

Институт математики, физики и информационных технологий
(наименование института полностью)

01.03.02 Прикладная математика и информатика

Системное программирование и компьютерные технологии
(направленность (профиль)/специализация)

Тольятти 2018

АННОТАЦИЯ

Тема: «Алгоритмы анализа и классификации патологий на рентгенограммах грудной клетки, с использованием нейронных сетей».

Целью ВКР является повышение точности алгоритмов распознавания и классификации патологий грудной клетки.

Объектом ВКР является диагностика патологий грудной клетки.

Предмет исследования – нейросетевые алгоритмы распознавания и классификации патологий грудной клетки.

В данной выпускной квалификационной работе исследуются патологии грудной клетки и методы их диагностики. Рассматриваются основные задачи систем распознавания и методы повышения точности распознавания.

В работе представлены результаты тестирования нескольких нейронных сетей. Одни обучены на сегментированных снимках, на которых удалена лишняя информация, другие – на обычных.

Структура ВКР состоит из введения, трех глав, заключения, списка литературы.

Во введении описывается актуальность проводимого исследования, формулируется цель и ставятся задачи, которые необходимо решить.

В первой главе приводится анализ дифференциации патологий грудной клетки и методов распознавания.

Во второй главе рассматривается процесс разработки нейронных сетей и подготовки исходных данных.

В третьей главе рассматривается процесс тестирования и внедрения нейронных сетей.

Бакалаврская работа состоит из пояснительной записки, включая 15 рисунков, 10 источников на иностранном языке и 1 приложения.

ABSTRACT

The title of the graduation work is “Algorithms for the analysis and classification of pathologies on the chest X-ray, using neural networks”.

The goal of the graduation work is to increase the level of algorithms for recognizing and classifying pathologies of the chest.

The object of the graduation work is the diagnosis of chest pathologies.

The subject of the study is neural network algorithms for recognizing and classifying pathologies of the chest.

In this graduation work, chest pathologies and methods for their diagnosis are investigated. The main tasks of recognition systems and methods for increasing recognition accuracy are considered.

The main tasks of recognition systems and methods for increasing recognition level are considered.

The paper presents the results of testing several neural networks. Some are trained in segmented images, on which excess information has been deleted, others – on ordinary pictures.

The structure of the graduation work consists of an introduction, three chapters, a conclusion, a list of literature.

The introduction describes the relevance of the ongoing research, formulates the goal and sets tasks that need to be addressed.

In the first part of the review of methods of analysis of recognition systems is provided.

The second chapter deals with the process of developing neural networks and preparing initial data.

The third chapter examines the process of testing and implementing neural networks.

The graduation work consists of an explanatory note, including 16 drawings, 10 sources in a foreign language and 1 appendixes.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	9
1 АНАЛИЗ СПОСОБОВ ДИФФЕРЕНЦИАЦИИ ПАТОЛОГИЙ ГРУДНОЙ КЛЕТКИ	10
1.1 Методы диагностики грудной клетки	10
1.2 Основные патологии легких	15
1.3 Анализ методов распознавания	17
1.3.1 Описание задач распознавания	17
1.3.2 Описание методов определения класса	19
1.3.3 Подходы реализации методов определения класса	20
1.3.4 Анализ нейросетевого подхода	22
1.4 Формирование требований к алгоритмам распознавания	30
1.4.1 Определение функционала нейронных сетей	30
1.4.2 Выбор архитектуры нейронных сетей	30
1.4.3 Анализ работы сверточной нейронной сети	35
2 РАЗРАБОТКА НЕЙРОННЫХ СЕТЕЙ	38
2.1 Обоснование выбора технических средств	38
2.2 Архитектура нейронных сетей	40
2.3 Подготовка исходных данных	42
2.4 Обучение сетей	43
3 ТЕСТИРОВАНИЕ И ВНЕДРЕНИЕ НЕЙРОННЫХ СЕТЕЙ	45
3.1 Тестирование сетей	45
3.2 Размещение сетей на сервере	46
ЗАКЛЮЧЕНИЕ	49
СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ	50
ПРИЛОЖЕНИЕ А	52

ВВЕДЕНИЕ

Заболевания, связанные с органами дыхания, представляют важную проблему для здоровья человека. Под влиянием неблагоприятных факторов среды у многих людей развиваются хронические болезни. Такие болезни очень опасны, но своевременное обращение к специалистам помогает избежать тяжелых осложнений.

Часто даже опытным специалистам тяжело поставить точный диагноз. Компьютерные технологии помогают ускорить обследования, получить более точные сведения о заболевании. Внедрение компьютерных технологий в медицину обеспечило высокую точность и скорость проведения различных исследований, а это дало возможность использовать ранее недоступные методы. Так например, стало доступным использование нейронных сетей.

Нейронные сети моделируют работу, протекающую в мозге, и могут обучаться и исправлять ошибки [3]. Это позволяет применять нейронные сети в различных задачах, таких как, распознавание образов, классификация, прогнозирование, оптимизация и анализ данных. Нейронная сеть позволяет хранить важную информацию, находить закономерности в образах.

Цель ВКР: повышение точности алгоритмов распознавания и классификации патологий грудной клетки.

Объект ВКР: диагностика патологий грудной клетки.

Предмет ВКР: нейросетевые алгоритмы распознавания и классификации патологий грудной клетки.

Основные задачи ВКР:

- исследовать методы диагностики грудной клетки;
- исследовать основные патологии грудной клетки (легких);
- исследовать основные методы систем распознавания;
- исследовать архитектуры нейронных сетей;
- разработать нейронные сети для классификации патологий легких.

1 АНАЛИЗ СПОСОБОВ ДИФФЕРЕНЦИАЦИИ ПАТОЛОГИЙ ГРУДНОЙ КЛЕТКИ

1.1 Методы диагностики грудной клетки

Для исследования органов дыхания применяются различные методы. Рассмотрим основные из них.

Рентгенологические исследования легких:

- рентгеноскопия;
- флюорография;
- рентгенография.

Рентгеноскопия – метод обследования органов дыхания, который заключается в получении изображения легких с помощью рентгеновских лучей на специальном экране. Такой метод позволяет увидеть различные патологии в легочной ткани, наличие жидкости или воздуха в плевре [1].

Основные преимущества метода:

- возможность проводить исследование в разных проекциях;
- возможность оценить локализацию некоторых изменений.

Основные недостатки метода:

- невозможно увидеть мелкие очаги патологий;
- высокая доза облучения.

Флюорография – метод исследования грудной клетки с применением рентгеновских лучей. В результате работы метода получается изображение, которое записывают на особо чувствительную пленку.

Метод используется для обнаружения наличия патологий на ранней стадии [1].

Флюорография позволяет обнаружить:

- воспаления тканей;
- опухоли;
- уплотнения и рубцовые преобразования;
- инородные предметы.

Преимущества метода:

- минимальные затраты времени;
- более низкое радиационное воздействие, по сравнению с другими методами рентгенологического исследования;

Недостатки метода:

- высокая степень погрешности, при использовании пленочного аппарата;
- не позволяет детально рассмотреть патологии.

Рентгенография – это один из основных методов рентгенологического исследования, состоящий в получении изображения с помощью рентгеновских лучей. Такой метод очень точный и информативный. Он позволяет увидеть даже самые мелкие изменения. Для более точного исследования снимки производят в двух проекциях. Метод применяется для обнаружения пневмонии, туберкулеза, различных травм и опухолей [1].

