УДК 00.009.43

ОПТИМИЗАЦИЯ АРХИТЕКТУРЫ CONVNEXT ДЛЯ ЗАДАЧИ КЛАССИФИКАЦИИ ДЕРМАТОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ

Н.К. Захаров[1], И.А. Матвеева[1]

[1]Самарский университет, г. Самара
e-mail: nk.z4kharov@gmail.com

Аннотация

Исследованы 49 конфигураций нейросетевой классификации дерматоскопических изображений из набора HAM10000. Исследовано влияние на эффективность классификации предварительной обработки данных, балансировки классов, контролируемого контрастивного обучения, использования фокальной функции потерь, калибровки вероятностей и подбора порога. Показано, что лучший общий баланс метрик дает сочетание ConvNeXt, весов классов и контрастивного обучения, а максимальная чувствительность идентификации злокачественной меланомы достигается при перевыборке редких классов.

Ключевые слова: дерматоскопия, классификация новообразований кожи, злокачественная меланома, ConvNeXt, калибровка вероятностей, подбор порога.

Введение

Набор HAM10000 содержит 10015 дерматоскопических изображений семи диагностических классов и широко используется для сопоставления алгоритмов компьютерной диагностики [1]. Соревнования ISIC показали, что для медицинских систем важны не только общая точность, но и устойчивость к дисбалансу классов, обобщающая способность и интерпретируемые вероятностные оценки [2].

Современные обзоры указывают, что практическую ценность определяют чувствительность, специфичность, коэффициент корреляции Мэттьюса (MCC) и калибровка [3]. Цель работы - выбрать конфигурацию для сценария, где пропуск злокачественной меланомы имеет более высокую цену, чем дополнительное обследование [4].

1. Материалы и методы

Использован набор HAM10000 с семью классами кожных новообразований [1]. Разбиение train/valid/test выполнялось по идентификатору очага в пропорции 70/10/20 со стратификацией по диагнозу, что снижает риск утечки данных между наборами.

Базовая архитектура построена на ConvNeXt-Large [5]. В реализованной модели общий сверточный извлекатель признаков дополняется классификационным и проекционным блоками. Проекционный блок формирует 128-мерные нормализованные признаки для контролируемого контрастивного обучения [6].

В серии из 49 экспериментов варьировались тип данных, функция потерь, веса классов, перевыборка/уменьшение выборки, сглаживание меток и контрастивное обучение. Качество оценивалось по сбалансированной точности, коэффициенту

корреляции Мэттьюса (MCC), макро-усредненной площади под кривой «точность–полнота» (PR-AUC), площади под ROC-кривой, а также по сценариям «злокачественные против доброкачественных» и «злокачественная меланома против остальных». После обучения выполнялись температурное масштабирование и подбор порога по индексу Юдена [7].

Фокальная функция потерь использовалась из-за дисбаланса классов: она уменьшает вклад легко классифицируемых примеров и переносит акцент на трудные случаи [8]:

$$FL(p_t) = -(1 - p_t)^{\gamma} \log(p_t), \quad (1)$$

где p_t – оцененная моделью вероятность истинного класса, а γ задает степень подавления вклада легко классифицируемых наблюдений.

2. Результаты и обсуждение

На уровне средних значений лучший результат дали эксперименты на исходных изображениях: средняя сбалансированная точность составила 0,770, тогда как для первой схемы предобработки она равнялась 0,747, для второй – 0,697, а для набора, расширенного синтетическими изображениями – 0,741. Следовательно, агрессивная коррекция цвета и контраста не гарантирует улучшения и может ослаблять тонкие признаки, значимые для дерматоскопии. Контролируемое контрастивное обучение оказалось наиболее устойчивым фактором улучшения: в среднем оно повысило коэффициент Мэттьюса с 0,676 до 0,720 и F1-меру при выявлении злокачественной меланомы с 0,566 до 0,625. Такой результат согласуется с [6], где показано, что контрастивный функционал улучшает разделимость классов и повышает устойчивость модели. Лучшую совокупность показателей показала конфигурация 24_raw_supcon_weighted: сбалансированная точность 0,8188, коэффициент Мэттьюса 0,7323 и макро-ROC-AUC 0,9526. Максимальный коэффициент Мэттьюса 0,7478 был получен у 20_raw_supcon, однако эта модель уступила лидеру по сбалансированной точности.

