

Сравнительный анализ стандартных архитектур глубокого обучения для классификации опухолей головного мозга на основе МРТ

Я. А. Низамли^{1*}, В. В. Фадел¹, А. Ю. Филатов¹,
Ю. А. Шичкина¹, К. З. Мрейш^{1,2}, С. Кархеили¹

¹Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

²Университет Аленпо

*yanizamli@stud.etu.ru

Аннотация. Магнитно-резонансная томография (МРТ) служит основным диагностическим инструментом для опухолей головного мозга из-за своей высокой точности. Однако ручная интерпретация может быть трудоемкой и подверженной человеческим ошибкам. Автоматизированные компьютерные системы помогают устранить эти ограничения, а последние достижения в области глубокого обучения предлагают многообещающую поддержку для диагностических рабочих процессов. В этом исследовании оценивается эффективность широкого спектра стандартных моделей глубокого обучения для классификации опухолей головного мозга. После пред-обработки сканы МРТ подвергаются извлечению признаков с использованием двенадцати эталонных сетей, включая варианты VGG (VGG16, VGG19), архитектуры ResNet (ResNet50, ResNet101, ResNet152), модели Inception (InceptionV3, InceptionResNetV2), конфигурации DenseNet (DenseNet121, DenseNet169, DenseNet201) и версии MobileNet (MobileNetV1, MobileNetV2). Полученные признаки затем классифицируются с использованием либо алгоритма SVM, либо алгоритма случайного леса. Мы оцениваем производительность системы по нескольким показателям, чтобы определить наиболее эффективную архитектуру. Результаты показывают, что стандартные модели глубокого обучения превосходят специально созданные сети, при этом оставляя место для дальнейшего улучшения.

Ключевые слова: медицинская визуализация, опухоли головного мозга, МРТ, глубокое обучение, предварительно обученные модели

I. ВВЕДЕНИЕ

Опухоли головного мозга представляют собой сложные неврологические расстройства, поражающие людей всех возрастных групп, часто приводящие к опасным для жизни осложнениям [1, 2]. Эти аномальные клеточные разрастания в головном мозге могут проявляться разнообразными симптомами, такими как постоянные головные боли, судороги, нарушения зрения, изменения настроения и тошнота [3]. Точная и быстрая диагностика имеет решающее значение для эффективного лечения [4]. Среди различных диагностических подходов магнитно-резонансная томография (МРТ) выделяется как эффективная и безопасная. Эта техника позволяет получать изображения с высоким разрешением за счет применения магнитных полей и радиочастотных

импульсов [5, 6]. Окончательный диагноз обычно требует оценки несколькими врачами-специалистами. Однако ручная интерпретация сканов МРТ остается трудоемкой и подверженной ошибкам, что потенциально ставит под угрозу результаты лечения пациентов [7]. Автоматизированные системы обнаружения для медицинской визуализации, включая анализ МРТ, предлагают многообещающее решение для устранения ограничений человеческой диагностики.

Последние разработки показывают, что автоматизированные системы все чаще объединяются с методами машинного обучения для повышения эффективности обработки изображений МРТ. Помимо обнаружения опухолей головного мозга (бинарная классификация), различение различных типов опухолей представляет большую сложность по мере увеличения числа классов. В [6] исследователи разработали нейронную модель, состоящую из трех блоков свертки, для дифференциации основных опухолей головного мозга: глиом, менингиом и опухолей гипофиза. Предложенная модель достигла точности 93,1%. По сравнению с традиционными методами, такими как локальный бинарный шаблон (LBP) и матрица совпадений уровней серого (GLCM), где признаки поступают в разные классификаторы, комбинация LBP+SVM достигла наилучшей точности 84,95%. Исследователи в [8] представили модель CNN с десятью последовательными слоями свертки, за которыми следует нейронный классификатор. Предложенная система протестирована для нескольких задач, включая классификацию опухолей головного мозга на изображениях МРТ, где она достигла точности 94,74%. В [9] исследователи предложили сеть извлечения признаков для изображений МРТ, состоящую из пяти блоков свертки. Использование нейронного классификатора или SVM для окончательной классификации дало точность 95,42% и 96% соответственно. В [10] и [11] примеры известной архитектуры You Only Look Once (YOLO) используются для выявления аномалий мозга. В [10] исследователи используют модифицированную версию YOLOv5m, в то время как в [11] обучается YOLOv8s. Из значений точности (precision) и полноты можно вывести оценку F1, которая достигла 89,86% и 92,47% соответственно. В [12] извлечение признаков выполняется с использованием двух стандартных моделей, InceptionV3

и Xception, по отдельности, в то время как классификация выполняется с использованием ансамблевого обучения. Точность 94,34% зарегистрирована с использованием InceptionV3, в то время как последняя достигла только 93,79%.

