

Классификация рака молочной железы с помощью сверточных нейронных сетей

Д.А.Сасов, А.В. Зубков, Ю.А. Орлова, А.В. Турицына

Волгоградский государственный технический университет

Аннотация: В данной работе рассмотрены актуальные архитектуры нейронных сетей, применяемые для классификации рака молочной железы по гистологическим микрофотографиям. Главной целью исследования было проведение расширенного тестирования обученных моделей на данных, которые принципиально будут отличаться от датасета для обучения. В качестве данных для обучения был выбран крупный российский датасет с классификацией гистологических микрофотографий по характеру поражения, а в качестве тестовых данных выступили данные из этого же датасета, но с измененными параметрами. Для обучения были выбраны модели ResNet 152, DenseNet 121, Inception_resnet_v2. Для обучения был применен подход transfer learning. Предобработка изображений перед обучением состояла в нормализации значений всех каналов изображений в диапазоне от 0 до 1. По итогу тестирования была выявлена проблема с классификацией изображений с изменениями в балансе цвета, контрастности и яркости. Эти изменения привели к значительному ухудшению метрик. Из проведенных тестов сделан вывод, что для качественного обучения моделей с целью классификации рака молочной железы, недостаточно элементарной нормализации, ведь она делает модели неустойчивыми к изменениям изображений, которые могут встретиться на практике.

Ключевые слова: нейронная сеть, модель, машинное обучение, рак молочной железы, классификация рака, ResNet 152, Inception_resnet_v2, DenseNet 121, искусственный интеллект, transfer learning, гистологическая микрофотография.

Введение

В 2020 году во всем мире у 2,3 миллиона женщин был диагностирован рак молочной железы (далее РМЖ) и 685 000 умерли от этого заболевания [1]. В России смертность от данного типа рака среди женщин составляет 15,9%, а распространенность – 21,2% [2]. Поэтому для России вопрос борьбы с РМЖ также актуален.

Ранняя диагностика и эффективное лечение имеют решающее значение для выживаемости. Для пациентов, обратившихся к врачу с небольшой опухолью на ранней стадии, выживаемость является более высокой, поэтому разрабатывается много новых технологий для выявления опухолей и эффективного их лечения [3].

Из множества методов диагностики РМЖ можно выделить гистопатологическую оценку биопсийного материала, которая является очень важной для диагностики и лечения РМЖ. Гистологический анализ является золотым стандартом для определения типа рака [4]. Однако гистопатологический анализ РМЖ трудоемок и нетривиален, поэтому даже между двумя опытными специалистами могут возникнуть расхождения в диагнозах [5].

Искусственный интеллект в медицинских задачах

Искусственный интеллект все больше проникает в нашу жизнь. Он позволяет решать самые разные по сложности задачи. Медицина не стала исключением. Уже сейчас применение искусственного интеллекта в области диагностики рака активно исследуется в различных работах. В исследовании [6] нейронные сети применяются для классификации новообразований на коже, авторам удалось достичь неплохих метрик в задаче бинарной классификации. Наивысшая точность составила 0,752, специфичность – 0,665, чувствительность – 0,665. В работе [7] нейронные сети применяются для локализации опухоли по данным микроволновой термометрии, однако наивысшая достоверность (ассигасу) составила только 61,6% на тестовых данных.

Отдельно стоит отметить сверточные нейронные сети (далее СНС), они сейчас очень часто появляются в исследованиях, связанных с анализом медицинских изображений [8]. Благодаря мощности современных компьютеров стало возможным обучение СНС на изображениях без предварительного выделения признаков [9]. СНС хорошо себя показывают в задачах классификации медицинских изображений с точки зрения различных метрик. Хотя на данный момент они представляются в виде черного ящика, который, обрабатывая изображение, не может дать четкого ответа, почему был поставлен именно такой диагноз [10].

Сейчас существует достаточно архитектур CNN: AlexNet, DenseNet, Xception, Inception, ResNet, MobileNet, VGG и др. Они часто встречаются в задачах классификации медицинских изображений [11].

