

**Алгоритм подавления ложных гипотез  
при сегментации экземпляров сложных объектов**

*И.И. Журавлев, А.В. Макаренко*

Институт проблем управления им. В.А. Трапезникова РАН

**Введение.** Модели сегментации экземпляров на выходе формируют кортежи координат и масок объекта с некоторой оценкой уверенности. Эти кортежи называют гипотезами. Фильтрация гипотез в большинстве случаев производится алгоритмом non maximum suppression (NMS) [1] или его модификацией soft-NMS [2], однако для некоторых задач требуется дополнительная фильтрация, в частности, при решении задачи сегментации экземпляров листочков томата в целях определения общей площади поражения растения. Листочек томата имеет сложную форму: на разных масштабах элементы листочка похожи на отдельный листочек (рис. 1), скручивание сухих листочков создает визуальные фиктивные границы, вследствие чего элемент листочка распознается как отдельный экземпляр (ложная гипотеза). Принцип работы NMS не позволяет в некоторых случаях отклонять такие гипотезы. Мы предлагаем алгоритм подавления ложных гипотез (ПЛГ), который делает проверку, является ли множество точек гипотезы (маски объекта) подмножеством другой гипотезы. Если является, то такая гипотеза отклоняется.

**Алгоритм.** Пусть  $B_1, B_2$  – наборы координат внутренней области ограничивающих рамок (ОР) экземпляров. Степень погруженности одной ОР в другую определим формулой:

$$MoU(B_1, B_2) = \frac{\max(|B_1|, |B_2|)}{|B_1 \cup B_2|}, \quad (1)$$

где  $|\cdot|$  означает площадь, занимаемую ОР. Максимальное значение метрики равно 1. Минимальное достигается при  $|B_1 \cup B_2| = |B_1| + |B_2|$ . Если метрика превышает заданный порог, то переходим на следующий этап алгоритма. Пусть  $B_{\min}, B_{\max}, M_{\min}, M_{\max}$  – минимальные и максимальные ОР и маски гипотез соответственно. Извлечем область максимальной маски, которую занимает минимальная ОР:

$$M_{\max}^c = cropMask(M_{\max}, M_{\min}). \quad (2)$$

Далее вычисляется метрика IoU [1] для  $M_{\min}, M_{\max}^c$ . Если метрика выше заданного порога, то принимаем меньшую гипотезу за ложное срабатывание и отклоняем ее. Отметим, что формула (1) не является обязательной в работе алгоритма. Она используется в качестве дополнительной проверки, чтобы избежать излишних вычислений за счет вызова функции (2).

**Эксперименты.** Для подтверждения работоспособности алгоритма было проведено сравнение NMS и NMS в связке с ПЛГ на 69 размеченных изображениях, содержащих листочки сложной формы (рис. 2). Значение метрики precision с использованием ПЛГ оказалось выше, чем при использовании «чистого» NMS. При этом recall у NMS + ПЛГ ниже. Это можно объяснить тем, что одна гипотеза может содержать в себе два экземпляра, объединенных в один, при этом для этих экземпляров также предсказаны гипотезы с меньшими ОР. Согласно описанному алгоритму, меньшие гипотезы будут отвергнуты. Однако, учитывая особенность задачи, а именно, определить долю пораженных листочков, ложные срабатывания алгоритма могут привести к завышению степени поражения томатов, поэтому в данном случае можно допустить незначительное увеличение ошибок первого рода взамен на уменьшение ошибок второго рода.

Теоретические рассуждения о вспомогательной роли метрики (1) подтверждаются рис. 3. При низких значениях метрики проверке подвергаются все гипотезы. При значениях, близких к 1, число обнаруженных дубликатов резко уменьшается, поскольку для отклонения гипотезы

требуется, чтобы ОР одной гипотезы полностью содержала другую ОР. Вследствие этого происходит уменьшение precision. Поведение precision на рис. 4 объясняется тем, что для отклонения гипотезы требуется значительно большее перекрытие между масками гипотез с ростом порога IoU. Такие случаи редки, из-за чего precision уменьшается. Рост recall на рис. 3 и рис. 4 объясняется теми же предположениями, что и результаты рис. 2.