Разработка и обучение глубоких нейронных сетей с многочисленными параметрами требуют значительных объемов данных, что является сложным требованием для приложений медицинской визуализации, особенно МРТ. Это ограничение обычно смягчается с помощью методов аугментации, в первую очередь с использованием геометрических преобразований. Однако такие методы увеличивают вычислительные затраты, предлагая ограниченное разнообразие данных. Трансферное обучение представляет собой альтернативное решение, обходя необходимость в разработке индивидуальной сетевой архитектуры и оптимизации параметров. Тем не менее, в современной литературе отсутствует тщательное исследование разнообразной модели трансферного обучения для классификации опухолей мозга. Наша работа устраняет этот пробел с помощью трех ключевых вкладов: 1) представить набор подходов для улучшения обнаружения опухолей мозга на изображениях МРТ; 2) сравнить производительность классификации между несколькими пред-обученными моделями и определить наиболее эффективные из них; 3) оценить производительность системы с использованием различных критериев оценки.

II. МАТЕРИАЛЫ И МЕТОДЫ

A. Набор данных и предобработка

Для обучения и оценки модели мы используем набор данных МРТ головного мозга Figshare [13]. Эта коллекция содержит 3064 изображения, неравномерно распределенных по трем классам, представляющим основные типы опухолей мозга: глиомы, менингиомы и опухоли гипофиза. Пред-обработка изображений минимальна и включает только изменение размера с 512×512 до 224×224 пикселей для снижения вычислительных требований. На рис. 1 показано распределение выборок в наборе данных.

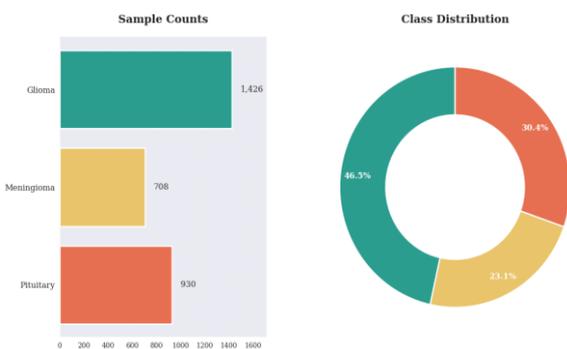


Рис. 1. Распределение данных.

B. Извлечение признаков

В этом исследовании модели глубокого обучения применяются в соответствии с подходом трансферного обучения для извлечения признаков из магнитно-резонансных изображений. Модели, пред-обученные на наборе данных ImageNet, имеют замороженные веса и удаленные окончательные слои классификации. Затем признаки выводятся путем распространения

изображений МРТ по сети. Набор эталонных моделей, используемых в этой работе, описан ниже:

1) *Сеть VGG*: VGG (Visual Geometry Group) — это простая, но высокоэффективная сверточная нейронная сеть, изначально разработанная для задач классификации изображений. Эта архитектура глубокого обучения состоит из нескольких блоков свертки и пулинга, за которыми следует нейронный классификатор. Ключевой характеристикой этой модели является ее последовательное использование небольших фильтров 3×3 во всех слоях. Исследователи разработали две основные версии этой модели: VGG16 и VGG19, которые отличаются только глубиной архитектуры. VGG16 включает 13 сверточных слоев и 3 полностью связанных слоя, в то время как VGG19 содержит 16 сверточных слоев с теми же 3 полностью связанными слоями [3, 14].

2) *Модель ResNet*: Остаточная сеть состоит из глубокой последовательности сверточных блоков, включающих пропускные соединения. Эти остаточные соединения, также называемые пропускными соединениями, устанавливают прямые пути, которые обходят один или несколько слоев, чтобы соединить выход слоя с входом последующего слоя, как показано на рис. 2. Такие соединения помогают улучшить градиентный поток во время обратного распространения, тем самым смягчая проблему исчезающего градиента [7]. Архитектура реализована в нескольких вариантах, включая ResNet50, ResNet101 и ResNet152. Числовой суффикс в каждом случае указывает общее количество обучаемых слоев, что соответствует 16, 33 и 50 остаточным блокам соответственно. Хотя более глубокие архитектуры демонстрируют повышенную точность на наборе данных ImageNet, это улучшение требует больших вычислительных затрат.

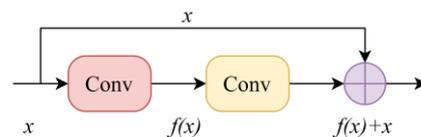


Рис. 2. Остаточный блок.

