



Московский государственный университет имени М.В. Ломоносова

Факультет вычислительной математики и кибернетики

Кафедра математической физики

Загайнов Сергей Дмитриевич

**Программная реализация автоматического метода  
выделения ткани на полнослайдовых  
гистологических изображениях**

Отчёт о преддипломной практике

**Научный руководитель:**

к.ф.-м.н., н.с

А.В.Хвостиков

Москва, 2023

# Оглавление

<b>1</b>	<b>Введение</b>	<b>3</b>
<b>2</b>	<b>Обзор литературы</b>	<b>5</b>
2.1	Методы, использующие пороговое значение . . . . .	5
2.2	Нейросетевые методы . . . . .	6
<b>3</b>	<b>Цель работы</b>	<b>7</b>
<b>4</b>	<b>Описание используемых наборов данных</b>	<b>8</b>
<b>5</b>	<b>Метод выделения ткани на полнослайдовых изображениях</b>	<b>9</b>
5.1	Описание нейронной сети . . . . .	9
5.2	Обучение нейронной сети . . . . .	10
5.3	Результаты обучения нейронной сети . . . . .	10
5.4	Метод классификации всех тайлов пирамиды полнослайдового изображения .	11
5.5	Визуализация работы предложенного метода . . . . .	12
5.6	Уточнение метода детекции ткани . . . . .	15
<b>6</b>	<b>Заключение</b>	<b>17</b>
6.1	Программная реализация . . . . .	17
<b>7</b>	<b>Дальнейшее развитие</b>	<b>18</b>

# 1. Введение

В последнее время научно-технический прогресс привел к тому, что во многие сферы нашей жизни произошло внедрение информационных технологий. Патология не стала исключением. Появились сканеры, позволяющие врачам-гистологам получать полнослайдовые изображения (англ. Whole Slide Image) тканей в очень высоком разрешении. Разрешение подобных изображений составляет порядка 100 тыс.×100 тыс. пикселей. Объем каждого из таких изображений составляет несколько гигабайт. Пример уменьшенного в 40 раз полнослайдового изображения можно увидеть на Рис. 1.

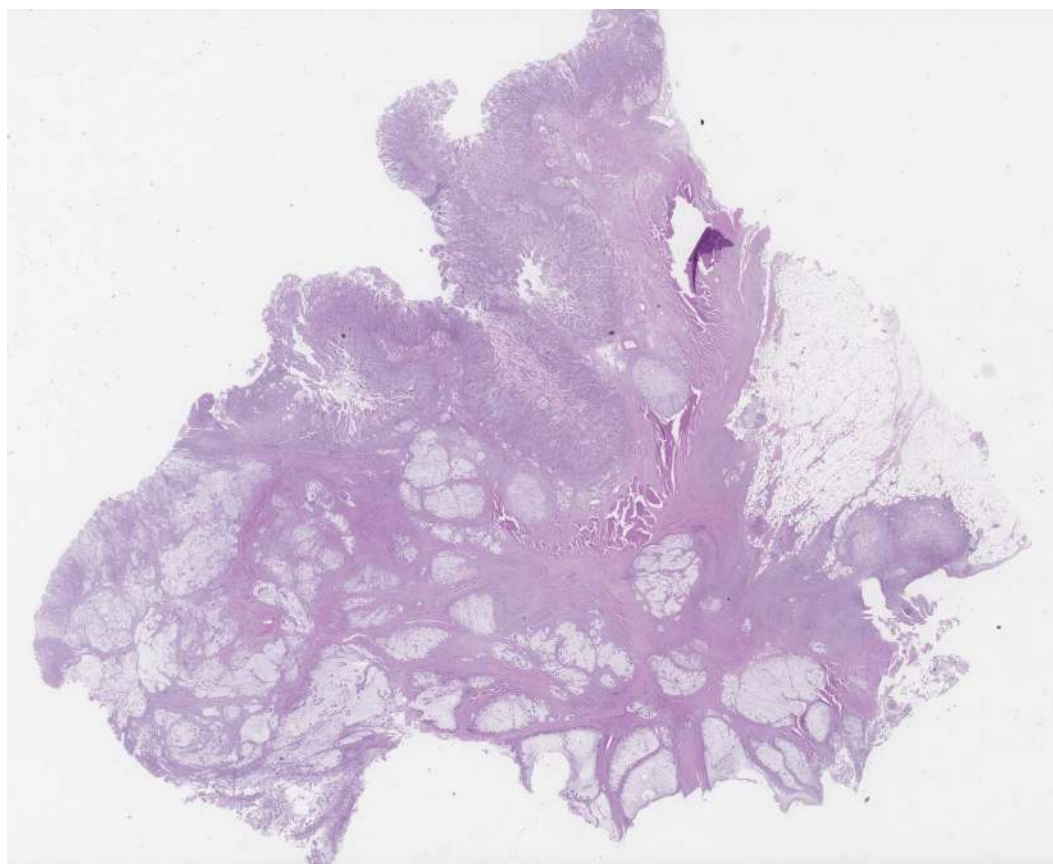


Рис. 1: Уменьшенное в 40 раз полнослайдовое изображение. Изначальное разрешение 90473×111552 пикселей. Пример из набора данных PATH-DT-MSU WSS2.

