Оптический журнал
Иконика

DOI: 10.17586/1023-5086-2022-89-02-25-35
УДК 004.932.4

Исследование способностей нейронных сетей к извлечению и использованию семантической информации при обучении восстановлению зашумлённых изображений

Михаил Алексеевич Титаренко[1], Роман Олегович Малашин[2]

[1], [2]*Институт физиология им. И.П. Павлова РАН, Санкт-Петербург, Россия*
[1]*Университет ИТМО, Санкт-Петербург, Россия*
[2]*Государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия*
[1]*mik199997@gmail.com; https://orcid.org/0000-0001-5978-4389*
[2]*malashinroman@mail.ru; https://orcid.org/0000-0002-2493-839X*
Адрес для переписки: Титаренко Михаил Алексеевич mik199997@gmail.com

Аннотация

Предмет исследования. В работе исследуются глубокие свёрточные нейронные сети, выполняющие задачу улучшения изображений в условиях шума, которые получают на вход дополнительную информацию об объектах на изображении в виде масок сегментации. Исследуется несколько способов использования семантической информации в процессе работы сети. Во-первых, путём подачи маски вместе с изображением на вход сети, а во-вторых, путём создания функции потерь с использованием маски сегментации. **Метод.** Проведено несколько серий экспериментов с различными типами использования семантической информации и без её использования. Эксперименты проводились с несколькими интенсивностями шума. Выполнен анализ качества восстановления изображения в целом и качества восстановления области изображений, соответствующих целевому классу. Класс дорожных знаков был выбран в качестве целевого, поскольку обладает меньшей изменчивостью, чем многие другие классы, что даёт преимущество сети при его восстановлении при наличии семантической информации, чем без таковой. В ходе исследования использовался набор данных COCO с размеченными картами сегментации. Для анализа семантических свойств всех объектов, содержащихся в наборе COCO, на восстановление изображений было разработано тестовое окружение с возможностью визуализации результатов тестирования, что позволило сделать некоторые полезные выводы о том, как различные свойства объектов влияют на точность их восстановления. **Основные результаты.** Мы показываем, что в условиях сильного шума сеть реконструкции, обученная с использованием дополнительной информации в виде масок сегментации, может лучше восстанавливать объекты, соответствующие маскам (на 3,5%), при этом способность сети восстанавливать все изображение существенно не падает (на 0,4%), однако для слабого и среднего шума такого прироста качества не получено. **Практическая значимость.** В работе мы ставили цель не создать законченный алгоритм и архитектуру нейронной сети, а лишь исследовать возможные свойства подобных алгоритмов, поэтому мы подавали на вход нейронных сетей эталонные семантические разметки. При доработке такой метод может быть дополнен сетью-сегментатором, которая бы извлекала информацию автоматически из зашумлённого изображения (в этом случае сам процесс улучшения может быть итеративным — после сегментирования изображение улучшается, а по улучшенному изображению строится уточнённая маска сегментации).

Ключевые слова: улучшение изображений, сегментация изображений, глубокие нейронные сети

Ссылка для цитирования: Титаренко М.А., Малашин Р.О. Исследование способностей нейронных сетей к извлечению и использованию семантической информации при обучении восстановлению зашумлённых изображений // Оптический журнал. 2022. Т. 89. №2. С. 25–35. DOI: 10.17586/1023-5086-2022-89-02-25-35

Коды OCIS: 150.1135, 100.2980.

ВВЕДЕНИЕ

Стандартный подход для улучшения изображений с использованием нейронных сетей заключается в обучении сети на наборе эталонных и искажённых изображений [1–3]. В этой работе предлагается модификация этого подхода за счёт использования дополнительной семантической информации. Один из известных подобных методов заключается в подготовке обучающего набора данных с выделением примеров, объединённых общими признаками. Например, в работе [4] исследователи показывают, что для сетей-реконструкторов использование для обучения подвыборки (из полного набора данных) определённых классов помогают лучше восстанавливать их без существенной потери общего качества восстановления.

Существует также подход использования дополнительной сети для извлечения признаков и их применение в функции потерь. Методы обучения нейронных сетей с функцией потерь, учитывающей точность восстановления признаков высокого уровня, называются обучением с ошибкой восприятия [5, 6].

Также есть метод с использованием дополнительной информации при подаче на вход сети [7–9]. В нашей работе выполнено исследование этого метода. На вход нейронной сети подаётся дополнительная семантическая информация в виде маски сегментации объектов.

Актуальность исследования вытекает из того, что в стандартном подходе улучшения изображений с использованием нейронных сетей при обучении только на наборе эталонных и искажённых изображений семантическая информация может быть использована не полностью. На рис. 1 изображён пример работы такой сети-реконструктора. При правильном использовании семантической информации сеть должна была бы перенести текстуру с тела жирафа на голову. Однако сеть неправильно окрасила область головы.

Предположительно информация о сегментации может помочь во время обучения составить более точное представление о структуре объектов, и это поможет восстановлению изображений этих объектов.