3) *Архитектура Inception*: Эта модель глубокого обучения вводит блок Inception, включающий несколько сверточных слоев, работающих одновременно. Такая конструкция позволяет захватывать разнообразные, многомасштабные признаки, которые улучшают производительность классификации [15]. InceptionV3 содержит 94 сверточных слоя, в то время как InceptionResNetV2 расширяется до 164 сверточных слоев, которые, как указано в названии, используют пропуски соединений. На рис. 3 показана начальная конфигурация блока Inception для обеих моделей.

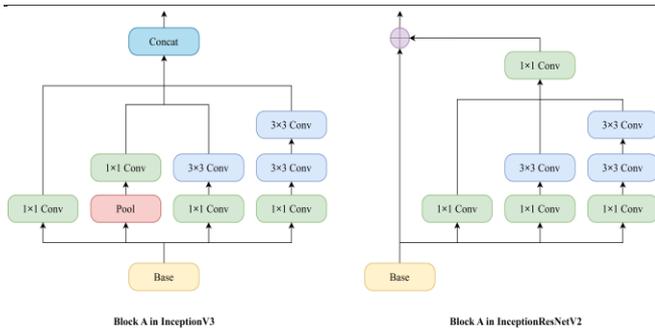


Рис. 3. Блок Inception.

4) *Модель DenseNet*: эта сверточная сеть использует плотные соединения, где каждый слой соединяется со всеми предыдущими слоями, обеспечивая надежное повторное использование признаков [16]. На рис. 4 показана визуализация плотного блока. Подобно пропускам соединений, плотные соединения решают проблему исчезающих градиентов, но достигают превосходного сохранения признаков, поскольку они используют конкатенацию, а не суммирование. Реализации DenseNet различаются по глубине, на что указывает их общий вес слоев. Три распространенных варианта включают DenseNet121, DenseNet169 и DenseNet201.

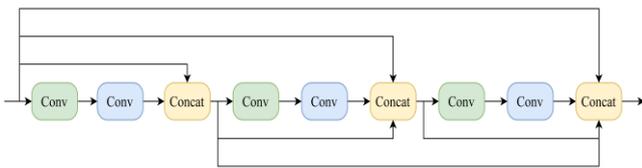


Рис. 4. Блок Dense.

5) *Структура MobileNet*: Эта сверточная сеть специально разработана для мобильных устройств и встроенных систем с ограниченными вычислительными ресурсами. Первоначальная версия использует блоки глубинно-разделимой свертки (DWS), объединяющие depthwise и pointwise свертки для оптимизации эффективности [15]. В MobileNetV2 слой расширения предшествует процессу DWS, улучшая представление признаков [17]. На рис. 5 сравниваются основные строительные блоки MobileNetV1 и MobileNetV2, демонстрируя включение пропускных соединений в обновленной версии.

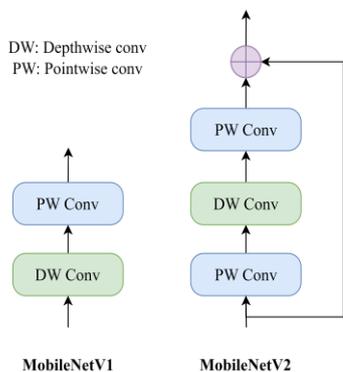


Рис. 5. Блок MobileNet.

С. Классификация

Для задачи классификации используются два различных алгоритма машинного обучения: Support

Vector Machine (SVM) и Random Forest (RF). Алгоритм SVM определяет оптимальную гиперплоскость для разделения выборок классов в пространстве признаков, тогда как RF действует как ансамблевый метод, который объединяет прогнозы из нескольких деревьев решений. Из-за их различных принципов работы один алгоритм может демонстрировать более высокую производительность обобщения по сравнению с другим. Оптимизация гиперпараметров выполняется эвристически; классификатор SVM использует ядро RBF с параметром регуляризации, установленным на 10, в то время как реализация случайного леса использует 100 оценщиков, ограниченных максимальной глубиной 25.