Преимущества метода:

- возможность рассмотреть детали изображения;
- меньшая лучевая нагрузка по сравнению с рентгеноскопией.

Недостатки метода:

- большие затраты времени для обработки изображений по сравнению с флюорографией;
- невозможность проводить исследование в разных проекциях без использования большой лучевой нагрузки.

Компьютерная томография – метод диагностики дыхательной системы, который позволяет детально рассмотреть исследуемый объект по слоям. В компьютерной томографии изображения получают с помощью рентген лучей. Полученные изображения передаются компьютеру для обработки и анализа специалистами [1].

Метод позволяет обнаружить:

- воспаления тканей;

- врожденные патологии;
- опухоли;
- онкологические заболевания;
- уплотнения и рубцовые преобразования;
- травмы грудной клетки.

Преимущества:

- более детальное исследование;
- более четкая визуализация структур органа.

Недостатки метода:

- при применении контрастирования требуется подготовка;
- наличие ограничений для проведения исследования с контрастированием;
- высокая стоимость проведения исследования.

Томографические исследования легких:

- магнитно-резонансная томография;
- позитронно-эмиссионная томография.

Магнитно-резонансная томография – вид медицинского исследования, который заключается в получении снимков с помощью радиоволн. В основе метода лежит использование мощного магнитного поля. С помощью этого метода точно определяется вид, форма и локализация патологии [1].

С помощью этого метода возможно обнаружить:

- наличие уплотнений и опухолей в легочной ткани;
- наличие воспалительных процессов в лимфоузлах;
- состояние плевральной полости;
- наличие патологий в области бронхов.

Преимущества метода:

- более высокая точность чем в радиологических исследованиях;
- магнитные поля менее вредные чем рентген излучение;
- метод исключает человеческий фактор;

- безопасное контрастное вещество.

Недостатки метода:

- противопоказания людям с кардиостимуляторами и металлическими имплантатами;
- высокая стоимость проведения исследования;
- большие временные затраты.

Позитронно-эмиссионная томография – медицинский вид исследования, позволяющий хорошо изучить состояние органов. Для получения изображения в методе применяется радиоактивное химическое вещество и специальный сканер [1].

Метод применяется для:

- обнаружения очагов;
- определения наличия метастаз;
- обнаружения изменений легочных структур.

Преимущества метода:

- позволяет получить информацию о качестве и природе патологических изменений.
- позволяет обнаруживать очень ранние образования.

Недостатки метода:

- необходимость введения радиационного вещества для качественного исследования;
- изображения содержат меньше информации чем другие томографические исследования;
- высокая стоимость проведения исследования.

Эндоскопические исследования легких:

- Бронхоскопия;
- Торакоскопия.

Бронхоскопия – метод исследования гортани, трахеи и бронхов с помощью медицинского оборудования. Оборудование состоит из прибора для

освещения, видеокамеры и хирургического инструмента. Видеокамера применяется для трансляции и сохранения изображений, а хирургический инструмент для удаления инородных тел и мелких опухолей.

Метод применяется для:

- диагностики хронического бронхита, пневмонии;
- удаления инородных тел из органов дыхания;
- удаления мелких опухолей;
- очищения слизистой.

Преимущества метода:

- возможность обследования нижних дыхательных путей;
- точность исследования порядка 97%;
- возможность применения для лечебных целей.

Недостатки метода:

- необходимость введения эндоскопа в дыхательные пути;
- возможность механического повреждения гортани.

Транскопия – метод эндоскопической диагностики, который заключается в исследовании плевральной полости.

Метод применяется для:

- исследования плевры;
- обнаружения патологий в области диафрагмы, средостения, перикарда;
- удаления избытка жидкости.

Преимущества метода:

- возможность вывода на экране отдельных структур с многократным увеличением;
- возможность проведения хирургического вмешательства.

Недостатки метода:

- необходимость прокола грудной клетки;
- множество противопоказаний;
- возможные осложнения.

Среди всех рассмотренных методов оптимальными являются методы, относящиеся к рентгенологическим исследованиям, а именно флюорография и рентгенография. Эти методы наиболее распространены, не требуют больших временных затрат и изображения, полученные с помощью этих методов, обладают достаточным набором информативных признаков для классификации. Далее будут рассмотрены основные патологии, которые могут быть обнаружены на изображениях.

1.2 Основные патологии легких

Отличительная особенность многих легочных заболеваний – сложная диагностика. Каждая причина отдельной патологии носит субъективный характер. Но существует ряд общих факторов, которые провоцируют возникновение и развитие процесса воспаления:

- неблагоприятная экологическая обстановка данного региона;
- наличие вредных производств;
- обострение других хронических заболеваний;
- нарушение электролитного баланса в организме, повышение уровня сахара;
- наличие вредных привычек.

Среди основных патологий легких выделяют: пневмонию, инфильтрацию, фиброз, гидроторакс, ателектаз, пневмоторакс, эмфизему.

Рассмотрим подробнее несколько патологий.

Пневмония – это воспалительный процесс, возникающий в легких в случае заражения инфекцией.

На рентгеновском снимке характеризуется:

- затемнением легочной ткани;
- контрастностью легочного рисунка;
- увеличением легочного корня на воспаленной стороне.

Инфильтрация – уплотнение легочной ткани за счет клеточных элементов, химических веществ. На рентгеновском снимке характеризуется

уплотнениями неправильной формы. Возникает в результате инфекции или травмы.

Фиброз лёгких – замещение легочной ткани соединительной тканью. Возникает по нескольким причинам:

- проникновение инфекции;
- наличие аллергии;
- продолжительное вдыхание пыли.

На рентгеновском снимке характеризуется изменениями в нижних отделах легких.

Гидроторакс – скопление жидкости в области плевры. Обычно протекает совместно с другими патологиями. Сжимает и уменьшает размер легочной ткани. На рентгеновском снимке характеризуется затемнениями средней интенсивности.

Ателектаз – патология, которая характеризуется отсутствием воздуха в легочной ткани. Возникает в результате изменения проходимости бронхов.

На рентгеновском снимке характеризуется:

- нечеткими границами сердечных отделов;
- трехгранным затемнением;
- подъемами легочных корней.

Пневмоторакс – патология, которая характеризуется скоплением воздуха в области плевры. Возникает в результате различных механических повреждений. На рентгеновском снимке характеризуется опущенным куполом диафрагмы, просветлением с отсутствием легочного рисунка.

Все патологии имеют набор отличительных характеристик, с помощью которых возможно произвести классификацию. Далее будут рассмотрены основные задачи и методы распознавания, которые применяются для классификации изображений.

1.3 Анализ методов распознавания

1.3.1 Описание задач распознавания

Распознавание образов – это наука, которая изучает классификацию объектов по нескольким категориям или классам. Распознавание применяется для идентификации объектов, систем безопасности и учета, обнаружения дефектов, диагностики в медицине и т.д. Объекты распознавания обладают набором признаков. Обучающая выборка состоит из нескольких эталонных объектов, для которых точно определен класс. Алгоритм распознавания основывается на известной информации и определяет класс для новых поступающих объектов.

Проведем анализ основных задач при разработке систем распознавания образов.