В работах [12, 13] сравниваются разные модели сверточных нейронных сетей в задачах бинарной и многоклассовой классификации гистологических микрофотографий с применением подхода transfer learning. В работе [12] сравниваются модели Inception_v3 и Inception_resnet_v2, вторая модель показывает более хорошие результаты: на аугментированных данных из датасета BraKHis accuracy не опускается ниже 99,10% в бинарной классификации и ниже 96,89% в многоклассовой. В исследовании [13] сравниваются модели архитектур EfficientNet, ResNet и MobileNet. Авторы выделили модель EfficientNetV2B0-21k, которая имела более высокие показатели, чем остальные модели, также отмечено, что все протестированные модели в целом хорошо справились с задачей классификации гистологических микрофотографий.

В работах [14, 15] предложены методы двухэтапной обработки гистологических микрофотографий. В исследовании [14] на первом этапе происходит классификация частей изображений путем голосования по трем моделям (Inception_resnet_v2, VGG16 и VGG19), а на втором этапе результаты классификации попадают в двухэтапную нейронную сеть для сопоставления информации на уровне частей изображения с предсказанием на уровне всего изображения. В работе [15] авторы предлагают использовать сверточные нейронные сети в паре со стандартными классификаторами (методом опорных векторов и логистической регрессией). На первом этапе сверточная нейронная сеть выделяет признаки у изображений, а на втором эти признаки передаются в классификатор для выбора одного из классов.

Есть исследования, в которых авторы предлагают собственные архитектуры на основе модулей из уже созданных архитектур [16, 17]. Еще

одним направлением можно считать построение ансамблей из моделей СНС [5].

Тестирование моделей в большинстве исследований проводится на том же датасете, на котором проходило ее обучение, но для обучения отводится большая часть данных, а для тестирования меньшая. Однако все датасеты для классификации РМЖ по гистологическим микрофотографиям отличаются. У изображений отличаются увеличение, яркость, оттенок, разрешение. Такая проблема возникает из-за разных микроскопов и методов окрашивания. Лишь в некоторых источниках это отмечают и используют нормализацию цвета для изображений [12, 18]. Чаще применяется лишь нормализация значений всех каналов изображения в определенном диапазоне. С учетом такого подхода неясно, как себя поведет модель, обученная и протестированная на данных с одним увеличением и цветом, на данных с другим увеличением и цветом.

Описание набора данных

В данном исследовании в качестве данных для обучения выбран российский датасет, представленный в работе [19]. Он содержит гистологические микрофотографии молочной железы с разным увеличением и разным разрешением. В таблице 1 приведено распределение изображений в датасете.

Таблица № 1

Распределение изображений в датасете по увеличению и разрешению

Разрешение (кол. пикселей) / Увеличение	x4	x10
300x300	6730	23235
500x500	3049	9890
300x300	6730	23235

В датасете представлено несколько видов классификации, однако нам понадобится лишь один вид – характер поражения (benign, in situ, invasive).

Для обучения были выбраны изображения в разрешении 300x300 пикселей и увеличением x4 и x10. Пример изображения приведен на рис. 1.

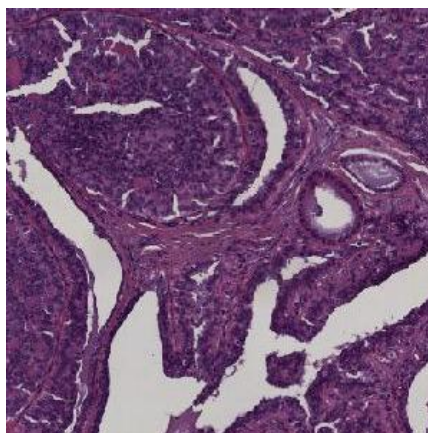


Рис. 1. – Пример изображения гистологической микрофотографии молочной железы

Для предобработки изображений использовалась стандартная нормализация $PyTorch$, суть которой состоит в приведении значений по всем каналам изображения к диапазону от 0 до 1.

Классы в датасете несбалансированны, поэтому была применена аугментация (вращение изображения на 90, 180, 270 градусов и отражение по горизонтали и вертикали) для выравнивания классов. В таблице 2 приведено распределение изображений до и после аугментации.

