

## Обзор моделей компьютерного зрения для определения наличия заболевания COVID-19

Лиманова Наталия Игоревна, доктор технических наук, профессор  
Морозов Дмитрий Александрович, аспирант  
Поволжский государственный университет телекоммуникаций и информатики (г. Самара)

В статье представлен обзор современных моделей компьютерного зрения на основе нейронных сетей для определения наличия заболевания COVID-19.

**Ключевые слова:** COVID-19, нейронные сети, компьютерное зрение.

Эпидемия нового коронавируса (COVID-19) переросла в глобальную пандемию. Более 17 миллионов человек были инфицированы во всем мире, что привело к более чем 667 000 смертельных случаев с 20 января по 30 июля 2020 года. Быстрая передача от человека к человеку, сопровождающаяся не выявленной природой вируса, привела к огромной вспышке. Чтобы победить COVID-19, во всем мире необходимо принять соответствующие и основанные на фактах действия. Для определения наличия заболевания, требуется предварительная обработка рентгеновских снимков. Данный процесс очень трудозатратный и с этой целью для его оптимизации стали активно применяться модели компьютерного зрения, основанные на нейронных сетях, которые позволяют выявлять патологии с более высокой точностью.

Компьютерное зрение, основанное на глубоком обучении, имеет очень впечатляющие показатели. В 2012 году AlexNet реализовали сверточную нейронную сеть с оптимизацией для графического процессора. Авторы обратили свое внимание на набор данных ImageNet. Результат значительно превзошел конкурентов с точностью 63,3% по сравнению с предыдущим показателем 50,9% при использовании функций изображения, созданных вручную. Данная модель вызвала дальнейший интерес к нейронным сетям для компьютерного зрения. Исследователи разработали новые архитектуры, новые способы представления обучения на основе необработанных данных и новую инфраструктуру для обучения более крупных моделей. В 2020 году, через восемь лет после AlexNet, модель Noisy Student EfficientNet-L2 достигла максимальной точности 88,4%, что является абсолютным улучшением показателя на 25,1%. Deep Computer Vision позволяет создавать фотореалистичные изображения, переносом художественного стиля с одного изображения на другое и обеспечением роботизированного управления исключительно с помощью визуального ввода.

Успех Deep Computer Vision во многом объясняется набором данных ImageNet [1]. Датасет ImageNet - это набор данных, содержащий 1,2 миллиона изображений, разбитых на 1000 категорий. Разрешение данных входных изображений для глубокого обучения является важным фактором с точки зрения затрат на вычисления и хранение.

Компьютерное зрение призвано улучшить качество здравоохранения во многих отношениях. Наиболее заметным и часто обсуждаемым применением к COVID-19 является диагностика медицинских

изображений. Глубокое обучение очень хорошо зарекомендовало себя при диагностике медицинских изображений и прошло клинические испытания при многих заболеваниях. Задачи медицинских изображений в основном включают классификацию и сегментацию, а также реконструкцию из двухмерных срезов в компьютерной томографии для получения окончательного трехмерного изображения.

Компьютерное зрение также помогает в тонких операциях в больницах. Оно позволяет камерам идентифицировать то, что они записывают. Например, они могут маркировать каждый пиксель дороги для беспилотных автомобилей, наносить спутниковые изображения на дорожные карты, могут представить себе, как бы выглядел эскизный рисунок, если бы это был сфотографированный объект. Эти приложения описывают потенциал компьютерного зрения для включения большого набора тонких приложений в ответ на пандемию мониторинг социального дистанцирования или инвентаризации больничного оборудования.

Экономический ущерб - одна из самых больших жертв COVID-19. Робототехника на основе зрения является многообещающим направлением для этого. Робототехника на основе машинного зрения предлагает более общее решение для управления, которое может адаптироваться к новой ориентации объектов или распределению веса ящиков.

**Анализ медицинских изображений**

Содействие и автоматизация анализа медицинских изображений - одно из наиболее часто обсуждаемых функций глубокого обучения. Эти системы быстро улучшаются и уже прошли клинические испытания. Мотивируя использование рентгенологической диагностики, Рейек с соавт. [2] обнаружили, что чувствительность КТ составляет 98% по сравнению с 71% при ОТ-ПЦР., делая вывод, что «КТ грудной клетки можно рассматривать как основной инструмент для текущего обнаружения COVID-19 в эпидемических зонах». Фредриксон с соавт. [3] выделяют некоторые причины, по которым рентген грудной клетки предпочтительнее компьютерной томографии, а именно то, что они дешевле и доступнее, а также имеют более низкое ионизирующее излучение, чем компьютерная томография.