Рассмотрим представление исходных данных распознаваемого объекта. Все измерения образа являются характеристиками распознаваемого объекта. Если количество элементов обозначить за n , то измерения можно представить в виде вектора образа:

$$x = (x_1, x_2, \dots, x_n)^T \quad (1)$$

Образы – это непрерывные функции. Если проводить измерения в дискретных точках, то вектор образа имеет вид:

$$x_1 = f(t_1), x_2 = f(t_2), \dots, x_n = f(t_n) \quad (2)$$

Измерение значений функций в дискретных точках показано на рисунке 1.

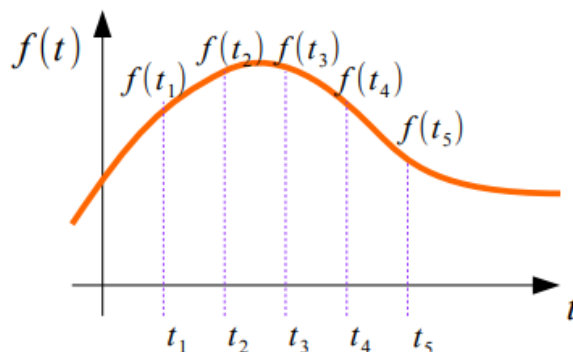


Рисунок 1 – Измерение значений функций в дискретных точках

Если измерения приводят к информации в форме действительных чисел, то векторы образов рассматривают как n -мерное евклидов пространство.

Образы, принадлежащие к одному классу, рассматривают как совокупность точек, лежащих в одной области. Тогда вектор образа имеет вид:

$$x = (x_1, x_2)^T \quad (3)$$

Демонстрация векторов образов в качестве точек n-мерного пространства показана на рисунке 2.

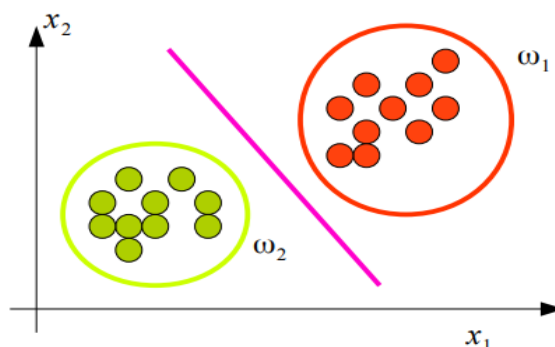


Рисунок 2 – Векторы образов в качестве точек n-мерного пространства

Обработка исходных данных заключается в выборе и выделении характерных признаков и снижении размерности векторов образов. Множество образов одного класса, имеет особые отличительные признаки. Такие признаки называются межклассовыми. Множество образов, состоящее из образов всех классов, содержит внутриклассовые признаки. Такие признаки не используются в задачах распознавания образов. Выделение признаков сложная процедура, но необходимая для большинства практических задач автоматического распознавания.

Для классификации и идентификации необходимо найти решающие функции, которые построят границы области решения и разделят точки в пространстве измерений. Классифицируемые объекты должны обладать набором измеримых характеристик. Для получения решающих функций можно воспользоваться различными методами. Если известны априорные сведения, то решающие функции определяются с помощью этой информации. Если известны только качественные сведения, то границы области решения могут существенно отклониться от истинных. Поэтому следует воспользоваться

дополнительной информацией, полученной с помощью лингвистических статистик и подкорректировать решающие функции.

Точность распознавания зависит от количества различающей информации, которая содержится в характеристиках объекта. Чтобы повысить точность распознавания, необходимо измерить все возможные характеристики и увеличить количество времени для обработки.

1.3.2 Описание методов определения класса.

Среди основных методов выделяют:

- определение класса перечнем входящих в него членов;
- определение класса свойствами, общими для всех его членов;
- определение класса образованием кластеров в пространстве признаков.

Рассмотрим подробнее каждый из методов.

Если класс определен перечнем входящих в него членов, то алгоритм распознавания основывается на сравнении объекта с эталоном. В системе сохраняется несколько объектов, принадлежащих определенному классу. При подаче системе новых объектов, она сравнивает их с эталонами. Если какой-либо объект совпадает с эталоном, то он относится к соответствующему классу. Такой метод подходит для небольших систем с малым количеством элементов и признаков.

Если класс определен классовыми свойствами входящих в него членов, то алгоритм распознавания основывается на выделении признаков. Объекты одного класса содержат общие свойства и признаки. Когда системе подается новый объект, то признаки этого объекта сравниваются с признаками, хранящимися в памяти системы распознавания. При наличии необходимых признаков, система относит объект к определенному классу.

В отличие от системы с определением перечня классов входящих в него членов, принцип сопоставления признаков допускает вариативность характеристик. Основная сложность такого метода заключается в отыскании полного набора признаков.

Если класс определен образованием кластеров в пространстве признаков, то классификация основывается на принципе геометрической близости. Объекты образуют область в пространстве признаков, называемую кластером. В таких системах, класс определяется на основе расположения объекта относительно кластера. В случае, если кластеры располагаются далеко друг от друга, то объекты можно отнести к классу, расстояние до которого минимально.

1.3.3 Подходы реализации методов определения класса

Для реализации методов распознавания, используются различные подходы. Рассмотрим основные из них.

Эвристические подходы основаны на опыте и интуиции человека. Они включают перечисление всех членов класса и их общих свойств. Эти подходы применяются только для конкретных задач распознавания. Точность распознавания сильно зависит от опыта разработчиков.

Математические подходы основываются на общности свойств объектов и кластеризации. Условно математические подходы можно разделить на детерминистские и статические.

Детерминистские подходы основываются на эмпирических и индуктивных методах. В подходах используется моделирование действий и математическая логика.

К детерминистским подходам относят:

- построение решающих правил;
- метод построения эталонов;
- метод дробящихся эталонов;
- линейные решающие правила;
- метод ближайших соседей;
- метод потенциальных функций;
- структурные (лингвистические) методы;
- метод кластерного анализа;

- критерии информативности признаков;
- метод отбора информативных признаков.

Статистические подходы основываются на стохастической аппроксимации, последовательном анализе, теории информации и других результатах математической статистики.

К статистическим подходам относят:

- параметрическое оценивание распределений;
- метод максимума правдоподобия;
- метод статистически независимых признаков;
- метод при неизвестных априорных вероятностях образов;
- метод потенциальных функций;
- метод ближайших соседей.

Некоторые детерминистские методы были статистически обоснованы, поэтому существует несколько методов, относящихся к двум подходам.

В лингвистическом подходе образы рассматриваются в виде структуры формальных языков. Сначала выбираются произвольные элементы объекта. Как в формальных языках образы составляют грамматику, которая содержит множество элементов и правила подстановки.

Среди грамматик выделяют:

- регулярные;
- бесконтекстные;
- грамматики составляющих.

Такой подход применяется для объектов, которые не могут быть описаны числовыми измерениями.

Нейронная сеть – программа, работающая с помощью особых алгоритмов. Нейронная сеть включает в себя математическую модель и может быть использована для различных задач, в том числе и для распознавания изображений. Нейронная сеть обладает способностью обобщать исходные данные и обучаться с помощью специальных математических алгоритмов.

В нейронной сети могут использоваться эвристические и эмпирические концепций.

