

Реферирование научной статьи

Реферат на научную статью А. Эстеви, Б. Купрела, Р.А. Новoa, Дж. Ко, С.М. Светтер, Х.М. Блау, С. Труна "Классификация рака кожи на уровне дерматолога с использованием глубоких нейронных сетей" // Nature. - 2017. - Т. 542. - С. 115-118. DOI: 10.1038/nature21056.

В статье А. Эстеви и соавторов "Классификация рака кожи на уровне дерматолога с использованием глубоких нейронных сетей" рассматривается возможность автоматического распознавания злокачественных и доброкачественных кожных образований по цифровым фотографиям. Авторы исследуют, способна ли единая глубокая сверточная нейронная сеть решать диагностические задачи, которые в клинической практике требуют подготовки врача-дерматолога, визуального осмотра и, при наличии показаний, проведения биопсии.

В начале статьи авторы отмечают, что визуальная диагностика заболеваний кожи представляет собой сложную задачу из-за большого числа патологий и сходства их внешних признаков. Доброкачественный невус может иметь отдельные признаки меланомы, а некоторые злокачественные новообразования внешне напоминают безопасные возрастные изменения кожи. При этом своевременное выявление меланомы и кератиноцитарных карцином существенно влияет на выбор лечения. Поэтому создание алгоритма, способного анализировать изображение поражения кожи и оценивать вероятность опасного процесса, имеет прямое медицинское значение.

Целью исследования стало обучение нейронной сети классификации кожных заболеваний непосредственно по изображениям, без предварительного ручного выделения формы, цвета, границ или других признаков. В качестве входных данных использовались только пиксели изображения и диагностические метки. Такой подход отличается от традиционных систем компьютерной диагностики, в которых набор

признаков заранее задается разработчиком. В предложенной модели значимые характеристики формируются автоматически в процессе обучения.

Основу исследования составил крупный набор данных, включавший 129 450 клинических изображений кожных поражений и 2 032 диагностических наименования. В массив вошли обычные клинические фотографии и дерматоскопические изображения. Данные были собраны из медицинских источников и открытых репозиториях, а диагнозы предварительно размечены специалистами. По объему этот набор значительно превосходил массивы, которые ранее применялись для обучения систем распознавания кожных заболеваний.

Для организации большого числа диагнозов авторы построили иерархическую таксономию кожных болезней. Она объединяла отдельные заболевания по клиническому и визуальному сходству. С помощью специального алгоритма исходные 2 032 диагноза были распределены по 757 обучающим классам. Такое решение позволило избежать двух крайностей: слишком узких классов с малым количеством примеров и чрезмерно широких категорий, которые не отражают медицинские различия между заболеваниями.

В верхней части созданной таксономии были выделены крупные группы поражений, имеющие значение для принятия клинического решения. К ним относились злокачественные, доброкачественные и неопухолевые заболевания. На более детальном уровне система различала меланоцитарные и эпидермальные поражения, а также отдельные типы опухолей. Авторы подчеркивают, что иерархическая организация диагнозов дает возможность обучать сеть на множестве частных заболеваний, а затем объединять полученные вероятности для решения более общей диагностической задачи.

В качестве вычислительной основы была использована архитектура Inception v3, ранее обученная на базе ImageNet для распознавания общих объектов. Последний классификационный слой сети был заменен и настроен для работы с медицинскими изображениями. Затем все слои дополнительно

обучались на собранном массиве фотографий кожи. Применение переноса обучения позволило использовать уже сформированные в сети способности распознавать контуры, текстуры, цветовые переходы и сложные пространственные сочетания признаков.

Каждое изображение приводилось к размеру, необходимому для работы выбранной архитектуры. Для повышения устойчивости модели применялось искусственное расширение данных: изображения поворачивались, отражались и кадрировались. Благодаря этому сеть должна была распознавать заболевание независимо от расположения поражения в кадре, угла съемки и части технических различий между фотографиями.

На этапе проверки авторы сначала оценили способность нейронной сети различать крупные диагностические категории. Результаты модели сравнивали с решениями двух дерматологов. Сеть, обученная на детализированной иерархии болезней, показала более высокую точность, чем вариант, который обучался только на нескольких широких категориях. По мнению исследователей, обучение на большом количестве конкретных классов заставляет модель выявлять более содержательные визуальные признаки, которые затем помогают и при укрупненной классификации.

Основная клиническая проверка была сосредоточена на задачах, связанных с выявлением рака кожи. Авторы сформулировали три практических направления тестирования:

- различение злокачественных кератиноцитарных карцином и доброкачественных себорейных кератозов по клиническим фотографиям;
- различение злокачественной меланомы и доброкачественных невусов по клиническим фотографиям;
- различение меланомы и доброкачественных невусов по дерматоскопическим изображениям.

Первое направление связано с выявлением наиболее распространенных форм рака кожи, второе и третье - с диагностикой меланомы, которая

относится к наиболее опасным злокачественным опухолям кожи. Для тестирования использовались ранее не предъявлявшиеся сети изображения, диагнозы которых были подтверждены биопсией. Это повышало надежность эталонной разметки и уменьшало риск того, что алгоритм сравнивается с субъективной предварительной оценкой.

Работу нейронной сети сравнивали с решениями практикующих дерматологов. В основных испытаниях участвовал 21 сертифицированный специалист. Врачам показывали те же изображения и предлагали выбрать одно из двух действий: рекомендовать биопсию или лечение либо успокоить пациента и не назначать вмешательство. Такое построение задания приближало эксперимент к реальному клиническому выбору, хотя не воспроизводило полноценный прием пациента.

