

# Комплекс прикладных решений по построению и обучению искусственных нейронных сетей для семантической сегментации аэрокосмических изображений произвольной канально-спектральной структуры в условиях дефицита обучающих данных

***Гвоздев О.Г. (1,2), Рихтер А.А. (1), Мурынин А.Б. (1,3)***

(1) НИИ "АЭРОКОСМОС", Москва, Российская Федерация

(2) МИИГАиК, Москва, Российская Федерация

(3) Федеральный исследовательский центр "Информатика и управление" РАН, Москва, Российская Федерация

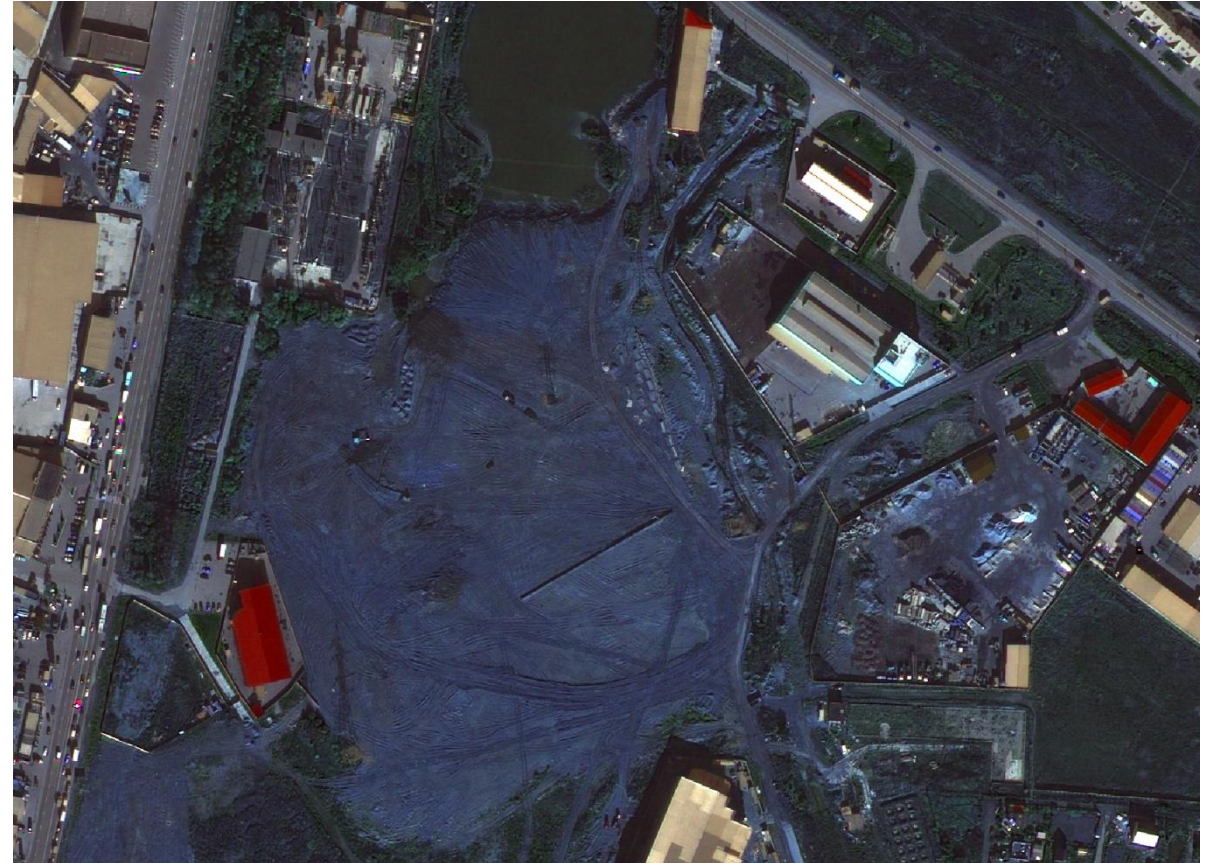
Докладчик – Рихтер Андрей Александрович

Исследования проведены в рамках проекта № RFMEFI58317X0061

# Описание проблемы



Открытый мусорный покров, Калужская область, Боровский район, Курьяново [фотографии, весна, 2019]