Преимущества нейронной сети:

- устойчивость к шумам входных данных;
- адаптация к изменениям;
- высокое быстродействие;
- наличие сверточной архитектуры для работы с изображениями.

Нейросетевой подход обладает рядом преимуществ для работы с изображениями и отлично подходит для задачи классификации. Далее нейросетевой подход будет рассмотрен более подробно.

1.3.4 Анализ нейросетевого подхода

Нейронная сеть является аналогом биологической нейронной сети. Она состоит из слоев нейронов, соединенных между собой. Соединения нейронов образуют различные структуры, способные выполнять различный круг задач. Технология нейронных сетей развивалась обособленно от классических методов распознавания, но оказывала значительное влияние на терминологический и методологический аппараты этих дисциплин [5].

Рассмотрим подробнее общую архитектуру нейронной сети. В сети имеются специальные числовые параметры, называемые веса. На вход сети подаются вектор данных. Выходной вектор является ответом для решаемой задачи. Работа сети заключается в преобразовании входного вектора в выходной вектор с помощью весов.

Условно нейронные сети можно разделить по совокупности критериев на статические и динамические.

Статических нейронные сети – это сети прямого распространения. В таких сетях между слоями используется связь в одном направлении, и выход сети зависит только от входных параметров.

Среди статических нейронных сетей выделяют:

- персептрон;
- сеть Кохонена;

- когнитрон;
- неокогнитрон;
- свёрточную нейронную сеть.

К динамическим нейронным сетям относят сети, в которых используются динамические переменные, реализована структура с обратной связью. В таких сетях новое состояние сети зависит от предыдущего.

Среди динамических нейронных сетей выделяют:

- нейронную сеть Хопфилда;
- нейронную сеть Коско;
- нейронную сеть Джордана;
- нейронную сеть Элмана.

Рассмотрим подробнее работу нейронной сети на примере однослойного персептрона.

Однослойный персептрон состоит из различных элементов. Входные данные передаются с помощью сенсорных элементов ассоциативным элементам, которые выступают как связующие элементы. Ассоциативные элементы активизируются при достижении определенного числа сигналов от сенсорных элементов и передают сигнал реагирующим элементам. В зависимости от полученного сигнала результирующие элементы выдают какой-либо результат. Логическая схема однослойного персептрона представлена на рисунке 3.

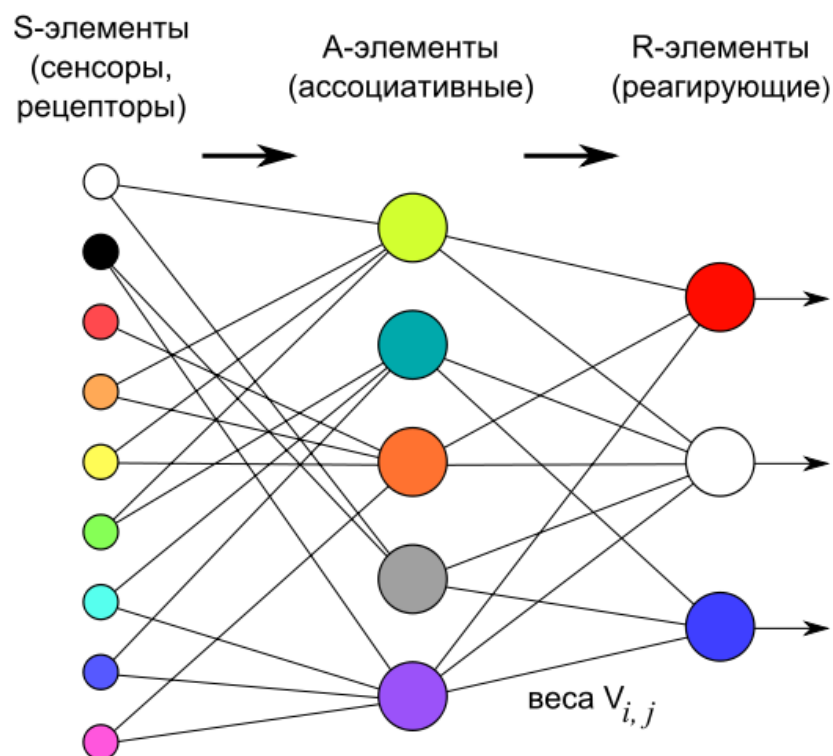


Рисунок 3 – Однослойный персептрон

Обучение нейронной сети заключается в изменении весовых коэффициентов между связями ассоциативных и реагирующих элементов. Обучение позволяет настроить персептрон на работу, связанную с задачами распознавания или обобщения. Для обучения применяются детерминированные и стохастические методы.

Детерминированные методы основываются на итеративной коррекции параметров сети. Каждая текущая итерация основывается на текущих параметрах. Примером детерминированного метода может служить метод обратного распространения ошибки, который далее будет рассмотрен.

К стохастическим методам относят методы, которые работают с помощью сравнения и изменения параметров сети. Такие сети изменяют параметры сети случайным образом, сохраняя важные изменения. В таких методах имеются решения для некоторых проблем, возникающих в результате обучения сети. Так например, ловушка локального минимума решается с помощью стохастических усложнений.

Методы обучения также условно можно разделить на обучение с учителем и обучение без учителя.

Метод обучения с учителем. В этом методе для обучения используются данные с заранее определенными для них классами. После получения результата алгоритм сравнивает его с правильным ответом, на основе чего происходит корректировка весов.

Метод обучения без учителя. В этом методе сети не известна принадлежность данных к классам. Корректировка весов происходит на основе поиска закономерностей, особых взаимосвязей между объектами.

Рассмотрим несколько основных методов обучения.

Метод обратного распространения ошибки. Этот метод применяется, если принадлежность объектов к классам известна заранее. Он основан на минимизации среднеквадратической ошибки. Путем движения входных сигналов к выходу формируется выходной сигнал, из которого вычисляется значение ошибки. Далее полученное значение начинает движение в обратном направлении, корректируя веса сети [4]. Метод имеет вид:

$$E = \frac{1}{2} \sum_{k \in OUT} (t_k - o_k)^2 \quad (4)$$

где E – функция ошибки;

$\omega_{i,j}$ – синаптический вес между нейронами i, j ;

OUT – множество выходов сети;

t_k – правильные ответы сети;

o_k – выход k -го нейрона.

На основе значений правильных результатов происходит вычисление весов по формуле:

$$\Delta \omega_{i,j} = -\eta \frac{\partial E}{\partial \omega_{i,j}} \quad (5)$$

где $0 < \eta < 1$ – множитель скорости движения

При использовании метода решаются задачи оптимизации с одним критерием. Условие использования метода – дифференцируемость функции активации. Метод является классическим и часто применяется для обучения.

Генетические алгоритмы. Генетические алгоритмы основаны на эволюции популяций и нахождении наибольшего или наименьшего значения многоэкстремальной функции. В методе используется эволюция популяций – изменение признаков хромосом так, чтобы улучшить приспособление популяции к среде. Хромосомы – векторы данных, состоящий из весовых коэффициентов. Наборы хромосом образуют популяции.

Осуществляется отбор по формуле:

$$P_i = \frac{f_i(x)}{\sum f(x)} \quad (6)$$

где P_i – нормализованная величина;

$f_i(x)$ – значение i -го вектора в популяции;

$\sum f(x)$ – сумма всех векторов в популяции.