Из-за огромных размеров полнослайдовых изображений необходим особый способ хранения и представления таких изображений. Один из вариантов это пирамида изображений. Такой способ представления полнослайдовых изображений используется в программном обеспечении PathScribe[1] [2]. Пирамида изображений - это своего рода многомасштабное представление изображения. Она представляет собой серию коллекций изображений, расположенных в форме пирамиды с постепенно уменьшающимся разрешением и полученных из одного и того же исходного изображения. Это достигается поэтапной понижающей

дискретизацией, и дискретизация не останавливается, пока не будет достигнуто определенное условие завершения (разрешение изображения  $\leq$  тайлу). Чем выше уровень, тем меньше изображение, а именно его разрешение понижается в два раза на каждом следующем уровне. Каждый уровень пирамиды состоит из маленьких фрагментов, называемых тайлами (англ. tile), в нашем случае размер тайла  $512 \times 512$  пикселей или, если тайл является граничным, то его размер может быть меньше  $512 \times 512$  пикселей. Схематичный вид пирамиды представлен на Рис. 2.

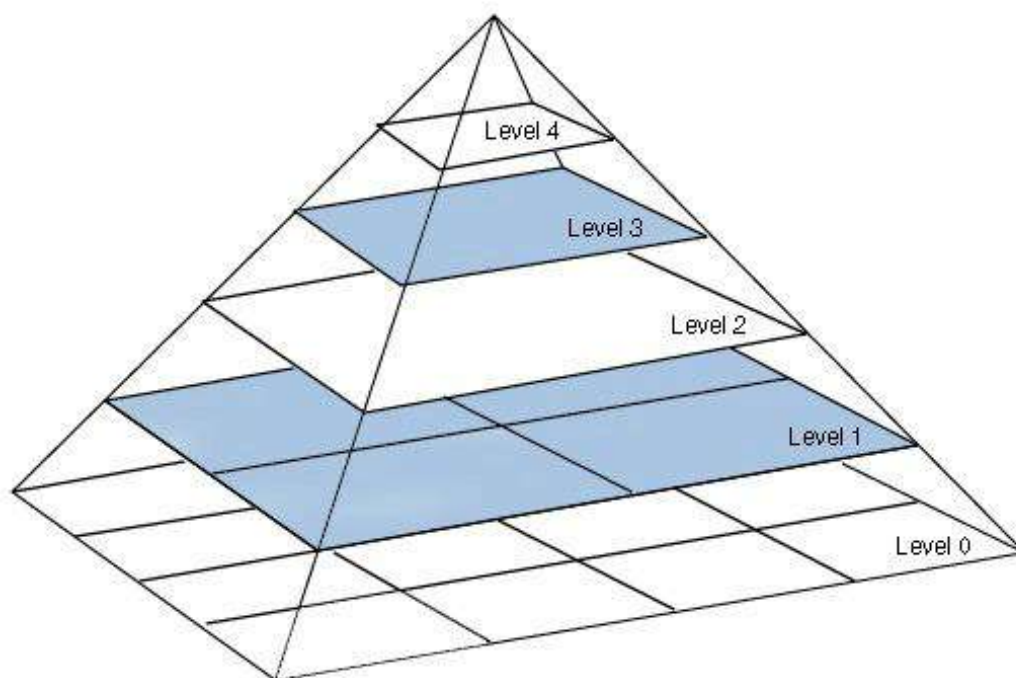


Рис. 2: Схематичный вид пирамиды полнослайдового гистологического изображения.

## 2. Обзор литературы

Необходимым шагом для применения автоматических методов анализа гистологических изображений является извлечение ткани. Следовательно, удаление фона является важным этапом предобработки для получения эффективных и точных результатов для многих алгоритмов. Идентификация участков ткани на полнослайдовых изображениях может быть сложной задачей, главным образом из-за наличия цветовых вариаций и артефактов. Более того, некоторые случаи, такие как альвеолярные ткани, жировые ткани и ткани с плохим окрашиванием, трудно обнаружить.

### 2.1. Методы, использующие пороговое значение

Обычно в обработке изображений задача отделения фона решается методами, использующими пороговое значение (англ. threshold-based methods), например, методом Оцу [3]. Одна из наиболее известных и широко используемых библиотек с открытым исходным кодом для гистопатологии, Histomics Toolkit (HistomicsTK)<sup>1</sup> также может выполнять обнаружение тканей на полнослайдовых гистологических изображениях. Процесс обнаружения содержит серию сглаживаний по Гауссу и выбор порогового значения методом Оцу. Кроме того, дополнительный порог используется для фильтрации областей, размер которых меньше заданного.

В статье [4] описывается не нейросетевой метод отделения ткани от фона, основанный на локальной энтропии (англ. local entropy). Однородные пиксели в локальных соседних областях (прямоугольники со стороной 10 пикселей) вносят меньший вклад в локальную энтропию, тогда как гетерогенные пиксели в локальных соседних областях вносят больший вклад в локальную энтропию. Кроме того, фон изображения имеет меньшую энтропию из-за более высокой яркости и более однородной текстуры, в то время как ткань имеет более высокую энтропию из-за разнообразия интенсивности пикселей в локальной окрестности. Другими словами, гетерогенные пиксели в ткани с высокой энтропией более информативны, чем однородные фоновые пиксели с низкой энтропией. После расчета локальных энтропий, строится гистограмма всех локальных энтропий. Поскольку фон изображения имеет более низкое значение энтропии, сначала определяется порог как локальный минимум гистограммы карты энтропии. Затем этот порог применяется для преобразования локальной карты энтропии в бинарную маску, которая отличает фоновый шум от ткани.

---

<sup>1</sup> <https://github.com/DigitalSlideArchive/HistomicsTK>

## 2.2. Нейросетевые методы

Обычно нейронные сети обучают на патчах (небольших квадратных фрагментах) с максимально доступным увеличением полнослайдового изображения, потому что в них нет потерь информации из-за сжатия и ресемплирования. В результате для обработки всего полнослайдового изображения в нейронную сеть необходимо отправить большое количество небольших патчей, что является трудоемкой задачей.

Авторы статьи [5] выделили четыре различных признака, характеризующих патч: цвет, внешний вид, текстуру и пространственные особенности. Они передавали эти признаки в двухслойную нейронную сеть, чтобы классифицировать патч как ткань или фон.

