



PATHONET

Реализация алгоритма PathoNet, предложенного в статье Negahbani, F., Sabzi, R., Pakniyat Jahromi, B. et al. "PathoNet introduced as a deep neural network backend for evaluation of Ki-67 and tumor-infiltrating lymphocytes in breast cancer." Sci Rep 11, 8489 (2021).

Докладчик: Дарья Рудась

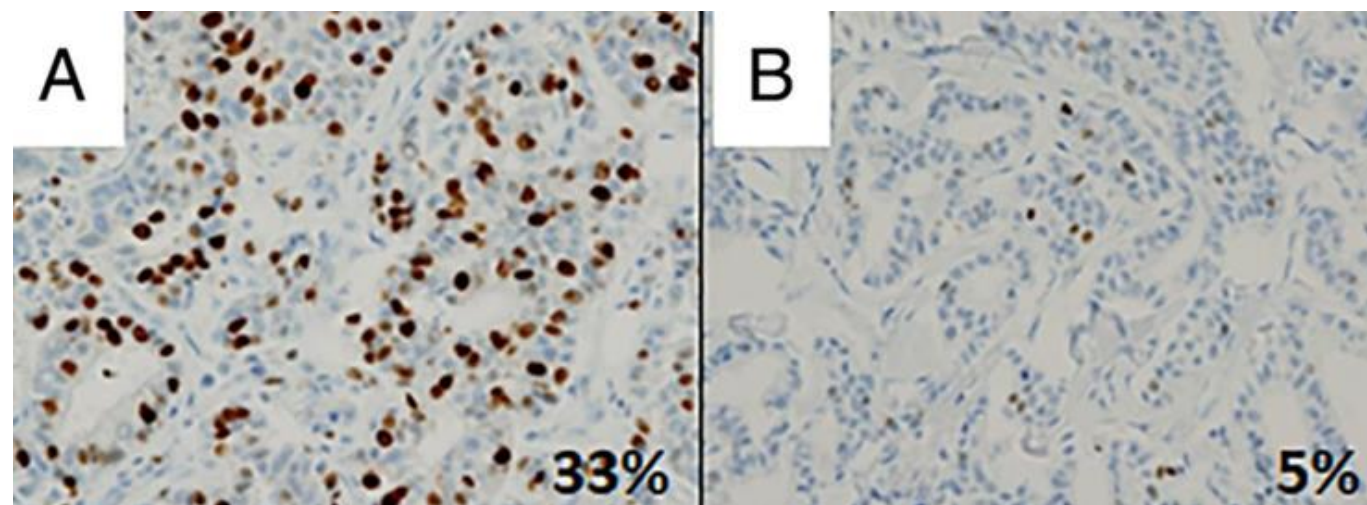
Собеседование на кафедре ИППИ 14.04.2025

Содержание

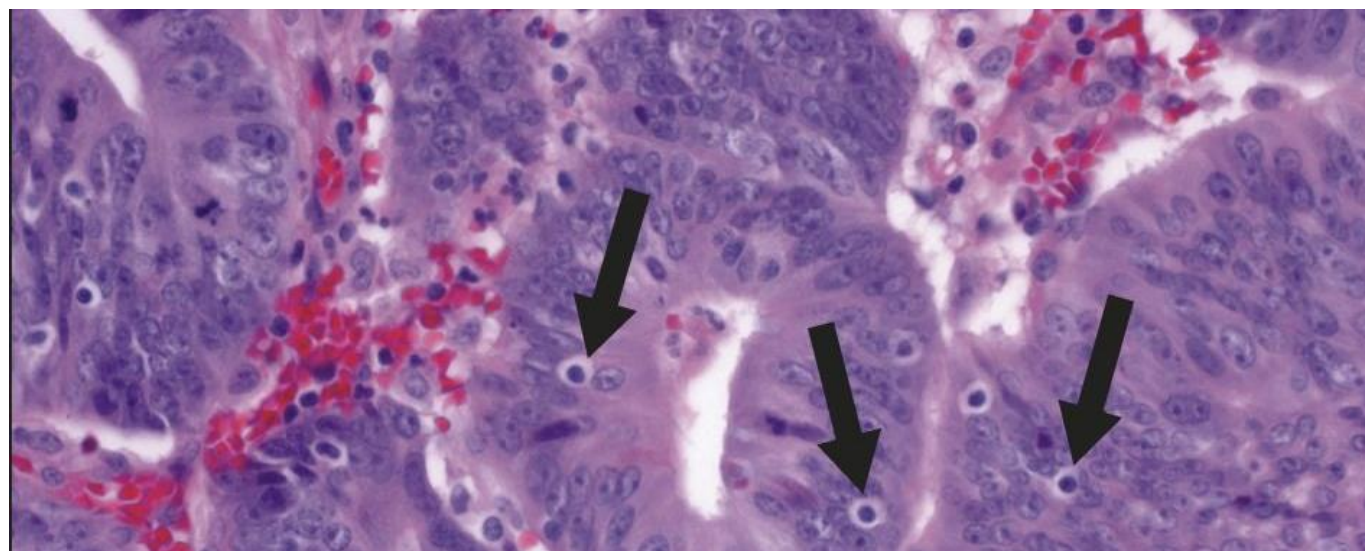
1. Введение
 - a. Описание задачи
 - i. Ki-67 и TIL
 - b. Актуальность
2. Задача сегментации
 - a. Виды сегментации
 - b. U-Net
3. Описание решения
 - a. Определение качества
 - i. Набор данных
 - ii. Метрики
 - b. Архитектура PathoNet
 - i. Dilated convolution
 - ii. Watershed algorithm
 - c. Экспериментальные результаты
4. Анализ