Чтобы определить активность нейрона применяется функция активации. Она определяет выход сигнала и необходима для придания нелинейности. Произведения входных параметров и соответствующих весов суммируются и подаются для активации, после чего формируется возбуждение нейрона. Возбуждение определяется значением от -1 до 1, в зависимости от выбранной функции. Рассмотрим несколько основных функций.

Пороговая функция Хевисайда. Она подходит для простых задач бинарной классификации и в целях обучения теории нейронных сетей [4].

Значения функции определяются по формуле:

$$\theta x = \begin{cases} 0, x < 0 \\ 1, x \geq 0 \end{cases} \quad (7)$$

Сигмоидальная функция активации. Применяется для усиления слабых сигналов и ослабления сильных. При любых входных числах выдает числа в диапазоне от 0 до 1.

Недостатки функции:

- затухание градиентов при насыщении;

- выход не отцентрован относительно нуля.

График функции показан на рисунке 4.

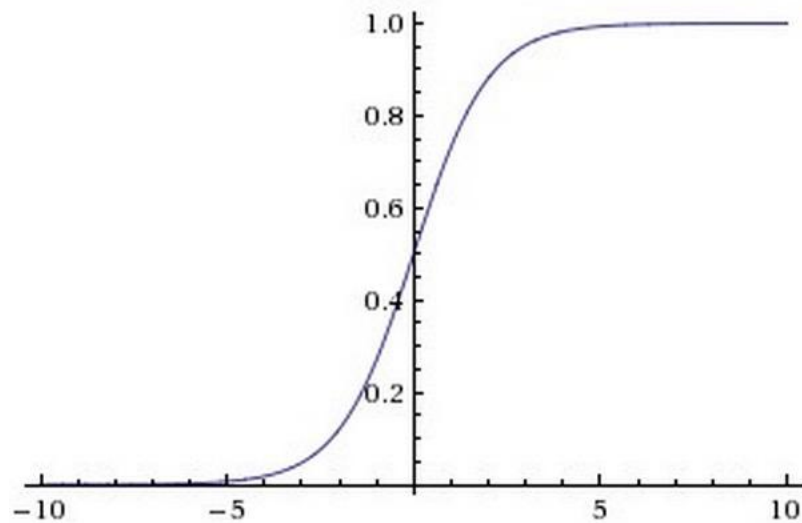


Рисунок 4 – Сигмоидальная функция активации

Формула функции имеет вид:

$$\sigma x = \frac{1}{1 + e^{-x}} \quad (8)$$

Гиперболический тангенс. Возрастающая функция, преобразующая поступающие данные в диапазоне от -1 до 1. Подвержена перенасыщению, но имеет отцентрованность относительно нуля. График функции показан на рисунке 5.

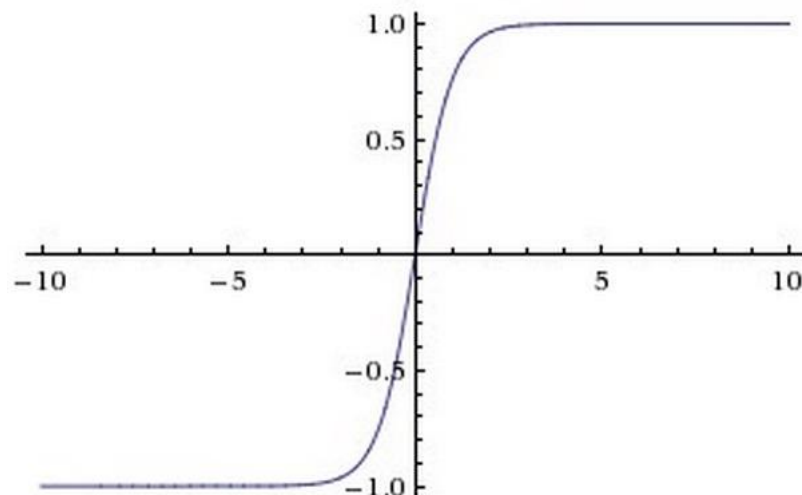


Рисунок 5 – Гиперболический тангенс

Формула функции имеет вид:

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (9)$$

Функция ReLU. Является модификацией линейной функции активации. Применяется для устранения проблем с затуханием или разрастанием градиентов.

Преимущества функции:

- ReLU не требует больших вычислений;
- не подвержена насыщению;
- повышает скорость нахождения локального экстремума.

Основным недостатком функции является – отключение большей части нейронов при неправильном подборе параметров. График функции показан на рисунке 6.

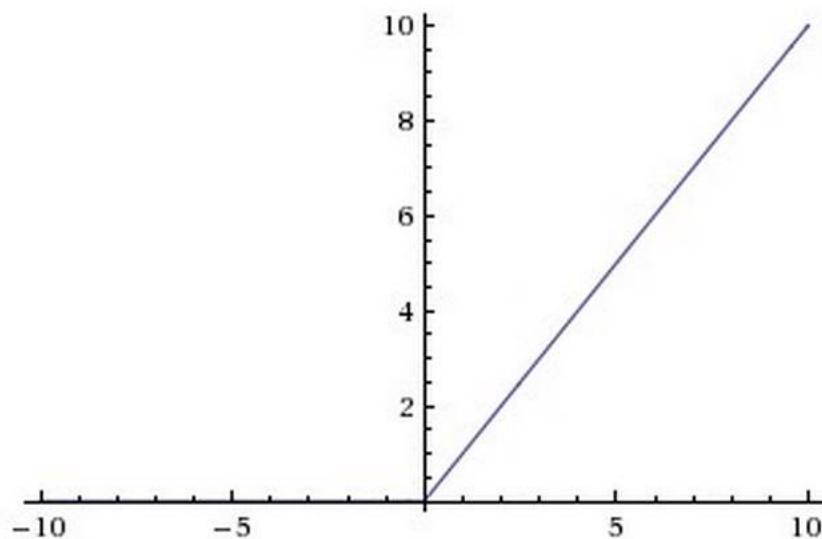


Рисунок 6 – Функция ReLU

Формула функции имеет вид:

$$f(x) = \max(0, x) \quad (10)$$

Существует несколько различных модификаций функции ReLU, устраняющих проблему отключения большей части нейронов.

Leaky ReLU. Функция увеличивает диапазон ReLU, значение функции не обнуляется при отрицательном аргументе [3]. График сравнения функции показан на рисунке 7.

