В марте 2020 года ВОЗ (всемирная организация здравоохранения) объявила пандемию короновирусной инфекции. Пандемия каснулась миллионов людей по всему миру. Самым важным симптомом является пневмония.

Было обнаружено на ранних стадиях исследований, что поражение можно увидеть на снимках КТ.

ПРЦ долгий, трудозатратный,сложный. Кроме того неточный.

Альтернативный метод скрининга - КТ грудной клетки. На снимке рентгенолог визуально определяет маркеры, соответствующие заболеванию. На ранних этапах исследований было установлено, что наличие на снимке аномалий является признаком инфицирования. Также в пользу применения в диагностике такого подхода говорит возможность

-быстрой диагностики в условиях большого количества пациентов в больших городах

-наличие необходимого рентген оборудования для проведения снимка

-возможность передачи снимков по сети для консультаций.

Следовательно, анализ рентгеновского изображения может быть проведен быстрее и является хорошим дополнением к ПРЦ тестированию. Так же предполагается, что сканирование может являться необходимостью, если у пациента отрицательный результат ПЦР-теста.

Таким образом, рентгеновское исследование является частью разработанной стратегии диагностики заболевания. Поскольку визуальные индикаторы могут быть нечеткими, у рентгенологов возникает необходимость в помощи от более точных и быстрых цифровых системах.

Рентгеновский снимок представляет собой изображение темных областей внутренних органов с белыми «тенями». Рентген – один из самых известных и старых методов диагностики. В сравнении с другими методами диагностики COVID-19, рентген изображение имеет несколько «+»ов – низкая стоимость, широкая доступность, неинвазивный метод исследования. Таким образом, этот метод диагностики может являться лучшим кандидатом на массовую, быструю и простую диагностику в условиях панемии и глобального кризиса системы здравоохранения, как с точки зрения необходимого количества медицинских работников, так и с точки зрения диагностического оборудования.

-1.1Описание развития ИИ.

Глубокое обучение продемонстрировало большое влияние за последнее десятилетие на решение многих сложных задач.

Модели глубокого обучения состоят из глубоких нейронных сетей, которые автоматически изучают важные функции из данных. Обучение глубоких моделей обычно выполняется с использованием оптимизации **стохастического градиентного спуска**.

За последнее десятилетие машинное обучение и ИИ получили значительное развитие. Нейронные сети превзошли традиционные модели по многим важным показателям. Зарекомендовали себя в широком спектре областей – распознавание речи, обработка изображений, распознавание речи.

Качество анализа и классификации медицинских изображений уже можно де факто признать стандартом. СNN исользовались для различных задач классификации медицинских изображений, таких как: болезни легких, определение переносчика малярии в образцах крови, рак груди, беспроводная эндоскопия, диагностик рака кожи по классификации изображения.

Another recent and relevant study has been conducted on validation and adaptability of Decompose-, Transfer-, and Compose-type deep CNN for COVID-19 detection using chest X-ray image classiﬁcation (A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classiﬁcation of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” 2020, http://arxiv.org/abs/2003.13815). The authors have reported the results of the study with an accuracy of 95.12%, sensitivity of 97.91%, and speciﬁcity of 91.87%.

На невысокие качественные характеристики ранних модель влияли недостаточность снимков с COVID-19 и низкое качество некоторых изображений в датсетах.

Искусственный интеллект, усиленный методами глубокого обучения может играть важную роль для скринига COVID-19.

Сверточная нейронная сеть

Нейронная сеть имитирует работу части человеческого мозга, отвечающую за распознавание образов. НС используется для распознавания и анализа изображений, классификации изображений, а так же распознавания речи. СНН это тип глубоких нейронных сетей, содержащий сверточные, объединяющие уровни, и уровень активации. Основным является уровень свертки, в котором расположены ядра, на входы которым подается изображение. Все выходы сверточного слоя составляют карту объектов.

-1.2Описание существующих подходов и реализаций

**1.2.1Одна из первых опенсорсных - COVID ‑Net(2020)**

Сверточная нейронная сеть, приспособленная для выявления КОВИДА на снимках КТ. Открытый проект.

