

УЛУЧШЕННОЕ ДЕТЕКТИРОВАНИЕ СВЯЗАННЫХ ОБЪЕКТОВ НА ПРИМЕРЕ ЗАДАЧИ РЕИДЕНТИФИКАЦИИ ЧЕЛОВЕКА

Русаков К.Д.

Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия

rusakov.msk@yandex.ru

Аннотация. Данная работа представляет методику улучшенного детектирования связанных объектов, таких как силуэты и их атрибуты, с использованием новой функции потерь в модели YOLOv8, на примере задачи реидентификации человека для повышения точности и эффективности распознавания.

Ключевые слова: детектирование объектов, реидентификация человека, распознавание силуэтов, компьютерное зрение, глубокое обучение.

Введение

Концепция связанных объектов играет важную роль в задачах детектирования и распознавания. Каждый объект, будь то человек или что-то другое, обладает набором атрибутов, которые помогают идентифицировать его с высокой точностью. Например, в контексте человека, лицо является одним из наиболее значимых атрибутов силуэта, но также могут быть важны другие атрибуты, такие как пол, одежда, аксессуары и поза. Эти атрибуты, будучи связанными между собой, создают комплексный и детализированный образ объекта, что позволяет системе лучше распознавать и реидентифицировать его в различных условиях.

Реидентификация человека, как одна из ключевых задач в системах визуального наблюдения, не ограничивается только распознаванием лиц. Важно учитывать весь силуэт и его атрибуты, что позволяет эффективно идентифицировать человека даже в условиях, когда лицо может быть скрыто или плохо видно. Таким образом, подходы, направленные на детектирование и распознавание связанных объектов и их атрибутов, обеспечивают более высокую точность и надёжность в задачах реидентификации. Использование комплексных признаков, включая лицо, одежду и другие атрибуты, позволяет преодолеть ограничения традиционных методов и повысить эффективность автоматических систем наблюдения.

В последнее время широкое распространение получили цифровые IP видеокамеры, подключенные к компьютерным сетям. Тем не менее, визуальное наблюдение в основном осуществляется людьми. В будущем автоматические системы визуального наблюдения могут сыграть важную роль в поддержке и, в конечном итоге, замене людей-наблюдателей. Чтобы быть эффективными, эти системы должны выполнять следующие основные задачи:

- обнаружение людей
- отслеживание людей (трекинг)
- распознавание людей.

В данной работе основное внимание уделяется задаче распознавания и реидентификации человека. Однако, поскольку распознавание человека часто подразумевает предварительный этап обнаружения, в работе также будут затронуты подходы к обнаружению людей. Существует огромное количество методов компьютерного зрения, которые можно использовать в автоматических системах визуального наблюдения: обнаружение лиц [1, 2, 3, 4, 5, 6] и распознавание лиц [7, 8, 9, 10] были тщательно изучены за последние 20 лет в области компьютерного зрения. В современных системах распознавания лиц изображения лиц по-прежнему являются первым выбором для распознавания личности человека. Однако изображения лиц доступны не всегда, что делает распознавание лиц неприменимым в некоторых сценариях. Распознавание человека направлено на идентификацию личности человека по данному изображению или видео человека, не требуя присутствия лица человека. Существует ряд подходов, основанных на анализе отличных от лица признаков людей для их успешного распознавания. Как показано в работах [11, 12] периодичность походки позволяет обнаруживать идущего человека в последовательности изображений. Походка также использовалась для распознавания людей в последовательностях изображений [13]. Авторами в работе [13] сообщалось о небольшой группе из 85 человек, на которой проходило тестирование подхода. Неясно, можно ли отличить большее количество людей только по походке.

Кроме вышеуказанных подходов в настоящее время набирают популярность методы реидентификации человека. Реидентификация человека – это процесс определения и подтверждения личности индивида на основе сравнения его характеристик, полученных из различных источников

данных или в разные моменты времени. Этот процесс часто используется в системах безопасности и наблюдения, а также в различных информационно-поисковых системах. В контексте компьютерного зрения и обработки изображений реидентификация обычно включает анализ визуальных данных, таких как фотографии или видеозаписи, для идентификации одного и того же человека на разных изображениях или видео, даже если условия съемки различаются (например, изменение освещения, ракурса, фона, позы или одежды человека).

Последние работы по поиску людей можно условно разделить на одноэтапные методы [14, 15, 16] и двухэтапные методы [17, 18, 19]. Типичный двухэтапный метод сначала обнаруживает людей, а затем применяет модель распознавания к обнаруженным изображениям. Напротив, одноэтапные методы оптимизируют задачи обнаружения и распознавания людей одновременно. Чтобы повысить различительную способность моделей, в некоторых работах вводятся вспомогательные задачи [17, 20] (например, оценка позы, распознавание атрибутов), чтобы предоставить информацию для поиска человека. Исследователи в [21] извлекают признаки из видимых частей тела и используют распознавание как процедуру сопоставления частичных признаков. Авторы в работе [17] исследуют влияние справочной информации на поиск людей с помощью семантической сегментации.