В статье [6] была использована сверточная нейронная сеть U-Net [7], которая обучалась на патчах размером  $892 \times 892$  пикселей. Патчи случайным образом извлекались из 54-х полнослайдовых изображений. Для каждого патча строилась бинарная маска, которая и использовалась для обучения сверточной нейронной сети.

Авторы статьи [8] решили отойти от идеи обучения нейронной сети на патчах с максимального увеличения и работать с миниатюрами (англ. thumbnails) то есть патчем с маленького увеличения 1X или 2.5X. Преимущество работы с миниатюрами заключается в скорости вычислений. Сегментация ткани для большого увеличения может быть получена из миниатюр с помощью специальных алгоритмов для пирамидальных структур полнослайдовых изображений. Также в статье был предложен алгоритм для автоматического построения масок, которые в дальнейшем подвергаются ручной обработке, чтобы убедиться, что выбраны все участки ткани, а шум и артефакты удалены настолько, насколько это возможно. Далее эти маски использовались для обучения сверточной нейронной сети U-Net [7].

Так же существуют подходы, основанные на применении слабого обучения с учителем (англ. weakly supervised learning). Один из таких подходов описан в статье [9]. В ней решалась задача сегментации двух видов аденокарцином на полнослайдовых изображениях, а именно малодифференцированных и хорошо или умеренно дифференцированных. Эта задача была сведена к задаче классификации на малодифференцированные и хорошо или умеренно дифференцированные аденокарциномы. Для решения задачи классификации из полнослайдовых изображений выделялись области интереса (прямоугольник с тканью), обучался классификатор ResNet-50 (при обучении не использовались изображения, содержащие в себе оба вида аденокарцином). Далее с помощью Gradient-weighted class activation mapping (Grad-CAM) [10] получалась тепловая карта для интересующей нас ткани, и с помощью выбора порога решалась исходная задача.

### **3. Цель работы**

Целью данной работы является разработка метода автоматической детекции ткани на полнослайдовых гистологических изображениях. При этом сама детекция должна осуществляться как потайлово (каждый небольшой фрагмент полнослайдового изображения должен классифицироваться как фон или не фон), так и попиксельно (в расширенной версии метода).

## 4. Описание используемых наборов данных

В данной работе использовался набор изображений PATH-DT-MSU<sup>2</sup>, включающий два поднабора полнослайдовых гистологических изображений (WSS1, WSS2).

PATH-DT-MSU WSS1 это 10 полнослайдовых изображений полипов толстой кишки, разбитых на обучающую и тестовую выборки. Было использовано 5 изображений для обучения нейронной сети и 5 для тестирования. Примеры изображений из PATH-DT-MSU WSS1 приведены на Рис. 3.

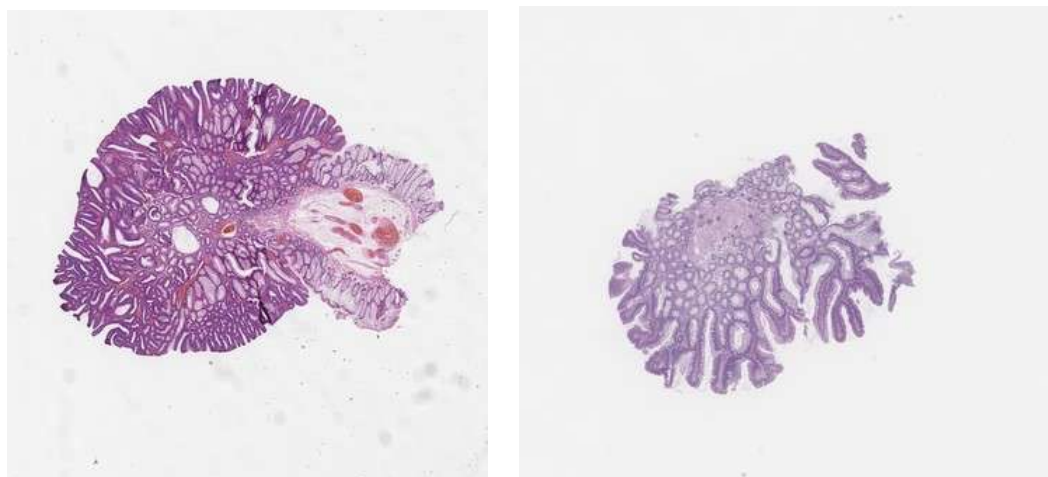


Рис. 3: Примеры изображений из PATH-DT-MSU WSS1 с верхнего уровня пирамиды.

PATH-DT-MSU WSS2 это 10 полнослайдовых изображений рака желудка, разбитых на обучающую и тестовую выборки. Было использовано 5 изображений для обучения нейронной сети и 5 для тестирования. Примеры изображений из PATH-DT-MSU WSS2 приведены на Рис. 4.

