**Abstract**

В этой статье предлагается неглубокая архитектура на основе капсульных сетей со сверточными слоями для обнаружения инфицированных COVID-19. Предлагаемый метод сочетает в себе способность капсульной сети понимать пространственную информацию со сверточными слоями для эффективного извлечения признаков. Из-за неглубокой архитектуры модели она имеет 23 миллиона параметров для обучения и требует меньшего количества обучающих выборок. Предлагаемая система является быстрой и надежной и правильно классифицирует рентгеновские изображения по трем классам: COVID-19, отсутствие результатов и вирусная пневмония. Экспериментальные результаты с набором данных X-Ray показывают, что наша модель работает хорошо, несмотря на меньшее количество выборок для обучения, и достигла средней точности 96,47% для мультиклассовой и 97,69% для бинарной классификации при 5-кратной перекрестной проверке.

**Introduction**

Технология рентгеновского сканирования дешевле, быстрее и имеет меньше излучения, чем процедура компьютерной томографии. Ручное исследование рентгеновских снимков позволяет поставить правильный диагноз COVID-19 примерно в 50-60%. Цель системы – повысить точность.

Проблема — это отсутствие единообразных, огромных и размеченных наборов данных. Для небольших наборов данных капсульные сети работают лучше, чем сверточные нейронные сети. Они имеют гораздо меньше параметров – быстрее обучаются. Они не требуют аугментации данных, при этом выдавая точность, конкурентно-способную с точностью глубоких сетей, обученных на больших наборах данных. В традиционных CNN пространственная информация теряется, а капсульные сети захватывают эту информацию.

Таким образом, мы предлагаем модифицированную капсульную сеть, которая включает дополнительные сверточные слои перед первичными слоями капсулы. Рентгенограмма грудной клетки передается в сверточные слои в процессе обучения. Они извлекают отличительные признаки из изображений и передают вектор признаков в первичные слои капсулы. Всего создается 32 капсулы, и векторы далее передаются на слой рентгеновской капсулы. Этот слой содержит 16D-векторы и используется для классификации рентгеновских изображений грудной клетки по трем классам (COVID-19, норма и пневмония).

**Related work**

В недавних исследованиях мы видим, что рентгеновских изображений грудной клетки очень мало, и большинство авторов либо использовали метод увеличения данных, либо GAN для создания новых изображений, чтобы можно было решить проблемы переобучения и недообучения моделей глубокого обучения. Но описанные выше методы требуют много времени для обучения модели, так как количество обучаемых параметров большое и не подходит для развертывания в реальном времени.

**3 Proposed work**

**3.1 Capsule networks**

Капсульные сети и концепция динамической маршрутизации между капсулами были представлены Хинтоном и его коллегами-исследователями в 2017 для набора цифровых данных MNIST(для классификации рукописных цифр). Эти сети показали улучшение производительности по сравнению со сверточной нейронной сетью. Основная причина этого может заключаться в том, что у CNN есть сло, используемые для понижения дискретизации карт признаков изображений. Таким образом, CNN не может фиксировать мелкие детали и пространственную информацию. Капсулы отличаются от нейрона тем, что описывают пиксели векторами. Естественно, вектор содержит гораздо больше информации, чем одно число.

**3.2 Proposed convolutional capsule network**

добавили 4 сверточных слоя. После экспериментов с количеством слоев с размерами ядра и фильтра четыре сверточных слоя оказались оптимальным количеством слоев для извлечения признаков. Мы изменили размер этих изображений до 150 × 150 для наилучшей производительности после экспериментов с изображениями разных размеров.

Первый сверточный слой обучает 16 ядер размера 5 × 5 с шагом 1. Затем выполняется операция максимального объединения с шагом 2. Тот же размер ядра используется в следующих двух последовательных слоях, а номера ядер равны 32 и 64 соответственно. Затем обучается ядро четвертого сверточного слоя 128 размером 9 × 9 с шагом 1. Во всех слоях используется функция активации ReLU.

Входные данные для слоя первичной капсулы представляют собой усовершенствованную функцию размера 16 × 16 × 128, а выходные данные — 32 капсулы размером 6 × 6 × 8 каждая. классификации рентгеновского изображения по трем классам COVID-19, нормальной и вирусной пневмонии . Слой X-Ray caps также используется для реконструкции изображения с использованием трех полносвязных слоев.