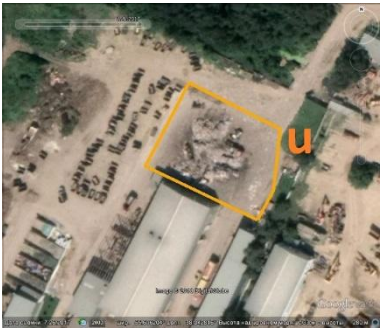
Открытый грунт, поселение Щербинка, Обводная дорога, НАО г. Москва [фрагмент космического снимка, GeoEye, лето, 2014]



# Проблема семантической сегментации аэрокосмических изображений

Пример верификации открытого замусоривания (свалка *u* в окрестности мусорного полигона Торбеево):  
а) фотосъёмка территории [июль 2017 г.]  
б) фотография свалки [июнь 2017 г.]  
в) спутниковый снимок (картографические данные Google, аккаунт Serge sp)

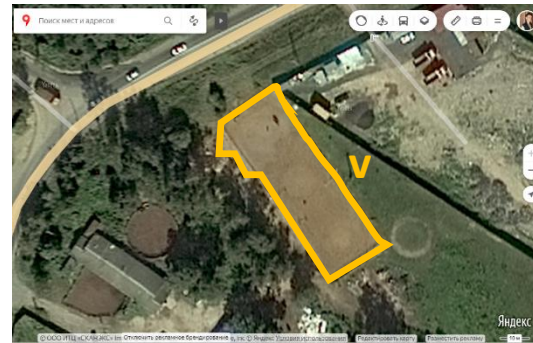
*Окрестность дер. Торбеево, Люберецкий район, Московский регион*



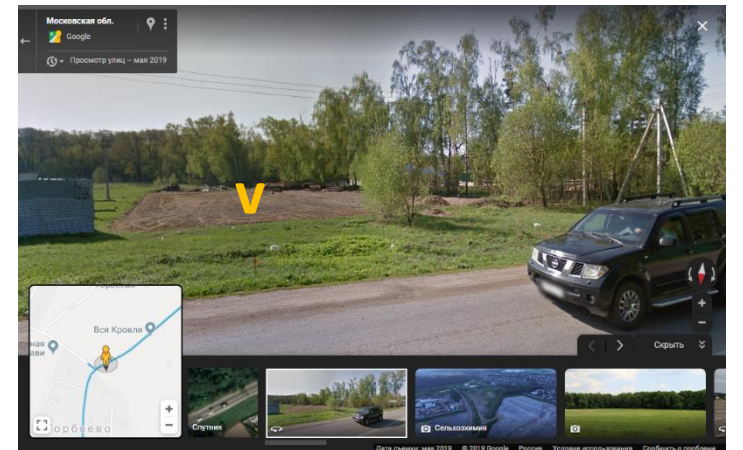
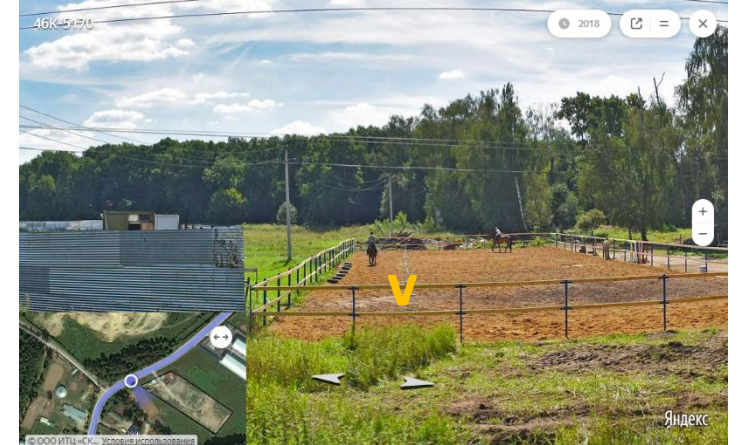
## Область открытого грунта (2)