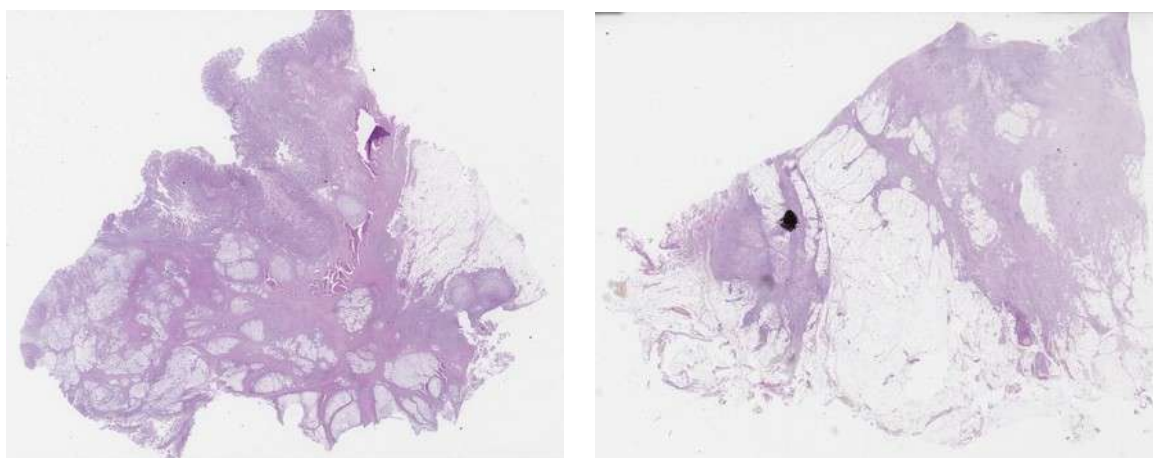


Рис. 4: Примеры изображений из PATH-DT-MSU WSS2 с верхнего уровня пирамиды.

<sup>2</sup><https://imaging.cs.msu.ru/en/research/histology/path-dt-msu>



## 5. Метод выделения ткани на полнослайдовых изображениях

Для решения поставленной задачи был выбран нейросетевой подход, основанный на обучении сверточной нейронной сети для классификации тайлов на нижнем уровне пирамиды полнослайдового изображения.

### 5.1. Описание нейронной сети

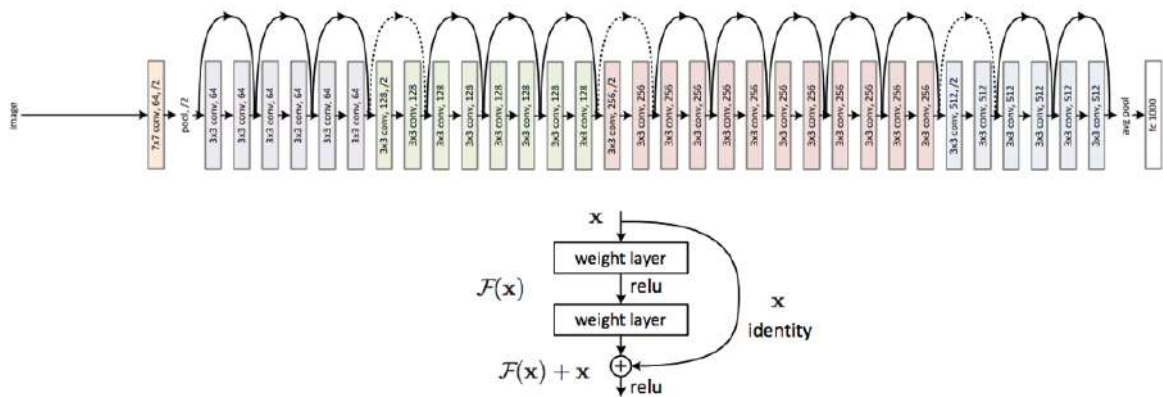


Рис. 5: Архитектура нейронной сети ResNet и остаточного блока (англ. residual block). Каждый прямоугольник соответствует многоканальной карте признаков, к которой применяются различные операции. Пунктирные линии разделяют последовательности из остаточных блоков.

В работе была обучена ResNet 32 [11] - сверточная нейронная сеть для классификации изображений на классы фона и ткани. На Рис. 5 приведена архитектура этой нейронной сети. Изначально к входному изображению применяется свертка размером  $3 \times 3$  со сдвигом (англ. stride) равным 2, функция нелинейной активации и слой субдискретизации (англ. pooling), понижающий пространственное разрешение полученного тензора в 2 раза. Далее идут 4 последовательности из остаточных блоков, состоящих соответственно из 3, 4, 6, 3 остаточных блоков. Остаточный блок приведен на Рис. 5. Остаточный блок - это последовательность сверток, организованная таким образом, что входной тензор остаточного блока добавляется к выходному тензору, получаемому как результат последовательности сверток. Затем к результату применяется нелинейность. Принцип работы остаточного блока можно описать формулой:

$$F(x) + x, \quad (1)$$

где  $\mathcal{F}$  - это результат последовательности сверток,  $x$  - это входной тензор.

Остаточные блоки позволяют решить проблему затухающего градиента в глубоких нейронных сетях. После последовательностей из остаточных блоков идет слой адаптивной субдискретизации с функцией усреднения (англ. adaptive average pooling), выходом которого является тензор с пространственной размерностью  $1 \times 1$ , и в конце к полученному тензору применяется полносвязный слой с числом нейронов, равным количеству классов (в нашем случае 2).

## 5.2. Обучение нейронной сети

Для получения обучающего и тестового набора данных полнослайдовые изображения из PATH-DT-MSU WSS1, WSS2 были разбиты на тайлы на нижнем уровне пирамиды (непересекающиеся фрагменты размера  $512 \times 512$  пикселей или, если тайл является граничным, то его размер может быть меньше  $512 \times 512$  пикселей). Далее для каждого полнослайдового изображения PATH-DT-MSU часть тайлов (примерно 10% от общего количества тайлов) было размечено вручную, а именно классифицировано как фон или ткань. Всего получилось 2262 изображений (тайлов) для обучения, 252 изображения для валидации и 978 для тестирования. Все изображения приводились к размеру  $512 \times 512$  с помощью билинейной интерполяции. Так же при обучении использовались аугментации: отражение относительно вертикальной оси (англ. HorizontalFlip). Размер обучающего пакета был выбран равным 16 (англ. batch size). Для обучения использовался оптимизатор Adam[7] со скоростью обучения (англ. learning rate)  $= 10^{-3}$  и  $L_{CCE}$  в качестве функции потерь, где  $L_{CCE}$  - это категориальная кросс-энтропия (англ. Categorical Cross-Entropy).

$$L_{CCE} = - \sum_{c=1}^M y_{o,c} \log(p_{o,c}),$$

здесь:  $M$  - количество классов;  $y_{o,c}$  - индикатор того, что метка класса  $c$  верна для наблюдения  $o$ ;  $p_{o,c}$  - предсказанная вероятность того, что наблюдение  $o$  принадлежит классу  $c$ .