Даже в своем отказе от трансферного обучения на основе ImageNet для анализа медицинских изображений Paqhu и др. [4] отмечают различия в производительности передачи при переработке вычислительных блоков в ResNet. Как упоминалось ранее, ар-

хитектура нейронной сети - одна из самых многообещающих тенденций в исследованиях Deep Computer Vision. Незначительные архитектурные изменения, такие как расположение Skip-соединений могут иметь значительный

Схемы обучения и архитектура нейронной сети являются наиболее важными областями для исследования. Однако многозадачное обучение может быть сложно реализовать. Сбор помеченных данных для обучения моделей семантической сегментации также может быть чрезвычайно утомительным. Шан и соавт. [5] представляют интересную стратегию «человек в петле» для маркировки пикселей при КТ-сканировании COVID-19.

**Окружающий интеллект**

Когда мы думаем о пересечении ИИ и здравоохранения или медицины, мы можем перейти к новаторскому предсказанию структуры белка или языковым моделям, которые определяют направления исследований. Однако не все проблемы в миссии настолько грандиозны. Одним из наиболее частых сюжетов пандемии COVID-19 был страх перед перегрузкой системы здравоохранения.

В отчете о врачебной ошибке Маккари и Даниил обнаружили, что «до 400 000 человек ежегодно умирают в Соединенных Штатах из-за ошибок и недостатков в принятии клинических решений и физических действий». Идея «окружающего интеллекта» состоит в том, чтобы интегрировать интеллектуальные датчики, которые могут отслеживать и собирать данные об этих физических действиях.

Одним из наиболее многообещающих направлений исследований глубокого обучения в окружающем интеллекте будет сжатие. Необходимо сжать эти нейронные сети, чтобы их можно было экономически эффективно встроить в интеллектуальные датчики. Некоторые из наиболее многообещающих подходов к сжатию включают смешанную точность, квантование и обрезку. Смешанная точность описывает использование 2, 4, 8 или 16-битной точности для весов нейронной сети, а не 32-битной. Квантование - более эффективная стратегия снижения этой точности. Это описывает кластеризацию по общим

значениям, а не просто усечение для соответствия требованиям точности. Многие веса в нейронной сети могут быть удалены или установлены на ноль, что не влияет на производительность модели. Некоторые интересные направления включают понимание успеха и феномена разреженных сетей, а также возможность дальнейшего изучения шума квантования во время обучения.

**Робототехника на основе зрения**

Роботизированное управление - одно из самых популярных приложений и областей исследований глубокого обучения. В качестве задачи глубокого обучения это включает отображение состояний пикселей и датчиков на двигательные действия. Наглядным примером робототехники, основанной на зрении, является тест на кофейной чашке Возняка. Он описывает, сможет ли робот зайти в любую комнату и приготовить чашку кофе. Стоит отметить, что этот уровень обобщения не нужен для большинства приложений для COVID-19.

Диапазон рассматриваемых роботизированных приложений сильно различается по сложности требуемых манипуляций. Например, роботу, распыляющему дезинфицирующий спрей на комнату, не нужно обладать тонкими навыками манипулирования. Однако робот, помогающий пациенту с респираторным заболеванием, справляется. Точно так же роботы, облегчающие производственную и экономическую деятельность, также различаются в этом масштабе. Сортировка бутылок требует больше зрения, чем ловкости, тогда как сборка требует невероятно ловких манипуляций.

Барфут [6] обрисовывает план будущего канадской робототехники в ответ на пандемию. Авторы подчеркивают потенциал применения в дезинфекции, удаленной сортировке, логистике и доставке. Авторы утверждают, что «пандемия стала переломным моментом для ускорения инвестиций в робототехнику».

Эти приложения исследуются в различных областях данных в области обработки естественного языка, компьютерного зрения, наук о жизни и эпидемиологии.

### **Литература:**

1. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. *Advances in neural information processing systems*, vol. 25. Curran Associates, Inc., 2012, p. 1097-1105.
2. Rieke N, Hancox J, Li W, Milletari F, Roth H, Albarqouni S, Bakas S, Galtier MN, Landman B, Maier-Hein K, Ourselin S, Sheller M, Summers RM, Trask A, Xu D, Baust M, Cardoso MJ. *The future of digital health with federated learning*, 2020.
3. Fredrikson M, Jha S, Ristenpart T. Model inversion attacks that exploit confidence information and basic countermeasures; 2015. p. 1322-1333.
4. Raghu M, Zhang C, Kleinberg J, Bengio S. Transfusion: understanding transfer learning for medical imaging, 2019.
5. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y. Lung infection quantification of COVID-19 in CT images with deep learning, 2020.
6. Barfoot T, Burgner-Kahrs J, Diller E, Garg A, Goldenberg A, Kelly J, Liu X, Naguib HE, Nejat G, Schoellig AP, Shkurti F, Siegel H, Sun Y, Waslander SL. Making sense of the robotized pandemic response: a comparison of global and canadian robot deployments and success factors, 2020.