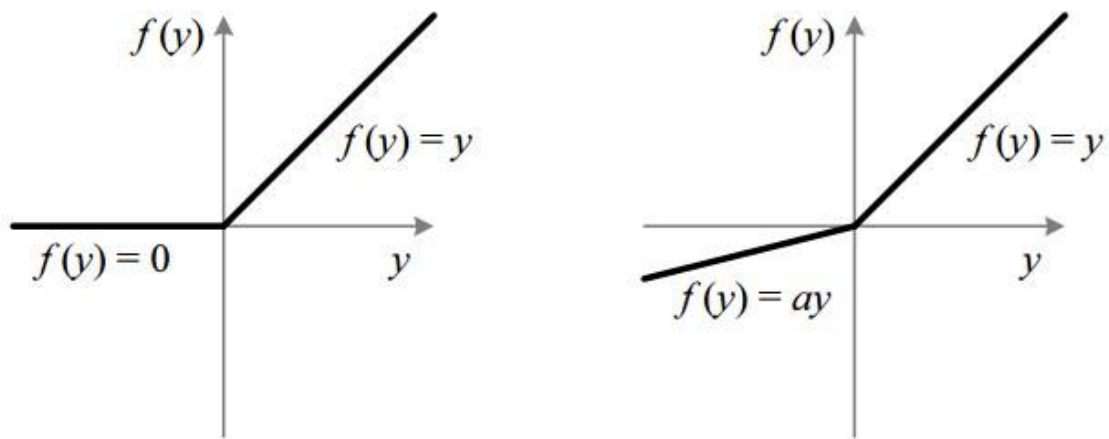


Рисунок 7 – Сравнение ReLU и Leaky ReLU

Parametric ReLU. В этой модификации на основе входных данных устанавливается угловой коэффициент для отрицательного интервала.

Randomized ReLU. В этой модификации угловой коэффициент устанавливается случайным образом из заданного интервала.

Часто во время обучения нейронной сети возникает проблема переобучения. Нейронная сеть может распознать только те примеры, которые использовались в процессе обучения. Для устранения этой проблемы используется регуляризация. К методам регуляризации относят методы, связанные с ограничением весов у признаков. Рассмотрим основные из них.

Метод Dropout. В сети создается подсеть, включающая некоторое количество случайных нейронов. Изменения весов происходят только внутри подсети. Сигнал не вошедших в подсеть нейронов равен нулю. Создается несколько таких подсетей, а затем усредняются полученные результаты.

Среди всех подходов был выбран нейросетевой. Нейронные сети могут выполнять различный круг задач и быть обучены для классификации. Они могут адаптироваться к изменениям и имеют специальную архитектуру для работы с изображениями.

1.4 Формирование требований к алгоритмам распознавания

1.4.1 Определение функционала нейронных сетей

Разработку нейронной сети можно разделить на несколько этапов:

- выбор модели сверточной нейронной сети;
- архитектура нейронной сети;
- обучение сети;
- тестирование сети;
- размещение сети на сервере.

Изображения легких обладают набором различных признаков класса. Одно изображение может одновременно относиться к нескольким классам патологий. Для достижения более высокой точности будет рассмотрена задача бинарной классификации патологий на рентгенограммах грудной клетки.

Всего будет представлено два класса:

- с патологиями;
- без патологий.

Для анализа нейросетевых алгоритмов классификации будет разработано несколько нейронных сетей с разными моделями. Чтобы повысить точность алгоритмов распознавания, для обучения будут применяться сегментированные изображения, на которых удалена вся лишняя информация, а также будут изменены и дополнены модели.

Сеть, которая покажет наилучшие результаты, будет размещена на сервере с возможностью удаленной классификации.

Далее будет рассмотрен процесс выбора архитектуры нейронной сети.

1.4.2 Выбор архитектуры нейронных сетей

Рассмотрим несколько архитектур нейронных сетей.

Нейронные сети прямого распространения и персептроны очень прямолинейны, они передают информацию от входного слоя к выходному. Нейронные сети часто описываются в виде нескольких слоев, нейроны соединены таким образом, чтобы связывать только соседние слои, а не нейроны одного слоя. В таких сетях используется обучение с методом обратного

распространения ошибки. В практических целях такие сети почти не применяются, но их можно использовать совместно с другими архитектурами.

Нейронные сети радиально-базисных функций применяются для решения различных задач, связанных с классификацией, аппроксимацией и кластеризацией. Основное свойство радиально-базисных функций – возрастание или убывание откликов, которое происходит симметрично относительно некоторой вертикальной оси. В сети определяется близость вектора выходных данных к центрам радиально-базисных функций. Все выходы в сети – это линейные комбинации выходов первого слоя.

Преимущества этой архитектуры:

- моделирование нелинейных зависимостей с помощью специального скрытого слоя;
- простой алгоритм для формирования весов;
- возможность нахождения глобального оптимума функции активации при наличии весов выходного слоя;
- быстрое обучение.

Недостатки нейронных сетей на основе радиально-базисных функций:

- необходимость настройки параметров радиально-базисных функций, сложность настройки при большом количестве скрытых элементов;
- заранее должно быть известно число эталонов для построения функций ошибки скрытого слоя.

Нейронная сеть Хопфилда – это полносвязная нейронная сеть, в которой матрица связи симметрична. Во время получения входных данных каждый узел является входом, в процессе обучения он становится скрытым, а затем становится выходом. Значения нейронов сохраняется в соответствии с установленным шаблоном. Сеть применяется для решения задач оптимизации, как фильтр или как автоассоциативная память. Сеть работает до тех пор, пока не будет достигнуто значение равновесия (следующее значение не будет равно предыдущему).

Недостатки нейронной сети Хопфилда:

- небольшой объем памяти;
- достижение состояния равновесия не гарантирует правильный ответ сети.

Машина Больцмана – особая модификация нейронной сети Хопфилда. Эта сеть имеет топологию схожую с сетью Хопфилда, но работает по другому алгоритму. Функция ошибки сети Хопфилда часто сводится к локальному минимуму не достигая нужного состояния, модификация Больцмана позволяет улучшить результаты для достижения глобального минимума.

Недостатки машины Больцмана:

- качество работы зависит от набора идеальных образов;
- работает медленнее обычной сети Хопфилда.

Ограниченная машина Больцмана – это модификация машины Больцмана, которая характеризуется особой архитектурой. Веса в сети принимают случайные значения. Связи присутствуют только между нейронами, принадлежащим к разным классам. При обучении скрытые слои запоминают признаки в данных. Сеть применяется в таких задачах как: выделение признаков, классификация, фильтрация признаков, тематическое моделирование. Такая архитектура может обучаться без учителя.

Автокодировщик – модификация сети прямого распространения. В этой архитектуре нейроны, находящиеся дальше от входных нейронов, могут влиять на активность тех нейронов, которые находятся ближе к входу. Данные в сети преобразовываются, чтобы уменьшить объем и устранить избыточность. Принцип работы сети заключается в получении на выходе отклика, наиболее близкого к входным данным. Автокодировщики применяются для обучения глубоких сетей.

Преимущества автокодировщика:

- возможность обучения без учителя;
- уменьшение шума в данных;
- уменьшение размерности многомерных данных для визуализации.

Недостатки автокодировщика:

- чтобы обучиться взаимосвязям и внутренним зависимостям, нужно иметь критерий обучения;
- не может использоваться при больших количествах мелких свойств.

Разреженный автокодировщик – модификация автокодировщика. В скрытом слое содержится большее количество нейронов, чем в входном и выходном слое. В сети действует ограничение, при котором нейроны входного и выходного слоя большую часть времени бездействуют. В сети используется фильтр разреженности, который пропускает только определенные ошибки. В гиперпространстве намного больше измерений, поэтому функция ошибки достигает нужного состояния.

Преимущества разреженного автокодировщика:

- позволяет работать с большим количеством различных свойств;
- возможность обучения без учителя.

Недостатки разреженного автокодировщика:

- требуется больше времени для обучения, чем для автокодировщика;
- активация нейронов скрытого слоя должна быть разреженной.