В работе мы ставили цель не создать законченный алгоритм и архитектуру нейронной сети, а лишь исследовать возможные свойства подобных алгоритмов, поэтому мы подавали на вход нейронных сетей эталонные семантические разметки. При доработке такой метод может быть дополнен сетью-сегментатором, которая бы извлекала информацию автоматически из зашумлённого изображения (в этом случае сам процесс улучшения может быть итеративным — после сегментирования изображение улучшается, а по улучшенному изображению строится уточнённая маска сегментации).

1. ИСПОЛЬЗОВАНИЕ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ ПРИ РАБОТЕ НЕЙРОННОЙ СЕТИ, ВЫПОЛНЯЮЩЕЙ УЛУЧШЕНИЕ ИЗОБРАЖЕНИЯ

Стандартный подход для улучшения изображений с использованием нейронных сетей заключается в обучении сети на наборе эталонных и искажённых изображений [1–3]. Сеть учится из искажённых изображений восстанавливать эталонные, используя только входное изображение и выученное представление

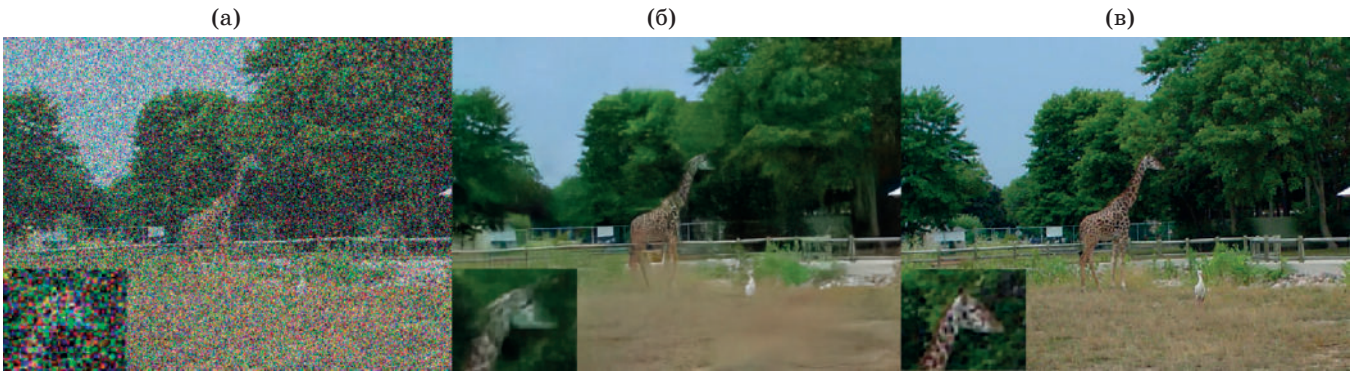


Рис. 1. Пример работы сети-реконструктора без дополнительного использования семантической информации: (а) зашумлённое изображение, (б) восстановленное изображение, (в) исходное изображение.

о том, как должны выглядеть изображения. В нашем подходе модель нейронной сети получает на вход кроме изображения дополнительную информацию в виде карты сегментации объектов на изображении. Маска сегментации подаётся на вход сети в четвёртом канале входного тензора, а три остальных соответствуют обычным цветовым каналам изображений.

На рис. 2 изображена схема исследуемой сети с использованием маски сегментации на входе сети.

Использование карт сегментации даёт нейронной сети информацию о положении объектов, что потенциально позволит ей лучше различать объекты и улучшать их.

В работе проверяется гипотеза, что информация о сегментации помогает во время обучения составить более точное представление о структуре объектов, и это помогает восстановлению изображений этих объектов. Исследование основано на сравнении результатов работы сетей, обученных с использованием семантических масок и без их использования.

Сравнение результатов работы сетей-реконструкторов выполнено с использованием критерия соотношения пикового сигнала к шуму (PSNR) двумя способами. Первый критерий — вычисление PSNR по всему восстановленному изображению. Второй критерий — вычисление PSNR в области масок сегментации определённых классов.

Мы предлагаем способ, который дополнительно стимулирует сеть использовать маску сегментации. Способ заключается в обучении сети с использованием специальной функции ошибки. Функция состоит из двух частей: среднеквадратичной погрешности (MSE) для выходного изображения и эталонного и MSE для выходного изображения и эталонного, но только для области изображения, соответствующего маске сегментации, дальше по тексту будем называть её MSE-mask. Погрешности взвешиваются параметром λ. Этот параметр даёт возможность контролировать, на что больше направлено внимание сети.

