть модель с двумя категориями, то при объединении прогнозируемой ситуации с фактическими результатами ситуации будет четыре ситуации, как показано в следующей таблице:

**P (positive):** представляет 1    **N (Negative):** представляет 0

**T(True):** предсказание верно    **F(False):** предсказание неверно

		True	
		1	0
predicted	1	TP	FP
	0	FN	TN

## 2.2. Точность (Precision)

Этот индикатор рассчитывается для результата прогнозирования, вероятности фактического положительного образца среди всех образцов, которые прогнозируются как положительные. То есть среди выборок, предсказанных как положительные, мы предсказываем вероятность быть правильными. Формула выглядит следующим образом:

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

## 2.3 Точность(accuracy)

Определение: процент правильных прогнозов в общей выборке.

Формула выглядит следующим образом:

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

Разница между показателем точности (precision) и показателем точности (accuracy) : показатель точности представляет собой степень точности прогноза в положительных результатах выборки. Показатель точности представляет собой общую точность предсказания.

## 2.4.Скорость отзыва (Recall)

Определение: Вероятность быть предсказанным как положительный образец среди образцов, которые действительно являются положительными.

Формула выглядит следующим образом:

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

## 2.5.Скорость точного отзыва (PR\_curve)

Мы надеемся, что и точность, и полнота одновременно очень высоки, но на самом деле эти два показателя представляют собой пару противоречий, одно из которых очень высокое, а другое обязательно будет очень низким. к реальным потребностям. Например, если нам нужна высокая скорость отзыва, мы пожертвуем некоторым показателем точности, в случае обеспечения максимальной скорости отзыва скорость точности не будет такой низкой.

## 2.6.Оценка F1 (F1\_curve)

Оценка F1 используется для баланса скорости точности и скорости отзыва. Он учитывал их одновременно, чтобы два коллеги могли достичь самого высокого уровня и взять баланс. Формула выглядит следующим образом:

$$F1\_curve = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

## 2.7 .Индикаторы обнаружения



mAP: среднее значение AP для каждого класса

AP: средняя точность при различных отзывах, площадь под кривой PR

TP: долговая расписка с наземной правдой  $> 0,5$

FP: IOU  $\leq 0,5$  с наземной истиной, избыточный кадр обнаружения той же наземной истины

FN: наземная правда не обнаружена

mAP@.5: IOU  $> 0,5$  считается TP

mAP@.5:.95: IOU составляет 0,5, 0,55, 0,6, 0,65 ... 0,95 соответственно, и получают несколько mAP, а затем усредняют

### 3. Экспериментально-сравнительный анализ (на примере распознавания номеров)

#### 3.1. Набор данных

Набор данных маркировки номерных знаков с открытым исходным кодом получен через roboflowx и разделен на 1414, 404, 202 изображения для поезда, действительного и тестового в соответствии с 7: 2: 1.

Размечены в соответствии с форматом набора текстовых данных YOLOv5, включая числа 0-9 и AZ, всего 36 меток.

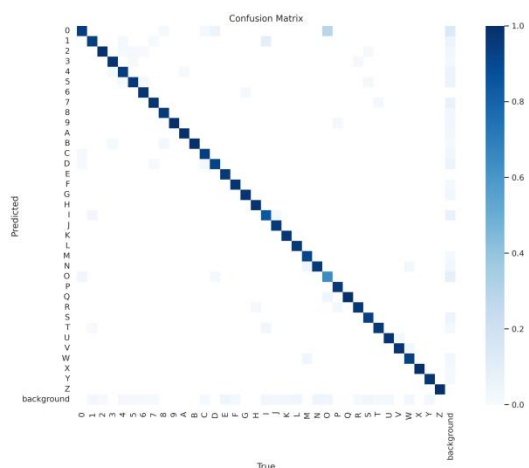
#### 3.2 Настройки параметров обучающей модели