(1)

(а)



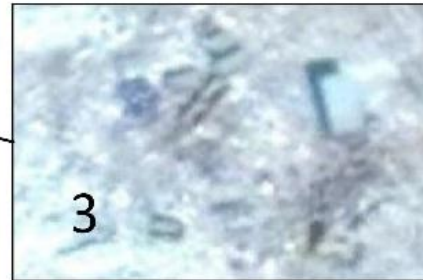
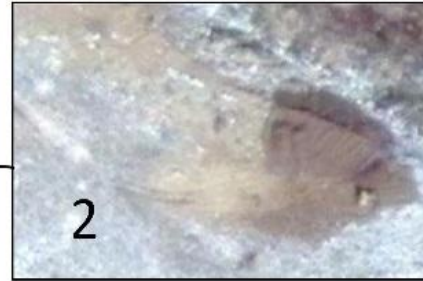
(б)



Пример космической (1) и панорамной (2) съёмки открытого грунта (ипподромное поле *v*) в Яндекс- [2018 г.] (а) и Google- [2019 г.] (б) картах



# Основные трудности при сегментации на примере мусорного полигона



1. Линейные (мусорные дороги и др.) объекты фона на свалке
2. Нечёткость области свалки на внутренних и внешних границах (с кавальерами, дорогами и т.д.)
3. Площадные (мусорная техника и др.) объекты фона на области свалки
4. Части области свалки на фоне (свалки рассредоточенной формы)

*Полигон твёрдых коммунальных отходов (ТКО) Воловичи, Коломенский район, Московский регион*



