

# Применение безопасного фреймворка на основе визуальной языковой модели для совместного анализа дерматологических изображений

Ц. Ян

Университет ИТМО  
tszyafengyang@gmail.com

А. Н. Субботин

Санкт-Петербургский  
государственный  
электротехнический  
университет «ЛЭТИ»  
им. В.И. Ульянова (Ленина)  
alesu1543@gmail.com

Н. А. Жукова

Санкт-Петербургский  
институт информатики и  
автоматизации Российской  
академии наук  
nazhukova@mail.ru

**Аннотация.** Обеспечение безопасного и конфиденциального анализа медицинских данных остается критической задачей в здравоохранении, особенно когда чувствительная информация о пациентах распределена между учреждениями. В данной статье применяется безопасный визуально-языковой фреймворк, разработанный для совместного анализа дерматологических изображений, с фокусом на задачу классификации DermMNIST. Используя вертикальное федеративное обучение и визуально-языковую модель CLIP, фреймворк позволяет распределенным учреждениям совместно обучать модели без обмена необработанными данными. Благодаря интеграции настройки промптов, фреймворк адаптирует CLIP для классификации медицинских изображений, сохраняя при этом конфиденциальность данных. Экспериментальные результаты демонстрируют высокую производительность на DermMNIST, достигая показателя AUC 0,933 и точности 0,793, превосходя существующие базовые показатели.

**Ключевые слова:** визуальная языковая модель, совместный анализ, дерматология, вертикальное федеративное обучение, анализ медицинских изображений

## I. ВВЕДЕНИЕ

Взаимодействие искусственного интеллекта и здравоохранения в последние годы стало свидетелем замечательного прогресса, особенно в области анализа медицинских изображений. Дерматология, с ее по своей сути визуальным диагностическим подходом, стала перспективной областью для применения ИИ. Модели глубокого обучения продемонстрировали впечатляющие возможности в классификации кожных поражений, в некоторых случаях достигая диагностической точности, сравнимой с сертифицированными дерматологами [1, 2]. Эти достижения открывают значительный потенциал для улучшения доступности дерматологической экспертизы, снижения диагностических вариаций и поддержки клинических решений как при наличии достаточных ресурсов, так и в условиях ресурсных ограничений.

Несмотря на эти многообещающие разработки, широкомасштабное внедрение систем ИИ в

дерматологическую практику сталкивается с несколькими критическими проблемами. Главная из них – сохранение конфиденциальности при анализе персональных данных пациентов. Медицинские изображения, включая дерматологические фотографии, содержат защищенную медицинскую информацию, попадающую под строгие нормативные требования. Традиционные подходы к разработке надежных моделей ИИ требуют централизованного сбора данных, что приводит к проблеме нарушения конфиденциальности. При этом обучение моделей на изолированных институциональных наборах данных приводит к ограниченной обобщаемости и потенциальным алгоритмическим смещениям, что ухудшает производительность при развертывании в различных клинических условиях [3].

Усугубляет эти проблемы сложность дерматологической диагностики, требующей визуального распознавания образов с контекстуальной лингвистической информацией. Хотя традиционные сверточные нейронные сети (CNN) показали впечатляющие результаты в задачах классификации, им не хватает способности включать текстовые клинические описания и медицинскую терминологию, которые дерматологи-люди регулярно используют в диагностическом процессе [4]. Это ограничение сужает интерпретируемость и клиническую полезность чисто визуальных моделей в дерматологических приложениях.

В данной статье мы применяем новый фреймворк, ранее предложенный и представленный на конференции ICCSA [20], для решения задачи дерматологической классификации. Результаты применения демонстрируют, как фреймворк, объединяющий семантическое богатство визуально-языковых моделей с возможностями сохранения конфиденциальности вертикального федеративного обучения, может быть эффективно использован для совместного анализа дерматологических изображений. Подход использует модель CLIP (Contrastive Language-Image Pre-training) [5], которая продемонстрировала высокие возможности в соединении визуального контента с текстовыми описаниями в различных областях. Интегрируя CLIP с архитектурой вертикального федеративного обучения, мы позволяем нескольким медицинским учреждениям

совместно улучшать диагностическую эффективность без обмена необработанными данными пациентов, тем самым сохраняя конфиденциальность, используя при этом коллективные знания, встроенные в распределенные наборы данных.