Model	YOLOv8	YOLOv5
Input img	640*640	640*640
Batch	16	16
epoch	100	100
model	Yolov8s	Yolov5s
weights	Yolov8s.pt	Yolov5s.pt
layers	225	214
parameters	11149562	7116721

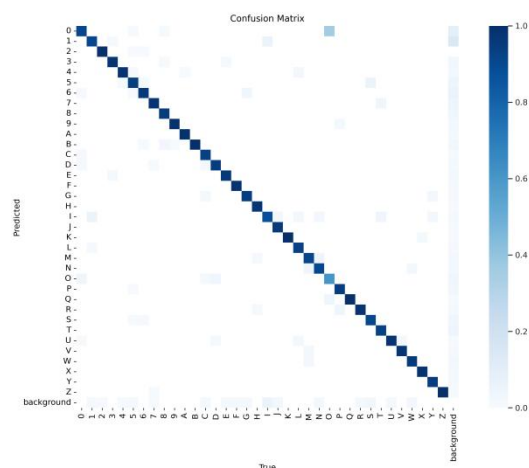
### 3.3. Сравнение результатов обучения

#### 3.3.1. confusion\_matrix

Из матрицы путаницы ясно видно, что количество ошибок предсказания YOLOv5 значительно больше, чем у YOLOv8.