Хотя вышеупомянутые методы показывают хорошие результаты в ограниченных сценариях применения, общая задача обнаружения и распознавания и реидентификации людей по-прежнему связана с рядом проблем: инвариантность к изменениям позы, инвариантность к изменениям освещения и выбор признаков изображения, которые позволяют достоверно идентифицировать людей.

Основные недостатки существующих методов заключаются в следующем:

- Большинство методов фокусируются исключительно на метках изображений, не учитывая атрибутивную информацию, что снижает их эффективность и точность
- Несмотря на высокую точность в лабораторных условиях, многие методы не готовы к практическому применению из-за нерешенных проблем, таких как изменчивость условий освещения, разнообразие поз и динамика движений
- На различных наборах данных, включая такие как Market-1501, результаты сильно колеблются, что указывает на низкую стабильность методов. Хотя методы глубокого обучения имеют свои преимущества, они сталкиваются с ограничениями, связанными с потребностью в больших объемах данных и высокими вычислительными затратами.

Новая методика, предложенная в работе направлена на преодоление этих ограничений с использованием современных подходов в области распознавания человека. Одним из ключевых предполагаемых подходов является метод автоматического взвешивания весов [22, 23], который позволяет интегрировать дополнительную информацию и повысить точность реидентификации в условиях неопределенности. Этот подход предполагает создание новой, более эффективной методики распознавания.

1. Методика детектирования силуэтов и лиц

Методика детектирования силуэтов и лиц является ключевым компонентом в процессе распознавания и реидентификации человека. Основная задача – эффективно определить и выделить силуэты человека на изображениях.

1.1. Детектор Yolov8

Для задачи обнаружения силуэтов и лиц в рамках разработки методики распознавания и реидентификации человека был выбран YOLOv8 – одна из самых современных моделей в сфере компьютерного зрения [24]. YOLOv8 выделяется своей способностью к быстрому и точному обнаружению объектов, что делает ее особенно привлекательной для задач, требующих реального времени обработки. Эта модель, хотя и не имеет на момент написания работы официального описания, уже зарекомендовала себя в научном сообществе как новый эталон обнаружения объектов. Основное ее преимущество – высокие значения mAP (средняя точность обнаружения) и низкая задержка обработки данных, особенно заметная при тестировании на известном датасете COCO.

YOLOv8, в качестве инновационной модели обнаружения объектов, была специально адаптирована для выполнения задач детектирования силуэтов и лиц. Основным отличием от предшественников является улучшенная способность к обучению на разнообразных данных, позволяющая модели выявлять более абстрактные и сложные визуальные паттерны. Это достигается за счет обучения на большом и многообразном наборе данных, который включает в себя различные категории объектов. Такой подход позволяет модели лучше адаптироваться к реальным условиям и различным сценариям

распознавания, что особенно важно при идентификации силуэтов и лиц в сложных и динамичных условиях.