III. РЕЗУЛЬТАТЫ, ОБСУЖДЕНИЕ И БУДУЩАЯ РАБОТА

Все эксперименты проводились на notebook Python с использованием графического процессора NVIDIA GeForce RTX 4070 и процессора Intel Core i7 12650H 12-го поколения. Данные были перетасованы, 80% выделено для обучения и 20% зарезервировано для тестирования. На рис. 6 представлена тепловая карта, иллюстрирующая точность оцениваемых систем, а в табл. 1 представлены отчеты о классификации, включая показатели точности (precision), полноты и оценки F1. При использовании классификатора SVM большинство моделей глубокого обучения достигли высокой точности, достигнув пика в 96.08% как для ResNet152, так и для DenseNet121. Модели Inception продемонстрировали самую низкую производительность с точностью всего 92.50% для InceptionV3 и 83.69% для InceptionResNetV2. Остальные сети показали сопоставимые результаты, в диапазоне от 94.13% до 95.92%. С классификатором RF производительность заметно упала, при этом наивысшая точность достигла 92.33% для DenseNet121. Наблюдаемое превосходство алгоритма SVM может быть обусловлено его эффективностью в многомерных пространствах признаков по сравнению с RF, особенно для таких моделей, как ResNet, которые производят более 100000 признаков. Для большинства моделей значения точности близко соответствовали оценкам F1, что предполагает минимальное смещение класса, несмотря на дисбаланс набора данных.

Несколько процедур и подходов могут быть использованы для повышения точности ранее исследованных систем для идентификации опухолей мозга, каждая из которых потенциально может служить основой для отдельных будущих исследований. К ним относятся: 1) обработка изображений МРТ, улучшение качества и шумоподавление, которые могут улучшить результаты модели; 2) расширение набора данных посредством аугментации, либо во входном пространстве (например, геометрические методы), либо в пространстве признаков (например, SMOTE и его варианты); 3) оптимизация извлеченных глубоких признаков для устранения избыточности и нерелевантной информации, потенциально используя такие методы, как анализ главных компонент (PCA), выбор признаков на основе дерева или регуляризация L1; 4) более тонкая настройка пред-обученных моделей и классификаторов, включая целевое извлечение признаков из определенных слоев или автоматическую оптимизацию гиперпараметров. Реализация одного или комбинации этих методов может повысить надежность и производительность предлагаемых систем.

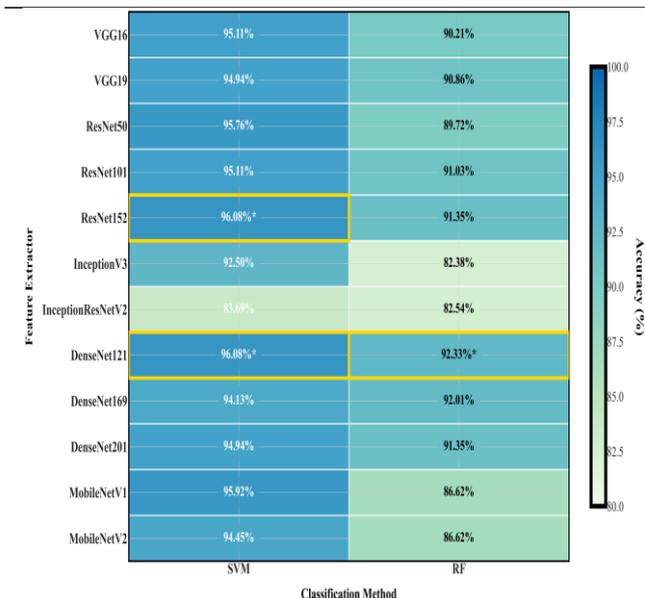


Рис. 6. Эффективность представленных подходов, измеренная по точности.

ТАБЛИЦА I. ОТЧЕТЫ О КЛАССИФИКАЦИИ

Model	Class	PRE	SEN	F1
VGG16	SVC	94.38	95.13	94.72
	RF	89.52	88.06	88.55
VGG19	SVC	94.16	94.89	94.50
	RF	90.24	88.99	89.38
ResNet50	SVC	95.05	95.74	95.37
	RF	88.54	88.19	88.28
ResNet101	SVC	94.57	94.41	94.47
	RF	90.16	89.15	89.48
ResNet152	SVC	95.26	96.21	95.69
	RF	89.97	90.09	89.99
InceptionV3	SVC	91.56	92.34	91.90
	RF	81.38	80.28	80.48
InceptionResNetV2	SVC	82.31	82.64	82.40
	RF	80.67	80.80	80.73
DenseNet121	SVC	95.38	95.93	95.65
	RF	92.05	91.03	91.47
DenseNet169	SVC	93.44	94.01	93.70
	RF	91.38	90.88	91.08
DenseNet201	SVC	94.02	95.18	94.52
	RF	90.68	89.78	90.12
MobileNetV1	SVC	95.34	95.71	95.51
	RF	86.48	83.11	83.94
MobileNetV2	SVC	94.03	93.69	93.84
	RF	86.36	83.50	84.14