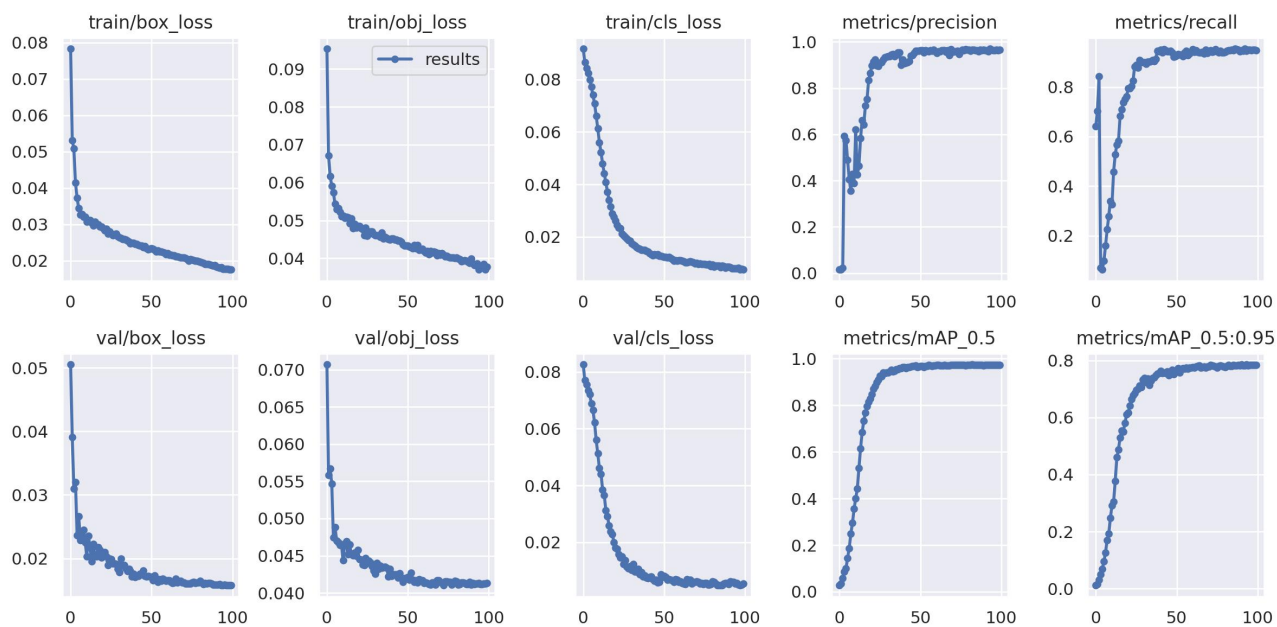


YOLOv8

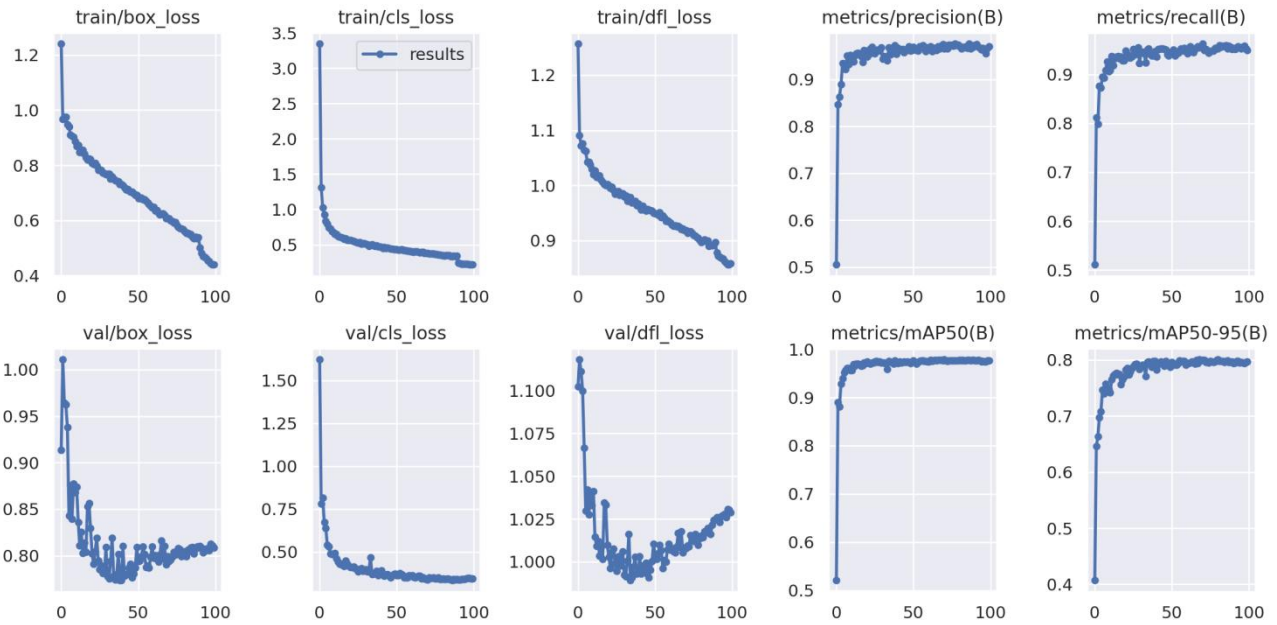


YOLOv5

#### 3.3.2 результаты



### YOLOv5 results

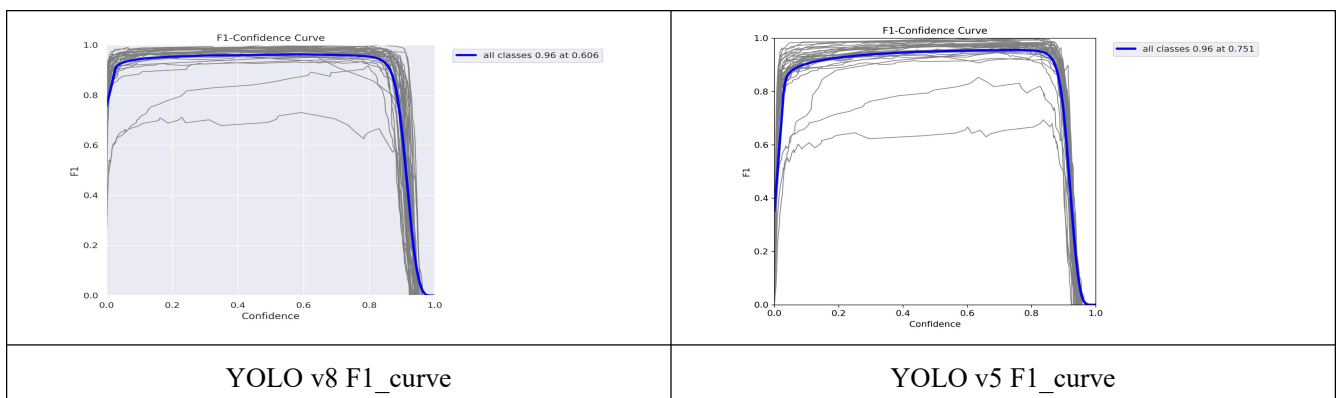


### YOLOv8 results

Model	YOLOv8	YOLOv5
Precision	0.97604	0.97039
recall	0.96264	0.95416
mAP50	0.97865	0.9749
mAP50-90	0.80059	0.78609

Из результатов видно, что в случае обучения 100 эпох mAP<sub>0,5:0,95</sub> YOLOv8 достигает максимума 0,80059, а YOLOv5 достигает только 0,78609. Остальные показатели YOLOv8 лучше, чем YOLOv5.

### 3.3.3. F1-curve



Видно, что YOLOv8 достигает своего пика раньше, чем YOLOv5.

### 3.3.4. Сравнение результатов определения номерных знаков

Из результатов распознавания видно, что когда номерной знак четкий, оба могут быть хорошо распознаны и точны, но YOLOv8 может получить более высокую точность, чем YOLOv5, но в случае размытого номерного знака и двойного изображения в данных обстоятельствах это трудно идентифицировать, но это только частный случай. В той же ситуации слабого освещения YOLOv8 работает лучше.



YOLOv8\_batch0\_pred



YOLOv5\_batch0\_pred





YOLOv8\_batch0\_pred



YOLOv5\_batch1\_pred

## 4. Заключение

В этой статье подробно анализируется и обобщается новейший алгоритм YOLOv8, сравнивается его с YOLOv5 и подробно объясняются от общего дизайна до структуры модели, расчета потерь, улучшения обучающих данных, стратегии обучения и процесса рассуждений. И возьмем в качестве примера распознавание номерных знаков, сравним точность и скорость обоих в инженерных