# Примеры изображений спорных областей

*Спорные с открытым мусорным покровом*



Разрушенное здание



Деревья



Строительные материалы



Строящиеся объекты



Скрытые (затенённые) участки



Склад

*Спорные с открытым грунтом*



Земляное полотно

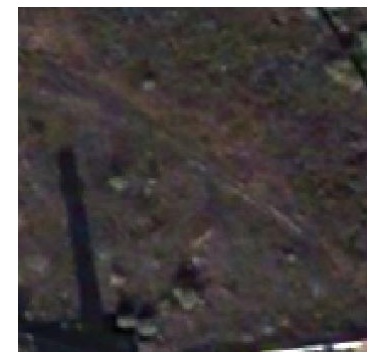


Скопления грунта, строительные материалы



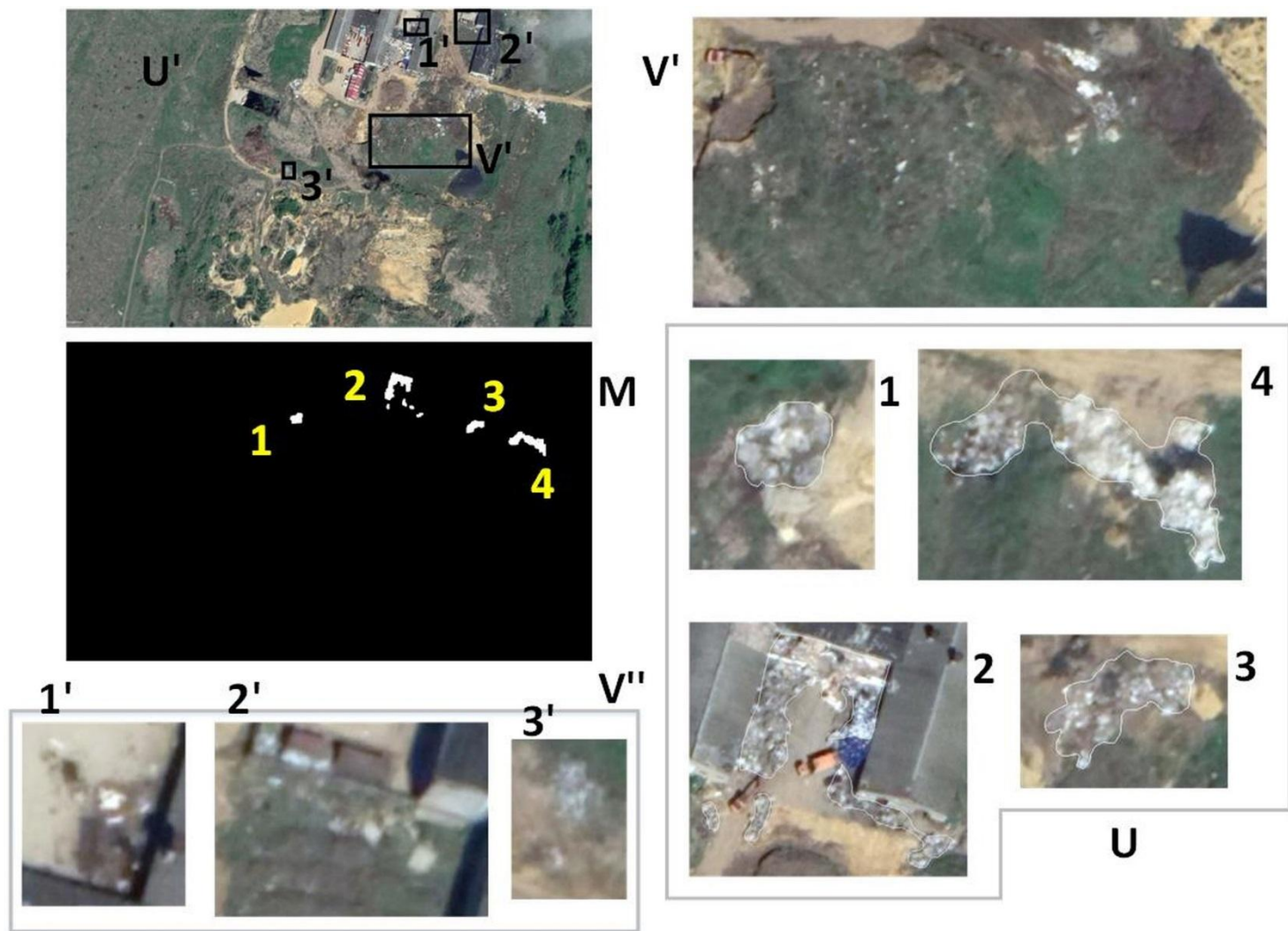
Дорожное полотно

Бетонное, асфальтированное покрытие



Редкая растительность

# Формирование обучающей и тестовой выборок (на примере открытого мусорного покрова)



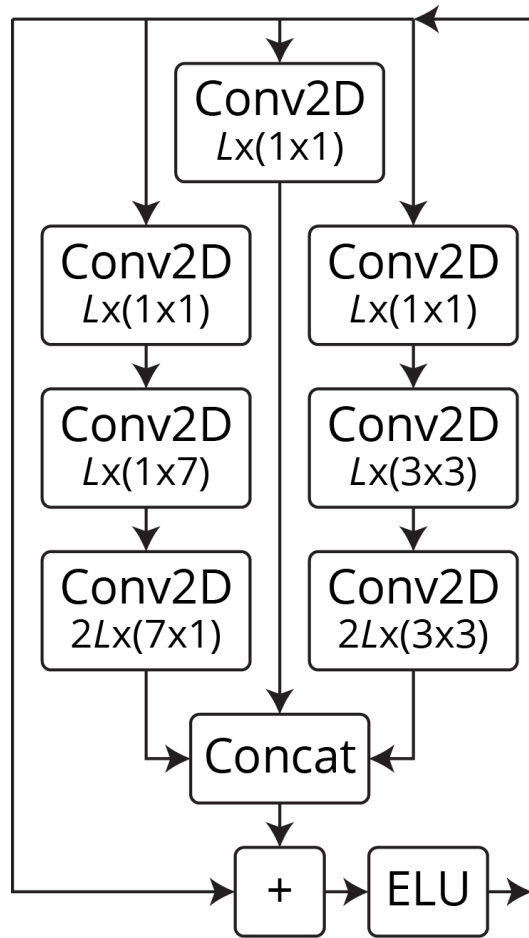
$n \sim 100$  - число снимков и соответствующих масок  
 $m \sim 5$  - среднее число непрерывных областей открытого мусорного покрова (ОМ)  
 $q \sim 100$  - число аугментированных изображений в окрестности каждой области ОМ  
 $N = mnq \sim 50000$  - размер обучающей выборки