Таблица № 2

Распределение данных в датасете для обучения до и после аугментации

Характер поражения	Всего до аугментации	Тестирование	Тренировка + валидация	Всего после аугментации
benign	9514	2160	12240	14400
in situ	899	2160	12224	14384
invasive	18687	2160	12240	14400

В датасете 15% от всего количества изображений было выделено для первого этапа тестирования моделей, оставшиеся изображения делились на тренировочные (85%) и валидационные (15%). Для второго этапа тестирования были взяты те же данные, что и для обучения, но в другом разрешении (изображения в разрешении 500x500 пикселей были сжаты до

разрешения 300x300 пикселей). Третьим этапом тестирования являлась проверка моделей на изображениях с разрешением 300x300 пикселей, но с изменением баланса цвета, яркости и контрастности. Для изменения параметров изображений использовалась библиотека PIL для Python. Для баланса цвета, контрастности и яркости были выбраны следующие значения: 1,2 и 0,8.

Обучение моделей

В данной работе для обучения взяты 3 модели, относящиеся к архитектурам, которые отлично себя зарекомендовали в других исследованиях и показали отличные результаты в анализе гистологических микрофотографий молочной железы: Inception_ResNet_v2, DenseNet 121, Resnet 152. Модели были предобучены на датасете ImageNet, для обучения был применен подход transfer learning.

В качестве фреймворка использовался PyTorch, вычисления производились на видеокарте. В качестве оптимизатора был выбран Adam, а в качестве функции ошибки – cross-entropy loss. В таблице 3 приведены значения гиперпараметров.

Таблица № 3

Значения гиперпараметров

Параметр	Значение
batch_size	8
epoches	20
learning_rate	0,001
gamma	0,1
step_size	8

В процессе обучения после каждой эпохи фиксировались значения метрик: train и validation loss, accuracy, macro average precision (далее macro avg precision), macro average recall (далее macro avg recall), macro average F1-score (далее macro avg F1). Для расчета метрик использовалась библиотека

scikit-learn. Лучшие веса модели определялись по наивысшему показателю макро avg F1 во время обучения. В таблице 4 указаны лучшие значения метрик по итогам обучения на валидационном наборе данных.

Таблица № 4

Метрики по итогам обучения

Показатель	ResNet 152	Inception_resnet_v2	DenseNet 121
Accuracy	0,986	0,998	0,994
Macro avg precision	0,986	0,998	0,994
Macro avg recall	0,986	0,998	0,995
Macro avg F1	0,986	0,998	0,994

Тестирование моделей

Во время тестирования оценивались следующие показатели: ассурасу, макро avg precision, макро avg recall, макро avg F1, average ROC-AUC score (далее ROC-AUC score). Для расчета метрик во время тестирования также использовалась библиотека scikit-learn.

Результаты первого этапа тестирования приведены в таблице 5. В целом можно считать, что модели успешно справились с задачей классификации, т.к. все метрики находятся в диапазоне от 0,890 до 0,954.

Таблица № 5

Значения метрик для первого этапа тестирования

Показатель	ResNet 152	Inception_resnet_v2	DenseNet 121
Accuracy	0,890	0,938	0,902
Macro avg precision	0,905	0,944	0,919
Macro avg recall	0,890	0,938	0,902
Macro avg F1	0,890	0,938	0,901
ROC-AUC score	0,917	0,954	0,927

Второй этап тестирования проводился на данных в другом разрешении. Результаты тестирования моделей на этих данных приведены в таблице 6. По результатам тестирования видно, что метрики макро avg recall, макро avg F1 и ROC-AUC score для всех 3-х моделей ухудшились. Лучше всего себя

показала модель Inception_resnet_v2, которая имела самые высокие значения метрик среди остальных моделей.

Таблица № 6

Значения метрик для второго этапа тестирования

Показатель	ResNet 152	Inception_resnet_v2	DenseNet 121
Accuracy	0,923	0,950	0,950
Macro avg precision	0,900	0,955	0,952
Macro avg recall	0,806	0,864	0,816
Macro avg F1	0,840	0,899	0,862
ROC-AUC score	0,878	0,917	0,891

Модели на последнем этапе тестирования модели показали себя хуже всего, результаты представлены в таблицах 7, 8. Заметна сильная просадка всех метрик. Все это демонстрирует, что модели опирались на конкретную цветовую гамму при классификации.