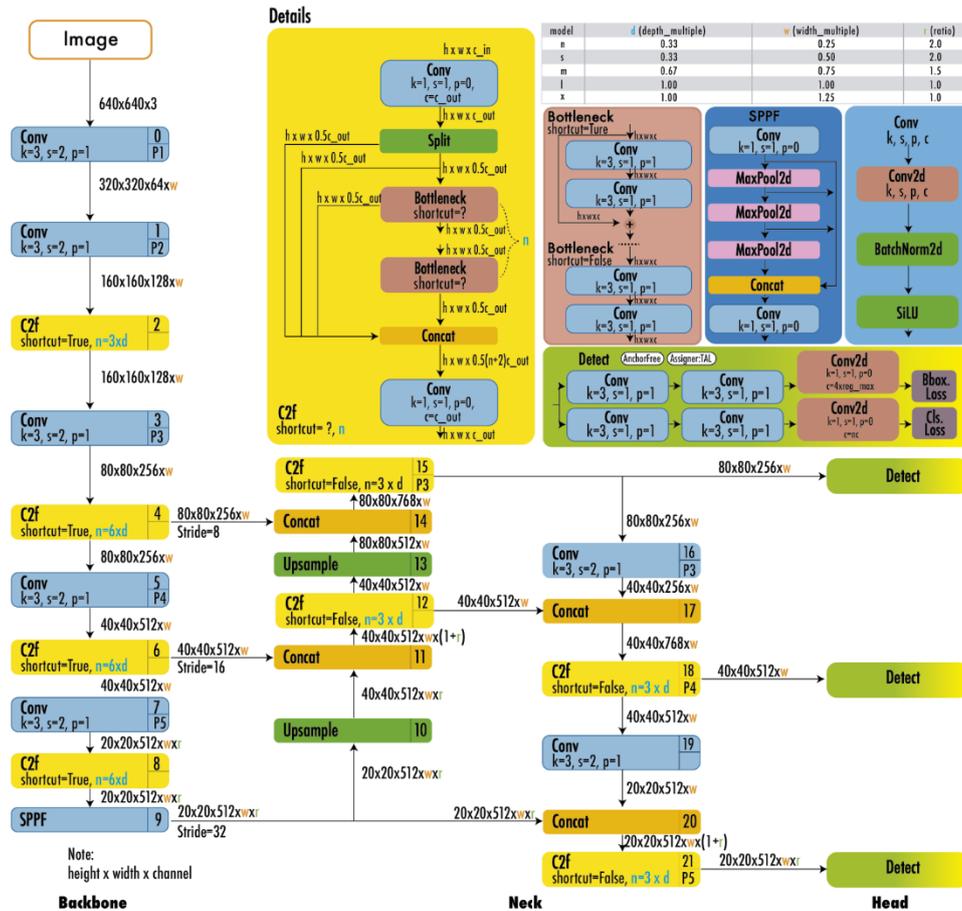


Рис. 1. Схема архитектуры YOLOv8

На Рисунке 1 представлена детальная архитектура YOLOv8. Эта модель развивает концепцию, заложенную в YOLOv5, но с некоторыми модификациями, в частности, в модуле CSPLayer, теперь именуемом как модуль C2f. Модуль C2f (cross-stage partial bottleneck с двумя сверточными слоями) объединяет высокоуровневые особенности с контекстуальной информацией, что способствует повышению точности детектирования. YOLOv8 представляет собой модель без якорей (anchor-free) с отдельными блоками обработки для определения наличия объекта, классификации и регрессии. Такой дизайн позволяет каждой ветви специализироваться на своей задаче, улучшая общую точность модели. В выходном слое YOLOv8 используется сигмоидная функция активации для оценки вероятности наличия объекта в ограничивающей рамке. Для вероятностей принадлежности объектов к классам применяется функция softmax.

YOLOv8 использует функции потерь CIoU и DFL для определения потерь ограничивающей рамки и бинарную кросс-энтропию для классификации. Эти функции потерь обеспечивают улучшенную производительность детектирования, особенно при работе с мелкими объектами. Кроме того, YOLOv8 предлагает модель семантической сегментации, YOLOv8-Seg. В ней в качестве основы используется извлекатель признаков CSPDarknet53, за которым следует модуль C2f вместо традиционной архитектуры "шеи" YOLO. За модулем C2f расположены две головы сегментации, обученные предсказывать маски семантической сегментации для входного изображения. YOLOv8-Seg обеспечивает выдающиеся результаты на различных бенчмарках по детектированию объектов и семантической сегментации, сохраняя при этом высокую скорость и эффективность.

Обобщенная функция потерь и процедура обновления весов:

$$L(\theta) = \frac{\lambda_{\text{box}}}{N_{\text{pos}}} L_{\text{box}}(\theta) + \frac{\lambda_{\text{cls}}}{N_{\text{pos}}} L_{\text{cls}}(\theta) + \frac{\lambda_{\text{dfl}}}{N_{\text{pos}}} L_{\text{dfl}}(\theta) + \frac{\phi}{2} |\theta|^2 \quad (1)$$

Уравнение для определения скорости:

$$V_t = \beta V_{t-1} + \nabla_{\theta} L(\theta_{t-1}). \quad (2)$$

Правило обновления весов:

$$\theta_t = \theta_{t-1} - \eta V_t \quad (3)$$

Конкретная функция потерь YOLOv8 определяется следующим образом:

$$L = \frac{\lambda_{\text{box}}}{N_{\text{pos}}} \sum_{x,y} 1_c^*(x,y) \left[\left(1 - q_{x,y} + \frac{|b_{x,y} - \widehat{b}_{x,y}|^2}{2\rho^2} \right) + \alpha_{x,y} \nu_{x,y} \right] +$$

$$\frac{\lambda_{\text{cls}}}{N_{\text{pos}}} \sum_{x,y} \sum_{c \in \text{classes}} [y_c \log(\widehat{y}_c) + (1 - y_c) \log(1 - \widehat{y}_c)] +$$

$$\frac{\lambda_{\text{dfl}}}{N_{\text{pos}}} \sum_{x,y} 1_c^*(x,y) [-(q_{(x,y)+1} - q_{x,y}) \log(\widehat{q}_{x,y}) + (q_{x,y} - q_{(x,y)-1}) \log(\widehat{q}_{(x,y)+1})] \quad (4)$$

где $q_{x,y}$ определяется как IoU для ячеек x, y :

$$q_{x,y} = \frac{\widehat{\beta}_{x,y} \cap \beta_{x,y}}{\widehat{\beta}_{x,y} \cup \beta_{x,y}} \quad (5)$$

и дополнительные выражения:

$$\nu_{x,y} = \frac{4}{\pi^2} \left(\arctan\left(\frac{w_{x,y}}{h_{x,y}}\right) - \arctan\left(\frac{\widehat{w}_{x,y}}{\widehat{h}_{x,y}}\right) \right)^2 \quad (6)$$

$$\alpha_{x,y} = \frac{\nu}{1 - q_{x,y}} \quad (7)$$

$$\widehat{y}_c = \sigma(\cdot) \quad (8)$$

$$\widehat{q}_{x,y} = \text{softmax}(\cdot) \quad (9)$$

где:

N_{pos} - общее количество ячеек, содержащих объект.

$1_c^*(x, y)$ - индикаторная функция для ячеек, содержащих объект.