$U$  — разметка  
 $U'$  — фон  
 $V = V' + V''$  — вычитаемая область (вводится ввиду сложности в разметке и неопределённости идентификации объектов)  
1–4 — размеченные области ОМ  
1'–3' — примеры вычитаемых областей  
 $M$  — маска, отделяющая размеченные объекты от фона и вырезанных участков

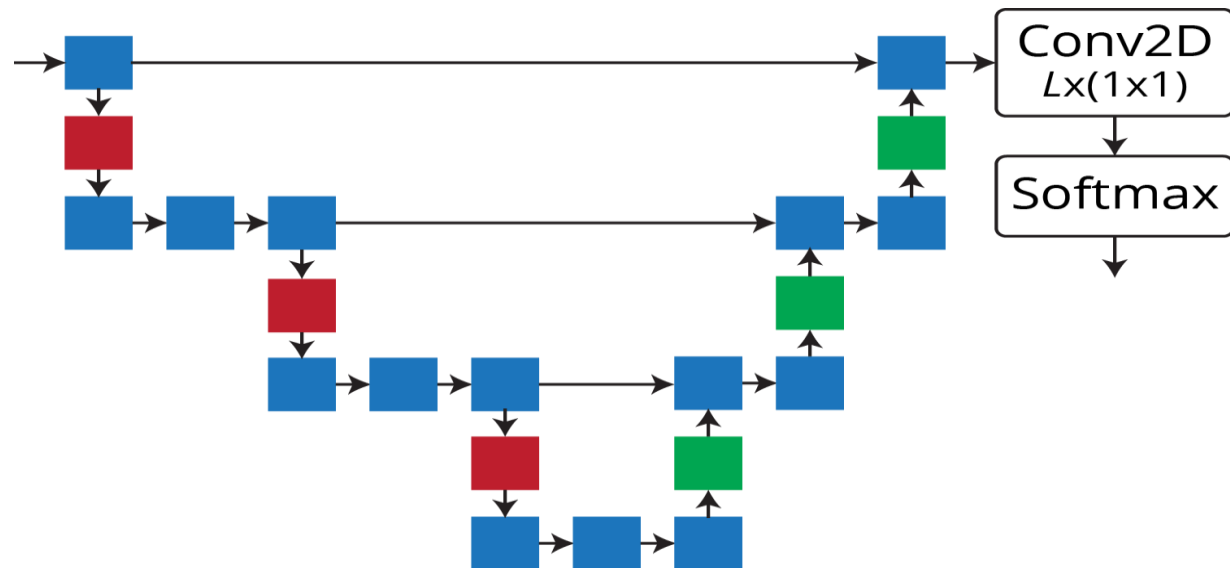
Построение обучающей выборки на примере изображения (Михайловская Слобода, Раменский район, Московский регион)