Открыла новое применение искусственного интеллекта для раннего, эффективного и масштабного выявления вирусов среди людей.

Использоаван датасет COVIDx 13 975 изображений от 13 870 пациентов. Датасет является компиляцией пяти открытых источников данных:

COVID-19 Image Data Collection

COVID-19 Chest X-ray Dataset Initiative

ActualMed COVID-19 Chest X-ray Dataset Initiative

RSNA Pneumonia Detection Challenge

COVID-19 radiography.

-Принципы разработки сети.

Детали реализации:

Предобучена на ImageNet, а затем обучена на COVIDx датасете с использованием Adam оптимизатора. learning rate = 2e−4, number of epochs = 22, batch size = 64, factor = 0.7, patience = 5. Finally, we introduce a batch re-balancing strategy to promote better distribution of each infection type at a batch level. The initial COVID-Net prototype was built and evaluated using the Keras deep learning library with a Tensor-

Flow backend. The proposed COVID-Net architecture was built using generative synthesis 41 , as described in

“Machine-driven design exploration”

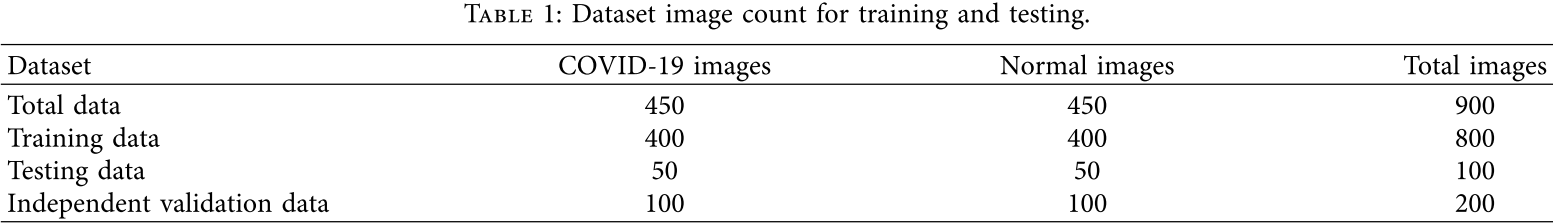
**1.2.2CNN(2021)**

ИИ и сверточные сети в частности достигли успеха в анализе медицинских изображений и классификации. В этой работе представлена нейронная сеть для диагностики COVID-19 по классификации рентгеновских снимков.

Модель тестировалась на 100 снимках для получения точности 100%. Вторым этапом модель была протестирована на несвязанном датасете и показала 99,5% точности.

Для обучения модели использовался датасет “Github Covid19 X-ray dataset” из репозитрия https://github.com/ieee8023/covid-chestxray-dataset, 2020. Online. Поскольку размер датасета мал и несбалансирован, то для решения этой проблемы была применена аугментация. Для генерации широкого набора вариантов применялись поворот изображения на разные углы. Для балансировки датасета по классам были добавлены изображения минорных классов.

Первый этап сбор первичного набора изображений двух классов: с COVID-19 и здоровых пациентов. Вторым этапом некачественные и недостаточно информативные для постановки диагноза изображения исключались медицинскими специалистами. Таким образом был сформирован качественный датасет. На третьем этапе датасет был аугментирован стандартными техниками для увеличения его размера. Результирующий датасет был использован для обучения сети на следующем этапе. После обучения модель была протестирована на качество распознавания на первичном датасете, а также на обособленном датасете. Каждый датасет имел свои характетиристики по количеству изображений, тестовый набор, сверочный набор, и разное соотношение изображений, принадлежащих одному классу.



--Материалы и методы.

Датасет

В первичный набор состоял из 178 рентгеновских изображений грудной клетки. Из которых 136 были снимками с подтвержденным заболеванием, и 42 изображения - обычные или с различными заболеваниями, такими как пневмония (J. P. Cohen, “Github Covid19 X-ray dataset,” 2020, https://github.com/ieee8023/covid-chestxray-dataset, 2020.). Таким образом, датасет состоял из изображений двух классов – COVID-19 и другие. Такое распределение показывает несбалансированность датасета. Следовательно, для достижения лучшего результата необходим предподготовка. На изначальном наборе данных модель показала недостаточную точность в 54%.