Вариационный автокодировщик – модификация автокодировщика, изучающего параметры распределения вероятностей, моделирующих входные данные. Такая сеть учитывает несвязанные события, происходящие в различных местах, и позволяет создавать сложные генерирующие модели данных. Во время работы сети формулируется описание распределения вероятности для каждого скрытого атрибута.

Недостатки вариационного автокодировщика:

- сложность реализации вероятностных выводов;
- обученные на изображениях выборки получаются размытыми.

Глубокая сеть доверия – архитектура нейронной сети, которая включает в себя несколько автокодировщиков, соединенных между собой. Входные данные преобразовываются таким образом, что каждый последующий скрытый слой определяет более высокоуровневые признаки изображения.

Преимущества глубокой сети доверия:

- способна запоминать обучающую выборку и нужные признаки;
- способна воссоздать информацию, которая была использована для обучения.

Недостатки глубокой сети доверия:

- необходимость предобучения;
- локальные решения, не гарантируют оптимального конечного результата.

Шумоподавляющий автокодировщик – модификация сети прямого распространения, созданная для выделения ярких признаков. Входные данные в сети подаются с шумом, и автокодировщик учится восстанавливать чистые данные. Для обучения используются методы с учителем и без учителя. Сеть применяется для предобучения более глубокой сети.

Недостатки архитектуры:

- сеть не запоминает мелкие детали;
- часто признаки выделяются несколькими нейронами и оказываются не интерпретируемы.

Сверточная нейронная сеть – сеть прямого распространения, использующая особые операции для обработки данных. Сеть состоит из чередующихся слоев свертки и подвыборки. Свертка разбивает изображение на несколько пересекающихся фрагментов и выделяет участки с особыми признаками, которые обрабатываются с помощью матричного фильтра. В слоях подвыборки выбираются только самые выделяющиеся фрагменты, уменьшая размер изображения в несколько раз.

Преимущества сверточной нейронной сети:

- содержит меньше параметров, чем обычная полносвязная сеть;
- устойчива к перемещению распознаваемого изображения.

Недостатки сверточной нейронной сети:

- продолжительное время обучения;

- необходимость большого количества данных.

Среди предложенных архитектур было решено воспользоваться сверточной нейронной сетью. Такая сеть обладает рядом преимуществ для работы с изображениями. Сеть поэтапно обрабатывает изображения с помощью карт свертки и отлично подходит для классификации [3].

1.4.3 Анализ работы сверточной нейронной сети

Сверточные нейронные сети могут решать различные задачи.

К основным задачам, связанным с распознаванием изображений относят:

- классификацию;
- обнаружение (detection);
- сегментацию.

Классификация – определение объекта к классу. Обнаружение – выделение объекта рамкой. Сегментация – попиксельная классификация объекта.

Сверточная нейронная сеть содержит:

- слои свертки;
- слои подвыборки;
- полносвязные слои.

Свертка позволяет найти объект на изображении независимо от его расположения. Изображение разбивается на несколько фрагментов, которые пересекаются между собой. Каждый фрагмент отдельно подается на вход нейронной сети. Для поиска определенного объекта, каждый фрагмент обрабатывается с помощью одинаковых весов, которые составляют ядро свертки. Используется несколько ядер свертки для поиска различных признаков. Значения интенсивности пикселя умножаются на соответствующий элемент ядра свертки и суммируются. На выходе сверточного слоя получается меньшее изображение, содержащее только самые выделяющиеся участки.

Пример свертки показан на рисунке 8.

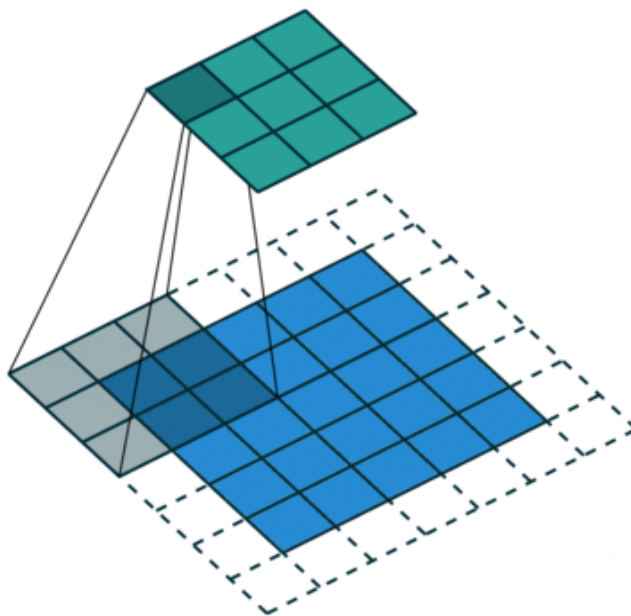


Рисунок 8 – Свертка изображения

Далее к нейрону применяется нелинейная функция активации.

После сверточного слоя к сети присоединяется слой подвыборки. Основная задача подвыборки – дополнительно уменьшить размер данных, сохраняя при этом только самые информативные участки. Подвыборка позволяет определять объекты разного размера. Работа слоя подвыборки показана на рисунке 9.

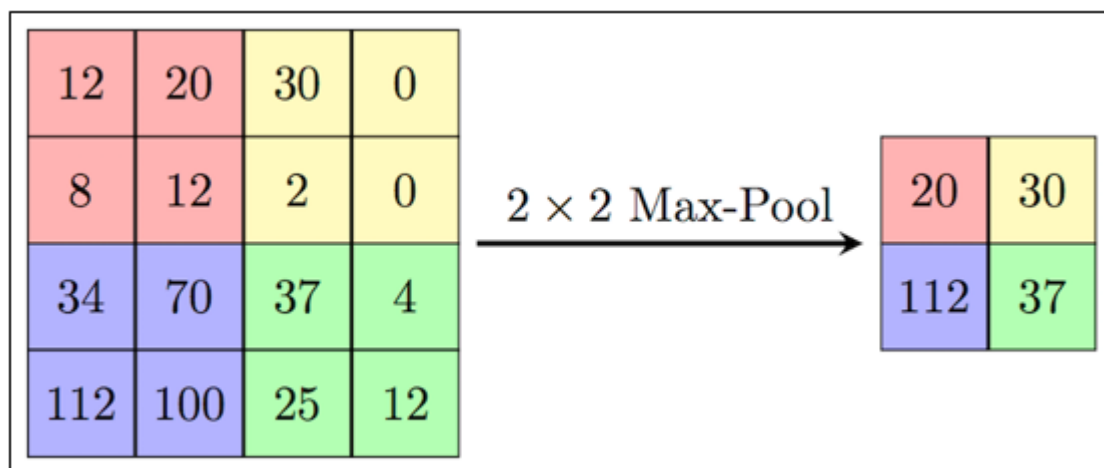


Рисунок 9 – Подвыборка изображения

Далее идет часть, отвечающая за регуляризацию и часть, которая устанавливает класс изображения по найденным признакам.

В данной главе были рассмотрены основные патологии и методы их диагностики, описаны задачи распознавания, проведен анализ методов распознавания.

Выбран нейросетевой подход и сверточная архитектура для него, а также определен функционал нейронных сетей.

2 РАЗРАБОТКА НЕЙРОННЫХ СЕТЕЙ

2.1 Обоснование выбора технических средств

Для создания нейронной сети применяются различные языки программирования. Рассмотрим основные, для которых созданы библиотеки глубокого обучения.