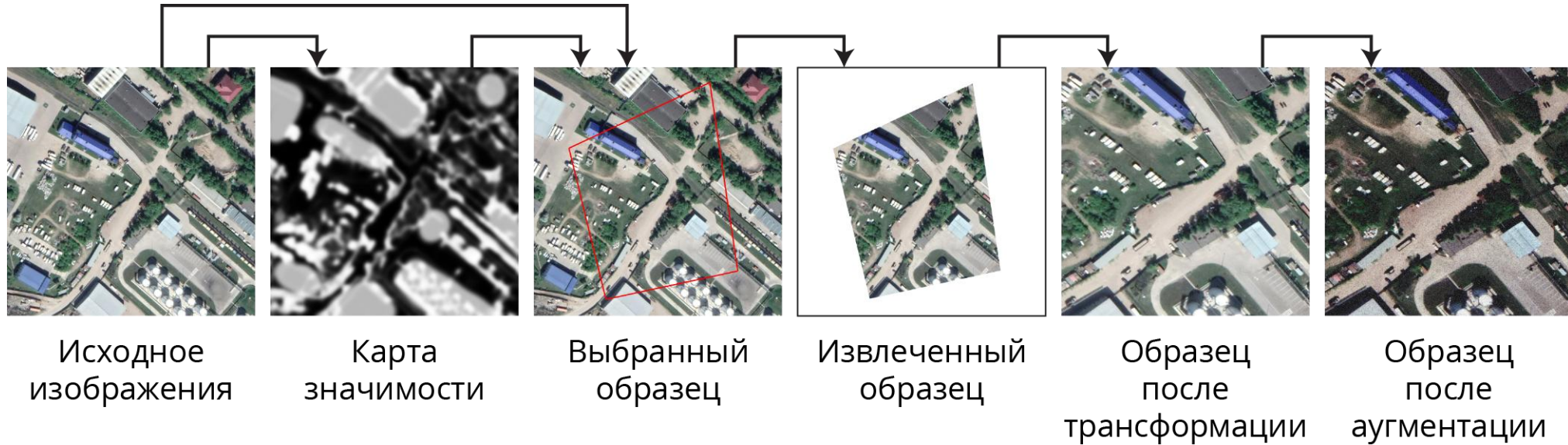
# Искусственная нейронная сеть



- Частный случай **семантической сегментации**
- Базовая топология – **U-Net**
- Базовый блок – на основе **Inception-ResNet**
- Активационная функция – **ELU**
- **Squeeze-and-Excitation**  
для учета глобальных особенностей снимка