Общий вид функции потерь:

$$L(x, \bar{x}, m) = L_R\left[F_R(x, m), \bar{x}\right] + \\ + \lambda L_M\left[F_R(x, m)^{\circ}m, \bar{x}^{\circ}m, m\right], \quad (1)$$

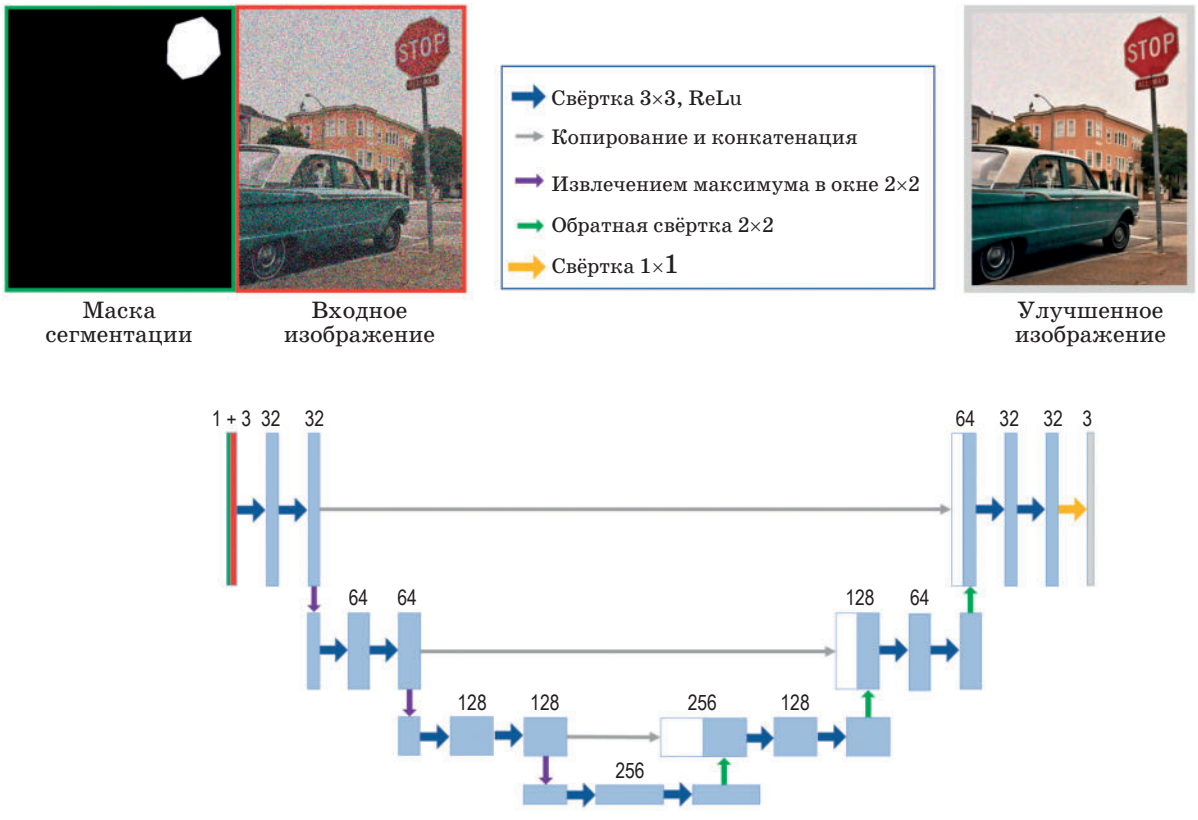


Рис. 2. Схема сети U-net с дополнительным входом в виде маски сегментации.

где x — изображение с шумом на входе, $\bar{x}$ — изображение без шума и m — маска сегментации, λ — параметр для взвешивания погрешности. F_R обозначает сеть реконструктор, L_R — погрешность реконструкции, которая вычисляется как среднеквадратичная погрешность между выходом сети-реконструктора и чистым изображением, $x^{\circ}m$ — это операция поэлементного произведения каждого канала изображения на маску сегментации. В результате выполнения операции все пикселы изображения станут нулевыми кроме тех, что соответствуют маске. L_M — это погрешность реконструкции в области масок сегментации, вычисляется как среднеквадратичная погрешность между выходом сети-реконструктора и чистым изображением в области масок сегментации, умноженная на поправку на площадь маски. Поправка вычисляется как отношение числа пикселов в маске к размеру всего изображения.

2. ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА РАЗРАБОТАННОГО МЕТОДА ИСПОЛЬЗОВАНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ ПРИ ПОДАЧЕ НА ВХОД СЕТИ

В качестве сети для улучшения качества изображений используется архитектура U-net [10]. Полносвёрточная нейронная сеть U-net — одна из наиболее эффективных архитектур, применяемых в последнее время для улучшения изображений [11–13]. U-net представляет собой свёрточную нейронную сеть для шумоподавления (в общем случае, она пригодна и для решения других задач обработки изображений), которая принимает зашумлённое изображение и выдаёт результат восстановления. На рис. 2 представлена схема сети.

Для обучения необходима выборка изображений с размеченными картами сегментации. Поэтому в ходе обучения нейронной сети использовался набор данных COCO [14]. На рис. 3 приведены примеры картинок из этого набора данных с подписями соответствующих классов. Из набора данных выделено 1803 изображения с объектами класса «дорожные знаки». 1603 изображений использованы для тренировочной выборки и 200 изображений для тестовой.