Кроме того, применяемый фреймворк использует передовой метод настройки промптов Context Optimization (CoOp) [6] для эффективной адаптации модели CLIP к специализированной области дерматологической диагностики. Этот подход позволяет нам точно настраивать только небольшой набор непрерывных векторов промптов, сохраняя предобученную модель замороженной, что значительно снижает вычислительные требования, сохраняя при этом высокую производительность.

Мы оцениваем наш фреймворк на наборе данных DermaMNIST [8], стандартизированной коллекции дерматоскопических изображений, полученных из набора данных HAM10000 [9]. Наши эксперименты демонстрируют высокую производительность на DermaMNIST, достигая AUC 0,933 и точности 0,793, превосходя существующие базовые показатели.

## II. ОБЗОР ПРЕДМЕТНОЙ ОБЛАСТИ

Анализ медицинских изображений претерпел существенную трансформацию с появлением методов глубокого обучения, устанавливая новые эталоны в различных областях здравоохранения, включая дерматологию, радиологию и патологию [10, 11]. Применение сверточных нейронных сетей (CNN) к классификации дерматологических изображений продемонстрировало особенно перспективные результаты, причем несколько исследований сообщают о диагностической точности, сравнимой или превосходящей точность сертифицированных дерматологов [1, 2, 12]. Стандартизация наборов данных дерматологических изображений, примером которых являются коллекция HAM10000 [9] и ее производные, такие как DermaMNIST [8], способствовала воспроизводимым исследованиям и значимым сравнениям производительности между различными алгоритмическими подходами. Набор данных DermaMNIST, часть коллекции MedMNIST v2, состоит из 10 015 дерматоскопических изображений, категоризированных по семи типам кожных поражений. Примечательно, что хотя MedMNIST стандартизирует эти изображения до единого низкого разрешения 28×28 пикселей для облегчения проведения экспериментов, данный датасет также предоставляет доступ к изображениям более высокого разрешения. Этот подход с несколькими разрешениями позволяет исследователям сбалансировать вычислительную эффективность с диагностическими деталями, предлагая гибкость для различных экспериментальных требований. Эталон включает стандартные разделения на обучающую, валидационную и тестовую выборки, что еще больше улучшает воспроизводимость и обеспечивает справедливое сравнение между исследованиями.

Несмотря на эти технологические достижения, широкомасштабное внедрение диагностических инструментов на основе ИИ в клиническую практику сталкивается со значительными барьерами, главным из

которых являются опасения по поводу конфиденциальности данных. Медицинские данные строго регулируются для защиты конфиденциальности пациентов, но эти нормы часто препятствуют обмену обучающими данными между учреждениями, ограничивая разнообразие и размер наборов данных, доступных для разработки модели [13]. Как следствие, модели, обученные на данных одного учреждения, могут испытывать трудности с обобщением при развертывании в различных клинических условиях.

Федеративное обучение возникло как перспективная парадигма для решения проблем конфиденциальности, позволяя обучать модели на распределенных наборах данных без обмена необработанными данными [14]. Традиционные подходы федеративного обучения можно разделить на горизонтальное федеративное обучение (HFL), где участники имеют одинаковое пространство признаков, но разные выборки, и вертикальное федеративное обучение (VFL), где участники имеют разные признаки для одних и тех же или перекрывающихся выборок [15]. В контексте анализа медицинских изображений VFL особенно актуально, поскольку оно позволяет учреждениям с дополнительными типами данных (например, данными визуализации и клиническими записями) совместно обучать модели, сохраняя конфиденциальность данных [16].