Прогнозирование развития рака молочной железы

- Ядерный белок Ki-67 является маркером пролиферации
- Чрезмерная клеточная пролиферация коррелирует с прогрессированием злокачественной опухоли
- Инфильтрирующие опухоль лимфоциты (TIL) полезны для улучшения состояния пациентов с раком молочной железы



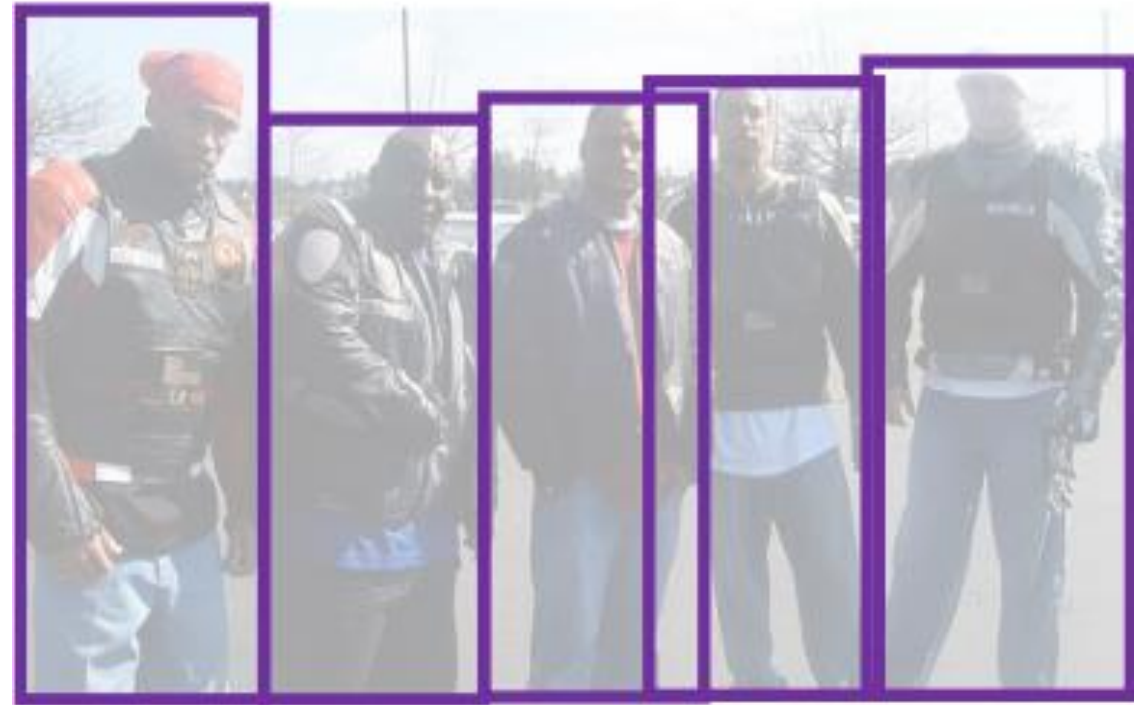
Иммуногистохимическое исследование Ki-67



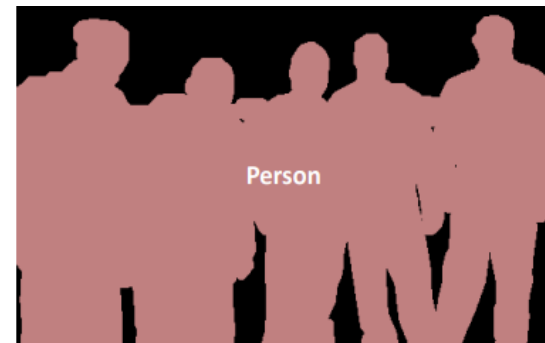
Эпителий , инфильтрованный лимфоцитами

Задача сегментации

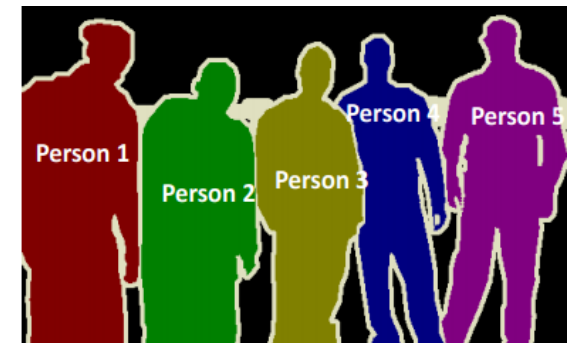
- Сегментация – это присвоение меток пикселям
- Модели сегментации предоставляют карты сегментов в качестве выходных данных, соответствующие входным данным