Класс дорожных знаков был выбран, поскольку обладает меньшей изменчивостью, чем многие другие классы (например, кот или поезд), что, по нашему мнению, даёт бóльшее преимущество сети при его восстановлении при наличии семантической информации, чем без таковой. В качестве алгоритма, выполняющего автоматическую оптимизацию сети методом обратного распространения ошибки, выбран стохастический метод оптимизации Adam [15]. Обучение производилось на протяжении 1280 эпох. Размер пакета изображений, единовременно проходящих через сеть в про-

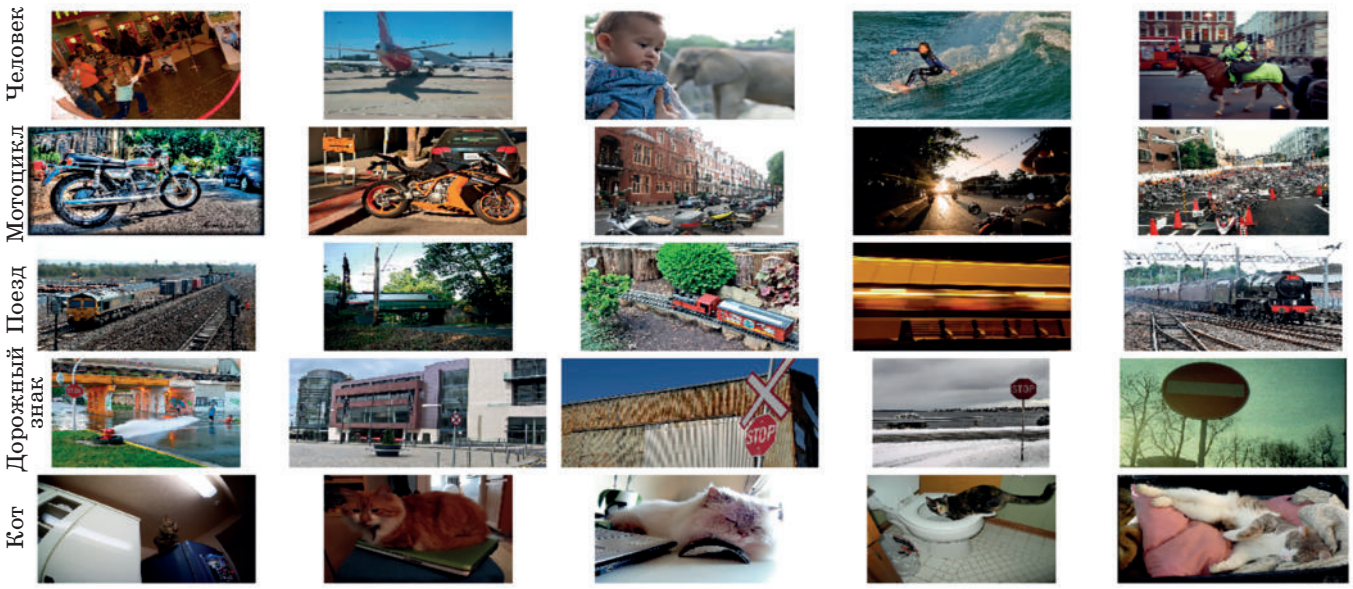


Рис. 3. Примеры из базы данных COCO.

цессе обучения, равен 4. Размер изображений в процессе обучения равен 256. Начальное значение скорости обучения равно 0,001. В процессе обучения шаг градиентного спуска сокращается в два раза после эпох 64, 128, 384, 896, 1024, 1152, 1216. Результатом обучения является результат эпохи с наибольшим значением PSNR.

Для обобщения наблюдаемых эффектов проводились серии экспериментов с тремя интенсивностями гауссова шума.

Эти параметры являются общими для всех экспериментов с изображениями дорожных знаков.

Изображения, подаваемые на вход U-net, переведены в вещественные числа в диапазоне от 0 до 1 (что важно для согласованности со значениями дисперсии шума).

В работе проведено исследование способности сети использовать информацию о сегментации для более качественного восстановления области изображений, соответствующих маскам сегментации. Для этого выполнено несколько серий экспериментов:

• Обучение сети на полном и урезанном наборе данных COCO, чтобы выяснить, как влияет размер выборки на качество восстановления различных классов.

• Обучение сети на части выборки изображений со знаками с использованием и без использования масок сегментации.

• Обучение сети на части выборки изображений со знаками с использованием дополнительной функции потерь в области масок сегментации.

Для сетей-реконструкторов разработан метод тестирования, показывающий качество восстановления изображений для всех классов, присутствующих в наборе данных COCO. Представлены графики с наглядной визуализацией результатов работы сетей-реконструкторов.

2.1. Обучение сети на полном и урезанном наборе данных COCO без использования масок сегментации

Так как в дальнейших экспериментах с использованием масок сегментации обучающая выборка довольно мала, надо убедиться в том, что сеть-реконструктор, обученная на таком небольшом наборе данных, способна научиться восстановлению изображений на таком небольшом наборе данных. Была проведена серия экспериментов по обучению сети на случайных подвыборках из полного набора данных COCO. При этом сохранялось число итераций в процессе обучения. Это выполнено пропорциональным увеличением числа эпох с уменьшением размера выборки.

Таблица 1. Результаты обучения сети-реконструктора на выборках различного размера из COCO без использования масок сегментации

Размер обучающей выборки	Число эпох	Средний PSNR на валидационной выборке, дБ
82783	20	25,63
40000	40	25,62
20000	80	25,68
10000	160	25,37
2500	640	23,91
1250	1280	25,68
625	2560	25,61

Во всех экспериментах обучение выполнялось методом обратного распространения ошибки со среднеквадратичной функцией потерь.