Параллельно с разработками в области больших языковых моделей, визуально-языковые модели, такие как CLIP (Contrastive Language-Image Pre-training), продемонстрировали замечательные способности в соединении визуального контента с текстовыми описаниями [5]. CLIP, обученный на 400 миллионах пар изображение-текст из интернета, изучает совместное пространство вложений для изображений и текста, обеспечивая классификацию с нулевым выстрелом и семантическое понимание в различных визуальных областях. Недавние исследования показали многообещающие результаты в адаптации CLIP для задач медицинской визуализации, включая интерпретацию рентгенограмм грудной клетки [17] и анализ патологических изображений [18], что предполагает его потенциальную полезность для дерматологических приложений.

Важным достижением в эффективной адаптации крупных предобученных моделей является настройка промптов, которая модифицирует небольшой набор непрерывных векторов (промптов), сохраняя параметры предобученной модели замороженными [19]. Этот подход значительно снижает вычислительные требования по сравнению с полной тонкой настройкой, сохраняя при этом конкурентоспособную производительность. Среди методов настройки промптов для визуально-языковых моделей Context Optimization (CoOp) [6] стал особенно эффективным подходом. Вместо использования текстовых промптов, созданных вручную, CoOp изучает набор непрерывных контекстных векторов, оптимизированных для конкретной нисходящей задачи. Эти обучаемые контекстные векторы служат префиксом к названиям классов в шаблоне промпта (например, «фото [контекстных векторов] [класса]»), позволяя модели

адаптироваться к доменно-специфическим визуальным концепциям без изменения параметров предобученной модели.

Интеграция визуально-языковых моделей, вертикального федеративного обучения и передового метода настройки промптов CoOp представляет собой многообещающий подход к совместному анализу медицинских изображений, который балансирует диагностическую производительность и обеспечит сохранение конфиденциальности. Эта работа применяет безопасный фреймворк на основе визуально-языковой модели к совместному анализу дерматологических изображений, фокусируясь на задаче классификации DermaMNIST. Мы используем вертикальное федеративное обучение для обеспечения сохранения конфиденциальности при сотрудничестве между учреждениями, собирающими разные типы данных, и используем методы настройки промптов CoOp для эффективной адаптации модели CLIP для классификации дерматологических изображений, сохраняя при этом сильные возможности обобщения.

### III. МЕТОДОЛОГИЯ

Применяемый фреймворк интегрирует три ключевых компонента: (1) модель контрастного предобучения изображений и языка (CLIP) в качестве базовой модели, (2) архитектуру вертикального федеративного обучения для сотрудничества с сохранением конфиденциальности, и (3) передовые подходы к настройке промптов (CoOp) для адаптации CLIP к классификации дерматологических изображений. На рис. 1 показана общая архитектура фреймворка [20].

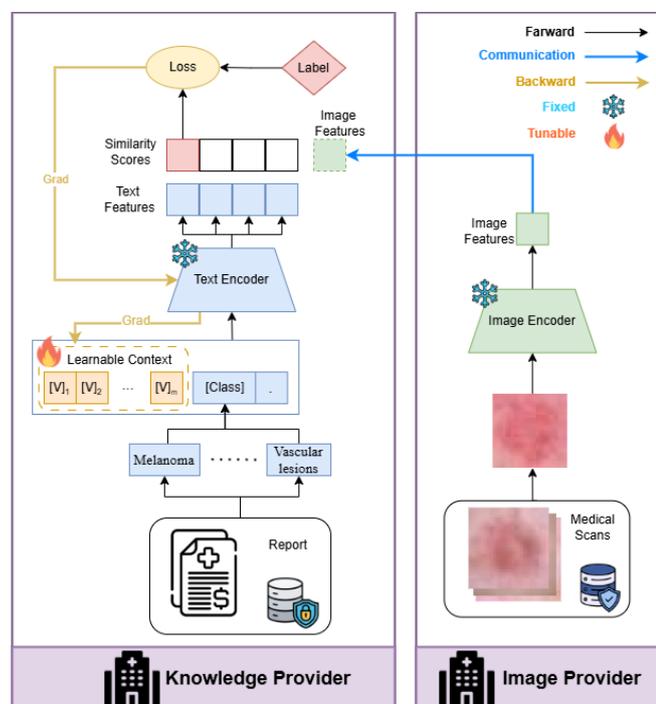


Рис. 1. Общая архитектура фреймворка VFL-CLIP [20]