приложениях. Видно, что YOLOv8 работает лучше, чем YOLOv5 по всем показателям. В этой статье в основном сравниваются результаты и коды, и они загружаются на [https://github.com/Chaochao690480/YOLOv8\\_licenseplate\\_detection](https://github.com/Chaochao690480/YOLOv8_licenseplate_detection), чтобы инженеры или учащиеся могли сделать справку по выбору в соответствии с их собственной ситуацией при выборе модели.

Проще говоря, YOLOv8 - это эффективный алгоритм, включающий классификацию изображений, обнаружение объектов без привязки и сегментацию экземпляров. Дизайн части обнаружения относится к большому количеству отличных последних улучшенных алгоритмов YOLO и реализует новый SOTA. Не

только это, но и совершенно новый фреймворк. Тем не менее, эта структура все еще находится на ранней стадии и нуждается в постоянном совершенствовании.

## Литература

- [1] YOLOv5, <https://github.com/ultralytics/yolov5>
- [2] YOLOX: Exceeding YOLO Series in 2021, <https://arxiv.org/abs/2107.08430>
- [3] PP-YOLOE: An evolved version of YOLO, <https://arxiv.org/abs/2203.16250>
- [4] CSPNet: A New Backbone that can Enhance Learning Capability of CNN,  
<https://arxiv.org/abs/1911.11929>
- [5] Path aggregation network for instance segmentation,  
<https://arxiv.org/abs/1803.01534>
- [6] OTA: Optimal Transport Assignment for Object Detection,  
<https://arxiv.org/abs/2103.14259>
- [7] Computer Architecture: A Quantitative Approach
- [8] SIOU Loss: More Powerful Learning for Bounding Box Regression,  
<https://arxiv.org/abs/2205.12740>