Object Detection



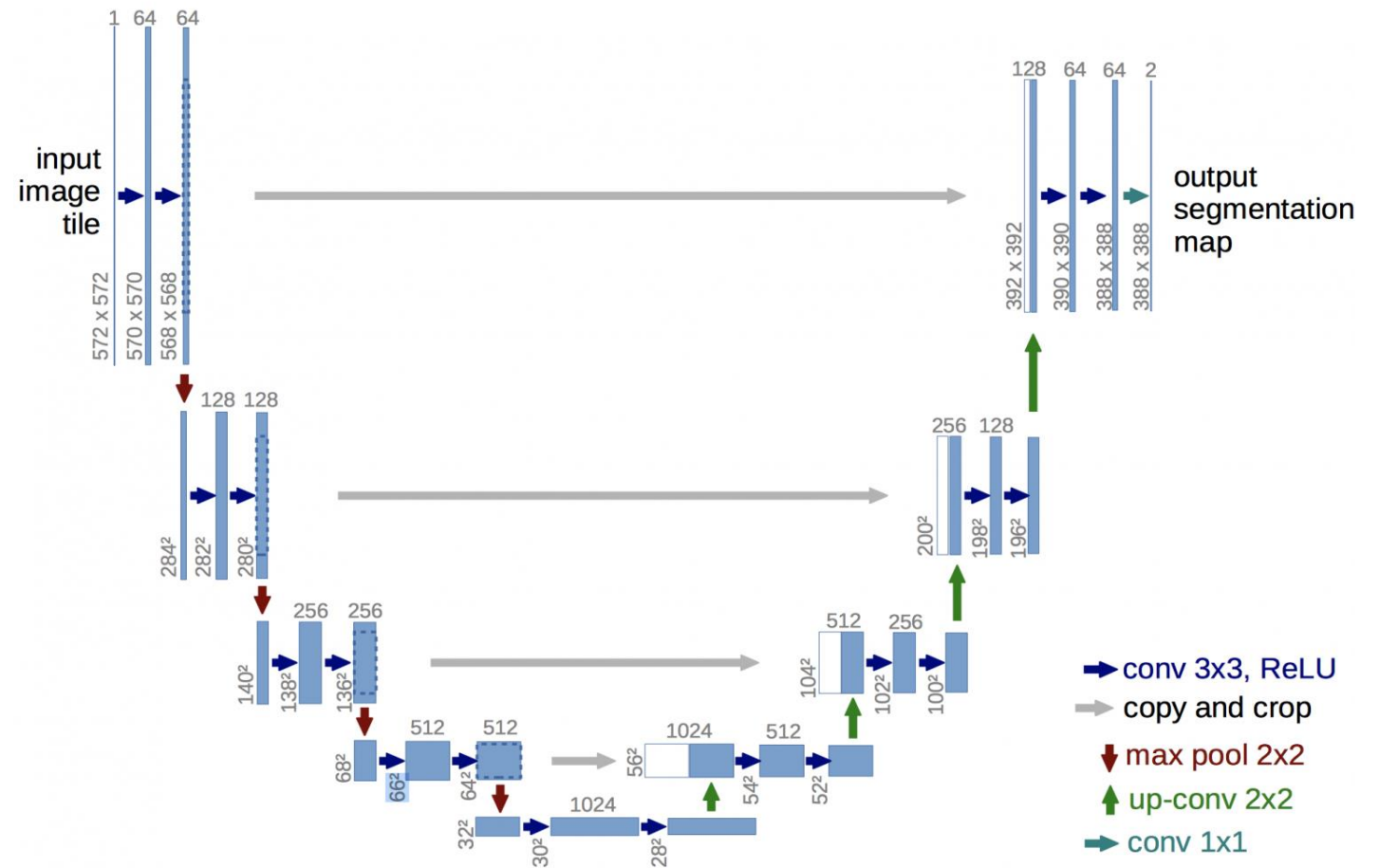
Semantic Segmentation



Instance Segmentation

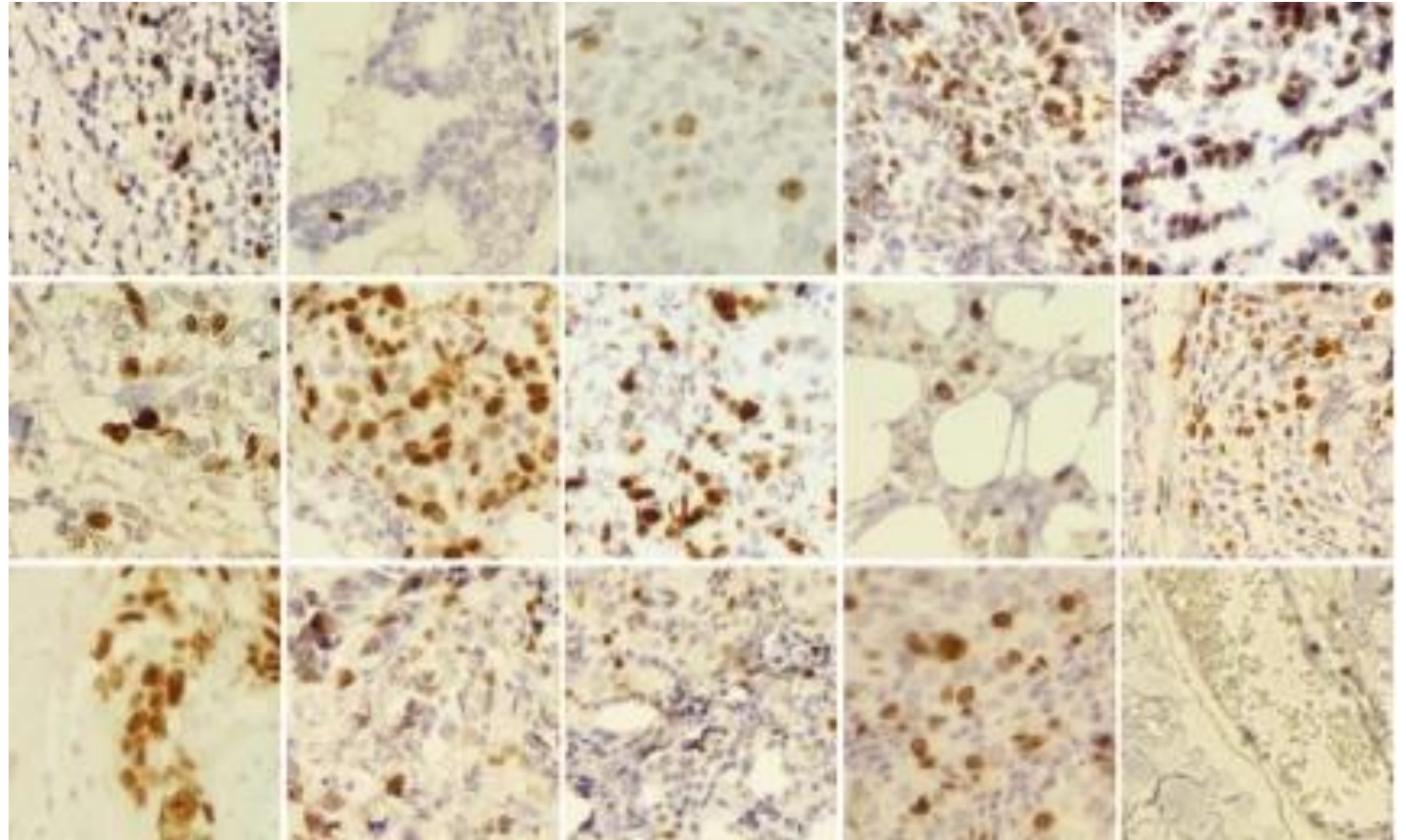
Архитектура U-Net

- U-Net состоит из двух частей : энкодера и декодера
- Downsampling
- Upsampling + Skip connection



Набор данных

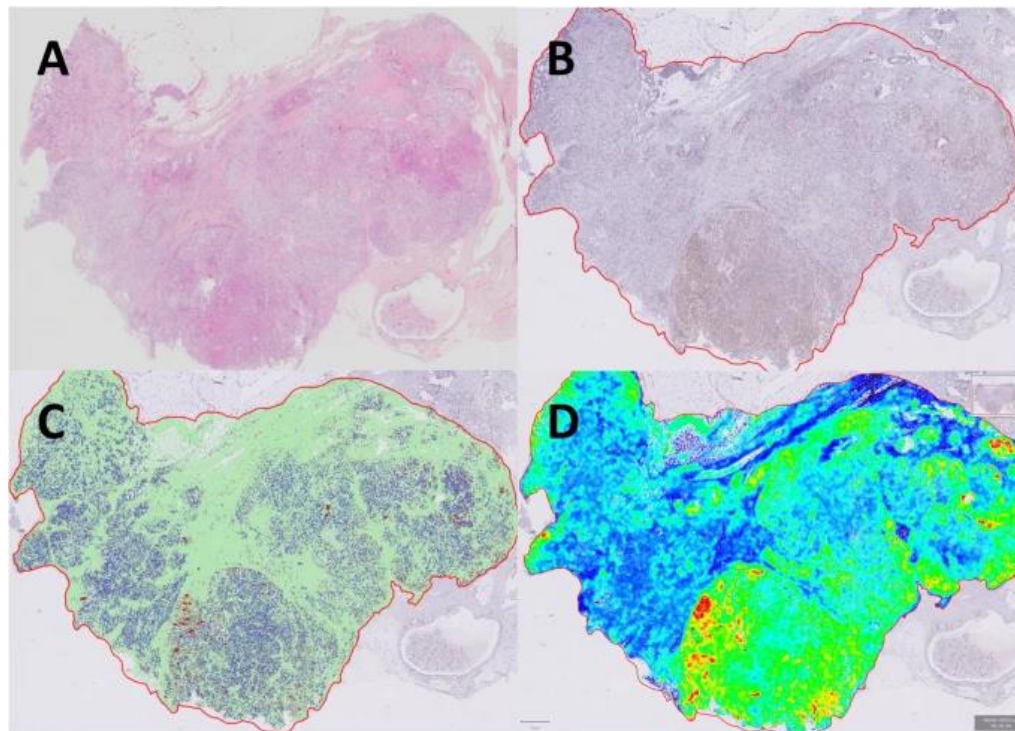