#### A. Визуально-языковая модель CLIP

CLIP (Contrastive Language-Image Pre-training) [5] состоит из двойных энкодеров: энкодера зрения,

который обрабатывает изображения, и энкодера текста, который обрабатывает текстовые описания. Во время предобучения CLIP учится выравнивать изображения с соответствующими текстовыми описаниями в общем пространстве вложений. Для нашей реализации мы используем различные варианты CLIP, включая модели на основе ResNet и ViT.

#### A. Фреймворк вертикального федеративного обучения

Наш фреймворк вертикального федеративного обучения (VFL) разработан для сценариев, когда несколько учреждений сотрудничают при анализе дерматологических изображений без обмена необработанными данными пациентов. Мы рассматриваем два основных типа участников:

- Поставщики изображений: Медицинские учреждения, обладающие дерматологическими изображениями, но которым может не хватать специализированной диагностической экспертизы или достаточных данных для надежного обучения модели.
- Поставщики знаний: Специализированные дерматологические центры с обширной диагностической экспертизой, которые могут предоставлять текстовые описания и знания по классификации.

В нашей настройке VFL, поставщики изображений используют энкодер зрения CLIP для извлечения вложений признаков из своих локальных изображений без обмена необработанными изображениями. Эти вложения, которые не содержат непосредственно идентифицируемой информации о пациенте, затем передаются безопасному агрегатору. Одновременно, поставщики знаний используют текстовый энкодер CLIP для обработки описаний дерматологических состояний и генерации соответствующих текстовых вложений.

#### B. Настройка промпта с помощью CoOp

Для адаптации CLIP к конкретным требованиям классификации дерматологических изображений мы используем передовой подход к настройке промптов - Context Optimization (CoOp).

CoOp заменяет текстовые промпты, формируемые вручную, обучаемыми непрерывными контекстными векторами, оптимизированными для конкретной нисходящей задачи. Для каждого дерматологического состояния в наборе данных DermaMNIST мы адаптируем шаблон промпта: «Дерматологическое изображение [КОНТЕКСТ] [КЛАСС]», где [КОНТЕКСТ] заменяется  $n$  обучаемыми контекстными векторами  $P = \{p_1, p_2, \dots, p_n\}$ , а [КЛАСС] – это название дерматологического состояния (например, «меланома», «невус» и т. д.). Эти контекстные векторы случайно инициализируются, а затем оптимизируются во время обучения, сохраняя предобученную модель CLIP замороженной.

Целью оптимизации для CoOp является максимизация сходства между вложениями изображений и соответствующими вложениями текста для конкретных классов:

$$\mathcal{L}_{\text{CoOp}} = -\log\left(\frac{\exp(\text{sim}(z_v, z_t^y))}{\sum_{k=1}^C \exp(\text{sim}(z_v, z_t^k))}\right), \quad (1)$$

где  $z_v$  – это вложение изображения,  $z_t^y$  – вложение текста для истинного класса  $y$ , а  $z_t^k$  – текстовые вложения для всех классов  $k$  в наборе из  $C$  классов. Функция  $\text{sim}(\cdot, \cdot)$  вычисляет косинусное сходство между вложениями.