$\beta_{x,y}$ - кортеж, представляющий истинную ограничивающую рамку, состоящую из $(x_{\text{coord}}, y_{\text{coord}}, width, height)$.

$\widehat{\beta}_{x,y}$ - предсказанная ячейкой ограничивающая рамка.

$b_{x,y}$ - кортеж, представляющий центральную точку истинной ограничивающей рамки.

y_c - метка истинности для класса c для каждой отдельной ячейки сетки (x, y) , независимо от наличия объекта.

$q_{(x,y)\pm 1}$ - ближайшие предсказанные IoU для соседних рамок (слева и справа).

$w_{x,y}$ и $h_{x,y}$ - ширина и высота соответствующих рамок.

ρ - диагональная длина наименьшей охватывающей рамки, покрывающей предсказанные и истинные рамки.

Эта функция потерь включает в себя потери CIoU для точности ограничивающей рамки, стандартную двоичную кросс-энтропию для многоклассовой классификации и распределенный фокусный потерь.

Протестированный на наборе данных MS COCO test-dev 2017, YOLOv8x показал точность AP 53.9% при размере изображения 640 пикселей (по сравнению с 50.7% у YOLOv5 при том же размере входа), при этом скорость составила 280 кадров в секунду на NVIDIA A100 и TensorRT.

1.2. Функция потерь для детектирования лиц и силуэтов людей

В рамках улучшения методики распознавания и реидентификации человека была разработана новая функция потерь, направленная на более точное детектирование лиц внутри силуэтов. Эта модификация функции потерь учитывает не только общие характеристики объекта, но и специфику расположения лица в пределах силуэта. Разработка усовершенствованной функции потерь для детектирования лиц в силуэтах включает в себя интеграцию нового компонента в стандартную функцию потерь модели YOLOv8. Формула новой функции потерь представлена следующим образом:

$$L' = L + \frac{\lambda_{face}}{N_{pos}} \sum_{x,y} 1_{x,y}^{face} L_{face}(\theta_{x,y}) \quad (10)$$

В этой формуле, L обозначает исходную функцию потерь, λ_{face} - коэффициент, который задает вес нового компонента в общей функции потерь, N_{pos} - общее количество ячеек с объектами, $1_{x,y}^{face}$ - индикаторная функция, которая определяет, содержит ли ячейка лицо, $L_{face}(\theta_{x,y})$ - компонент функции потерь для детекции лица внутри ограничивающей рамки силуэта.

Основной принцип работы новой функции потерь заключается в повышении точности идентификации лиц в сложных сценариях. Это достигается за счет штрафования модели в случаях неправильного обнаружения лиц внутри силуэта или их отсутствия, когда они ожидалась. Таким образом, модель становится более чувствительной к положению лица в рамках силуэта, что повышает общую точность системы распознавания. Новая функция потерь была разработана на основе понимания динамических условий визуального распознавания лиц. Она учитывает ряд критически важных факторов, таких как частичное закрытие лица, нестандартные углы наблюдения и переменная освещенность. Теоретически, такой подход должен повысить чувствительность модели к мелким деталям и улучшить её способность к распознаванию лиц в менее очевидных или прямых сценариях.

Внедрение новой функции потерь предполагается улучшить общую производительность системы в следующих аспектах:

- Основное предназначение новой функции потерь - улучшение точности идентификации лиц в условиях, где традиционные модели могут испытывать затруднения (например, при частично закрытом лице)
- В условиях быстрого изменения сцен (например, в местах массового скопления людей), где лица могут появляться и исчезать в кадре, новая функция потерь помогает модели быстрее адаптироваться к изменениям и эффективнее идентифицировать лица.
- Путем штрафования за ошибки детектирования вне силуэтов или пропуска лиц, когда они должны быть обнаружены, функция потерь способствует снижению ложных положительных и отрицательных срабатываний.

2. Экспериментальная проверка

2.1. Описание экспериментальной установки и используемых данных

Для эксперимента по детекции лиц и силуэтов людей была подготовлена экспериментальная установка, включающая в себя несколько ключевых этапов, таких как сбор данных, их аннотирование, подготовка и предобработка, а также применение модели YOLOv8 для прогнозирования границ объектов (bounding boxes).

Были использованы изображения и аннотации из набора данных WIDER FACE [26], который является одним из самых обширных и разнообразных наборов данных для задачи детекции лиц. Данный набор данных был скачан с официального сайта и распакован в соответствующие директории на локальной машине. Все файлы были распакованы и структурированы в отдельные директории для тренировочных, валидационных и тестовых данных. Аннотации в исходном формате WIDER FACE были преобразованы в формат, совместимый с моделью YOLO. В этом формате каждая аннотация включает в себя координаты центра объекта, его ширину и высоту, нормализованные относительно размеров изображения. Этот процесс включал автоматическое считывание исходных аннотаций, их преобразование и сохранение в новом формате.

Все изображения и аннотации были объединены в единую структуру каталогов, подходящую для обучения модели YOLO. Были созданы отдельные директории для тренировочных и валидационных данных, содержащие подкаталоги для изображений и соответствующих аннотаций.