**4 Experimental results**

**4.1 Dataset and setup**

Набор данных COVID-19 Radiography [16, 42], который был разработан исследователями из Университета Катара. Набор данных включает изображения каждого класса — COVID-19 (3616), нормальный (10192) и вирусная пневмония (1345), как показано в таблице 2.

**2 Experimental discussion**

Мы не развернули модель в режиме реального времени для клинического использования, но ее можно развернуть для клинических испытаний. XraySetu, платформа, разработанная в Индии для интерпретации рентгенографии грудной клетки с помощью What sApp. После обучения модели на данных рентгенографии грудной клетки ее можно сохранить в облаке и использовать для диагностики.

**4.2.1 Image pre-processing**

Исходный размер изображения в наборе данных составлял 300 × 300 пикселей. Затем мы изменяем размер изображения до 150 × 150 пикселей. Кроме того, мы преобразовали изображение RGB в изображение в градациях серого, уменьшив количество каналов с 3 до одного канала. Затем мы проанализировали интенсивность их пикселей и применили метод адаптивного выравнивания гистограммы с ограниченным контрастом (CLAHE). Полная процедура показана на рис. 7.

**4.2.2 Evaluation metrics**

Мы оценили предложенную нами модель в отношении точности и отзыва и выполнили 5-кратную перекрестную проверку, чтобы получить среднюю точность. Мы случайным образом разделили набор данных на 80% обучающих выборок и 20% тестовых выборок, а затем обучили модель в 5 раз и записали среднюю точность предложенной модели.

**4.3 Results of the proposed model**

В мультиклассе используются три класса: COVID-19, нормальная и вирусная пневмония, а в бинарном классе используются только два класса: COVID-19 и не-COVID-19, как показано на рис. 8.

Мы провели сравнительный анализ предложенной модели Conv-CapsNet с оригинальной капсульной сетью. Производительность предлагаемой Conv-CapsNet анализируется на основе точности классификации каждого класса и сравнения матриц путаницы для получения дополнительной информации. Как видно на рис. 13, Conv-CapsNet работает лучше, чем исходная капсульная сеть, с точки зрения точности по классам. Более того, показатели класса COVID-19 значительно улучшились в предложенной нами модели (94,95% по сравнению с исходной капсулой, что очень мало (78,01%). Основное внимание в этом исследовании уделяется поиску максимального количества положительных случаев. так что дальнейшее распространение вируса может быть остановлено. в случае класса COVID-19 в предлагаемой Conv-CapsNet значительно сократилось количество ложноотрицательных результатов.

Для классификации двоичных классов с использованием предложенной Conv-CapsNet точность обучения и проверки и кривые потерь показаны на рис. 15. Сравнительная классовая точность исходной капсульной сети и Conv-CapsNet показана на рис. 16. Здесь точность для COVID-19 класс значительно улучшился с 78,7% до 94,6%. Матрицы путаницы для обеих моделей показаны на рис. 17 и 18. Здесь количество неправильно предсказанных случаев без COVID-19 сократилось с 21,3% до 2,67%. В случае классификации бинарного класса мы получили точность более 94% в обоих классах и достигли средней точности 97,69%, как показано в таблице 6.

**4.4 Comparative analysis with the literature**

Сравнение с литературой для бинарного класса показано в таблице 7. В бинарной классификации большинство авторов получили точность более 95%, а предложенный нами подход получил отличную точность 97,69%, что сравнимо с работами [6, 35, 47]. В этих документах использовалась более глубокая сверточная нейронная сеть со сложными структурами с более поддающимися обучению параметрами и выполнялись методы увеличения данных для получения большего количества выборок данных. Предлагаемая нами модель Conv-CapsNet имеет неглубокую структуру с меньшим количеством параметров и большей точностью, и она работает на меньшем количестве обучающих выборок. Это делает его быстрым и надежным, что подходит для приложений реального времени.

это неглубокая сеть, для обучения которой требуется меньше времени, а размер обученной модели компактен.

**5 Conclusion and future work**

В исследовании сделан вывод о том, что капсульные сети эффективнее классифицируют рентгеновские снимки по сравнению со свёрточными нейронными сетями для небольших наборов данных. Предлагаемый нами Conv-CapsNet может обеспечить точность 96,47% для диагностики COVID-19 у пациентов с нормальной и вирусной пневмонией и 97,49% для классификации бинарных классов. Недостаток оригинальной капсульной сети по сравнению с предлагаемой системой заключается в том, что она недостаточно эффективна для класса COVID-19, а это означает, что модель превосходит класс большинства.