#### IV. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

##### A. Набор данных

Мы оцениваем данный фреймворк, используя набор данных *DermaMNIST* из *MedMNIST v2* [8], который содержит 10 015 дерматоскопических изображений по 7 категориям кожных поражений: актинический кератоз и внутриэпителиальная карцинома (*akiec*), базальноклеточная карцинома (*bcc*), доброкачественные кератозоподобные поражения (*bkl*), дерматофиброма (*df*), меланома (*mel*), меланоцитарные невусы (*nv*) и сосудистые поражения (*vasc*). Следуя официальному разделению в *MedMNIST v2*, мы используем 7 007 изображений для обучения, 1 003 для валидации и 2 005 для тестирования. Хотя стандартный набор данных *DermaMNIST* предоставляет изображения размером  $28 \times 28$  пикселей, мы используем изображения более высокого разрешения ( $224 \times 224$ ,  $336 \times 336$ ,  $488 \times 488$ ) для наших экспериментов, чтобы позволяло обеспечить более эффективное извлечение признаков с помощью модели CLIP.

##### B. Детали реализации

Для настройки промптов CoOp мы используем 16 обучаемых контекстных токенов, инициализированных случайным образом, и оптимизируем их с помощью оптимизатора Adam с скоростью обучения 0,001. Все эксперименты проводились на сервере с GPU NVIDIA A100 и 64 ГБ ОЗУ. Для нашей визуально-языковой модели в качестве основы мы экспериментировали с несколькими вариантами CLIP, включая архитектуры на основе ResNet (RN50x4, RN50x16, RN50x64) и Vision Transformer (ViT-B16, ViT-B32, ViT-L14, ViT-L14-336).

##### C. Метрики оценки

Следуя протоколу оценки, установленному в *MedMNIST v2* [1], мы используем площадь под ROC-кривой (AUC) и точность (ACC) в качестве наших основных метрик оценки. Кроме того, мы сообщаем макро-среднее значение (вычисляя метрики для каждого класса отдельно, а затем берем среднее значение), чтобы учесть дисбаланс классов в наборе данных.

##### D. Сравнение с базовыми методами

Мы сравниваем наш подход с несколькими базовыми методами, которые рассматриваются в статье *MedMNIST v2* [8], включая ResNet-18, ResNet-50, Auto-sklearn, AutoKeras и Google AutoML Vision. В табл. I представлены полученные результаты.

Наши предложенные подходы превосходят все базовые методы по обоим метрикам AUC и ACC. Модель ViT-L14-336 достигла наивысшего AUC (0,934), в то время как модель ViT-L14 достигла наивысшей ACC

(0,793). Интересно, что модели на основе ResNet (RN50x4 и RN50x16) показали конкурентоспособные результаты со значениями AUC 0,931 и 0,929 соответственно, но более крупная модель RN50x64 показала сниженную производительность (AUC: 0,884, ACC: 0,675), возможно, из-за переобучения на относительно небольшом наборе данных *DermaMNIST*.

ТАБЛИЦА I. СРАВНЕНИЕ С БАЗОВЫМИ МЕТОДАМИ

Методы	Метрики	
	AUC	ACC
ResNet-18 (28)	0.917	0.735
ResNet-18 (224)	0.920	0.754
ResNet-50 (28)	0.913	0.735
ResNet-50 (224)	0.912	0.731
Auto-sklearn	0.902	0.719
AutoKeras	0.915	0.749
Google AutoML Vision	0.914	0.768
VFL-CLIP_RN50x4	0.931	0.776
VFL-CLIP_RN50x16	0.929	0.776
VFL-CLIP_RN50x64	0.884	0.675
VFL-CLIP_ViT_B16	0.923	0.771
VFL-CLIP_ViT_B32	0.914	0.755
VFL-CLIP_ViT_L14	0.933	<b>0.793</b>
VFL-CLIP_ViT_L14_336	<b>0.934</b>	0.780

Рис. 2 и 3 иллюстрируют динамику обучения нашей лучшей модели ViT-L14 с AUC 0,933 и ACC 0,793, позволившей получить улучшения на 2,0 % и 3,3 % соответственно по сравнению с лучшим базовым методом (Google AutoML Vision). Это демонстрирует эффективность нашего предложенного фреймворка для классификации дерматологических изображений. На рис. 4 представлена матрица несоответствия для нашей лучшей модели на тестовом наборе. Модель демонстрирует высокую производительность для большинства классов, особенно преуспевая в сосудистых поражениях (*vasc*) с 1 251 правильным предсказанием из 1 341 образца. Модель испытывает наибольшие трудности с дерматофибромой (*df*), правильно идентифицируя только 8 из 23 образцов, вероятно, из-за ее недостаточной представленности в обучающих данных.