Техническое оснащение для обучения модели Core i7 7th поколения, 8 GB RAM, Microsoft Windows 10 операционная система,написана на языке программирования Python с использованием библиотеки Anaconda 3 и среды Jupyter Notebook.

Препроцессинг данных

Балансировка классов

Для балансировки набора данных использовались изображения 136 изображений без признаков заболевания из набора данных (P. Mooney, “Kaggle X rays dataset,” 2020, https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia Online.)

После балансировки датасета модель показала точность в 69%, что еще не достаточно эффективной системы диагностики COVID-19.

Анализ экспертами

Тщательный анализ изображений медицинскими специалистами 135 изображений выделил 90 изображений, подходящих для обучения модели. Результирующий датасет содержал 90 изображений COVID-19 и 90 обычных изображений. Полученный результат показал улучшение качества распознавания заболевания в 72%. Однако, невысокий прирост качественной характеристики обусловлен небольшим количеством изображений.

Аугментация

Аугментация данных, это способ существенно увеличить объем данных для обучения модели. Для изображений используются базовые операции такие как переворачивание, вращение, обрезка или дополнение. Первоначальный набор данных дополняется измененными таким образом изображениями, существенно увеличивая размер датасета для обучения нейронной сети. Изменение изображений дает дополнительные данные для обучения модели.

Поскольку изображения являются нелинейными по природе, то применяются также нелинейные функции активации.

Пример нелинейной функции активации – функция ReLU, где для всех неположительных аргументов, значение функции равно 0. График функции изображен на рисунке 1.

(1)

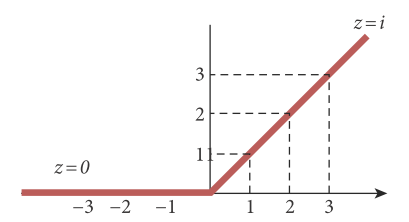


Рис График функции ReLU

Уровень объединения или уровень подвыборки также важный слой в рассматриваемом типе нейронной сети. Каждый объект, извлеченный сверточным слоем, обрабатывается независимо в этом слое. Чтобы минимизировать переобучение и количество выделенных объектов, размер карты уменьшается, и возвращаются только важные объекты. Значения объединения могут быть максимальными (max), средними (average) и суммированными (sum).

Эта сеть использует MAX пуллинг, более подходящий для определения границ контраста.

Дополнительно используется слой пакетной нормализации (batch normalization) для включения глубокой нейронной сети в обучение.

Модель настраивает масштабирование и активацию, чтобы нормализовать входной слой и ускорить процедуру обучения между скрытыми блоками.

Для минимизации переобучения использовался слой дропаута с параметром 20%.

Выравнивающий слой создает вектор из пикселей изображения, полученного после сверточных слоев, и передает на вход полносвязных слоев.

В полносвязном слое каждый нейрон предыдущего слоя напрямую связан с каждым нейроном последующего слоя. Основная задача полносвязного слоя это классификация изображения по полученным сглаженным результатам от предыдущего слоя. Каждое значение сглаженного набора объектов представляет вероятность принадлежности объекта к определенному классу. Таким образом, на основе этих вероятностей, полносвязная сеть в конечном итоге определяет решение о классификации.

Архитектура

Модель состоит из 38 слоев, в каждом из которых 6 сверточных, 6 объединяющих, 6 дропаута, 8 слоев с функциями активации, 8 слоев батч нормализации, 1 сглаживающий и 3 полносвязных слоя. Размер входных изображений 150\*150\*3 – 150\*150 в цвете RGB…

После обработки изображения получается 3-размерная матрица, которая преобразуется в вектор, который передается на вход 3 полносвязным слоям.

Используется бинарная классификация, поэтому используется бинарная кроссэнтропия (BCE) функции потерь. В двоичной классификации, поскольку для отнесения данных к одному из двух заданных классов требуется только один выходной узел, поэтому в случае функции потерь BCE выходное значение присваивается сигмовидной функции активации. Выходное значение, выдаваемое сигмовидной функцией активации, находится в диапазоне от 0 до 1. Он находит ошибку между предсказанным классом и фактическим классом. Был использован оптимизатор “Adam”, который изменяет вес атрибута и скорость обучения, чтобы уменьшить потери модели обучения.