Для детекции людей на изображениях валидационного набора данных была использована предобученная модель YOLOv8. Модель была настроена для предсказания границ объектов на изображениях. Полученные предсказания включали координаты границ и классы объектов. Предсказанные моделью YOLOv8 bounding boxes людей (класс "человек") были добавлены к существующим аннотациям лиц, при этом bounding boxes людей были помечены как класс "1". Это позволило объединить информацию о лицах и силуэтах в едином формате, готовом для дальнейшего использования в задачах распознавания и реидентификации.

Экспериментальная установка включает в себя высокопроизводительный сервер с установленными библиотеками для глубокого обучения, необходимыми для работы модели YOLOv8. На сервере была настроена среда выполнения, обеспечивающая быстрый и эффективный процесс обучения и

тестирования модели. Все эксперименты проводились с использованием графических процессоров (GPU), что позволило значительно ускорить процесс обработки изображений и обучения модели.

Таким образом, подготовка данных и настройка экспериментальной установки обеспечили прочную основу для проведения экспериментов по детекции лиц и силуэтов людей, что является важным шагом в разработке методов распознавания и реидентификации лиц.

2.2. Реализация подхода к детектированию лиц и силуэтов

Основным изменением в модели YOLOv8 стало добавление нового компонента в функцию потерь, который учитывает детекцию лиц внутри силуэтов. Этот компонент штрафует модель за неправильное обнаружение лиц внутри силуэтов. Для оценки влияния параметра λ_{face} на производительность модели было проведено несколько экспериментов с различными значениями λ_{face}

Каждая из моделей была обучена на одном и том же наборе данных, и результаты были оценены по метрикам Precision (P), Recall (R), mAP@0.5 и mAP@0.5:0.95.

Полученные результаты для различных значений λ_{face} приведены в таблице 1:

Таблица 1. Сравнительная таблица результатов обучения модели при различном λ_{face}

λ_{face}	Precision (Face)	Recall (Face)	mAP@0.5 (Face)	mAP@0.5:0.95 (Face)	Precision (Human)	Recall (Human)	mAP@0.5 (Human)	mAP@0.5:0.95 (Human)
0,1	0,9	0,603	0,714	0,395	0,829	0,902	0,939	0,837
0,4	0,899	0,606	0,716	0,397	0,825	0,905	0,94	0,837
0,8	0,901	0,607	0,719	0,399	0,83	0,902	0,94	0,836
1	0,896	0,609	0,718	0,399	0,825	0,905	0,94	0,835
1,3	0,903	0,603	0,718	0,4	0,832	0,9	0,94	0,833
1,6	0,9	0,608	0,72	0,401	0,826	0,904	0,939	0,834
2	0,907	0,607	0,723	0,402	0,829	0,9	0,939	0,832
2,5	0,903	0,61	0,722	0,402	0,827	0,903	0,94	0,833
2,7	0,901	0,611	0,723	0,402	0,825	0,9	0,938	0,83

Из таблицы видно, что значение $\lambda_{face} = 2.0$ обеспечивает наилучшие метрики для класса "face", при этом метрики для класса "human" остаются стабильными и высокими:

- Precision (Face): 0.907 (самое высокое значение)
- Recall (Face): 0.607 (стабильное значение)
- mAP@0.5 (Face): 0.723 (самое высокое значение)
- mAP@0.5:0.95 (Face): 0.402 (самое высокое значение)
- Precision (Human): 0.829 (стабильное значение)
- Recall (Human): 0.900 (стабильное значение)
- mAP@0.5 (Human): 0.939 (стабильное значение)
- mAP@0.5:0.95 (Human): 0.832 (стабильное значение)

Для лучшего понимания зависимости различных метрик от значения λ_{face} были построены следующие графики:

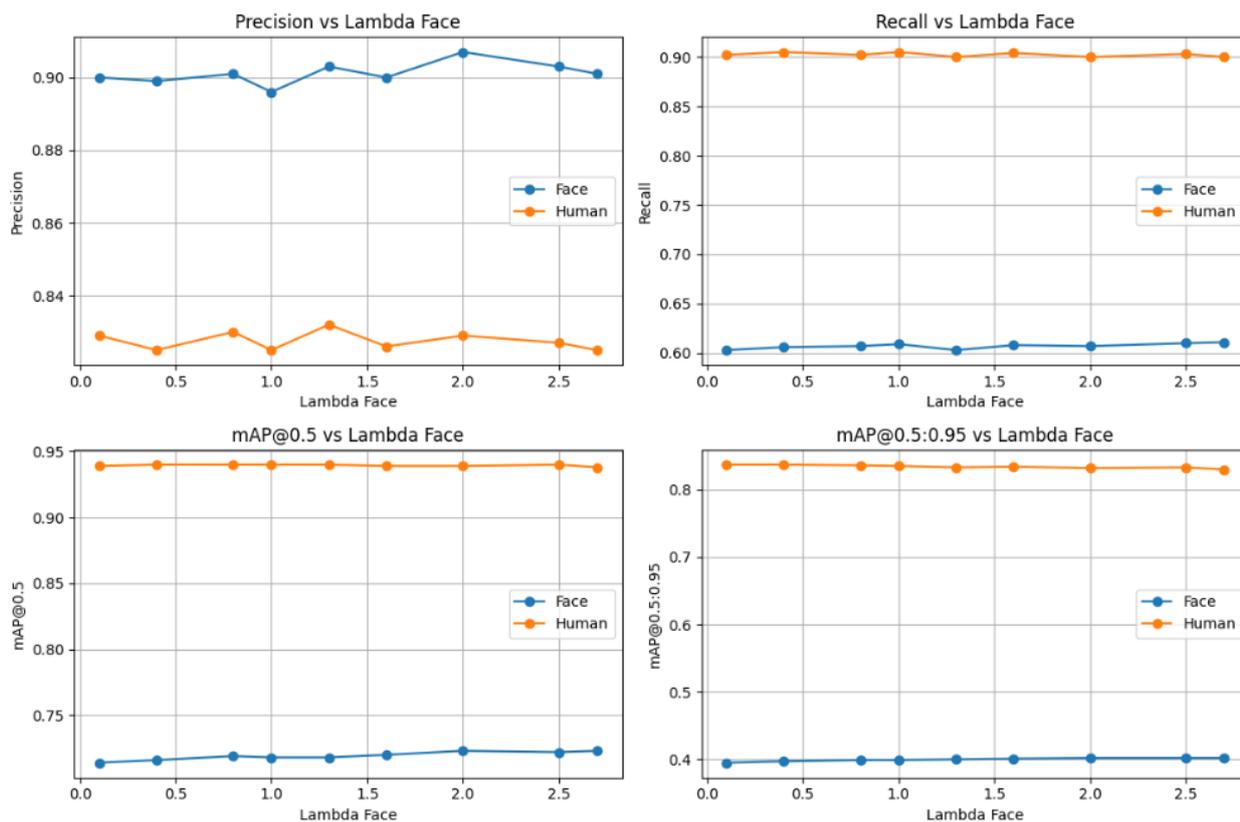


Рис. 2. Графики зависимости показателей качества от λ_{face}

В данном разделе был подробно рассмотрен процесс реализации подхода к детектированию лиц и силуэтов с использованием измененной функции потерь модели YOLOv8. Основной целью изменений было улучшение точности детекции лиц внутри силуэтов, что достигалось добавлением нового компонента в функцию потерь. Этот компонент учитывал ошибки детекции лиц, что способствовало более точному обучению модели в сложных сценариях, таких как частичное закрытие лица или нестандартные углы обзора.

Для оценки эффективности предложенного подхода были проведены детальные эксперименты с моделью YOLOv8 при различных значениях параметра λ_{face} . В данном разделе приведены подробные результаты обучения модели при $\lambda_{face} = 2$, так как это значение обеспечило наилучшие метрики для класса "face", при этом метрики для класса "human" остались стабильными и высокими.

Для проведения экспериментов использовались следующие основные параметры обучения: модель использовалась "yolov8m", обучение проводилось в течение 100 эпох с размером батча 16. Размер изображений составлял 640 пикселей. Начальная скорость обучения была равна 0.01.

На рисунке 3 представлены ключевые метрики обучения модели, включая потери на обучающей и проверочной выборках, а также значения Precision и Recall для классов "face" и "human".

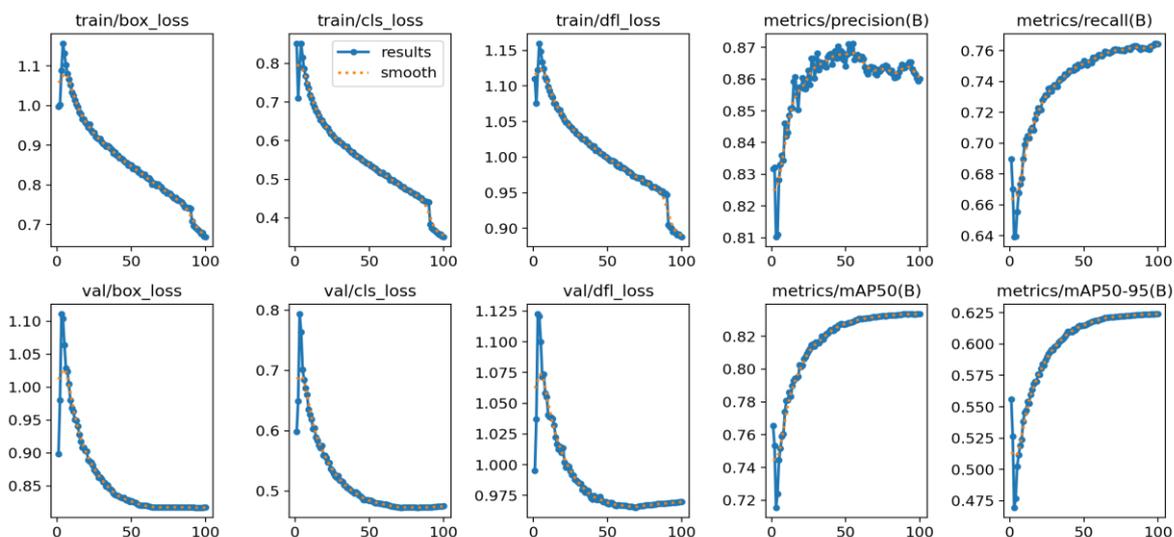


Рис. 3. Схема архитектуры YOLOv8

Матрица ошибок (рисунок 4) предоставляет детальный обзор точности модели в отношении классов "face" и "human". Она демонстрирует, как часто модель правильно распознавала объекты каждого класса и где допускала ошибки.