**Выводы.** Эксперименты свидетельствуют об эффективности предложенного алгоритма в рамках решаемой задачи сегментации экземпляров листочков томата. ПЛГ позволяет произвести более тонкую настройку фильтрации ложных срабатываний алгоритмов сегментации.



Рис. 1. Гипотезы mask R-CNN после фильтрации NMS

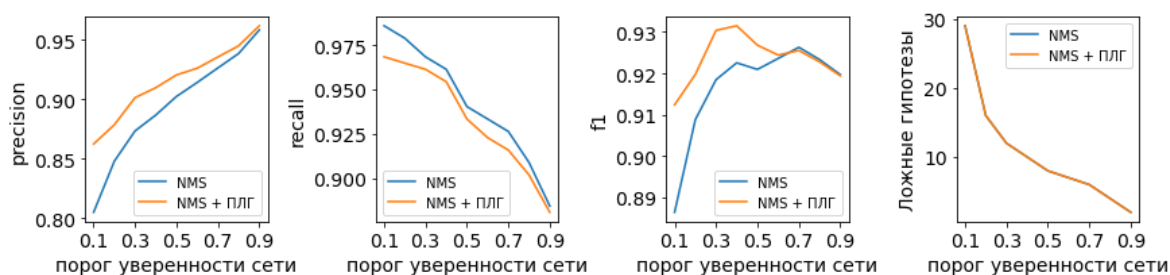


Рис. 2. Значения основных метрик и число дубликатов в зависимости от порога уверенности Mask R-CNN

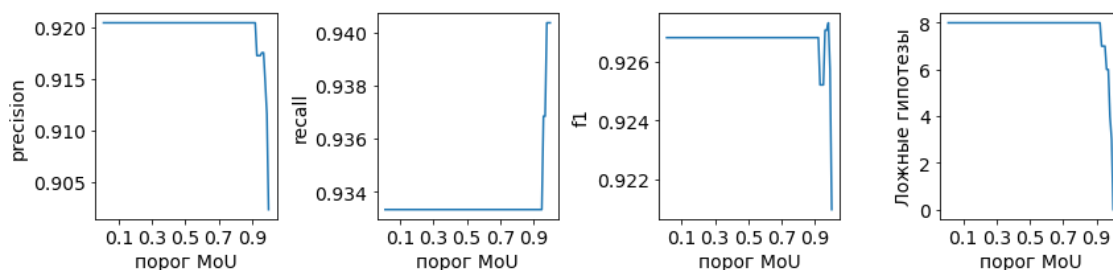


Рис. 3. Значения метрик в зависимости от порога MoU

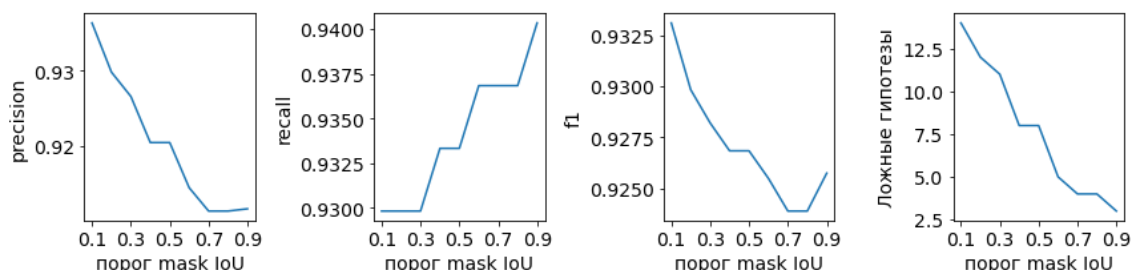


Рис. 4. Значения метрик в зависимости от порога для перекрытия масок

## Литература

1. He K. et al. Mask r-cnn //Proceedings of the IEEE international conference on computer vision. 2017. С. 2961-2969.
2. Bodla N. et al. Soft-NMS – improving object detection with one line of code //Proceedings of the IEEE international conference on computer vision. 2017. С. 5561-5569.