## 5.3. Результаты обучения нейронной сети

На тестовой выборке удалось достичь следующих показателей. Матрица ошибок (англ. confusion matrix) приведена в таблице 1.

Значения метрик классификации приведены в таблице 2. Введем обозначения: TP (True Positive) - истинно положительные предсказания, FP (False Positive) - ложно положи-

истинно \ предсказано	фон	ткань
	фон	ткань
фон	429	0
ткань	73	486

Таблица 1: Матрица ошибок.

тельные предсказания, FN (False Negative) - ложно отрицательные предсказания, TN (True Negative) истинно отрицательные предсказания.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN},$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN},$$

Название метрики	Значение
Accuracy	0.926
Precision	0.87
Recall	1
F1	0.93

Таблица 2: Значения метрик.

Анализируя результаты, можно понять, что нейронная сеть ошибается на тайлах, где присутствует ткань. Это связано с тем, что бывают тайлы, на которых присутствуют небольшие фрагменты ткани, и тайлы, на которых присутствует полупрозрачная ткань. Примеры представлены на Рис. 6.

## 5.4. Метод классификации всех тайлов пирамиды полнослайдового изображения

Так как в рамках разрабатываемого подхода детекции ткани на полнослайдовых изображениях необходимо было классифицировать каждый тайл на всех уровнях пирамиды, решено было воспользоваться тем свойством пирамиды, что тайлы верхних уровней состоят из 4 тайлов нижнего уровня (или из меньшего количества от 1 до 4 в случае если тайл граничный). Изначально инициализируется пирамидообразная структура данных, аналогичная пирамиде полнослайдового изображения (каждому тайлу соответствует метка 0, если этот

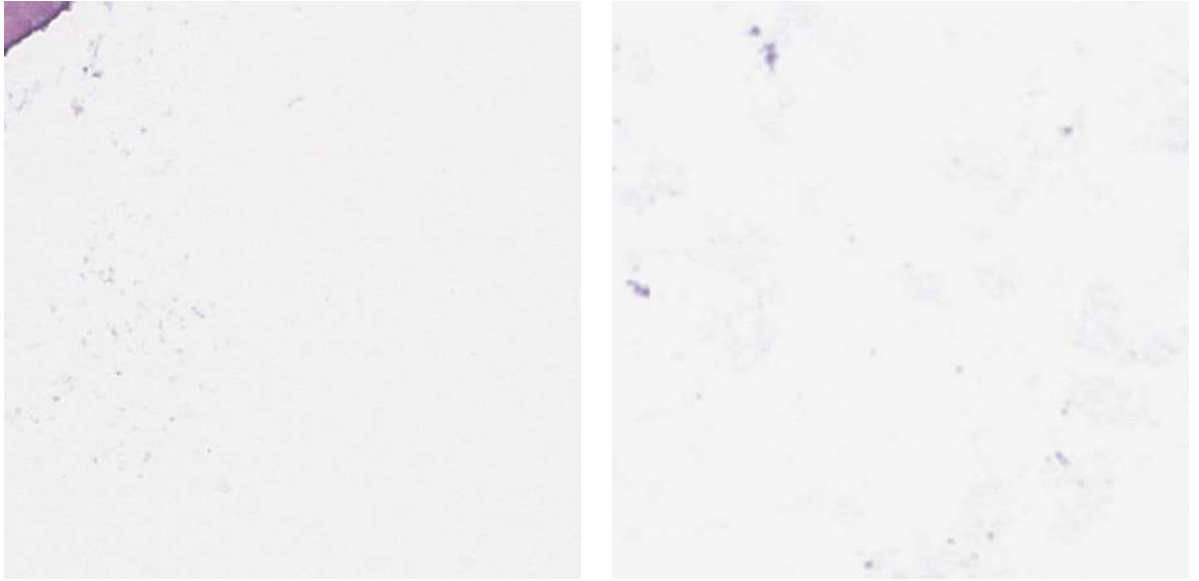


Рис. 6: Примеры ошибочной классификации. На данных фрагментах присутствует ткань, но фрагменты были ошибочно классифицированы как фоновые.

тайл - фон, или метка 1, если этот тайл - ткань, на соответствующем уровне) и каждый уровень пирамиды дополнялся нулями снизу, если его высота нечетная, и нулями справа, если его ширина нечетная. Далее на тайлах нижнего уровня пирамиды полнослайдового изображения запускалась обученная нейронная сеть для классификации, ставящая в соответствие тайлу его метку (0 - фон, 1 - ткань). Следующим шагом метода является распространение меток на уровень выше. Для этого по текущему уровню "проходит" окно размера  $2 \times 2$  и сопоставляет 4-м меткам текущего уровня одну метку на уровне выше по правилу:

$$a_{(z+1)yx} = \max(a_{zyx}, a_{z(y+1)x}, a_{zy(x+1)}, a_{z(y+1)(x+1)}),$$

где  $a_{zyx}$  - это элемент пирамидообразной структуры,  $z$  - номер уровня пирамиды,  $y$  - номер строки,  $x$  - номер столбца. Причем дополнение нулями до четных высоты и ширины гарантирует то, что при каждом положении окна будут учитываться 4 метки, а так как дополнение происходит нулями, то они не повлияют на итоговую метку. Метки распространяются пока не будет достигнут последний уровень пирамиды.

