

Определение расположения сельскохозяйственных полей на спутниковых снимках с помощью нейросети архитектуры U-Net

М.А. Павлова¹, А.В. Савчик¹, Л.М. Тепляков¹, М.А. Загарев²,
И.Ю. Кукоев², А.С. Григорьев¹

¹Институт проблем передачи информации им. А.А. Харкевича Российской академии наук (ИППИ РАН), Большой Каретный переулок 19 стр. 1, Москва, Россия, 127051

²ООО "Диджитал Агро", Пресненская набережная 6, стр 2, Москва, Россия, 123317

Аннотация. В данной работе рассматривается решение задачи автоматического определения расположения сельскохозяйственных полей на спутниковых изображениях на основе полноцветных оптических изображений и карт индекса вегетации NDVI со спутников Sentinel-2 с пространственным разрешением 10 м. Задача решается с применением сверточной нейронной сети с архитектурой из семейства U-Net. Исследуется постановка задачи как детекция масок полей либо их границ; одновременное (мультиклассовое) обучение эффективности не показало. Показана применимость данного подхода для автоматизации картирования земель в задачах сельскохозяйственного мониторинга.

1. Введение

В данной работе рассматривается задача автоматизации картирования земель сельскохозяйственного назначения на основе спутниковых данных 10-метрового пространственного разрешения. Данная задача актуальна как в кадастровом учете, так и в сельскохозяйственном мониторинге, поскольку высокоточное картирование полей является трудоемким процессом, а знания о границах полей и других земельных участках часто являются базовым элементом для решения других задач сельхозмониторинга, в частности, оценки различных показателей урожайности и состояния земель при использовании подходов точного земледелия.

Многие существующие работы по автоматизации картирования в сельском хозяйстве опираются на данные высокого разрешения [1, 2]. Вместе с тем, данные более низкого разрешения являются более широко доступными и регулярно обновляются благодаря таким исследовательским программам, как Sentinel-2 [3], предоставляющей регулярно обновляемые (порядка 2 раз в неделю) данные мультиспектральной спутниковой съемки в разрешениях от 10 до 60 метров на пиксель в зависимости от спектрального канала. Выявление спектрального интервала и/или их комбинаций, наиболее подходящих для картирования, является нетривиальной задачей. В данной работе используется тот факт, что для решения задач сельскохозяйственного мониторинга давно известны и широко используются так называемые индексы вегетации - изображения, рассчитываемые по изображениям в различных диапазонах длин волн [4, 5].

В данной работе, помимо изображения в видимом диапазоне, исследуется использование индекса NDVI (Normalized Difference Vegetation Index), нормализованного относительного индекса вегетации, в качестве входных данных для автоматизации картирования. NDVI вычисляется по соотношению наблюдаемых интенсивностей красного (RED) и ближнего инфракрасного (NIR) каналов:

$$NDVI = \frac{NIR - RED}{NIR + RED}.$$

В настоящее время во всевозможных задачах анализа изображений получили широкое распространение алгоритмы, основанные на обучении искусственных нейронных сетей, в особенности с применением сверточных слоев [1, 2, 6, 7]. Частный случай полносвёрточной архитектуры, показывающей хорошее качество в задачах сегментации, — семейство нейросетей U-Net [8], реализующее многомасштабный подход к анализу изображений. Также можно отметить один из универсальных алгоритмов сегментации MaskRCNN [9], в котором детекция объектов сопровождается дальнейшей попиксельной сегментацией с помощью свёрточных слоёв. Это позволяет детектировать отдельные объекты и их точные границы, в том числе, когда их образы пересекаются. Однако для случая определения масок или границ полей, когда объекты не пересекаются и нет проблемы частичного заслонения, данный более сложный метод не требуется.

2. Предлагаемый подход

Для автоматического детектирования границ полей был применен нейросетевой подход. Решение в этом случае имеет вид функции $f_w : X \rightarrow Y$. Функция f_w отображает (потенциально многоканальные) изображения $x \in X = [0, 1]^{h \times w \times c}$ в одноканальные изображения $y \in Y = [0, 1]^{h \times w}$, кодирующие поточечную оценку уверенности $y_{i,j} \in [0, 1]$ в том, что пиксель i, j на исходном изображении относится к полю или к границе поля (в зависимости от варианта постановки задачи для обучения — поиск границ или маски полей). Вид функции f_w определяется конкретным типом решаемой задачи и задает архитектуру искусственной нейронной сети.