- Изображения были получены из образцов биопсии, собранных в ходе клинического исследования с 2017 по 2020 год.
- SHIDC-B-Ki-67 содержит 1656 обучающих и 701 тестовый набор данных



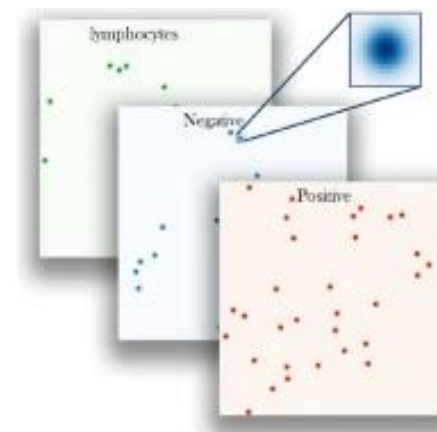
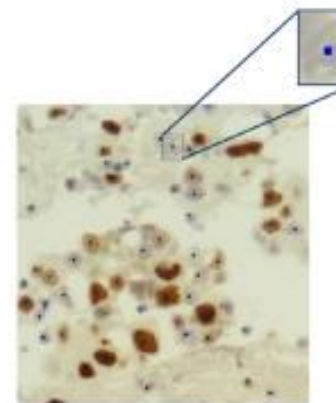
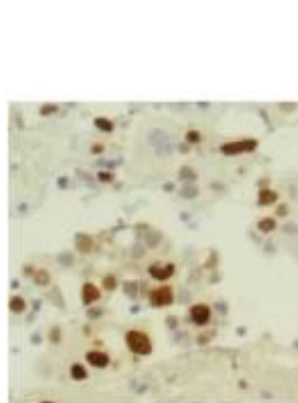
В SHIDC-B-Ki-67 каждое изображение содержит в среднем 69 клеток, а всего 162 998 клеток