Подход к определению числа слоев состоял в постоянном увеличении количества слоев, до достижения необходимых результатов, начиная с одного слоя.

**1.2.3AC-CovidNet(2022)**

Метод изучает устойчивые и дискриминационные признаки с помощью контрастивных потерь. Более того, предлагаемый метод придает большее значение инфицированным регионам, управляемым механизмом внимания.

attention guided contrastive CNN

Отличия:

1 Новый фреймворк для распознавания заболеванияна рентгеновских снимках

2 Используется модуль внимания, помогающий обучению.

3 Контрастные потери увеличивают способность к классификации изображения и надежность модели.

Использовались возможности механизма внимания и контрастного обучения, чтобы справиться с обучением с ограниченными данными для распознавание Covid-19 по изображениям.

Механизм внимания в глубоком обучении облегчает изучение локализованных признаков, что более важно в контексте проблемы распознавания Covid-19 на рентгенограммах.

Также обнаружено, что модель, основанная на внимании, может превзойти модели простой нейронной сети (Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L.,Polosukhin, I.: Attention is all you need. arXiv preprint arXiv:1706.03762 (2017)).

Сравнительное обучение - это свежая тенденция к изучению сходства и различия между похожими и непохожими образцами в абстрактном пространстве признаков для визуальных представлений. Как правило, контрастивное обучение зависит от сходства признаков между положительными и отрицательными парами.

Описание модели

Основана на модели COVID-Net. В модели широко используется легкий механизм расширения проекции остаточного проецирования (PEPX). В этой архитектуре также используется селективная связь на большие расстояния в модели, что улучшает репрезентативность. Это также облегчает обучение модели. Однако широкое использование этих соединений дальнего действия может привести к появлению множества избыточных низкоуровневых функций. Для решения этой проблемы в предлагаемой модели используется механизм внимания. Внимание помогает модели расставить приоритеты в важных областях. Внимание также полезно для подавления активации избыточных функций из начальных слоев и помогает сосредоточиться на важных функциях, необходимых для решения задачи. Используются шлюзы внимания в предлагаемой архитектуре на различных уровнях архитектуры COVID-Net, где используется много соединений дальнего действия. Это повышает чувствительность, поскольку модель лучше учитывает важные визуальные особенности зараженных областей на изображениях CXR из-за Covid-19. Поскольку разница между признаками Covid-19 и пневмонии очень тонкая, предлагается использовать контролируемую контрастную потерю. Контрастная потеря помогает сети максимально увеличить расстояние между выученными представлениями классов.

Архитектура

Архитектура модели представляет собой расширенную архитектуру модели COVID-Net. Используются ворота внимания, где сходятся несколько соединений дальнего действия. Используются уровни PEPX и шлюзы внимания. Также используется функция контрастных потерь во время обучения. Модель сначала обучается с использованием контролируемого контрастного обучения, а затем завершается точной настройкой с использованием контролируемого обучения.

PEPX слой

Архитектура проекции-расширения-проекции-расширения (PEPX). Задача модуля состоит в том, чтобы проецировать объекты в более низкое измерение, используя первые два слоя conv1x1, затем расширять эти объекты с помощью слоя свертки по глубине (DWConv3x3) и снова проецировать в более низкое измерение, используя два слоя conv1x1. Таким образом, уровень PEPX приводит к эффективной модели за счет уменьшения количества параметров и операций.

Ворота внимания

Используются на различных уровнях. Функции из нескольких слоев проходят через conv1x1 и складываются вместе.Затем агрегированные функции передаются через функцию активации Relu, за которой следует conv1x1, а затем функция активации сигмовидной формы. Затем выходные данные карты объектов сигмовидного слоя проходят через ресемплер. Выходные данные ресемплера добавляются к функциям из ближайшего входного слоя к модулю внимания, чтобы получить выход ворот внимания (рисунок 2).



Рис Схема модуля Ворота внимания

Детали нейронной сети. Описан алгоритм

Используется контролируемый контрастный метод обучения для обучения сети кодировщика извлечению признаков. Сеть классификатора обучается с использованием функции кросс-энтропийных потерь после замораживания сети кодировщика.