Простым примером нейронной сети является случай последовательной нейронной сети. Такая сеть состоит из входов $x = h_0$, к которым последовательно применяются некоторые обучаемые функции, или слои $h_i f_i(h_{i-1}), i = \{1, \dots, n\}$. Выход последнего слоя является и результатом работы сети $y = h_n$. Для обработки изображений применяют нейронные сети, использующие сверточные слои. В этом случае вход и выход слоя являются изображениями, т.е. таблицами (тензорами) размерности 3

$$h_i^{x,y,ct} = ReLU \left(\sum_{dx,dy,cf,ct} w_{dx,dy,cf,ct}^i h_{i-1}^{x+dx,y+dy,cf} \right),$$

где $ReLU(x) = \max(x, 0)$. Идея таких фильтров в том, чтобы локально обработать различные части одним и тем же образом, что, в частности, позволяет подбирать одни и те значения весов w для разных частей изображения. В основном применяются свертки небольшого размера, например 3×3 , (другие через них выражаются). Для того, чтобы выход свёрточного слоя был изображением того же размера, что и вход изображения, вокруг исходного изображения рисуется рамка из нулей (padding).

Параметры w функции f_w оптимизируются таким образом, чтобы на собранном наборе данных $\{x_i, y_i\}$ с экспертно размеченным положением границ полей (и соответственно восстанавливаемыми по границам масками полей) минимизировать функцию потерь (эмпирического риска) $L(w) = \sum_i l(f_w(x_i), y_i)$, где в качестве l была выбрана перекрёстная энтропия для двух классов.