Таблица 1 показывает два режима применения. Для минимизации пропусков меланомы предпочтительна перевыборка редких классов и кросс-энтропийная функция потерь с чувствительностью 0,9178. Для сбалансированного решения подходит контрастивное обучение с весами классов, где специфичность при выявлении злокачественной меланомы достигла 0,9051.

Таблица 1 – Ключевые конфигурации и их диагностические свойства

Конфигурация	**Сбал. точность**	**Коэф. Мэттьюса**	**Чувствительность / специфичность**	**Вывод**
Кросс-энтропия	0,7676	0,7341	0,8995 / 0,8310	Базовый уровень
Контрастивное обучение	0,7574	0,7478	0,8676 / 0,8748	Максимум коэффициента Мэттьюса
Контрастивное обучение и веса	0,8188	0,7323	0,8128 / 0,9051	Лучший баланс
Перевыборка и кросс-энтропия	0,7873	0,7206	0,9178 / 0,8276	Макс. чувствительность

Совокупное качество – совместный учет коэффициента корреляции Мэттьюса и сбалансированной точности по семи классам; соотношение чувствительности и специфичности - выбор рабочего порога при выявлении меланомы.

На рисунке 1а видно, что конфигурация 24_raw_supcon_weighted формирует лучший баланс между коэффициента корреляции Мэттьюса и сбалансированной точностью, тогда как 20_raw_supcon остается сильной альтернативой при максимизации коэффициента корреляции Мэттьюса. На рисунке 1, б специфичности обычно снижает.

Температурное масштабирование снизило ожидаемую ошибку калибровки с 0,0819 до 0,0670. Для базовой модели подбор порога увеличил специфичность при выявлении меланомы с 0,8007 до 0,8310 при снижении чувствительности с 0,9269 до 0,8995; поэтому оба этапа следует считать частью диагностической системы.