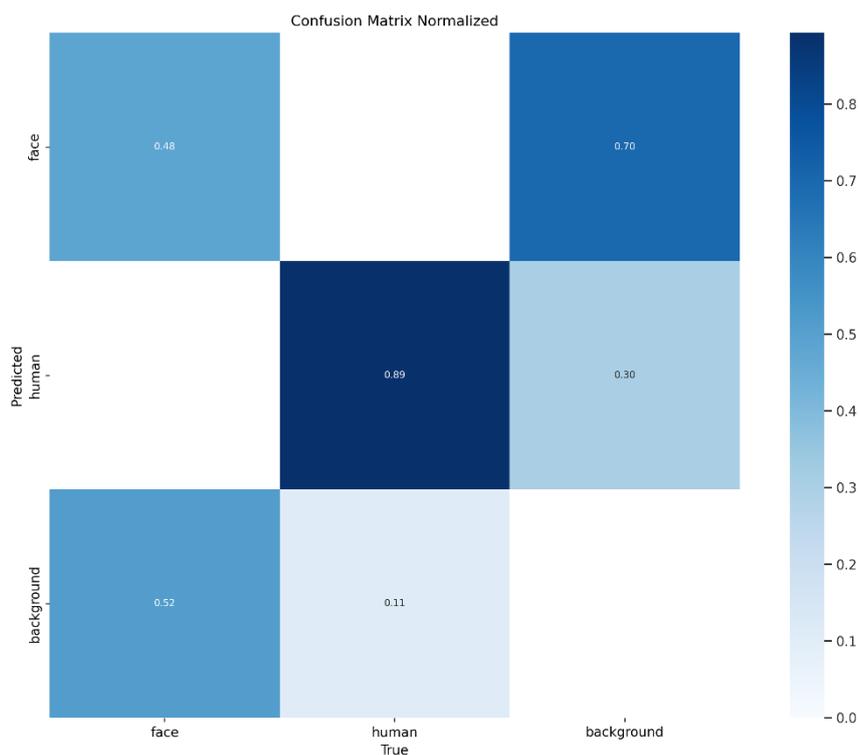


Рис. 4. Матрица ошибок

Ниже (рисунок 5) приведены примеры распознавания объектов моделью на изображениях из проверочного набора данных. Эти примеры иллюстрируют способности модели точно детектировать лица и силуэты даже в сложных условиях.



Рис. 5. Примеры работы модели

3. Заключение

В данной работе была предложена и реализована новая методика детектирования связанных объектов, таких как лица и силуэты, с использованием модели YOLOv8 и измененной функции потерь. Основное внимание было уделено добавлению нового компонента в функцию потерь, который учитывает детекцию лиц внутри силуэтов. Эксперименты показали, что предложенный подход значительно улучшает точность распознавания лиц в сложных сценариях, таких как частичное закрытие лица или нестандартные углы обзора. Оптимальные результаты были достигнуты при значениях параметра λ_{face} в диапазоне от 0.8 до 2.0, что позволило повысить метрики Precision и Recall для класса "face" без ухудшения производительности для класса "human".

Основные результаты работы включают улучшение метрик детектирования лиц в условиях, где традиционные модели могут испытывать трудности, а также демонстрацию эффективности предложенного подхода в задаче реидентификации человека. Эти результаты подтверждают целесообразность использования новых компонентов функции потерь для повышения точности и надежности систем визуального наблюдения. Более того, подход, представленный в данной работе, может быть адаптирован и применен к другим задачам детектирования связанных объектов, что открывает новые возможности для разработки более эффективных и универсальных систем компьютерного зрения.

Направления дальнейшей работы включают исследование других атрибутов, связанных с силуэтами, таких как пол, одежда и аксессуары, с целью улучшения общей точности и надежности систем реидентификации. Кроме того, стоит рассмотреть возможность применения предложенного подхода к другим моделям глубокого обучения и задачам компьютерного зрения. Также важно проводить дальнейшие исследования по оптимизации параметров модели и функции потерь для

различных наборов данных и сценариев, что позволит сделать системы визуального наблюдения более гибкими и эффективными в реальных условиях.