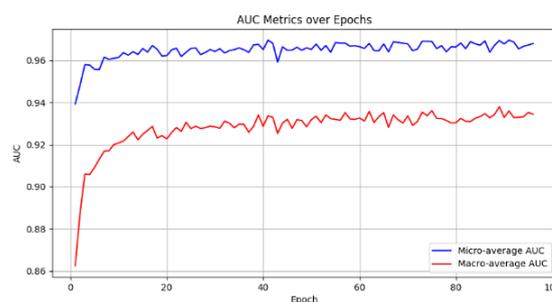


Рис. 2. Диаграмма метрики AUC по эпохам

Примечательно, что не всегда наблюдается хорошее разделение между тесно связанными классами, такими как меланоцитарные невусы (*nv*) и меланома (*mel*), а также между доброкачественными кератозоподобными поражениями (*bkl*) и сосудистыми поражениями (*vasc*). Этот шаблон соответствует известным диагностическим проблемам в дерматологии, где эти состояния могут иметь визуальные сходства.

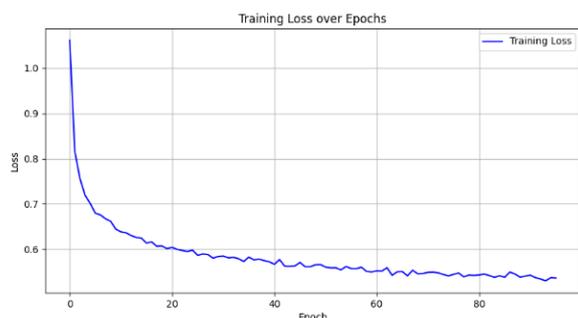


Рис. 3. Диаграмма значения потерь по эпохам



Рис. 4. Матрица несоответствий

## V. ЗАКЛЮЧЕНИЕ

В этой статье мы применили безопасный фреймворк, интегрирующий визуально-языковую модель CLIP с вертикальным федеративным обучением для совместного анализа дерматологических изображений. Наш подход позволяет решать критическую задачу обеспечения конфиденциальности данных пациентов, позволяя в то же время медицинским учреждениям совместно использовать распределенные наборы данных для улучшения дерматологической диагностики. Экспериментальная оценка на наборе данных DermaMNIST демонстрирует, что наш фреймворк достигает высокой производительности при его применении для классификации по сравнению с существующими методами, с AUC 0,933 и точностью 0,793, сохраняя при этом гарантии конфиденциальности.

Интеграция передового метода настройки промптов CoOp оказалась эффективной в адаптации модели общего назначения CLIP к специализированной области классификации дерматологических изображений. Анализ матрицы несоответствий выявил, что наша модель особенно хорошо работает с распространенными кожными состояниями, сохраняя при этом приемлемую производительность на более редких состояниях, что соответствует клиническим диагностическим проблемам в дерматологии.

Существует несколько перспективных направлений для будущей работы. Во-первых, расширение фреймворка для включения дополнительных

модальностей, таких как клинические метаданные, история пациента и дерматоскопические особенности, может еще больше повысить диагностическую точность и обеспечить более комплексную поддержку принятия решений. Во-вторых, исследование более сложных методов безопасной агрегации может снизить вычислительные накладные расходы, сохраняя или укрепляя гарантии конфиденциальности. В-третьих, оценка фреймворка на более крупных и разнообразных дерматологических наборах данных дала бы представление о ее обобщаемости на разные группы пациентов и клинические условия.