Контрастная потеря чаще всего используется в неконтролируемом и самоконтролируемом обучении.

\*\*формулы и описание

Использованный датасет

Сеть обучалась на данных датасета COVIDx.Это саамы большой публичный набор рентгенограмм специализированный под COVID-19. Этот датасет является компиляцией нескольких открытых наборов изображений. Содержит изображения трех классов: COVID-19, пневмония, и без патологий. В наборе 14 258 изобрежений, распределенных следующим образом:

7966-без патологий

5475 –пневмония

517 – COVID-19

в обучающем наборе и по 100 изображений каждого класса для тестирования.

На основе датасета было сформировано три датасета с различным распределением изображений по классам (рис 3)

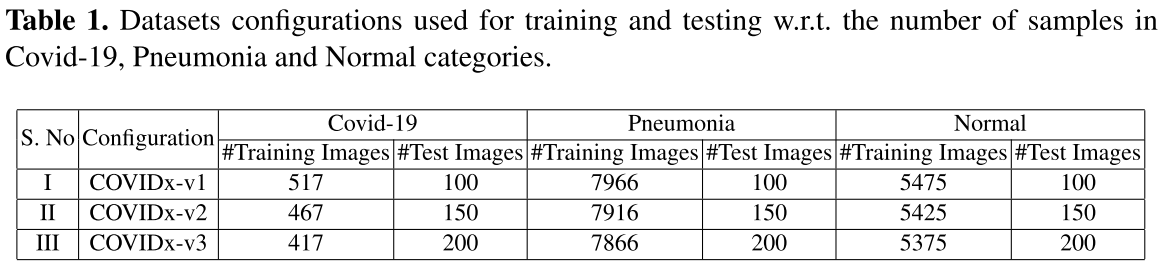


Рис Конфигурации наборов данных

Обучающие настройки

Модель была обучена на датасете. Последующее обучение модели состоит из двух этапов. На первом этапе сеть обучается с использованием функции потерь, для извлечения объектов. На втором этапе экстрактор признаков фиксируется и обучается путем добавления классификатора с функцией перекрестной потери энтропии.

Модель обучалась на трех созданных вариантах датасета. Использовался оптимизатор Adam. Значение уровня обучения составляло 1,7\*10^-4. Размер батча – 64. На каждом слое использовалась функция активации ReLU и софтмакс на последнем слое. Пулнг с параметром max используется после каждого PEPX слоя.

Модель обучалась с использованием вычислительных ресурсов Google Colab. Использовалась библиотека глубокого обучения Keras с tensorﬂow.

Использование ансамбля.

В основном исследователи используют одну архитектуру нейронной сети и с помощью подбора количества слоев и настройки коэффициентов пытаются достичь необходимых показателей качества распознавания (классификации) изображения. Некоторые исследователи создают новые архитектуры для решения задачи. Ни одно из ранних исследований не пыталось комбинировать модели с целью улучшения качества классификации модели в целом. Несколько исследований показывают, что ансамблевые модели лучше предсказывают, чем индивидуальные модели, а также помогают предотвратить переобучение. Одним из методов ансамблирования является вычисление средневзвешенного значения выходной вероятности каждой нейронной сети. Такой метод является более качественным, по сравнению с основанными на простом среднем значении.

Описание метода

Для принятия взвешенного решения в реальной жизни во многих сферах основываются на несколько мнений. Основываясь на этот подход, были адаптированы несколько моделей нейронных сетей. Каждая модель обучалась независимо. После обучения модели были скомбинированы, с использованием средневзвешенной средней для предсказания класса.

В модели используются три сверточные нейронные сети - DenseNet201, Resnet50V2, Inceptionv3.

Особенностью DenseNet является меньшее количество настраиваемых параметров, чем в других сверточных нейронных сетях. Каждый слой сети принимает на вход карты объектов от всех предыдущих слоев. Такая передача позволяет качественней передавать объекты между слоями.

Сеть ResNet50V2 с помощью остаточных блоков решает проблему исчезновения градиента. (что за поблема?). В резидентной сети блоки составляют стэк.