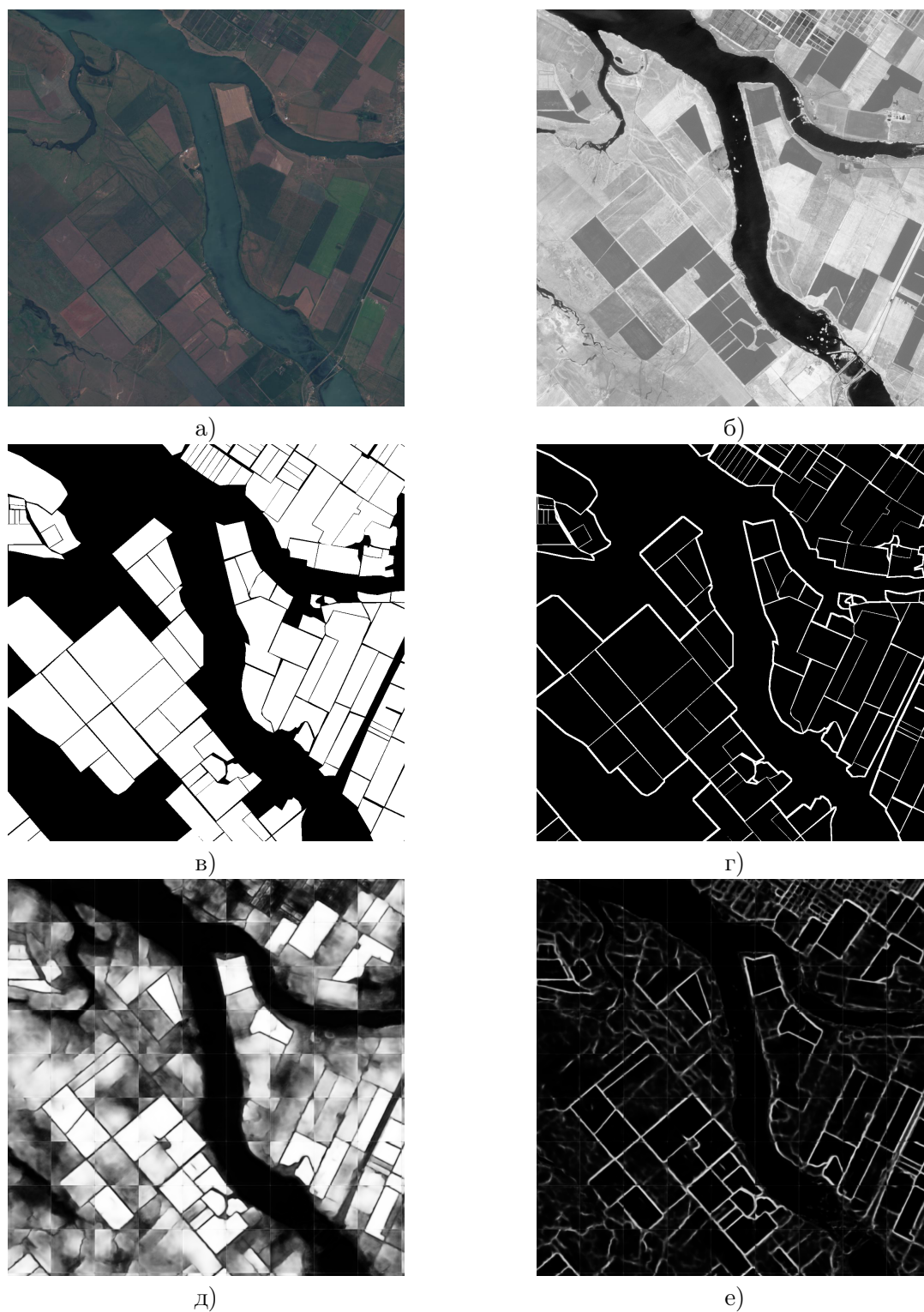


Рисунок 1. Пример изображений в наборе данных и результирующих предсказаний: а) цветные данные (TCSI), б) индекс вегетации NDVI, в) маска полей в разметке, г) маска границ в разметке, д) предсказание полей, е) предсказание границ полей.

$$l(y, \hat{y}) = \begin{cases} -\log(y) & , \hat{y} = 1 \\ -\log(1 - y) & , \hat{y} = 0. \end{cases}$$

Одна из архитектур нейронных сетей, применяемых для задач сегментации и попиксельной классификации изображений является семейство архитектур U-Net [8]. Нейронные сети такого вида не являются строго последовательными, а состоят из нескольких слоёв разного уровня детальности (масштаба обработки). Преимуществом архитектур данного семейства является достаточно большое рецептивное поле (область внимания), позволяющая учитывать для каждого пикселя результата относительно широкую пространственную окрестность входного изображения, большая выразительная способность с большим количеством весов и, одновременно, сравнительно малое время работы за счет многомасштабного подхода. Сеть такой архитектуры работает с малым количеством весов на изображении исходного разрешения, затем с большим количеством весов на изображении сильно меньшего разрешения и, в конце, объединяет полученные на низком разрешении результаты с данными высокого разрешения от исходной картинкой для более точного попиксельного предсказания.

В данной работе была использована нейронная сеть с 32 фильтрами на первом слое и наименьшим масштабом в сети равным $1/8$ (см. рисунок 2).

2.1. Методика обучения

Был подготовлен набор данных, состоящий из 122 4-канальных изображений, соответствующих 22 участкам поверхности Земли за период 05.04.2018–11.11.2018 из архива снимков спутников Sentinel-2. Размер каждого изображения составляет 1030×1030 пикселей, разрешение съемки 10 м/пиксель. Первые 3 канала каждого изображения в наборе соответствуют 3-канальным цветным изображениям (*TCI* в номенклатуре Sentinel-2), а в качестве четвертого канала используется карта индекса вегетации *NDVI* (вычисляется по полосам 4 и 8 исходного мультиспектрального изображения), см. рисунок 1аб.

На каждом полноцветном изображении были вручную размечены поля и схожие с ними структуры в виде полигональных контуров, в результате на всех изображениях было отмечено по 400 объектов в среднем. Маска границ полей была построена следующим образом: маска полей была морфологически дилатирована с квадратным окном 10×10 , после чего точки, попавшие в дилатированную маску, но не в исходную, считались точками границы (см рисунок 1вг).

Набор данных был разбит на обучающую и тестовую части, состоящую из 17 участков поверхности с 94 снимками и 5 участков с 28 снимками соответственно.