Особенности медицинских изображений

- Разнообразие форм и размеров анатомии между пациентами
- Возможна низкая контрастность с окружающими тканями
- Наложение пикселей во многих клетках
- Сложность поиска размеченных данных



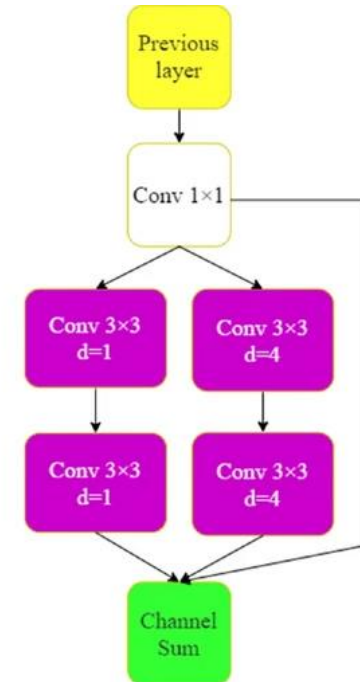
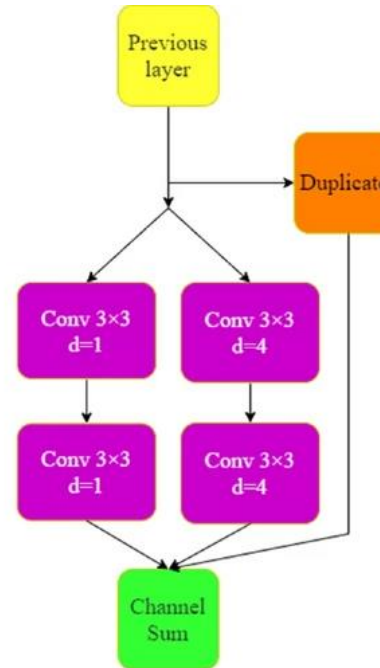
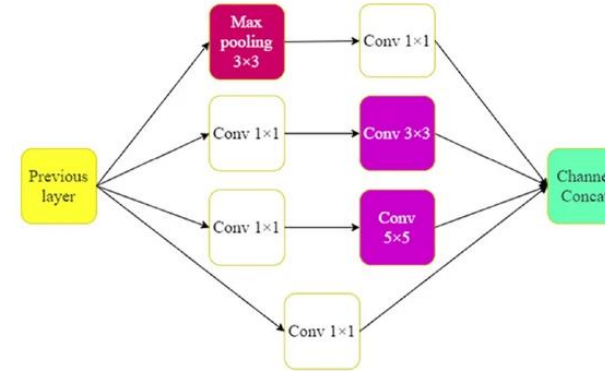
Тепловая карта распределения клеток



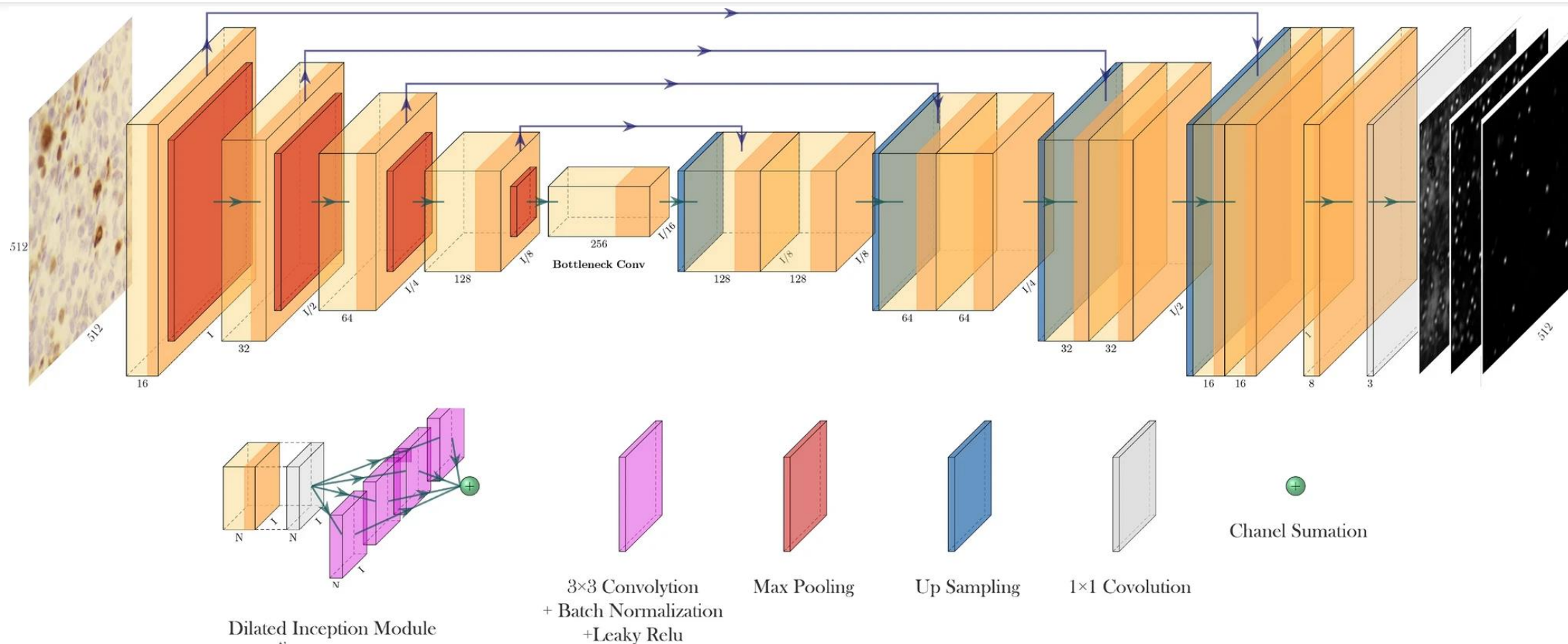
Выбор центрального пикселя клетки для классификации