Inception мощная модель, с высокой точностью извлечения признаков и классификации изображения на основе этих признаков. Модель состоит из 48 слоев, в каждом из которых 11 модулей. Каждый модуль состоит из сверточных фильтров, пулинг слоев, и модуля функции активации ReLU. Занечение дропаута – 0.6.

При распределении весов в ансамбле больший вес присваивается модели с меньшей ошибкой валидации.

-предоработка изображений

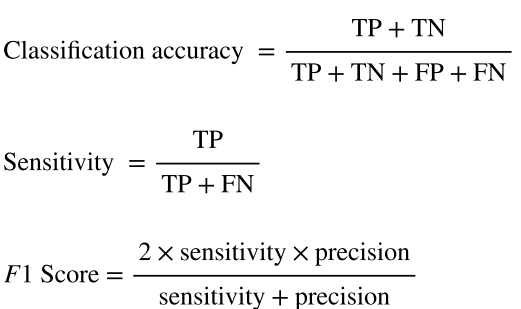
Датасет собран из других открытых наборов данных. Используются фронтальные снимки легких, поскольку на такой профиль снимка может быть лучше изучен. Собранные набор данных содержал три класса изображений – Ковид, пневмония, нормальный. Собранный датасет был разделен на два класса – Ковид+ и Ковид-. В итоге изображений класса Ковид+ в датасете 538 и 468 класа Ковид-.

-техническое описание

Обучение и тестирование проведено на платформе Google Colab, с использованием языка Python, библиотеки глубокого обучения Tensorflow.

-метрики качества сети

Оценка эффективности модели основывается по следующим параметрам: точность классификации, чувствительность и цена F1.



where TP stands for True Positive, FP for False Positive, FN for False Negative and TN for True Negative.

-сравнение с бенчмарком

Результат эксперимента

Модель принятия решения, основанная на технике средневзвешенного решения показала точность 91,62%, и чувствительность 95%.

Развитие разработки нейронных сетей осуществляется по направлениям создания новых архитектур под конкретные задачи, или же путем комбинирования и настройки нескольких моделей, и выработки результата на основе взвешенного подхода.

1.3Проблематика данных

Один из этапов обучения нейронной сети является предобработка имеющихся данных. Необходимость предобработки связана с недостатком качественных, размеченных и подтвержденных данных.

При разработке каждой модели нейронной сети требуется размеченный тренировочный набор данных.

Общей проблемой данных является их доступный объем. Имеющиеся открытые датасеты имеют разные характеристики по объему, по качеству и размеру изображений, сбалансированность по классам.

Также существенным барьером для развития искусственного интеллекта является наличие качественной верификации данных. Процесс разработки качественной разметки тренировочного набора данных требует вовлечение экспертов в своей области. Разметка набора данных трудоемкий процесс, поэтому многие публично доступные наборы данных небольшого объема.

Каждый медицинский центр может создавать датасеты различного качества разметки и формата, что затрудняет процесс обучения.

Следовательно, качество обучения алгоритмов анализа медицинских изображений связано с качеством тренировочного датасета.

Таким образом, существует проблема сбора, верификации, накопления и предоставления данных.

В России для решения проблемы качественных данных в 2020 году приступили к формированию федеральных (общероссийских) датасетов экспертами Центра диагностики и телемедицины. Первый датасет будет ориентирован на рентгенограммы и снимки компьютерной томографии пациентов с признаками COVID-19. Целью проекта является создание эталонного датасета, с помощью которого специалисты смогут оценивать корректность и точность работы сервисов по выявлению патологий в исследуемых областях.

Важной особенностью датасета является полная деперсонализация данных, то есть полное исключение персональных сведений о пациенте. Такой подход снимает вопрос защиты персональных данных при передаче накопленного массива исследователям и разработчикам.