Валидация сети производилась на одном и том же наборе данных из COCO размером 40504 изображений.

В табл. 1 приведены результаты обучения сетей восстановлению изображений с высокой интенсивностью шума $\sigma = 0{,}4$.

Результаты экспериментов дают близкие результаты, однако для размера обучающей выборки 2500 качество снижено по сравнению с остальными. Это можно объяснить неудачной инициализацией весов (что случается достаточно редко).

Результаты экспериментов говорят о том, что сети, обученные на полном и урезанном наборе данных, показывают близкие результаты. Значит, сеть способна обучиться качественно восстанавливать изображения даже на небольшой выборке. Важно при этом сохранять большое число итераций обучения.

2.2. Обучение сети на части набора данных COCO со знаками с использованием и без использования масок сегментации

Выполнено несколько серий обучения сети. Каждая серия экспериментов состояла из трёх экспериментов по восстановлению изображе-

ний в условиях гауссова шума с интенсивностями $\sigma = 0{,}04$, $\sigma = 0{,}1$, $\sigma = 0{,}4$.

В первой серии экспериментов сеть обучалась на случайной выборке из набора данных COCO с размером, соответствующем размеру выборки, содержащей объекты целевого класса. Этот случай считается базовым. Его результат показывает, как сеть может восстанавливать полное изображение без привязки обучающего набора данных к целевому классу и даёт оценку того, насколько хорошо такая неспециализированная сеть может восстанавливать объекты целевого класса.

Во второй серии экспериментов сеть обучалась на наборе данных, состоящем только из изображений, содержащих объекты целевого класса, но в процессе обучения маски не использовались.

В третьей серии экспериментов сеть обучалась на наборе данных, состоящем только из изображений, содержащих объекты целевого класса, при этом процессе обучения использовались маски сегментации объектов.

Во всех описанных выше экспериментах обучение выполнялось со среднеквадратичной функцией потерь.

В четвёртой серии экспериментов сеть обучалась с использованием специальной функции потерь согласно формуле (1). Для определения оптимального значения параметра λ производилась серия обучений с различными фиксированными значениями параметра. Критерием качества выступал PSNR в области маски на валидационной выборке.

Во всех случаях сети подаётся на вход тензор глубиной четыре. Три первых канала соответствуют обычным цветовым каналам изображений, а в четвёртый канал может подаваться маска сегментации или пустая маска — матрица, заполненная нулями.

Размер выборки и число итераций во всех экспериментах совпадал.

Валидация сети производилась на одном и том же наборе данных из COCO размером 40504 изображений. В каждой серии валидации использовались общие начальные параметры псевдослучайных чисел, значит, для всех изображений шум накладывался один и тот же. Это необходимо для достоверного сравнения результатов.

В табл. 2 приведены результаты серий экспериментов. Представлен средний PSNR для валидационной выборки и средний PSNR в области масок сегментации изображений из валидационной выборки с поправкой на площадь.

Во-первых, можно увидеть, что для областей восстановленных изображений, соответствующих маскам, PSNR ниже, чем для полного изображения. Это можно объяснить присутствием в выборке достаточно малых дорожных знаков. В условиях шума объекты, соответствующие этим маскам, сливаются

Таблица 2. Отношение пикового сигнала к шуму для сети U-net, обученной с использованием масок сегментации и без их использования (жирным выделены лучшие значения с учётом десятых и сотых долей децибел, не представленных в таблице)

	Функция ошибки	MSE				MSE + MSE-mask	
	Набор данных	COCO	знаки			знаки	
	Маска при обучении	–	–	+	+	+	+
Интенсивность шума	Маска при тестировании	–	–	–	+	–	+
$\sigma = 0{,}04$	Вся картинка, дБ	**35,8**	35,7	35,7	35,7	35,7	35,7
	В области маски, дБ	34,1	34,2	34,2	**34,2**	34,2	34,2
$\sigma = 0{,}1$	Вся картинка, дБ	**31,5**	31,4	31,4	31,4	31,4	31,4
	В области маски, дБ	29,9	30,2	30,2	30,2	30,2	**30,2**
$\sigma = 0{,}4$	Вся картинка, дБ	**25,8**	25,7	25,7	25,7	25,7	25,7
	В области маски, дБ	23,2	23,7	23,7	23,8	23,6	**24,0**

с шумом и при восстановлении сеть их удаляет. Подробнее этот эффект будет разобран в следующем разделе.

Из табл. 2 видно, что разница в качестве восстановления при использовании слабого шума $\sigma = 0{,}04$ пренебрежимо мала для всех экспериментов. Для $\sigma = 0{,}1$ при использовании выборки из дорожных знаков наблюдается незначительное повышение точности распознавания области маски при незначительном понижении точности распознавания всего изображения. При этом использование маски во время обучения даёт лишь пренебрежимо малый прирост восстановления образов дорожных знаков (при обучении на базе данных с дорожными знаками). Для сильного шума эта разница становится существенной. Соотношение пикового сигнала к шуму в этом случае повышается на 0,5 дБ при использовании модифицированного набора данных, и ещё на 0,3 дБ при использовании масок с модифицированной функцией потерь.