Residual dilated inception module

- Размер клеток может отличаться, поэтому выбор подходящего размера ядра важен
- Dilated convolution для расширения reception field и сокращения количества параметров
- Выходные данные двух путей суммируются с входными данными (аналогия с ResNet)



Архитектура PathoNet



Алгоритм водораздела

- Каждая точка соответствует пикселю(высота зависит от интенсивности)
- Начинаем "заполнять" области, как только какие-то сливаются в одну, проводим разделяющую границу

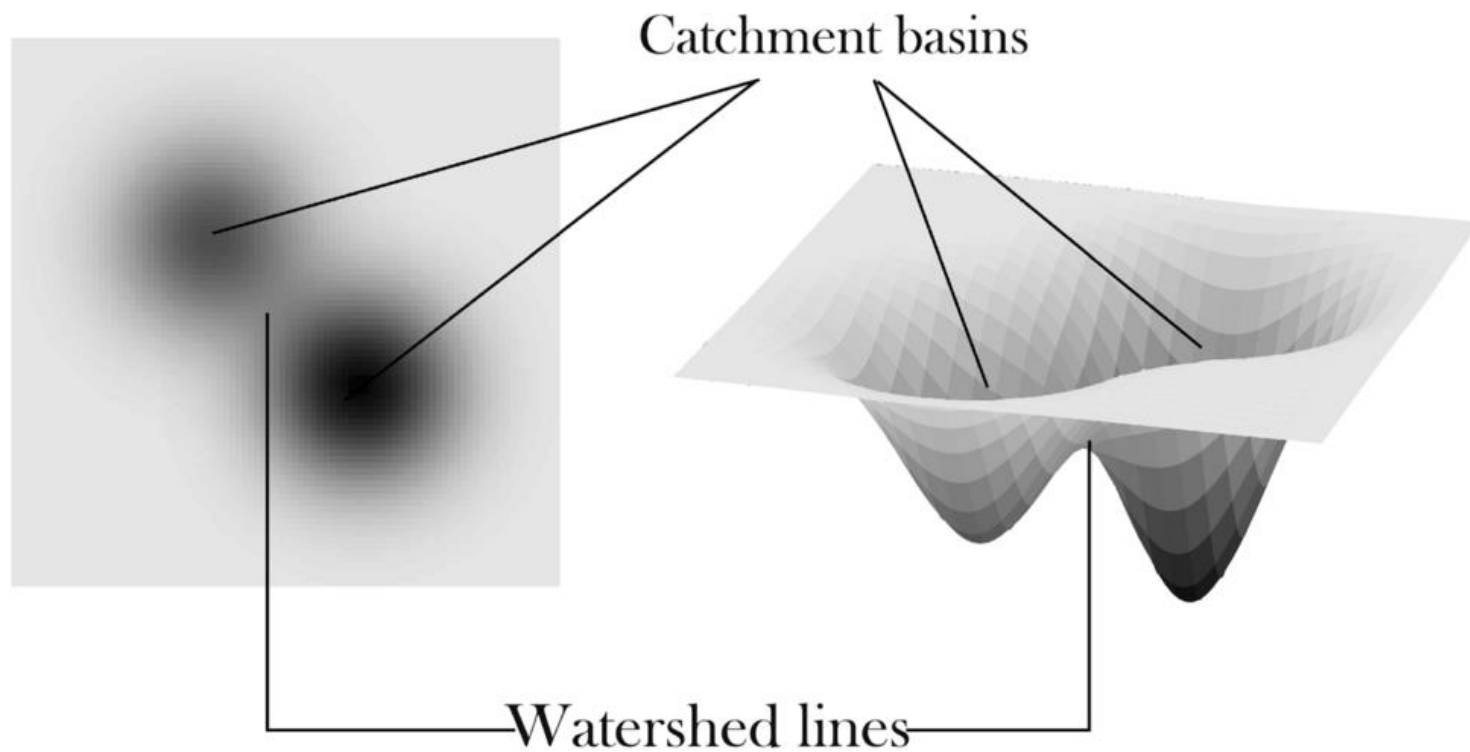
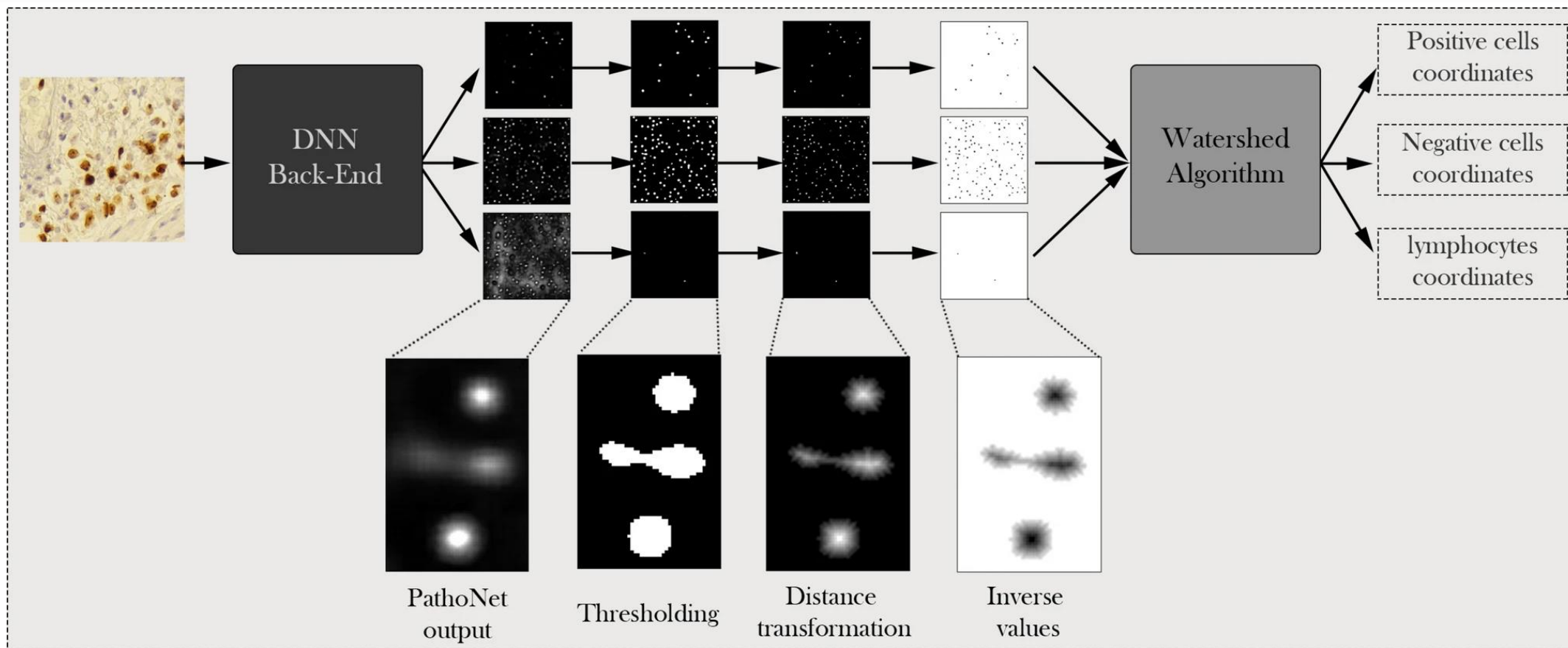