Сеть обучалась в течение 25 эпох, состоящих из 500 батчей. Каждый батч представлял собой 32 случайных фрагмента размером 128×128 , вырезанных случайным образом из исходных изображений. Оптимизация проводилась с помощью метода Adam [10] — одной из разновидностей методов стохастического градиентного спуска. В качестве функции потерь использовалась кросс-энтропия с l_2 — регуляризацией с коэффициентом 0.0001 для предотвращения переобучения.

3. Результаты

Для оценки качества классификатора применялись попиксельные метрики. Мы использовали метрику AUC-ROC, считающую площадь под графиком в координатах FPR (доля пикселей фона, получивших оценку классификатора ниже некоторого порога) и TPR (аналогично для пикселей объекта). Однако эта метрика может быть малоинформативна в случаях преобладания одного из классов. Так происходит и в нашем случае, когда количество

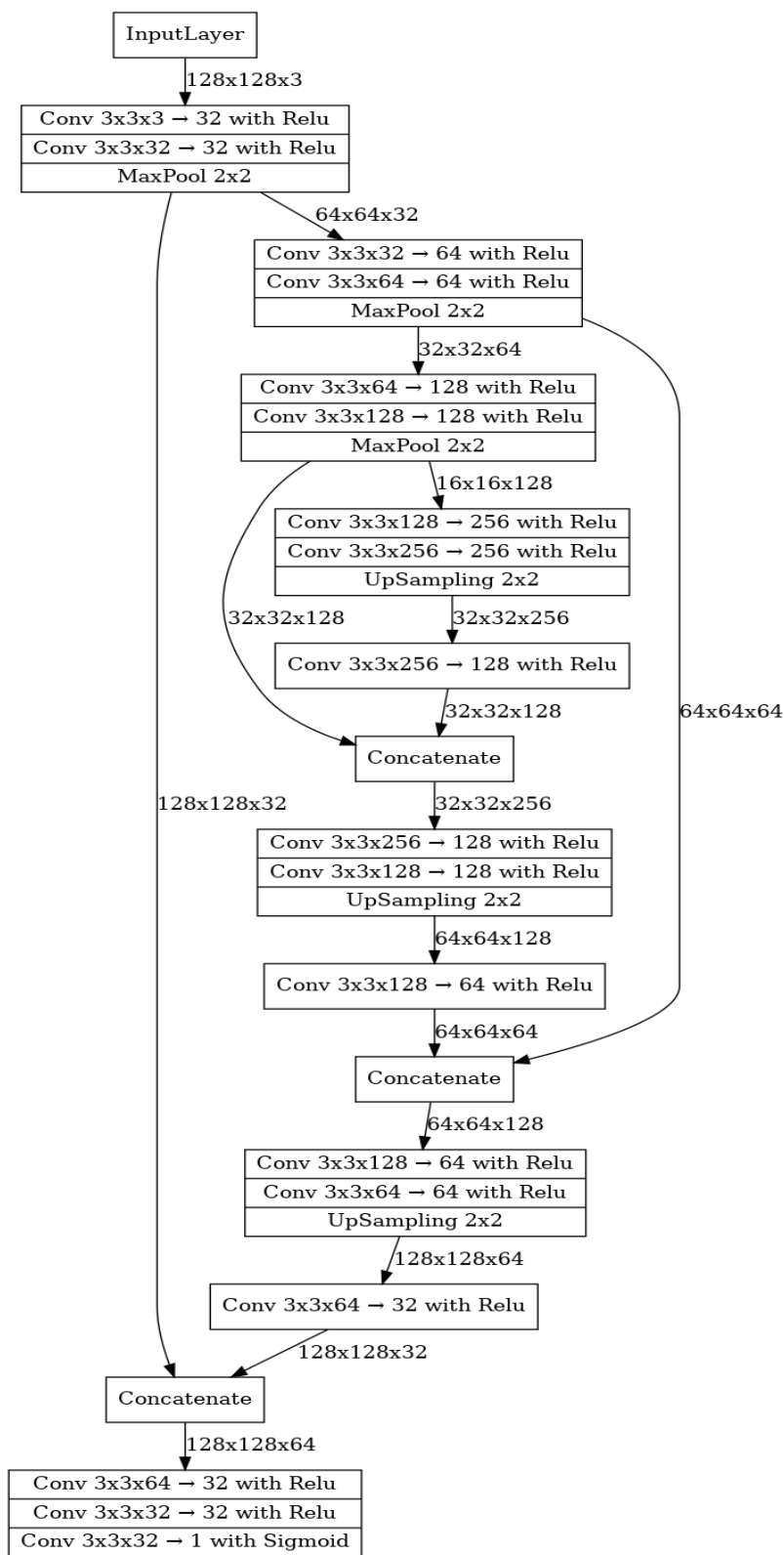


Рисунок 2. Архитектура используемой сети.

пикселей фона (не-полей или, особенно, не-границ) значительно меньше, чем количество пикселей объекта (полей или границ). Для уменьшения эффекта неравномерных классов мы также применяли метрику AUC-PR, построенную аналогичным образом по значениям точности (precision) и полноты (recall), посчитанным при разных значениях порога классификатора.

Было выполнено несколько вычислительных экспериментов. Отдельно исследовалась зависимость от входных каналов: TCI, NDVI, TCI+NDVI. Кроме того, был проведён эксперимент по применению мультиклассового обучения, когда нейронная сеть обучается одновременно предсказывать и поля, и границы. В каждом случае эксперимент повторялся трижды для более точной оценки качества и для оценки дисперсии полученных значений, см. таблицу.

Таблица 1. Сравнение разных методов обучения сети по метрикам AUC-ROC и AUC-PR. Среднеквадратичное отклонение указано по 3-м экспериментам. Для сравнения приведено значение метрик для случайного классификатора.