Таблица № 7

Значения метрик для третьего этапа тестирования для изображений с увеличенными показателями

Показатель	ResNet 152	Inception_resnet_v2	DenseNet 121
Accuracy	0,176	0,098	0,295
Macro avg precision	0,352	0,231	0,272
Macro avg recall	0,401	0,371	0,429
Macro avg F1	0,164	0,103	0,223
ROC-AUC score	0,534	0,508	0,545

Таблица № 8

Значения метрик для третьего этапа тестирования для изображений с уменьшенными показателями

Показатель	ResNet 152	Inception_resnet_v2	DenseNet 121
Accuracy	0,411	0,265	0,343
Macro avg precision	0,327	0,572	0,483
Macro avg recall	0,329	0,475	0,430
Macro avg F1	0,260	0,299	0,334
ROC-AUC score	0,489	0,600	0,582

Заключение

В ходе многоэтапного тестирования были выявлены проблемы с устойчивостью моделей к изменениям параметров изображений. Изменение разрешения повлияло на метрики в меньшей степени, а вот уже изменение баланса цвета, контрастности и яркости смогло заметно снизить метрики.

Все это подводит к выводу, что для качественного обучения моделей недостаточно базовой нормализации значений всех каналов изображения в диапазоне от 0 до 1. Модели, обученные с таким способом предобработки данных, плохо адаптируются к изменению входных данных, что в реальных условиях может привести к серьезным неточностям в классификации РМЖ.

Литература

1. Wild C.P., Weiderpass E., Stewart B.W. World Cancer Report: Cancer Research for Cancer Prevention. International Agency for Research on Cancer, 2020. pp. 611.
2. Мерабишвили, В. М. Состояние онкологической помощи в России: рак молочной железы среди женского населения. Заболеваемость, смертность, достоверность учета, детальная локализационная и гистологическая структура. (Популяционное исследование на уровне федерального округа) // Вопросы онкологии. 2022. Т. 68, №3. С. 286–293.
3. Bhushan A., Gonsalves A., Menon J.U. Current State of Breast Cancer Diagnosis, Treatment, and Theranostics // Pharmaceuticals. 2021. Vol. 68, №3. pp. 723.
4. Zeiser F. André da Costa C., Roehe A., Righi R., Marques N. Breast cancer intelligent analysis of histopathological data: A systematic review // Applied Soft Computing. 2021. Vol. 113. P. 107886.
5. Hameed Z., Zahia S., Garcia-Zapirain B., Javier Aguirre J., María Vanegas A. Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models // Sensors. 2020. Vol. 20, №16. P. 4373.

6. Параскевопуло К.М., Наркевич А.Н. Применение классических нейронных сетей для распознавания злокачественных новообразований на цифровых изображениях кожи // Инженерный вестник Дона. 2021. №5. URL: <http://www.ivdon.ru/ru/magazine/archive/n5y2021/6998>.

7. Глазунов В.А. Локализация опухоли при раке молочной железы по данным микроволновой термометрии с использованием искусственных нейронных сетей // Инженерный вестник Дона. 2021. №11. URL: ivdon.ru/ru/magazine/archive/n11y2021/7296.

8. Tang X. The role of artificial intelligence in medical imaging research // BJR Open. 2019. Vol. 2, №1. P. 20190031.

9. Araújo T., Aresta G., Castro E., Rouco J., Aguiar P., Eloy C., Polónia A., Campilho A. Classification of breast cancer histology images using Convolutional Neural Networks // PLoS One. 2017. Vol. 12, № 6. P. e0177544.

10. Sarvamangala D.R., Kulkarni R.V. Convolutional neural networks in medical image understanding: a survey // Evol. Intel. 2022. Vol. 15. pp. 1–22.

11. Different types of CNN models. URL: iq.opengenus.org/different-types-of-cnn-models/, (дата обращения: 05.04.2023).

12. Xie J., Liu R., Luttrell J., Zhang C. Deep learning based analysis of histopathological images of breast cancer // Frontiers in Genetics. 2019. Vol. 10. P. 80.

13. Voon W., Hum Y., Tee Y., Yap W.S., Salim M., Tan T., Mokayed H., Lai K.W. Performance analysis of seven Convolutional Neural Networks (CNNs) with transfer learning for Invasive Ductal Carcinoma (IDC) grading in breast histopathological images // Scientific Reports. 2022. Vol. 12. P. 19200.