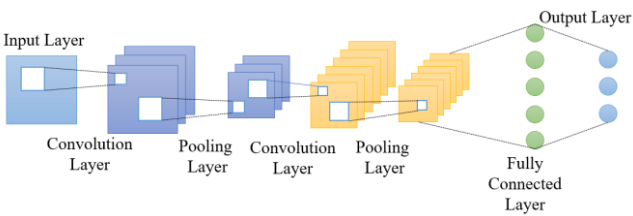
В период пандемии исследователи Центра диагностики и телемедицины собрали самую большую в мире базу данных пациентов с признаками коронавирусной пневмонии. База собиралась в сжатые сроки, что было вызвано необходимостью автоматизации процесса распознавания патологических изменений в легких при коронавирусной инфекции. На основе датасета, представленного в открытом доступе, разработчики всего мира могли "дообучать" и тестировать алгоритмы своих сервисов.

Выводы

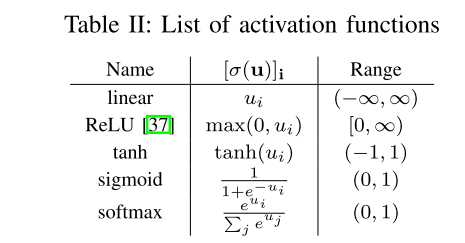
Создание больших по объему, качественных данных со стандартными характеристиками, предоставление свободного доступа к имеющимся наборам данных, позволит ускорить разработку и тестирование новых моделей. Качественный набор данных должен подкрепляться разметкой изображений экспертами. Применение единых стандартов к качеству и размеру изображений, позволит избежать трудоемкой разработки методик предварительной обработки изображений, а разрабатывать архитектуру нейронной сети.

1.4 Архитектура сверточных нейронных сетей

Фактически, архитектура сверточной нейронной сети представляет собой нейронную сеть, содержащую несколько скрытых слоев, каждый из которых имеет несколько двумерных плоскостей, состоящих из нескольких нейронов. Кроме того, все нейроны предполагаются независимыми. Его входные данные можно рассматривать как двумерное изображение, а модульизвлечения объектов встроен в архитектуру CNN. На рисунке представлена базовая архитектура сверточной нейронной сети.

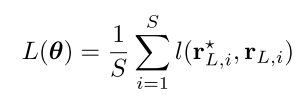
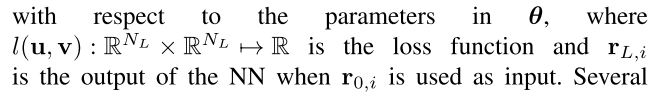


Функция активации вводит нелинейность в нейронную сеть, имитирующую функцию, при которой только нейронные электрические сигналы, превышающие определенный порог, могут быть переданы следующему нейрону. Без этой нелинейности не было бы большого преимущества в соединении нескольких слоев друг с другом. Как правило, функция активации применяется индивидуально к каждому элементу его входного вектора. Некоторые часто используемые функции активации перечислены в таблице 1

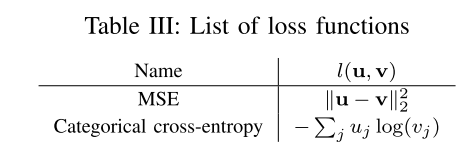


Нейронную сеть обычно обучают на размеченных данных, то есть, пар векторов ввод-вывод  

желаемый выходной сигнал нейронной сети, когда  используется в качестве входных данных. Цель обучения минимизировать значение функции потерь.

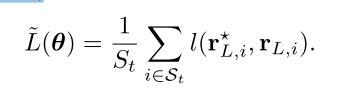
Примеры функций потерь указаны в таблице 2



Различные нормы (например, L1, L2) параметров или активаций могут быть добавлены к функции потерь в пользу решений с малыми или разреженными значениями (форма регуляризации). Наиболее популярным алгоритмом для поиска хороших наборов параметров θ является стохастический градиентный спуск (SGD), который начинается с некоторых случайных начальных значений θ = θ 0, а затем итеративно обновляет θ как



где η > 0 - скорость обучения и ˜ L(θ) - это аппроксимация функции потерь, которая вычисляется для случайной мини- партии обучающих примеров S t ⊂ {1, 2, . . . , S} размера S t на каждой итерации, т.е.,



Выбирая S t малым по сравнению с S, сложность вычисления градиента значительно снижается при одновременном снижении дисперсии обновления веса. Cуществует множество вариантов алгоритма SGD, которые динамически адаптируют скорость обучения для улучшения сходимости.