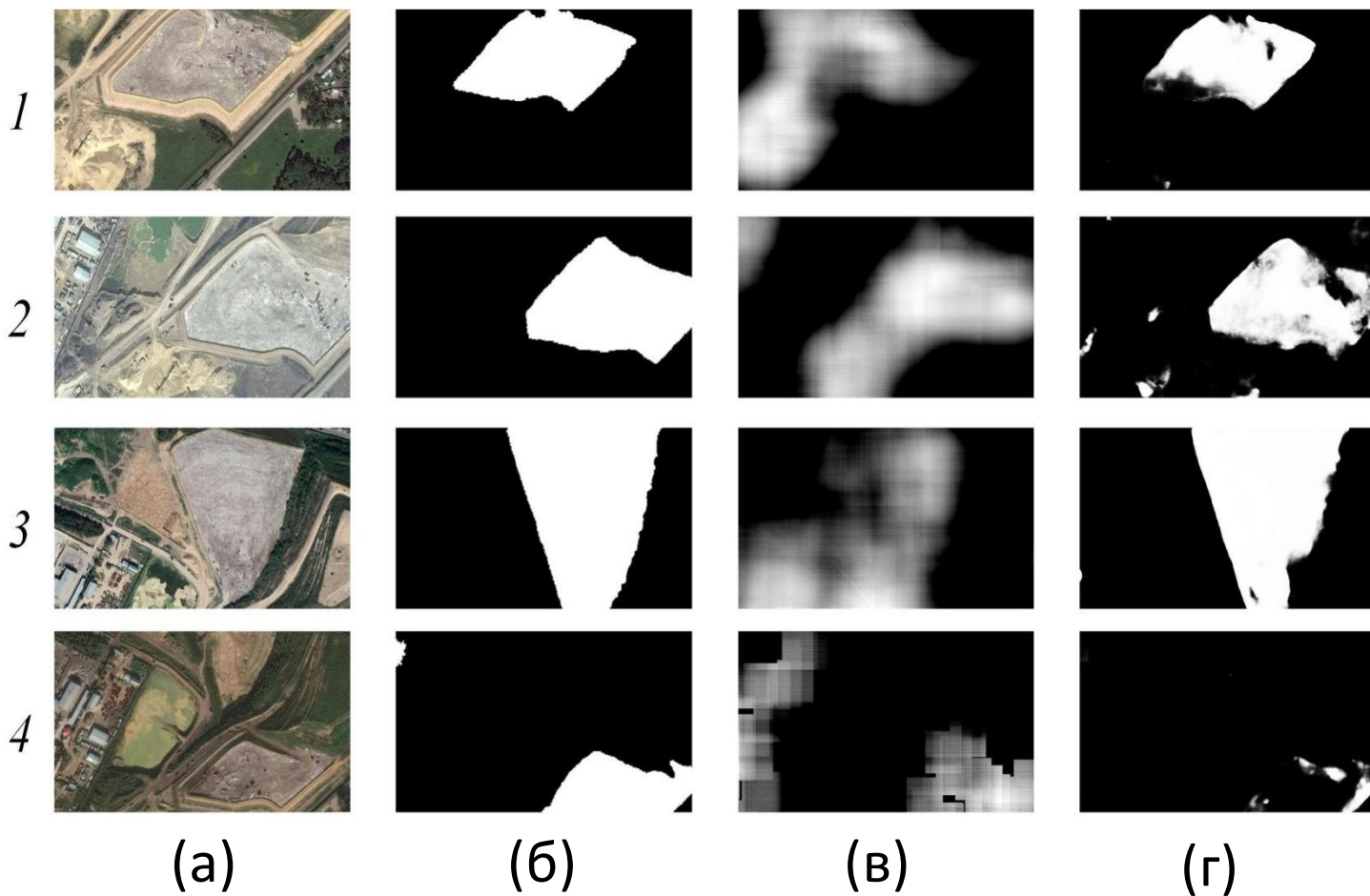
# Подготовка данных для обучения



- **Карта значимости** – функция веса класса, локальной энтропии изображения и дополнительных пометок эксперта
- Случайная генерация множества **образцов** – полигонов на снимке
- **Оптимизационная задача** – поиск подмножества образцов, оптимальным образом покрывающих снимок пропорционально **карте значимости**
- Извлечение и аугментация каждого образца по-отдельности



# Результаты обучения



Результаты бинарной классификации для классического метода (фрактальный анализ) и с использованием ИНС

Параметр	Классический метод, ~%	Применение ИНС, %
TP (True Positive)	60	80.4
TN (True Negative)	90	87.9
FP (False Positive)	20	12.1
FN (False Negative)	10	19.5

1	0.4084 – 2.0486
2	1.9242 – 2.1091
3	1.6479 – 2.0122
4	0.6449 – 1.2783

Значения фрактальной размерности  $f$   
9

Детектирование открытого мусорного покрова на участках (1–4) мусорного полигона Торбеево:

а) входные изображения; б) эталонные разметки;

в) детектирование классическими методами (на базе оценки фрактальной размерности); г) детектирование с применением ИНС

# Пример сегментации на тестовой выборке



Исходное  
изображение

Маска,  
полученная  
экспертно



Работа  
нейронной сети

Значения метрик качества, выбранных для предложенного метода (вероятностный метод)

Название	Значение
Качество алгоритма (Accuracy)	0.92
Точность (Precision)	0.66
Сходство (Jaccard)	0.38



## Примеры хорошей и ошибочной сегментации на тестовой выборке



Сегментация областей открытого мусорного покрова с посторонними объектами (участок мусорного полигона )