14. Bagchi A., Pramanik P., Sarkar R. A Multi-Stage Approach to Breast Cancer Classification Using Histopathology Images // Diagnostics. 2023. Vol. 13, №1. P. 126.

15. Gupta K., Chawla N. Analysis of Histopathological Images for Prediction of Breast Cancer Using Traditional Classifiers with Pre-Trained CNN // *Procedia Computer Science*. 2020. Vol. 167. pp. 878–889.
16. Jiang Y., Chen L., Zhang H., Xiao X. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module // *PLoS One*. 2019. Vol. 14, № 3. P. e0214587.
17. Li X., Shen X., Zhou Y., Wang X., Li T.-Q. Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet) // *PLoS ONE*. 2020. Vol. 15, №5. P. e0232127.
18. Wakili M.A., Shehu H.A., Sharif M.H., Sharif M., Umar A., Kusetogullari H., Ince I.F., Uyaver S. Classification of Breast Cancer Histopathological Images Using DenseNet and Transfer Learning // *Comput Intell Neurosci*. 2022. Vol. 2022. P. 8904768.
19. Борбат, А.М., Лищук С.В. Первый российский набор данных гистологических изображений патологических процессов молочной железы // *Врач и информационные технологии*. 2020. № 3. С. 25-30.

References

1. Wild C.P., Weiderpass E., Stewart B.W., editors. *World Cancer Report: Cancer Research for Cancer Prevention*. Lyon: International Agency for Research on Cancer, 2020. 611 p.
 2. Merabishvili V.M. *Voprosy onkologii*. 2022. Vol. 68, №3. pp. 286–293.
 3. Bhushan A., Gonsalves A., Menon J.U. *Pharmaceutics*. 2021. Vol. 68, №3. 723 p.
 4. Zeiser F. André da Costa C., Roehe A., Righi R., Marques N. *Applied Soft Computing*. 2021. Vol. 113. 107886 p.
 5. Hameed Z., Zahia S., Garcia-Zapirain B., Javier Aguirre J., María Vanegas A. *Sensors*. 2020. Vol. 20, №16. 4373 p.
-

6. Paraskevopulo, K.M., Narkevich A.N. Inzhenernyj vestnik Dona. 2021. №5. URL: ivdon.ru/ru/magazine/archive/n5y2021/6998.
 7. Glazunov V.A. Inzhenernyj vestnik Dona. 2021. №11. URL: ivdon.ru/ru/magazine/archive/n11y2021/7296.
 8. Tang X. BJR Open. 2019. Vol. 2, №1. 20190031 p.
 9. Araújo T., Aresta G., Castro E., Rouco J., Aguiar P., Eloy C., Polónia A., Campilho A. PLoS One. 2017. Vol. 12, № 6. e0177544 p.
 10. Sarvamangala D.R., Kulkarni R.V. Evol. Intel. 2022. Vol. 15. pp. 1–22.
 11. Different types of CNN models. URL: iq.opengenus.org/different-types-of-cnn-models (accessed 05.04.2023).
 12. Xie J., Liu R., Luttrell J., Zhang C. Frontiers in Genetics. 2019. Vol. 10. 80 p.
 13. Voon W., Hum Y., Tee Y., Yap W.S., Salim M., Tan T., Mokayed H., Lai K.W. Scientific Reports. 2022. Vol. 12. 19200 p.
 14. Bagchi A., Pramanik P., Sarkar R. Diagnostics. 2023. Vol. 13, №1. 126 p.
 15. Gupta K., Chawla N. Procedia Computer Science. 2020. Vol. 167. pp. 878–889.
 16. Jiang Y., Chen L., Zhang H., Xiao X. PLoS One. 2019. Vol. 14, №3. e0214587 p.
 17. Li X., Shen X., Zhou Y., Wang X., Li T.-Q. PLoS ONE. 2020. Vol. 15, №5. e0232127 p.
 18. Wakili M.A., Shehu H.A., Sharif M.H., Sharif M., Umar A., Kusetogullari H., Ince I.F., Uyaver S. Comput Intell Neurosci. 2022. Vol. 2022. 8904768 p.
 19. Borbat A.M., Lishchuk S.V. Vrach i informacionnye tekhnologii. 2020. №3. pp. 25–30.
-