Схема работы(постобработка + watershed)



Метрики качества

- True Positive – в размеченных данных существует центр того же типа с расстоянием менее R от нашего
- Несколько центров в полученных данных находятся ближе, чем R пикселей от размеченного центра, ближайший – TP, остальные – False Positive
- FN – не определен центр клетки, который есть в размеченных данных

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

$$Ki - 67 - score = \frac{Immunopositive}{Immunopositive + Immunonegative}$$

$$TIL - score = \frac{Lymphocyte}{Lymphocyte + Immunopositive + Immunonegative}$$

Оценка качества реализации

- Модель обучалась 10 эпох
- Learning rate scheduler
- 20% тренировочных данных для валидации
- Результаты получились сопоставимыми с результатами авторов
- Подбор гиперпараметров, отличных от авторских (порог бинаризации, learning rate)



Test: 100%|██████████| 700/700 [00:49<00:00, 14.01it/s]
Class 0 => Prec=0.780, Rec=0.890, F1=0.831
Class 1 => Prec=0.869, Rec=0.494, F1=0.630
Class 2 => Prec=0.344, Rec=0.341, F1=0.343
Overall => Prec=0.808, Rec=0.615, F1=0.699



Test: 100%|██████████| 700/700 [00:49<00:00, 14.16it/s]
Class 0 => Prec=0.732, Rec=0.912, F1=0.812
Class 1 => Prec=0.703, Rec=0.774, F1=0.737
Class 2 => Prec=0.266, Rec=0.037, F1=0.065
Overall => Prec=0.712, Rec=0.797, F1=0.752