В ходе экспериментов методом грубого перебора были подобраны оптимальные значения параметра λ: для слабого шума ($\sigma = 0{,}04$) $\lambda = 0{,}001$, для среднего шума ($\sigma = 0{,}1$) $\lambda = 0{,}008$, для сильного шума ($\sigma = 0{,}4$) $\lambda = 0{,}02$.

Наблюдается прямая зависимость оптимального значения λ от интенсивности шума и различия в качестве использования маски. Это подтверждает, что сети в условиях слабого шума не выгодно использовать маски в процессе обучения и лучше сконцентрироваться на общей задаче улучшения всего изображения. А вот для случая сильного шума маска оказывается очень полезной дополнительной информацией. Она как минимум сообщает о форме и положении объекта интереса.

Таким образом, во-первых, подтверждаются результаты работы [4], где исследователи показывают, что для сетей-реконструкторов использование на обучении подвыборки набора данных определённых классов помогает лучше их восстанавливать без существенной потери общего качества восстановления. Однако, как видно из табл. 2, в условиях проведённых нами экспериментов это наблюдается лишь при использовании сравнительно интенсивного шума. Также результаты показывают, что в условиях сильного шума сеть может с пользой научиться использовать информацию, заключённую в семантической разметке.



Рис. 4. Пример работы сетей, обученных в разных режимах. В левом нижнем углу каждого изображения приведено укрупненное изображение дорожного знака. (а) — зашумлённое изображение ($\sigma = 0{,}4$), (б) — исходное изображение, (в–е) — восстановленные изображения, (в) — сеть обучена на полном наборе COCO, (г) — сеть обучена с использованием маски, но на вход сети подана нулевая маска, (д) — сеть обучена с использованием маски, на вход сети подана маска, (е) — сеть обучена с использованием модифицированной функции потерь (1).

Визуальная демонстрация полученных результатов приведена на рис. 4.

Важно отметить, что сеть действительно реагирует на маску и её отсутствие (рис. 3г и 3д).

Хорошо видно, что сеть при подаче на вход маски лучше восстановила границу знака, и, значит, она её действительно использует.

2.3. Влияние семантических свойств объектов на восстановление изображений

Для анализа семантических свойств всех объектов, содержащихся в наборе COCO, на восстановление изображений было разработано

тестовое окружение, которое производит оценку качества восстановления изображений и областей изображений, соответствующих объектам всех классов в наборе данных COCO. Всего для этого использовалось 40504 изображений. Всегда использовались общие начальные параметры псевдослучайных чисел, значит, для всех изображений шум накладывался один и тот же.

Была добавлена возможность визуализации результатов тестирования, что позволяет сделать некоторые полезные выводы о том, как различные свойства объектов влияют на точность их восстановления.

На рис. 5–8 результаты тестирования представлены в виде точечного графика с зависимостью среднего размера объектов (диаметр области) от среднего PSNR. Каждая «точка» имеет свой цвет и форму, которые соответствуют определённому классу в базе данных COCO (рис. 5). Целевой класс (дорожный знак) обозначен розовым треугольником.

На графиках также показан средний PSNR для всех классов (красная пунктирная линия) и средний PSNR по полным изображениям, включая неразмеченный фон (чёрная пунктирная линия).

Ниже на рис. 6–8 приведены результаты тестирования для экспериментов с интенсивностью шума $\sigma = 0{,}4$.

Видно, что все графические элементы наложились друг на друга, за исключением целевого класса дорожных знаков (розовый треугольник, область выделена красным эллипсом), точность восстановления которого повышается при использовании маски сегментации.

Для экспериментов, в которых сеть обучена использовать маски сегментации (рис. 8), эксперимент выполнен с использованием и без использования масок сегментации. В этих случаях «точка» целевого класса (дорожный знак), обозначенная розовым треугольником, изображена два раза с использованием и без использования масок сегментации.

По приведённым рисункам можно наглядно показать, что подход с использованием масок сегментации в процессе обучения даёт прирост в качестве восстановления целевых объектов. При этом, качество восстановления объектов других классов практически не уступает подходу без использования масок сегментации (розовый треугольник сдвигается вправо по абсолютной шкале и относительно всех остальных графических элементов, обознача-



Рис. 5. Обозначения условных символов классов из набора данных COCO.