Качество классификации оценивалось через чувствительность и специфичность. Чувствительность отражала долю правильно распознанных злокачественных поражений, а специфичность - способность не относить доброкачественные образования к раку. Эти показатели имеют различное практическое значение. Слишком низкая чувствительность повышает риск пропуска опасного заболевания, а недостаточная специфичность приводит к лишним биопсиям, тревоге пациента и дополнительной нагрузке на систему здравоохранения.

Полученные результаты показали, что нейронная сеть достигла уровня, сопоставимого с результатами участвовавших дерматологов. На графиках соотношения чувствительности и специфичности модель демонстрировала возможность выбирать разные пороги принятия решения. При одних настройках она могла обеспечивать более высокую чувствительность, при других - увеличивать специфичность. В каждой из трех диагностических задач показатели сети находились на уровне специалистов или превышали средний результат группы.

Авторы не представляют алгоритм как систему, выдающую единственный неизменный ответ. Результатом обработки изображения

является распределение вероятностей по классам заболеваний. Затем вероятности отдельных классов могут суммироваться в соответствии с диагностической иерархией. Например, вероятности разных вариантов меланомы объединяются в общую оценку злокачественного меланоцитарного поражения. Такой механизм позволяет использовать одну сеть для нескольких уровней классификации.

С научной точки зрения значимым результатом работы является доказательство того, что глубокая сверточная сеть может обучаться не только на ограниченном наборе заранее отобранных опухолей, но и на широком спектре кожных заболеваний. Это приближает алгоритм к реальной дерматологической практике, где врачу приходится отличать опасную опухоль от множества доброкачественных и неопухолевых состояний. Расширение числа классов делает задачу сложнее, но одновременно уменьшает вероятность того, что система будет работать только в искусственно узких условиях.

Особое внимание авторы уделяют потенциальной доступности технологии. Клинические фотографии могут быть получены с помощью обычных цифровых камер и мобильных устройств. Поэтому в перспективе нейронная сеть может использоваться как средство предварительной оценки, особенно в местах, где доступ к дерматологу ограничен. Пациент или медицинский работник мог бы передать изображение для автоматического анализа и получить рекомендацию о необходимости очной консультации.

Вместе с тем из содержания статьи следует, что достигнутый результат нельзя прямо приравнять к готовой автономной диагностике. Эксперимент оценивал распознавание отдельных изображений, тогда как врач учитывает анамнез, возраст пациента, локализацию образования, скорость его изменения, сопутствующие симптомы и результаты осмотра всей поверхности кожи. Кроме того, окончательное подтверждение злокачественного процесса в клинической практике часто требует гистологического исследования.

Ограничением является и зависимость модели от состава обучающего массива. Если определенные варианты заболеваний, типы кожи, условия освещения или особенности съемки представлены недостаточно, качество распознавания для соответствующих групп пациентов может снижаться. Авторы указывают на необходимость дальнейшей проверки системы в реальных клинических условиях. Для медицинского применения требуется проспективное исследование, оценка безопасности, контроль качества изображений и понятный порядок взаимодействия алгоритма с врачом.

Еще один существенный вопрос связан с интерпретируемостью результата. Нейронная сеть формирует решение на основе большого числа внутренних параметров, но не объясняет его так, как это делает специалист, описывая асимметрию, границы, неоднородность цвета или изменение размера образования. Поэтому высокая точность на тестовом наборе должна дополняться методами контроля ошибок и анализом причин неверной классификации.

По мнению авторов, наиболее перспективным является использование искусственного интеллекта для расширения возможностей медицинской помощи, а не простая замена дерматолога. Алгоритм может применяться для первичного отбора пациентов, поддержки решения врача и повышения доступности специализированной оценки. При подозрительном результате пациент должен направляться на профессиональный осмотр и, при необходимости, на биопсию.

Практическая ценность исследования состоит в том, что в нем предложена единая схема от подготовки большого массива данных до сравнения модели с сертифицированными специалистами. Важными элементами этой схемы являются медицинская таксономия, перенос обучения, использование подтвержденных биопсией тестовых изображений и оценка баланса чувствительности и специфичности. Эти решения стали ориентиром для последующих работ по применению глубокого обучения в диагностике по медицинским изображениям.

Статья также показывает, что результат алгоритма зависит не только от архитектуры нейронной сети, но и от качества постановки задачи. Авторы не ограничились общей оценкой точности, а выбрали клинически значимые пары диагнозов и сравнили решения системы с действиями врачей. Благодаря этому технический показатель был связан с понятным медицинским вопросом: следует ли проводить биопсию или лечение конкретного поражения.

В заключение авторы делают вывод, что глубокая сверточная нейронная сеть способна классифицировать рак кожи по изображениям с компетентностью, сопоставимой с уровнем дерматологов, участвовавших в исследовании. Работа подтверждает возможность использования искусственного интеллекта в визуальной медицинской диагностике и создает основу для дальнейших клинических испытаний. При этом внедрение такой системы требует проверки на разнообразных группах пациентов, учета клинического контекста и сохранения врачебного контроля.

Таким образом, исследование А. Эстеви и соавторов демонстрирует переход от экспериментального распознавания отдельных типов изображений к более универсальной модели анализа кожных заболеваний. Главный результат статьи заключается не только в сравнении сети с дерматологами, но и в разработке подхода, который объединяет крупный обучающий массив, иерархию медицинских диагнозов и клинически осмысленную оценку качества. Предложенная технология может повысить доступность раннего выявления рака кожи, однако должна рассматриваться как инструмент поддержки, эффективность и безопасность которого необходимо подтверждать в реальной медицинской практике.

Ключевые слова: глубокое обучение, сверточная нейронная сеть, искусственный интеллект, дерматология, меланома, рак кожи, медицинская диагностика, классификация изображений.