Входные данные	Выходные данные	Полученный результат AUC-ROC/AUC-PR
TCI	поля	$0.82 \pm 0.01 / 0.53 \pm 0.02$
TCI	границы	$0.80 \pm 0.00 / 0.39 \pm 0.00$
TCI	поля и границы	(границы) $0.78 \pm 0.00 / 0.36 \pm 0.00$
TCI	поля и границы	(поля) $0.77 \pm 0.02 / 0.47 \pm 0.06$
NDVI	поля	$0.80 \pm 0.01 / 0.51 \pm 0.04$
NDVI	границы	$0.79 \pm 0.00 / 0.39 \pm 0.01$
NDVI	поля и границы	(границы) $0.77 \pm 0.01 / 0.34 \pm 0.02$
NDVI	поля и границы	(поля) $0.78 \pm 0.02 / 0.43 \pm 0.05$
TCI и NDVI	поля	$0.81 \pm 0.01 / 0.45 \pm 0.08$
TCI и NDVI	границы	$0.79 \pm 0.02 / 0.36 \pm 0.03$
TCI и NDVI	поля и границы	(границы) $0.78 \pm 0.01 / 0.37 \pm 0.00$
TCI и NDVI	поля и границы	(поля) $0.81 \pm 0.02 / 0.53 \pm 0.01$
	Случайная маска ответов	(границы) $0.50 \pm 0.00 / 0.06 \pm 0.00$
	Случайная маска ответов	(поля) $0.50 \pm 0.00 / 0.39 \pm 0.00$

По представленным результатам видно, что результаты заметно выше, чем у случайного алгоритма, равного 0.5 / 0.4 на полях и 0.5 / 0.06 на границах по метрикам AUC – ROC и AUC – PR. Результаты мультиклассового обучения получились такими же как и обычного в пределах погрешности. Качество детекции границ по метрике AUC-PR получилось значительно ниже, чем AUC-ROC, что связано с сильно меньшей долей пикселей у границ по сравнению с долей пикселей у полей. Результаты сети выглядят сильно скоррелированными с правильным ответом (см. рисунок 1де), что, собственно, наиболее важно в данной задаче: как сами метрики, так и на самом деле и правильный ответ были выбраны относительно произвольно, поскольку не было надежного способа проверить, есть ли именно граница поля в данной точке, или это другая визуально похожая структура. Главное, что ясно, что подобный подход применим при построении систем сельскохозяйственного мониторинга и при соответствующей оптимизации под конкретный сценарий применения и при использовании более масштабного набора данных нет препятствий к получению высоких результатов.