## 5.5. Визуализация работы предложенного метода

Для визуализации полнослайдовые изображения тканей PATH-DT-MSU WSS1 уменьшались в 8 раз, а полнослайдовые изображения тканей PATH-DT-MSU WSS2 уменьшались в 16 раз. Примеры классификации тайлов на разных уровнях пирамиды можно увидеть на

Рис. 7 и Рис. 8.

На каждом последующем уровне "квадраты" становятся крупнее. Это происходит из-за того, что тайлы с верхних уровней содержат в себе 4 тайла с уровня ниже (или меньшее количество, если тайл граничный), и если хотя бы один из этих 4 тайлов является тканью, то тайл с верхнего уровня также будет тканью. Таким образом, на каждом последующем уровне квадраты увеличиваются в два раза. Самый верхний уровень всегда состоит из одного тайла и он всегда должен детектироваться как не фоновый.

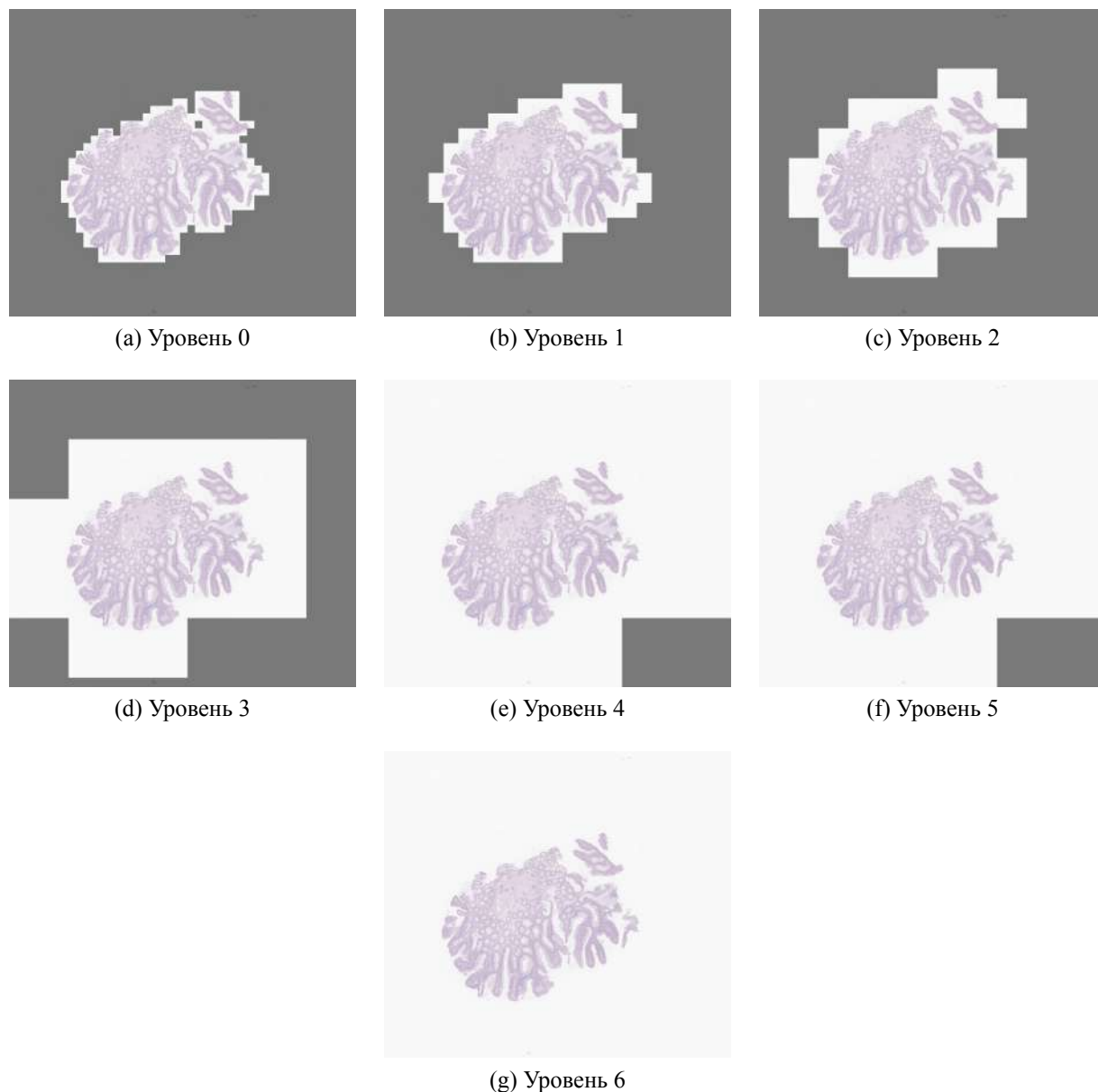


Рис. 7: Визуализация классифицированных тайлов на различных уровнях пирамиды для полнослайдового изображения из PATH-DT-MSU WSS1. Светлым выделены тайлы классифицированные как ткань, темным выделены тайлы классифицированные как фон. Размер изображений  $2641 \times 2988$  пикселей. Исходный размер  $21125 \times 23904$  пикселей.