Литература

1. *Quanyou Zhao, Shujun Zhang*. A face detection method based on corner verifying, 2011 International Conference on Computer Science and Service System (CSSS), 2011, pp. 2854-2857, doi: 10.1109/CSSS.2011.5974784
2. *Dang K., Sharma S*. Review and comparison of face detection algorithms, 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017, pp. 629-633, doi: 10.1109/CONFLUENCE.2017.7943228.
3. *Shanmugavadivu P., Kumar A*. Rapid face detection and annotation with loosely face geometry, 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016, pp. 594-597, doi: 10.1109/IC3I.2016.7918032.
4. *Hofer P., Roland M., Schwarz P., Schwaighofer M., Mayrhofer R*. Importance of different facial parts for face detection networks, 2021 IEEE International Workshop on Biometrics and Forensics (IWBF), 2021, pp. 1-6, doi: 10.1109/IWBF50991.2021.9465087.
5. *Tian Q., Zhao S*. A fast face detection method for JPEG image, 2012 IEEE 11th International Conference on Signal Processing, 2012, pp. 899-902, doi: 10.1109/ICoSP.2012.6491725.
6. *Nehru M., Padmavathi S*. Illumination invariant face detection using viola jones algorithm, 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017, pp. 1-4, doi: 10.1109/ICACCS.2017.8014571.
7. *He R., Cao J., Song L., Sun Z., Tan T*. Adversarial Cross-Spectral Face Completion for NIR-VIS Face Recognition, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 5, pp. 1025-1037, 1 May 2020, doi: 10.1109/TPAMI.2019.2961900.
8. *Li S. Z. et al*. A near-infrared image-based face recognition system, 7th International Conference on Automatic Face and Gesture Recognition (FGR06), 2006, pp. 455-460, doi: 10.1109/FGR.2006.13.
9. *Zhang Wenchao, Shan Shiguang, Gao W., Chang Yizheng, Cao Bo, Yang Peng*. Information fusion in face identification, Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., 2004, pp. 950-953 Vol.3, doi: 10.1109/ICPR.2004.1334686.
10. *Liu C*. The development trend of evaluating face-recognition technology, 2014 International Conference on Mechatronics and Control (ICMC), 2014, pp. 1540-1544, doi: 10.1109/ICMC.2014.7231817.
11. *Cutler R., Davis L*. Robust periodic motion and motion symmetry detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2000, pp. 615-622.
12. *Heisele B., Woehler C*. Motion-based recognition of pedestrians, Proceedings of International Conference on Pattern Recognition and Image Processing, 1998, pp. 1325-1330.
13. *Nixon M.S., Carter J.N., Nash J.M., Huang P.S., Cunado D., Stevenage S.V*. Automatic gait recognition, IEE Colloq. Motion Anal. Track., 1999, pp. 1-6.
14. *Xiaojun Chang, Po-Yao Huang, Yi-Dong Shen, Xiaodan Liang, Yi Yang, Alexander G. Hauptmann*. RCAA: Relational context-aware agents for person search, European Conference on Computer Vision, 2018, pp. 84-100.
15. *Di Chen, Shanshan Zhang, Wanli Ouyang, Jian Yang, Bernt Schiele*. Hierarchical online instance matching for person search, AAAI Conference on Artificial Intelligence, 2020, pp. 10518-10525.
16. *Di Chen, Shanshan Zhang, Jian Yang, Bernt Schiele*. Norm-aware embedding for efficient person search, Computer Vision and Pattern Recognition, 2020, pp. 12615-12624.
17. *Di Chen, Shanshan Zhang, Wanli Ouyang, Jian Yang, Ying Tai*. Person search by separated modeling and a mask-guided two-stream CNN model, IEEE Transactions on Image Processing, 2020, vol. 29, pp. 4669-4682.
18. *Wenkai Dong, Zhaoxiang Zhang, Chunfeng Song, Tieniu Tan*. Instance guided proposal network for person search, Computer Vision and Pattern Recognition, 2020, pp. 2585-2594.
19. *Chuchu Han, Jiacheng Ye, Yunshan Zhong, Xin Tan, Chi Zhang, Changxin Gao, Nong Sang*. Re-id driven localization refinement for person search, International Conference on Computer Vision, 2019, pp. 9814-9823.
20. *Di Chen, Andreas Doering, Shanshan Zhang, Jian Yang, Juergen Gall, Bernt Schiele*. Keypoint message passing for video-based person re-identification, AAAI Conference on Artificial Intelligence, 2022.
21. *Yingji Zhong, Xiaoyu Wang, Shiliang Zhang*. Robust partial matching for person search in the wild, Computer Vision and Pattern Recognition, 2020, pp. 6827-6835.
22. *Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, Andrew Rabinovich*. GradNorm: Gradient normalization for adaptive loss balancing in deep multitask networks, International Conference on Machine Learning, 2018, pp. 794-803.
23. *Michelle Guo, Albert Haque, De-An Huang, Serena Yeung, Li Fei-Fei*. Dynamic task prioritization for multitask learning, European Conference on Computer Vision, 2018, pp. 270-287.
24. *Jacob Solawetz and Francesco*. What is yolov8? the ultimate guide., 2023. 04-30-2023.
25. *Shuo Yang, Ping Luo, Chen Change Loy, Xiaoou Tang*. WIDER FACE: A Face Detection Benchmark, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 5525-5533, doi: 10.1109/CVPR.2016.596.