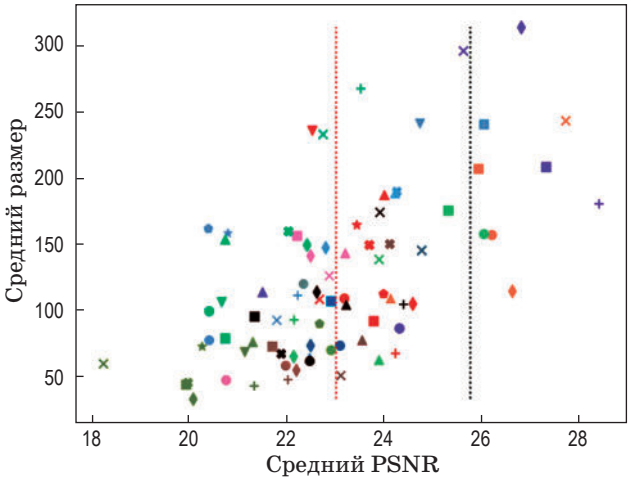


Рис. 6. Результат тестирования для модели сети, обученной на случайной выборке из набора данных COCO без использования масок сегментации.

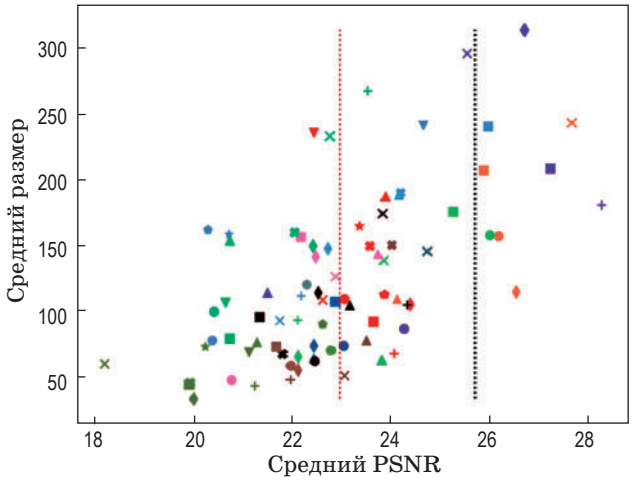


Рис. 7. Результат тестирования для модели сети, обученной на выборке, содержащей знаки без использования масок сегментации.

ющих различные объекты, которые остаются на своих местах).

Также заметна корреляция между размером объекта и качеством восстановления изображения. Когда объекты чрезвычайно малы, то в условиях шума они сливаются с фоном и при восстановлении сеть их удаляет. Это объясняет большую эффективность использования масок для сильного шума. Маска подсказывает сети, где нужно восстановить объект вместо того, чтобы его удалить.

Также можно заметить, что все размеченные объекты с точки зрения соотношения пикового сигнала к шуму в среднем сложнее поддаются

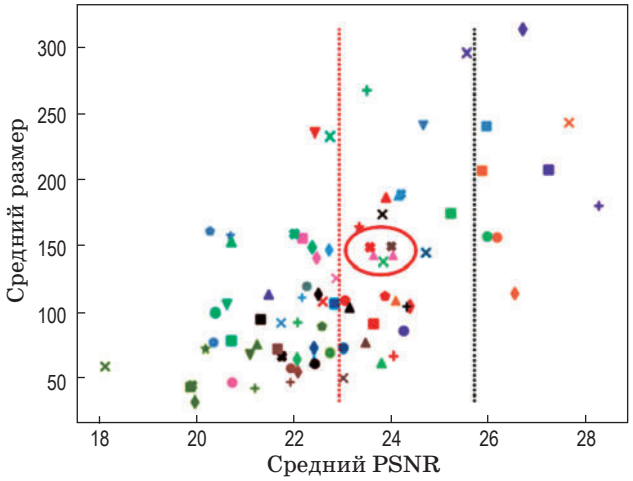


Рис. 8. Результат тестирования для модели сети, обученной на выборке, содержащей знаки с использованием масок сегментации с функцией ошибки (1). Тестирование выполнено с использованием и без использования масок сегментации.

восстановлению, чем фон (красная пунктирная линия правее, чем чёрная пунктирная линия). Можно заключить, что объекты, которые представляют интерес для человека, также характеризуются более сложной для восстановления структурой, чем фоновые.

Интерес также может представлять наблюдение, что наилучшим образом восстановились большие белые объекты (унитаз, холодильник), что можно объяснить в том числе эффектом обрезания максимальных яркостей самых ярких объектов в результате работы предварительных алгоритмов улучшения изображения, встроенных в фотоаппарат.

В целом можно заключить, что основным фактором, который сказывается на точность восстановления изображения с точки зрения отношения пикового сигнала к шуму, является однородность объекта — чем более однородный объект, тем точней он восстанавливается. Так, например, класс «пицца» восстановился менее точно в среднем, чем класс «банан» несмотря на то, что последний объект в среднем имел намного меньшую площадь.

ЗАКЛЮЧЕНИЕ

Проведено исследование метода применения нейронных сетей для улучшения изображений при использовании в явном виде информации об их смысловом содержании.

Эксперименты с обычными входными данными сети показывают, что сеть-реконструктор может лучше восстанавливать объекты при использовании обучающей выборки, содержащей эти объекты. Это подтверждает результаты работы [4].

Эксперименты с использованием масок сегментации при работе сети показывают, что в условиях сильного шума сеть-реконструктор, обученная с использованием дополнительной информации в виде маски сегментации, может лучше восстанавливать объекты, соответствующие маскам, и при этом способность сети восстанавливать всё изображение падает незначительно. Для слабого и среднего шума наблюдается пренебрежимо малый эффект.