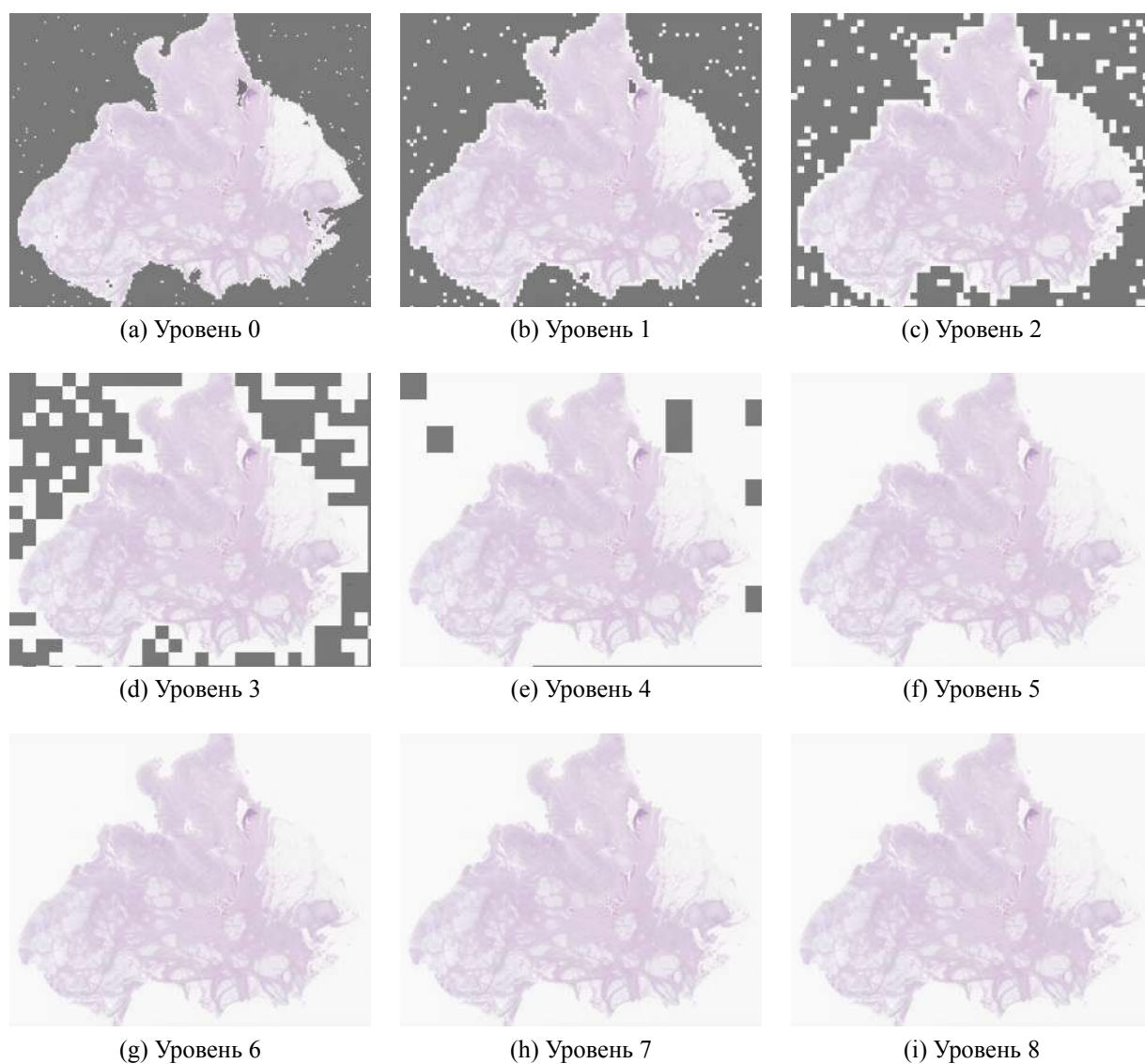


Рис. 8: Визуализация классифицированных тайлов на различных уровнях пирамиды для полнослайдового изображения из PATH-DT-MSU WSS2. Светлым выделены тайлы классифицированные как ткань, темным выделены тайлы классифицированные как фон. Размер изображений  $5655 \times 6972$  пикселей. Исходный размер  $90473 \times 111552$  пикселей.

## 5.6. Уточнение метода детекции ткани

Как было показано в секции 5.3. Результаты обучения нейронной сети, сеть ошибается на тайлах, на которых присутствует маленький кусочек ткани. Для улучшения классификации таких тайлов можно добавить дополнительный параметр сдвиг (англ. *stride*), отвечающий за расстояние между тайлами, раньше он был равен размеру тайла (в нашем случае 512), то есть тайлы не пересекались.

Введем обозначения: *tile\_size* - размер тайла, *stride* - расстояние между тайлами. Будем считать, что *tile\_size* делится нацело на *stride*. Можно сделать сдвиг, равным половине размера тайла (в нашем случае 256), и каждый тайл разобьется на  $(\frac{tile\_size}{stride})^2$  частей (в нашем случае 4 части) или меньшее число частей, если тайл граничный. Таким образом, для каждого фрагмента размера *stride*×*stride* пикселей будет сделано  $(\frac{tile\_size}{stride})^2$  предсказаний, усредняя которые, мы получим степень уверенности в том, что данный фрагмент является тканью (значение из [0, 1]). В итоге мы получим карту уверенности. Карта уверенности для нижнего уровня пирамиды полносладового изображения из PATH-DT-MSU WSS1 приведена на Рис. 9. Можно заметить, что на границе, где и были случаи неверной классификации, теперь есть возможность верной классификации. Для этого нужно выбрать порог (англ. *threshold*) и получить бинарную маску без ошибок классификации на краях. Сравнение способов классификации тайлов со сдвигом, равным размеру тайла, и способом классификации со сдвигом, равным половине размера тайла, можно посмотреть на рисунке Рис. 10