В заключение, через практическое применение предложенного авторами ранее фреймворка, это исследование дополнительно подтверждает его вклад в продвижение дерматологической диагностики с применением ИИ. По мере того, как здравоохранение продолжает цифровизироваться, а системы ИИ становятся более интегрированными в клинические рабочие процессы, такие фреймворки, как примененный в этом исследовании, будут иметь решающее значение для совмещения преимуществ медицины, основанной на данных, с необходимостью защиты конфиденциальности пациентов.

## СПИСОК ЛИТЕРАТУРЫ

- [1] H.A. Haenssle, C. Fink, R. Schneiderbauer, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists[J]. // Annals of oncology, 2018, 29(8): 1836-1842.
- [2] P. Tschandl, C. Rosendahl, B. N. Akay, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks[J]. // JAMA dermatology, 2019, 155(1): 58-65.
- [3] A.J. Larrazabal, N. Nieto, V. Peterson, et al. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis[J]. Proceedings of the National Academy of Sciences, 2020, 117(23): 12592-12594.
- [4] M. Kohli, L.M. Prevedello, R.W. Filice, et al. Implementing machine learning in radiology practice and research[J]. // American journal of roentgenology, 2017, 208(4): 754-760.
- [5] A. Radford, J.W. Kim, C. Hallacy, et al. Learning transferable visual models from natural language supervision[C] // International conference on machine learning. Pmlr, 2021: 8748-8763.
- [6] K. Zhou, J. Yang, C. C. Loy, et al. Learning to prompt for vision-language models[J]. // International Journal of Computer Vision, 2022, 130(9): 2337-2348.
- [7] K. Zhou, J. Yang, C. C. Loy, et al. Conditional prompt learning for vision-language models[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 16816-16825.
- [8] J. Yang, R. Shi, D. Wei, et al. Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification[J]. Scientific Data, 2023, 10(1): 41.
- [9] P. Tschandl, C. Rosendahl, H. Kittler. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions[J]. // Scientific data, 2018, 5(1): 1-9.
- [10] G. Litjens, T. Kooi, B. E. Bejnordi, et al. A survey on deep learning in medical image analysis [J]. // Medical image analysis, 2017, 42: 60-88.
- [11] S.K. Zhou, H. Greenspan, C. Davatzikos, et al. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises[J]. // Proceedings of the IEEE, 2021, 109(5): 820-838.
- [12] N. C. F. Codella, D. Gutman, M. E. Celebi, et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)[C] // 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE, 2018: 168-172.

- [13] W.J. Gordon, C. Catalini. Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability[J]. // Computational and structural biotechnology journal, 2018, 16: 224-230.
- [14] Q. Yang, Y. Liu, T. Chen, et al. Federated machine learning: Concept and applications[J]. // ACM Transactions on Intelligent Systems and Technology (TIST), 2019, 10(2): 1-19.
- [15] Y. Liu, Y. Kang, C. Xing, et al. A secure federated transfer learning framework[J]. // IEEE Intelligent Systems, 2020, 35(4): 70-82.
- [16] J. Xu, B. S. Glicksberg, C. Su, et al. Federated learning for healthcare informatics[J]. // Journal of healthcare informatics research, 2021, 5: 1-19.
- [17] D. Wang, X. Wang, L. Wang, et al. A real-world dataset and benchmark for foundation model adaptation in medical image classification[J]. Scientific Data, 2023, 10(1): 574.
- [18] F. Shakeri, Y. Huang, J. Silva-Rodríguez, et al. Few-shot adaptation of medical vision-language models[C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2024: 553-563.
- [19] B. Lester, R. Al-Rfou, N. Constant. The power of scale for parameter-efficient prompt tuning[J]. arXiv preprint arXiv:2104.08691, 2021.
- [20] The 25th International Conference on Computational Science and Its Applications. URL: <https://iccsa.org/>.