IV. ЗАКЛЮЧЕНИЕ

В этом исследовании изучалась эффективность различных пред-обученных глубоких сетей для классификации опухолей головного мозга на изображениях МРТ. Исследование включало все варианты архитектур VGG, ResNet, Inception, DenseNet и MobileNet, охватывая в общей сложности 12 моделей. Для окончательной классификации использовались два классификатора: SVM и RF. Эксперименты проводились с использованием набора данных Figshare для эталонных тестов, содержащего случаи глиомы, менингиомы и

опухоли гипофиза. Оценка с использованием нескольких метрик показала высокую производительность для глубоких моделей в паре с классификаторами SVM, в то время как классификаторы Random Forest дали сравнительно более низкие результаты. Эти результаты указывают на сильную способность систем определять аномалии на изображениях МРТ, что предполагает их потенциальную пригодность для клинических приложений. В исследовании также было предложено несколько дополнительных подходов, которые могли бы повысить производительность систем на основе глубокого обучения для классификации медицинских изображений, особенно при обнаружении опухолей головного мозга.

СПИСОК ЛИТЕРАТУРЫ

- [1] Y. Nizamli, A. Filatov, "MRI brain tumor classification using HOG features selected via impurity-based importances measure," IJEER, vol. 12, no. 4, pp. 1251–1257, 2024.
- [2] H. A. Alahmed, G. A. Al-Suhail, "Exploring transfer learning techniques for brain tumor diagnosis in MRI data," 2024 ICETI, pp. 1–8, 2024.
- [3] Y. Nizamli, A. Filatov, "Improving transfer learning performance for abnormality detection in brain MRI images using feature optimization techniques," 2024 SCM, pp. 432–435, 2024.
- [4] D. Rohini, A. Pooja, Sindhu, and S. H. Naaz, "Advancing brain tumor classification with deep learning," Recent Trends in Computer Graphics and Multimedia Technology, vol. 7, no. 2, pp. 1–10, 2025.
- [5] S. Jain, V. Jain, and J. M. Chatterjee, "Ensemble based brain tumor classification technique from MRI based on K fold validation approach," J. Integr. Sci. Technol., vol. 13, no. 5, 2025.
- [6] S. R. Sowrirajan, S. Balasubramanian, "Brain tumor classification using machine learning and deep learning algorithms," IJEER, vol. 10, no. 4, pp. 999–1004, 2022.
- [7] Y. Nizamli, A. Filatov, W. Fadel, and Yu Shichkina, "Feature processing techniques for enhancing deep brain MRI classification systems," 2025 EICon, 2025. In press.
- [8] W. Ayadi, W. Elhamzi, I. Charfi, and M. Atri, "Deep CNN for brain tumor classification," Neural Processing Letters, vol. 53, pp. 671–700, 2021.
- [9] A. Biswas, M. S. Islam, "A Hybrid deep CNN-SVM approach for brain tumor classification," Journal of Information Systems Engineering and Business Intelligence, vol. 9, no. 1, pp. 1–15, 2023.
- [10] S. Muksimova, S. Umirzakova, S. Mardieva, N. Iskhakova, M. Sultanov, and Y. I. Cho, "A lightweight attention-driven YOLOv5m model for improved brain tumor detection," Computers in Biology and Medicine, vol. 188, 2025.
- [11] R. S. Passa, S. Nurmaini, and D. P. Rini, "YOLOv8 based on data augmentation for MRI brain tumor detection," Scientific Journal of Informatics, vol. 10, no. 3, 2023.
- [12] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Alassafi "Brain tumor classification based on fine-tuned models and the ensemble method," Computers, Materials & Continua, vol. 67, no. 3, pp. 3967–3982, 2021.
- [13] J. Cheng, "brain tumor dataset," Figshare, 2017. https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5 (accessed 12.02.2025).
- [14] N. Pokhriyal, A. Jain, "Brain tumor classification through MRIs using transfer learning with VGG16 model," IJRASET, vol. 13, no. 1, 2025.
- [15] Y. Nizamli, A. Filatov, W. Fadel, Yu Shichkina, K. Mreish, and T. Alnajjar, "Impact of augmentation techniques on transfer learning performance for brain MRI classification: A comparative study of spatial and feature space transformations," 2025 SCM, 2025. In press.
- [16] S. U. R. Khan, M. Zhao, S. Asif, and X. Chen, "Hybrid-NET: A fusion of DenseNet169 and advanced machine learning classifiers for enhanced brain tumor diagnosis," International Journal of Imaging Systems and Technology, vol. 34, no. 1, 2024.
- [17] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," arXiv, 2019.