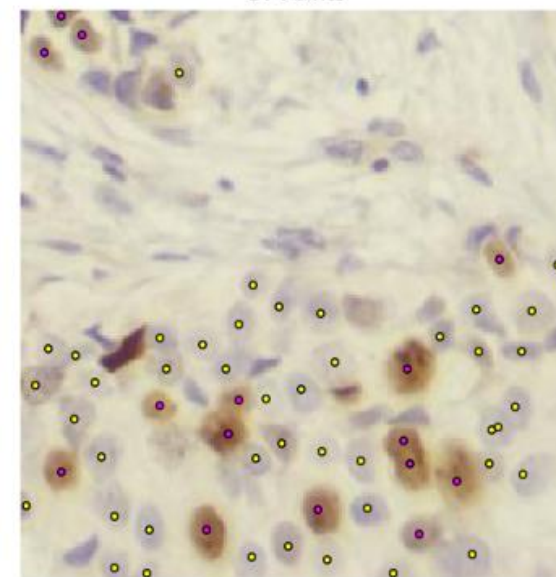
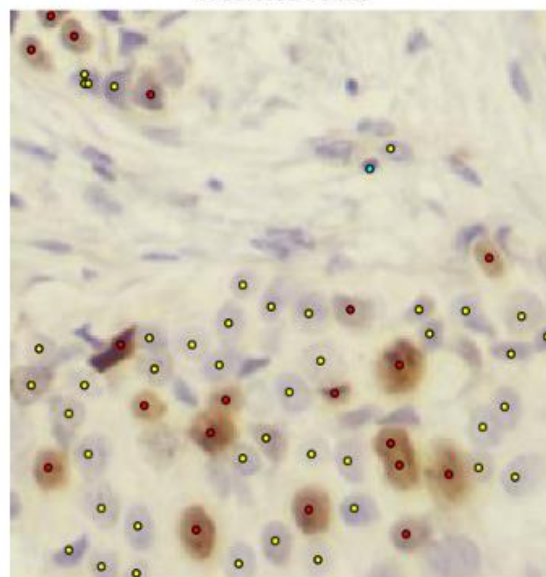
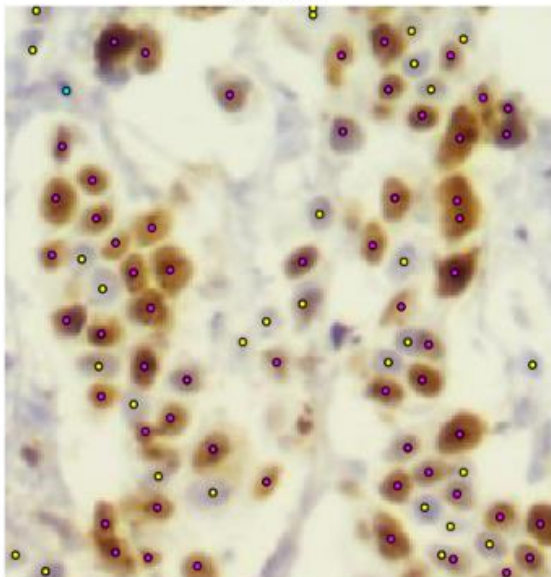
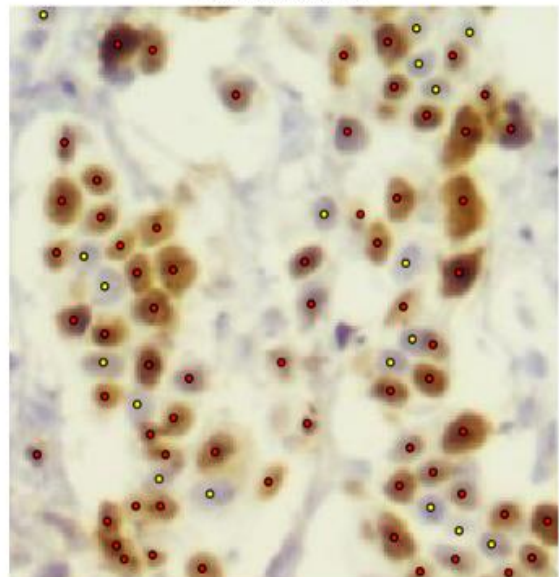
Пример работы на тестовых данных

Predicted Points

GT Points

Predicted Points

GT Points



Сравнение PathoNet с другими решениями

Backend	Ki-67 index (RMSE)	TILs score (RMSE)	Ki-67 cut-off accuracy	TILs cut-off accuracy
Modified DeepLabv3-Mobilenetv2 ⁴⁵	0.05123	0.05568	0.9565	0.8260
Modified DeepLabv3-Xeption ⁴⁵	0.06341	0.01639	0.9130	0.9565
Modified FCRN-A ⁴⁷	0.05836	0.01580	0.9565	1.0
Modified FCRN-B ⁴⁷	0.06285	0.01674	0.9565	1.0
Modified U-Net ³⁵	0.05206	0.01497	0.9556	0.9565
Ours(PathoNet)	0.04803	0.02530	0.9565	0.9565

Сравнение PathoNet с другими решениями

Backend	# Parameters	FPS	Avg. F1
Modified DeepLabv3-Mobilenetv2 ^{45}	3,236,907	20.52	0.7508
Modified DeepLabv3-Xeption ^{45}	41,253,587	8.76	0.7871
Modified FCRN-A ^{47}	2,142,019	22.03	0.7710
Modified FCRN-B ^{47}	1,365,888	14.6	0.7810
Modified U-Net ^{35}	31,036,323	12.34	0.7783
Ours (PathoNet)	3,142,208	14.86	0.7928

Анализ решения

- Различия в размере и цвете опухолевых клеток у разных пациентов усложняют маркировку и аннотирование клеток
- Использование картинок с маркировкой Ki-67 для оценки TIL приводит к маскировке Ki-67-положительных TIL
- PathoNet показал высокую точность в детекции Ki-67 и TIL по сравнению с другими решениями