Ошибочная сегментация (антропогенный участок другого типа)

# Обсуждение результатов

В качестве каркаса топологии ИНС семантической сегментации используется U-Net:

- Базовые блоки кодировщика и декодировщика, в котором заменяются на ResNet-подобный (выше скорость обучения) или Inception-подобный (выше точность предсказания);
- Активационная функция заменяется на ELU (для преодоления эффекта «умирания» нейронов);
- В отдельных случаях (например, при существенных колебаниях характера освещённости в выборке) к блокам кодировщика применяется техника Squeeze-Excitation;
- Для особенно малых обучающих наборов, для регуляризации применяется техника Dropout;
- При обучении используется плавное уменьшение learning rate;
- Опционально применение техник active learning и self-supervised learning;
- Последняя реализуется путём добавления в обучающую выборку разметки, полученной путём усреднения результатов многократной аугментации и распознавания каждой неразмеченной области;
- Совокупность рассмотренных техник позволяет достичь F1-меры  $>0.7$  для наборов в 20-40 исходных изображений и  $>0.8$  для наборов в 50-150 исходных изображений, при количестве свободных параметров ИНС в пределах  $15e7$ .



# Обсуждение результатов

- При распознавании свалок с применением ИНС описанной архитектуры отклонения от эталонной разметки объясняются, очевидно, качеством обучающей выборки. Несмотря на то что показатели ИНС (TR – 80.4%, TN – 87.9%, FP – 12.1%, FN – 19.5%) существенно превосходят показатели классических методов, проведённые эксперименты обнаружили ряд проблем, определивших направления для дальнейших исследований.
- Свалки не склонны иметь регулярные внутренние структуры, которые хорошо распознаются ИНС, в то же время, их цвет зависит от освещения. С одной стороны, организованным свалкам свойственно иметь в своём составе объекты, в равной степени свойственные фону: сухой грунт (грунтовые дороги), бетонные покрытия, технику, ограждения и строения. Свалки имеют существенный разброс по размерам и могут занимать область от нескольких десятков пикселей до 80% и более ширины снимка. С другой стороны, используемые процедуры обучения ИНС и оценки точности её работы опираются на распознавание отдельных пикселей изображения, а не площадных объектов, которыми являются свалки.
- Обозначенные проблемы, в совокупности, приводят к необходимости более точного раскрытия и количественной оценки возможности ИНС для данной задачи. Их решение требует значительно большего количества обучающих образцов, а также большей детализации классов объектов в них.

# Выводы

1. Разработан комплекс прикладных решений, позволяющих проводить разметку и обучение ИНС по космической съёмке для семантической сегментации аэрокосмических изображений произвольной канално-спектральной структуры в условиях дефицита обучающих данных. Проведена подготовка данных, разметка и обучение ИНС для таких объектов, как открытый грунт и открытый мусорный покров.
2. Перспективными направлениями развития представленного технологического процесса являются:
  - предобучение ИНС с помощью метода deep cluster, обобщённого для отдельных пикселей изображения;
  - решение задачи сегментации экземпляров объектов с помощью методов семейства deep watershed;
  - исследование альтернативных базовых топологий (U-Net++, DeepLab и др.);
  - применение более сложных стратегий обучения, в частности, возврата к ранее обученным моделям, в случае выявления деградации;
  - генерация обучающих образцов в реальном времени.
3. По результатам проведённых исследований ведётся дальнейшая разработка алгоритмов интеллектуального дешифрирования свалок наряду с другими объектами земной поверхности, с использованием ИНС. Для такой разработки требуется решение следующих задач: разработка методик построения обучающих выборок для обучения создаваемых алгоритмов на основе реальных данных, получаемых с космических аппаратов и с беспилотных летательных аппаратов, а также методик, обеспечивающих оптимальное обучение ИНС для решения поставленных задач; разработка методов и алгоритмов предобработки и аугментации аэрокосмических изображений, обеспечивающих выделение выбранных информативных признаков.