Определения и обучение нейронных сетей различных архитектур реализованы во многих существующих библиотек глубокого обучения.

В настоящее время существует множество инструментов для развертывания и обучения нейронных сетей. Такими инструментами являются библиотеки TensorFlow, Torch, MXNet и многие другие. В реестр отечественного программного обеспечения внесена библиотека PuzzleLib. Среди преимуществ PuzzleLib - высокая скорость работы, автоматическая оптимизация вычислений, большой набор нейросетевых модулей, которых насчитывается более 60 видов и ряд других.

Входной слой

Входной набор необработанных данных может быть непосредственно передан во входной слой. Одно изображение фактически вводится по его пиксельному значению во входной слой.

Сверточный слой

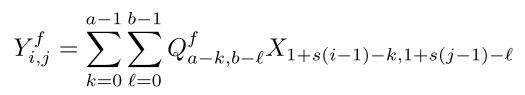
Также известен как уровень с расширенной выборкой, предназначен для извлечения объектов из входных данных. Свертка - это ключевой шаг для извлечения объектов. Каждый сверточный слой имеет свое собственное сверточное ядро, которое извлекает различные объекты из входных данных. Количество извлеченных объектов растет по мере увеличения числа сверточных ядер, включенных в слой. Выходные данные свертки называются картами объектов.

Слои сверточной нейронной сети были описаны в работе Y. LeCun et al., “Generalization and network design strategies,” Connectionism in perspective, pp. 143–155, 1989., чтобы обеспечить эффективный метод обучения для двумерных изображений. Связывая смежные сдвиги одних и тех же весов вместе способом, аналогичным тому, как фильтр скользит по входному вектору, сверточные слои способны принудительно изучать объекты с инвариантностью к сдвигам во входном векторе. Они также значительно снижают сложность модели, необходимую для представления эквивалентных объектов, инвариантных к сдвигу, с использованием полностью связанных слоев, уменьшая сложность оптимизации SGD и улучшая обобщение на соответствующих наборах данных.

В общем случае сверточный слой состоит из набора весов F фильтров

F называется глубиной

которые генерируют для каждого так называемую карту объектов из входной матрицы  исходя из следующей свертки:



Где s ≥ 1 это целое значение, называемое **шаг** число пикселей, на которое смещается окно после каждой операции, n’ = 

****, и предполагается что Х дополняется нулями  

Выходные размеры могут быть уменьшены либо путем увеличения шага s, либо путем добавления объединяющего слоя (Слой пуллинга).

Слой пулинга

Его основная функция состоит в завершении второго извлечения данных объектов, и передачи следующему сверточному слою. При нормальных условиях архитектура СНН содержит, по меньшей мере, два сверточных слоя и два слоя с понижающей дискретизацией соответственно. Чем больше уровней архитектуры задано, тем больше вероятность того, что извлечение объектов из входных данных поможет решению задачи классификации.

Объединяющий слой разбивает карту объектов на области p × p, для каждой из которых он вычисляет одно выходное значение, например, максимальное или среднее значение или L2-норму.

Полносвязный слой

Все карты объектов соединяются вместе в качестве входных данных и представляют вектор размерности n\*1, где n это количество фильтров финального слоя. В общем случае узлы нейронов в более позднем слое соединены с узлами нейронов в предыдущем слое, но узлы в каждом слое разъединены. Этот слой интегрирует и нормализует абстрагированные функции предыдущих сверток, чтобы получить вероятность для различных условий.

Выходной слой

Количество нейронов в этом слое задается в соответствии с требуемыми условиями. Если требуется классификация, то количество нейронов обычно соотносится с количеством категорий, подлежащих классификации.

1.5 Методы построения ансамбля

БУстинг

Бэггинг

….

1.6 Техническая реализация модели

Модель реализована на языке Python, поскольку используемые библиотеки и алгоритмы машинного обучения реализованы также на этом языке программирования.

Модель использует библиотеки: tensorflow для построения и обучения алгоритмов машинного обучения.

Вычисления производились на графической карте Nvidia GeForce GTX 1660 TI С 6 гигабайтами видеопамяти и на удалённом сервере предоставляемого Google Colaboratory.