Эксперименты с предложенным методом использования масок сегментации в функции потерь показывают его преимущество в условиях сильного шума. В этом случае существует возможность контролировать, на что будет больше направлено внимание сети в процессе обучения. Для сильного шума сети оказалось полезным использование бóльшего коэффициента, контролирующего вклад ошибки в области масок сегментации, чем при использовании слабого шума.

Метод может быть расширен добавлением сети-сегментатора. Тогда процесс улучшения можно сделать итеративным — после сегментирования изображение улучшается, а по улучшенному изображению строится уточнённая маска сегментации. Такой метод может быть полезен в задачах, где необходимо выполнять сегментацию на сильно зашумлённых изображениях, и при этом объекты на изображениях узкоспециализированны. Например для задачи улучшения и сегментации медицинских изображений, таких как снимки образцов крови, МРТ или УЗИ.

СПИСОК ИСТОЧНИКОВ

1. Zhang K., Zuo W., Chen Y., Meng D., Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising // IEEE Transactions on Image Processing. 2017. V. 26. P. 3142–3155.
2. Mao X.J., Shen C., Yang Y.-B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections // Advances in Neural Information Processing Systems. 2016. P. 2802–2810.
3. Anwar S., Barnes N. Real image denoising with feature attention // IEEE International Conference on Computer Vision. 2019. P. 1–10.
4. Remez T., Litany O., Giryes R. Deep class aware denoising // International conference on sampling theory and application. 2017. P. 138–142.
5. Lucas A., Lopez-Tapia S., Molina R. et al. Generative adversarial networks and perceptual losses for video super-resolution // IEEE Transactions on Image Processing. 2019. P. 3312–3327.
6. Titarenko M.A., Malashin R.O. Image enhancemen by deep neural networks using high-level information // Journal of Optical Technology. 2020. V. 87. № 10. P. 604–610.
7. Jin L., Zhang W., Ma G., Song E. Learning deep CNNs for impulse noise removal in images // Journal of Visual Communication and Image Representation. 2019. V. 62. № 7. P. 193–205.
8. Zhou S., Hu Y., Jiang H. Multi-view image denoising using convolutional neural network // Sensors (Basel). 2019. № 6. P. 1–24.
9. Zhang K., Li Y., Zuo W., Zhang L., Gool L.V., Timofte R. Plug-and-play image restoration with deep denoiser prior // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. 14 June. P. 1–17.
10. Heinrichheinrich M., Stille M., Buzug T. Residual U-Net convolutional neural network architecture for low-dose CT denoising // Current directions in biomedical engineering. 2018. P. 297–300.
11. Reymann M., Würfl T., Ritt P., Stimpel B., Cachovan M., Vija A.H., Maier A. U-Net for SPECT image denoising // IEEE Nuclear Science Symposium and Medical Imaging Conference. 2019. P. 1–2.
12. Crespo J.M., Moreno1 V., Juan Ramon Rabunal, Pazos A., Carbia M.C. Fringe pattern denoising using U-Net based neural network // EPJ Web of Conferences. 2020 P. 1–2.
13. Feng Z., Li Z., Cai A., Li L., Yan B., Tong L. A preliminary study on projection denoising for low-dose CT imaging using modified dual-domain U-net // 3rd International Conference on Artificial Intelligence and Big Data. Chengdu, China. 28-31 May 2020. P. 1–4.

14. Lin T., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., Zitnick L. Microsoft COCO: Common Objects in Context // European Conference on Computer Vision. Zurich, Switzerland. 6–12 September 2014. P. 740–755.
15. Kingma D., Ba J. Adam: A method for stochastic optimization // URL: https://arxiv.org/pdf/1412.6980.pdf (accessed: 2.11.2021).

АВТОРЫ

Титаренко Михаил Алексеевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; старший лаборант, Институт физиологии им. И.П. Павлова РАН, Санкт-Петербург, Российская Федерация; Scopus ID: 57225010914; https//orcid.org/0000-0001-5978-4389; mik199997@gmail.com

Роман Олегович Малашин — кандидат технических наук, доцент, Государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация; старший научный сотрудник, Институт физиологии им. И.П. Павлова РАН, 199034, Санкт-Петербург, Российская Федерация; Scopus ID: 55789083900; https://orcid.org/0000-0002-2493-839X; malashinroman@mail.ru

AUTHORS

Mikhail A. Titarenko — Postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation; Major assistant, Pavlov Institute of Physiology, Saint Petersburg, 199034, Russian Federation; Scopus ID: 57225010914; https//orcid.org/0000-0001-5978-4389; mik199997@gmail.com

Roman O. Malashin – Ph.D, associate professor, Saint-Petersburg State University of Aerospace Instrumentation, Saint Petersburg, 190000, Russian Federation; Senior researcher, Pavlov Institute of Physiology, Saint Petersburg, 199034, Russian Federation; Scopus ID: 55789083900; https://orcid.org/0000-0002-2493-839X; malashinroman@mail.ru

Статья поступила в редакцию 07.10.2021, одобрена после рецензирования 04.11.2021, принята к печати 09.12.2021