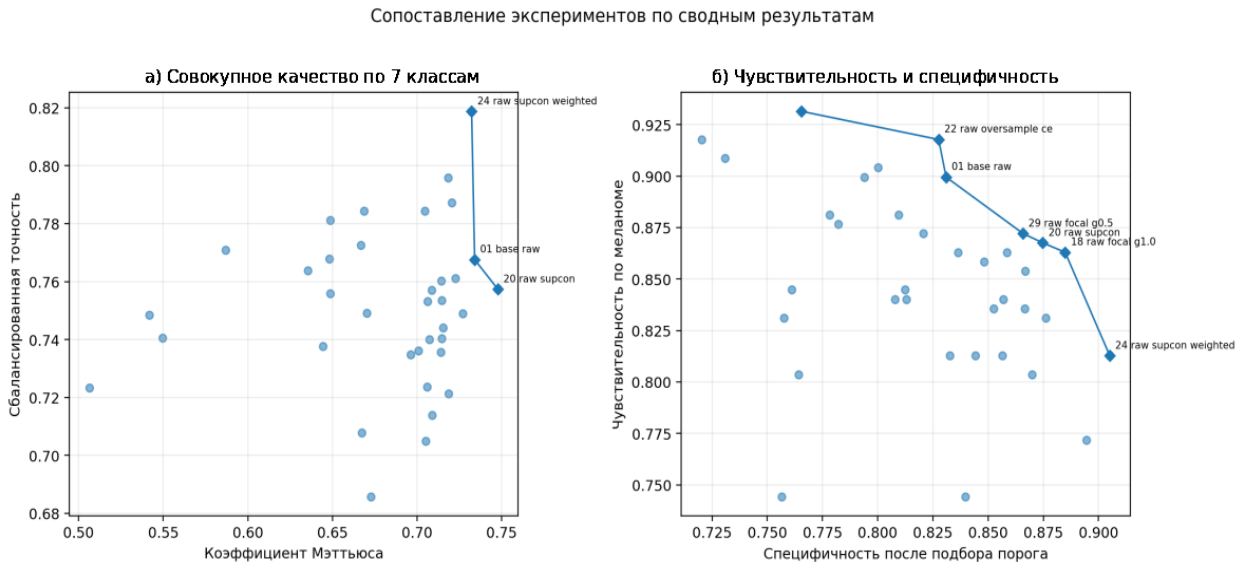

Рисунок 1 – Сопоставление экспериментов по совокупному качеству многоклассовой классификации и соотношению чувствительности и специфичности при выявлении злокачественной меланомы

Результаты согласуются с клинической ролью дерматоскопии: при первичном отборе важна высокая чувствительность, а в системах поддержки решений – специфичность и калиброванная вероятностная оценка [4].

Заключение

Сравнение 49 экспериментов показало, что на HAM10000 лучшие результаты обеспечивает не усложнение предобработки, а работа с дисбалансом классов, контрастивное обучение и калибровка. ConvNeXt с контролируемым контрастивным обучением и весами классов дал лучший баланс метрик, а перевыборка редких классов оказалась полезна для максимальной чувствительности при выявлении злокачественной меланомы. Практически целесообразны два режима: скрининговый с повышенной чувствительностью и экспертный с более высокой специфичностью. Далее необходима внешняя валидация и учет метаданных пациентов.

Список использованных источников

1. Tschandl, P. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions / P. Tschandl, C. Rosendahl, H. Kittler // Scientific Data. – 2018. – V. 5, № 180161.

2. Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC) / N. Codella, V. Rotemberg, P. Tschandl [et al.] // arXiv. – 2019. – № 1902.03368.

3. Skin Cancer Classification With Deep Learning: A Systematic Review / Y. Wu, B. Chen, A. Zeng [et al.] // Frontiers in Oncology. – 2022. – V. 12, № 893972.

4. Using Dermoscopy to Identify Melanoma and Improve Diagnostic Discrimination / G. A. Holmes, J. M. Vassantachart, B. A. Limone [et al.] // Federal Practitioner. – 2018. – V. 35, № 4. – P. S39–S45.

5. A ConvNet for the 2020s / Z. Liu, H. Mao, C.-Y. Wu [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2022. – P. 11976–11986.

6. Supervised Contrastive Learning / P. Khosla, P. Teterwak, C. Wang [et al.] // Advances in Neural Information Processing Systems. – 2020. – T. 33. – C. 18661–18673.

7. Focal Loss for Dense Object Detection / T.-Y. Lin, P. Goyal, R. Girshick [et al.] // Proceedings of the IEEE International Conference on Computer Vision. – 2017. – P. 2980–2988.

8. On Calibration of Modern Neural Networks / C. Guo, G. Pleiss, Y. Sun, K. Q. Weinberger // Proceedings of the 34th International Conference on Machine Learning. – 2017. – P. 1321–1330.

OPTIMIZATION OF THE CONVNEXT ARCHITECTURE FOR DERMATOSCOPIC IMAGE CLASSIFICATION

N.K. Zakharov[1], I.A. Matveeva[1]

[1]Samara University, Samara
e-mail: nk.z4kharov@gmail.com

Abstract

49 ConvNeXt configurations were evaluated on HAM10000 with varying preprocessing, class balancing, supervised contrastive learning, focal loss, calibration, and thresholds. Class weights with contrastive learning gave the best metric balance; rare-class oversampling gave the highest melanoma sensitivity.

Keywords: dermatoscopy, skin neoplasm classification, malignant melanoma, ConvNeXt, probability calibration, threshold tuning.