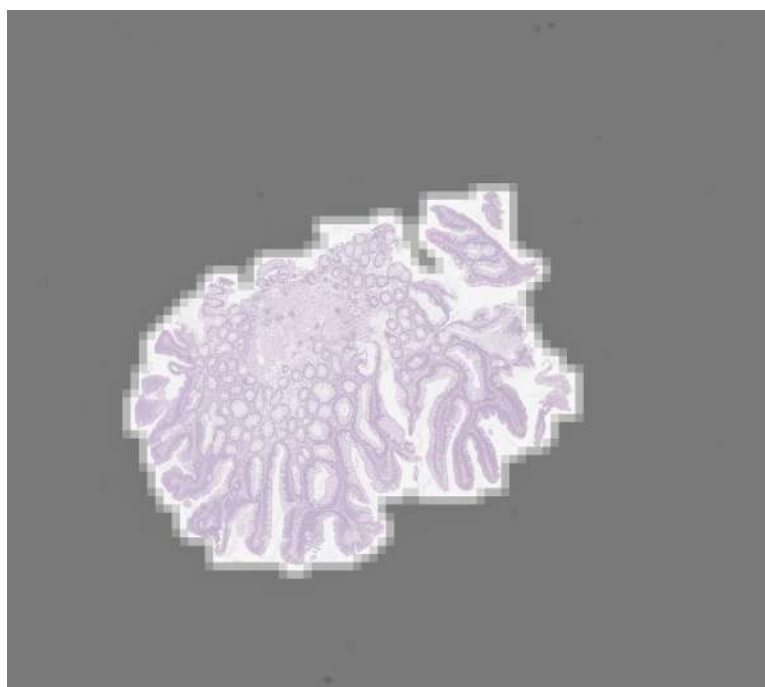


Рис. 9: Карта уверенности. Ткань выделяется светлым, фон выделяется темным.



(a) stride = 512



(b) stride = 256

Рис. 10: Различие между способами классификации тайлов. Ткань выделяется светлым, фон выделяется темным.



## 6. Заключение

Была обучена сверточная нейронная сеть ResNet для классификации тайлов на наборах данных PATH-DT-MSU WSS1, WSS2. Для этого был размечен набор данных, состоящий из 3523 изображений тайлов. Разработан алгоритм для распространения предсказаний на верхние уровни пирамиды полнослайдового изображения и способ уточнения классификации тайлов нижнего уровня пирамиды.

### 6.1. Программная реализация

Предлагаемый метод был программно реализован на языке Python3 с использованием следующих библиотек:

1. NumPy [12] - для работы с массивами данных;
2. PyTorch [13] - для работы с нейронной сетью;
3. psimage<sup>3</sup> - для работы с полнослайдовыми изображениями;
4. Pillow [14], Skimage [15] - для работы с изображениями.

---

<sup>3</sup><https://github.com/xubiker/psimage>

## **7. Дальнейшее развитие**

Дальнейшим развитием работы является увеличение точности нахождения ткани. Этого можно добиться с помощью уменьшения размера тайлов или с помощью перехода от классификации к сегментации.

## Список литературы

- [1] Visualization of whole slide histological images with automatic tissue type recognition / Khvostikov A., Krylov A., Mikhailov I., and Malkov P. // Pattern Recognition and Image Analysis. — 2022. — Vol. 32, no. 3. — P. 483–488.
- [2] PathScribe. — URL <https://pathscribe.ru/>.
- [3] Otsu N. A threshold selection method from gray-level histograms // IEEE transactions on systems, man, and cybernetics. — 1979. — Vol. 9, no. 1. — P. 62–66.
- [4] An automatic entropy method to efficiently mask histology whole-slide images / Song Y., Cisternino F., Mekke J. M., de Borst G. J., de Kleijn D. P., Pasterkamp G., Vink A., Glastonbury C. A., van der Laan S. W., and Miller C. L. // Scientific Reports. — 2023. — Vol. 13, no. 1. — P. 4321.
- [5] Localization of tissues in high-resolution digital anatomic pathology images / Alomari R. S., Allen R., Sabata B., and Chaudhary V. // Medical Imaging 2009: Computer-Aided Diagnosis / SPIE. — 2009. — Vol. 7260. — P. 349–358.
- [6] Comparison of different methods for tissue segmentation in histopathological whole-slide images / Bándi P., van de Loo R., Intezar M., Geijs D., Ciompi F., van Ginneken B., van der Laak J., and Litjens G. // 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) / IEEE. — 2017. — P. 591–595.
- [7] Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 / Springer. — 2015. — P. 234–241.
- [8] A comparative study of U-net topologies for background removal in histopathology images / Riasatian A., Rasoolijaberi M., Babaei M., and Tizhoosh H. R. // 2020 International Joint Conference on Neural Networks (IJCNN) / IEEE. — 2020. — P. 1–8.
- [9] A Weakly Supervised Tissue-level Approach for Histological Differentiation of Gastric Adenocarcinomas / Hyeon D.-W., Kim D., Ryu W.-S., Sung S. H., Bhattacharjee S., Kim C.-H., Cho N.-H., Moon D., Lee D.-H., Kim H., et al. — 2022.

- [10] Grad-cam: Visual explanations from deep networks via gradient-based localization / Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., and Batra D. // Proceedings of the IEEE international conference on computer vision. — 2017. — P. 618–626.
- [11] Deep residual learning for image recognition / He K., Zhang X., Ren S., and Sun J. // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — P. 770–778.
- [12] Array programming with NumPy / Harris C. R., Millman K. J., Van Der Walt S. J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N. J., et al. // Nature. — 2020. — Vol. 585, no. 7825. — P. 357–362.
- [13] Pytorch: An imperative style, high-performance deep learning library / Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T., Lin Z., Gimelshein N., Antiga L., et al. // Advances in neural information processing systems. — 2019. — Vol. 32.
- [14] scikit-image: image processing in Python / Van der Walt S., Schönberger J. L., Nunez-Iglesias J., Boulogne F., Warner J. D., Yager N., Gouillart E., and Yu T. // PeerJ. — 2014. — Vol. 2. — P. e453.
- [15] Clark A. et al. Pillow (pil fork) documentation // readthedocs. — 2015.