4. Заключение

В работе рассматривается решение задачи определения расположения сельскохозяйственных полей по мультиспектральным спутниковым снимкам с представлением ответов и обучающих данных в виде масок либо границ полей и их сочетания. Для повышения информативности входных данных помимо изображений видимого диапазона вычислялись также карты индекса вегетации NDVI.

Анализ результатов показал применимость сети U-Net для решения данной задачи со значением метрики AUC-ROC порядка 0.7 для полей и границ. Сравнение различных вариантов постановки задачи обучения показало осмысленность использование карт вегетационных индексов в данной задаче, однако при отсутствии инфракрасных изображений использование одного только изображения в видимом диапазоне показывает сравнимый результат. Мультиклассовое обучение преимуществ не показало.

В целом, полученные результаты показывают перспективность применения нейросетей архитектур типа U-Net для решения задач широкомасштабного автоматизированного сельскохозяйственного мониторинга с использованием свободно доступных спутниковых данных среднего пространственного разрешения (10 м/пкс в рассмотренном случае Sentinel-2). Дальнейшим развитием данной работы может служить построение более релевантных метрик (и, соответственно, функций потерь) для решаемых задач, более полное использование мультиспектральной информации, а также исторических снимков для повышения точности и релевантности результатов распознавания.

5. Литература

- [1] Musyoka, G.M. Automatic delineation of small holder agricultural field boundaries using fully convolutional networks / G.M. Musyoka // University of Twente, 2018.
- [2] Alemu, M.M. Automated farm field delineation and crop row detection from satellite images / M.M. Alemu // University of Twente, 2016.
- [3] Drusch, M. Sentinel-2: Esa's optical high-resolution mission for gmes operational services / M. Drusch, U. Del Bello, S. Carlier // Remote sensing of Environment. – 2012. – Vol. 120. – P. 25-36.
- [4] Pettorelli, N. Using the satellite-derived ndvi to assess ecological responses to environmental change / N. Pettorelli, J.O. Vik, A. Mysterud // Trends in Ecology & Evolution. – 2005. – Vol. 20(9). – P. 503-510.
- [5] Xue, J. Significant remote sensing vegetation indices: A review of developments and applications / J. Xue, B. Su // Journal of Sensors. – 2017. – Vol. 2017.
- [6] Davidse, J. Semi-automatic detection of field boundaries from high-resolution satellite imagery, 2015.
- [7] Garcia-Pedrero, A. A machine learning approach for agricultural parcel delineation through agglomerative segmentation / A. Garcia-Pedrero, C. Gonzalo-Martin, M. Lillo-Saavedra // International journal of remote sensing. – 2017. – Vol. 38(7). – P. 1809-1819.
- [8] Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical image computing and computer-assisted intervention / Springer, 2015. – P. 234-241.
- [9] He, K. Mask R-CNN / K. He, G. Gkioxari, P. Dollár, R. Girshick // Proceedings of the IEEE international conference on computer vision, 2017. – P. 2961-2969.
- [10] Kingma, D.P. Adam: A method for stochastic optimization / D.P Kingma, J. Ba // arXiv preprint arXiv:1412.6980, 2014.

Agricultural parcel localization on satellite images using U-Net-based neural network

M.A. Pavlova¹, A.V. Savchik¹, L.M. Teplyakov¹, M.A. Zagarev², I.Yu. Kukoev²,
A.S. Grigoryev¹

¹Kharkevich Institute for Information Transmission Problems (ИИТП RAS),
Bolshoy Karetny per. 19/1, Moscow, Russia, 127051

²Digital Agro LLC, Presnyenskaya emb. 6/2, Moscow, Russia, 123317

Abstract. This work considers the problem of automatic delineation of agricultural parcels on satellite images based on true color images and NDVI vegetation index maps from Sentinel-2 satellites (10 m ground sampling distance). The problem is solved using a U-Net-based convolutional neural network. We consider problem formulation either as parcel mask or boundary detection; the multiclass (simultaneous) training did not prove to be effective. The approach looks promising and applicable for automated land mapping for agricultural monitoring systems.