

УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ПРИЗНАКОВ ОБЪЕКТОВ

Титаренко М.А. (Университет ИТМО, Санкт-Петербург)
Научный руководитель – к.т.н. Малашин Р.О.
(Университет ИТМО, Санкт-Петербург)

Общий подход в компьютерном зрении состоит в том, чтобы отделить проблемы со зрением низкого уровня, такие как восстановление и улучшение изображения, от проблем со зрением высокого уровня и решать их независимо. Однако сегодня появляются архитектуры нейронных сетей, состоящие из последовательного объединения сети для шумоподавления и другой сети для задачи компьютерного зрения высокого уровня. Работа посвящена улучшению работы сети, выполняющей задачу высокого уровня для изображений плохого качества, с использованием аналогичной архитектуры каскадной нейронной сети.

Введение. В компьютерном зрении принято отделять проблемы со зрением низкого уровня от проблем со зрением высокого уровня и решать их независимо. В статье [1] предложен метод обучения глубокой нейронной сетью при использовании высокоуровневой информации в качестве дополнительного контроля качества шумоподавления. Высокоуровневой информацией извлекается из другой нейронной сети, решающей задачу компьютерного зрением высокого уровня.

Возникает вопрос: можно ли в такой системе улучшить качество работы сети, выполняющей задачу высокого уровня для изображений плохого качества, например с шумом? Целью данной работы является проведение серии экспериментов по обучению модели состоящую из двух сверточных нейронных сетей (первая решает задачу улучшения качества изображения, вторая классификатор) для решения задачи классификации в условиях шума.

Основная часть. Стратегию обучения каскадной сети такова, сначала инициализируется сеть высокоуровневой задачи из сети, которая хорошо обучена на изображениях без шума. Каскад двух сетей обучается сквозным образом, одним проходом получается очищенное изображение и высокоуровневая информация. Только веса в шумоподавляющей сети обновляются ошибкой, распространяемой обратно из следующей сети, которая аналогична минимизации потери восприятия для задачи высокого уровня. Веса второй сети не изменяются на протяжении обучения.

Важной частью обучения сетей является функция ошибки. Ошибка при восстановлении шумоподавляющей сети — это среднеквадратическая ошибка (MSE) между выходом шумоподавляющей сети и чистым изображением. Функция ошибки сетей сегментации являются кросс-энтропии между предсказанной меткой и достоверной меткой. Совместная ошибка является предметом исследования данной работы. В работе опробованы несколько возможных вариантов функции ошибки. Ошибку можно считать, как сумму ошибки при реконструкции и ошибки для задачи высокого уровня, а можно брать каждую по отдельности. В ходе экспериментов по обучению моделей нейронных сетей использовался датасет cifar10. Unet использована в работе, как архитектура для улучшения качества изображений. В качестве задачи высокого уровня выбрана классификация. Для классификации использована в архитектура resNet20. Сеть resNet20 была предварительно обучена на датасете cifar10 с 92,4% точности. В дальнейших экспериментах с каскадом двух сетей использовались веса этой обученной модели.

Среди проведенных экспериментов любопытным является результат обучения на зашумлённых изображениях каскада сетей с ошибкой от классификатора.

Модель на всем протяжении обучения увеличивает точность классификации, однако после 40 эпохи качество изображений начинает снижаться. Это значит, что Unet помимо решения

задачи денойзинга начинает вносить в изображения искажения, которые помогают сети классификатору выполнять свои обязанности.

С целью проверить способна ли сеть U-net обучиться повторять изображение, не искажая его при использовании высокоуровневой информации, выполнен еще один эксперимент с каскадом двух сетей обучающихся на чистых изображениях только с ошибкой классификации. В процессе обучения модель очень быстро переобучилась. В итоге максимальная точность в 36.4% достигнута после первой эпохи.

Выводы. Для всех экспериментов с шумом каскад сетей показал результаты ниже, чем у единственной сети классификатора.

В экспериментах с каскадом двух сетей максимальная точность классификации достигается при обучении сети препроцессора непосредственно на поврежденных парах изображений с помощью MSE. Добавление ошибки классификации во время обучения не улучшает результат.

Обратное распространение высокоуровневой информации из выхода сети классификатора при обучении сети препроцессора может привести к тому, что последняя приведет к искажению изображения. Возможное объяснение состоит в том, что препроцессор учится достигать некоторого пространства изображения, которое не охватывается обучающим датасетом. Использование этого эффекта может помочь найти некоторые новые идеи о том, как улучшить классификаторы.

1. Ding Liu, Bihan Wen. Connecting Image Denoising and High-Level Vision Tasks via Deep Learning. URL: <https://arxiv.org/pdf/1809.01826.pdf>.

Титаренко М.А. (автор)

Подпись

Малашин Р.О. (научный руководитель)

Подпись