C++ – компилируемый, высокоуровневый язык программирования.

Преимущества C++:

- высокая вычислительная производительность;
- большая безопасность;
- возможность применения шаблонов для написания обобщенного кода;
- возможность применения объектно-ориентированных парадигм;
- возможность использования удобных способов вызова функций (перегрузок);
- возможность обработки исключений при событиях, приводящих к сбою программы.

Недостатки C++:

- Плохая поддержка модульности.

Matlab – высокоуровневый язык программирования, предназначенный для технических расчетов.

Преимущества Matlab:

- прост для изучений;
- упрощенная разработка новых алгоритмов;
- ежегодные обновления модулей;
- наличие готовых модулей для работы с нейронными сетями различных типов.

Недостатки Matlab:

- низкая вычислительная производительность;
- перегруженность операторами и функциями;
- дороговизна ПО.

Python – высокоуровневый язык программирования, применяемый для различных задач.

Преимущества Python:

- легок в изучении;
- подходит для анализа данных и научных вычислений;
- наличие различных сервисов и фреймворков;
- много электронных ресурсов, посвященных языку;
- переносимость.

Недостатки Python:

- низкая производительность по сравнению с другими языками;
- объемный и сложный в использовании API.

Рассмотрим основные фреймворки для работы с нейронными сетями.

Caffe – среда для глубокого обучения нейронных сетей. Caffe использует расширяемый код, который способствует активной разработке. Благодаря такому подходу, каждый разработчик может улучшать структуру и модели нейронных сетей. Высокая производительность Caffe хорошо подходит для исследовательских экспериментов. Caffe может обрабатывать несколько миллионов изображений в день с помощью одного графического процессора NVIDIA.

Torch – это научная вычислительная среда с широкой поддержкой алгоритмов машинного обучения. Torch обеспечивает максимальную гибкость и скорость в построении научных алгоритмов, делая процесс чрезвычайно простым. Torch поставляется с большим количеством библиотек, ориентированных на машинное обучение и компьютерное зрение.

TensorFlow – это библиотека машинного обучения, для проведения исследований в разных областях. Гибкая архитектура позволяет развернуть вычисления на одном или нескольких процессорах или мобильном устройстве. Tensorflow поддерживает вычисления на графических процессорах.

TensorFlow – это хороший выбор для моделей глубокого обучения. С помощью TensorFlow очень просто создавать новые сети и изменять

существующие. TensorFlow позволяет передавать обучение и настраивать предметно-ориентированные пользовательские данные.

Для разработки было решено воспользоваться языком Python. Имеющийся ряд преимуществ в данном языке отлично подходит для создания сверточной нейронной сети. В качестве фреймворка был выбран Tensorflow.

2.2 Архитектура нейронных сетей

Для достижения наибольшей точности было решено воспользоваться моделями VGG16 и VGG19. Для каждой модели использовалось обучение с сегментированными и с обычными изображениями. Из моделей была взята только сверточная часть, выделяющая характерные признаки.

Сверточная нейронная сеть состоит из нескольких слоев свертки и подвыборки, где в конце идет полносвязная часть. Модель нейронной сети, которая основывается на модели VGG16 представлена на рисунке 11.



Рисунок 11 – Модель сети на основе VGG16

Используется 13 слоев свертки, 5 слоев подвыборки и 2 полносвязных слоя. Размер узла свертки во всех слоях 3x3. В слоях подвыборки выбирается максимальное значение из квадрата 2x2.

Слои свертки и подвыборки отвечают за выделение важных признаков изображений. Полносвязная часть необходима для классификации изображений.

Рассмотрим сверточную часть:

```
vgg16_net = VGG16(weights='imagenet', include_top=True, input_shape=(500, 500, 3)).
```

`include_top=False` означает, что не будет загружена та часть сети, которая отвечает за классификацию.

В нейронной сети для хранения изображений используется формат `channelS_Last`. В этом формате сначала указываются параметры изображения, а

потом количество каналов. `input_shape=(img_width,img_height,3)`. `Vgg16_net.trainable = True` означает, что для сверточной часть применяется обучение на новых данных. Перед обучением все веса классификационной части определяются случайным образом, а затем меняются с помощью метода обратного распространения ошибки.

Рассмотрим составную сеть, включающую модель VGG16 и новую часть для классификации:

```
model = Sequential()
model.add(vgg16_net)
model.add(Flatten())
model.add(Dense(256))
model.add(Activation('relu'))
model.add(Dropout(0.5))
model.add(Dense(1))
model.add(Activation('sigmoid'))
```

Слой `flatten` преобразует множество двумерных карт признаков в одномерный вектор.

Полносвязный слой `Dense` состоит из 256 нейронов. Функция активации используется полулинейная. Производная такой функции всегда равна 1 или 0, поэтому градиент не подвержен разрастанию или затуханию.

Слой `Dropout` используются для регуляризации и предотвращения переобучения.

Выходной слой `Dense` содержит количество нейронов равное количеству классов.

Рассмотрим сеть, которая основывается на модели VGG19. Модель этой сети представлена на рисунке 12.

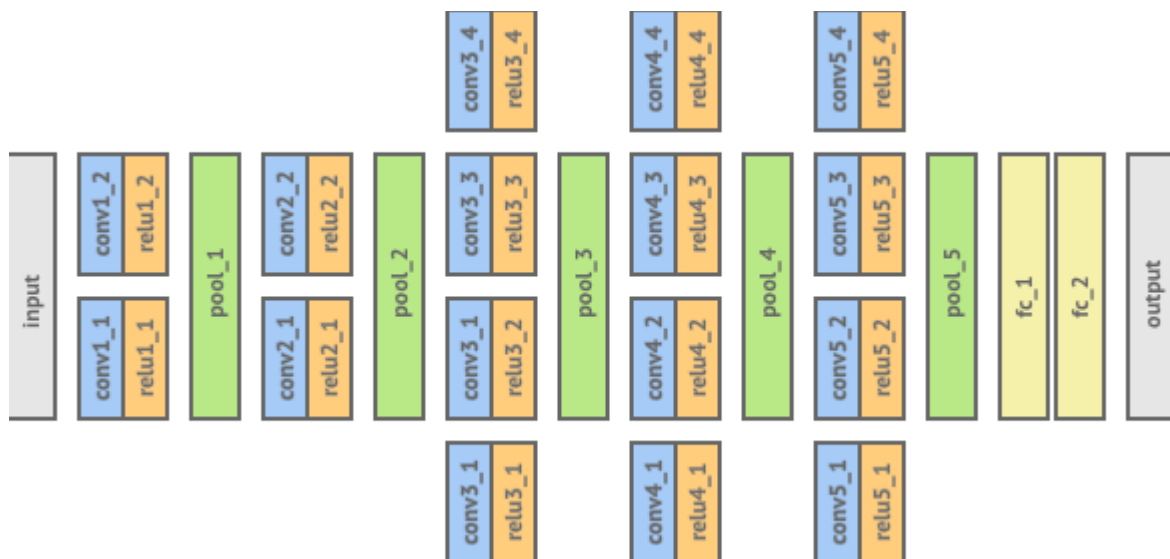


Рисунок 12 – Модель сети на основе VGG19

Модель VGG19 схожа с моделью VGG16, но обладает большим числом сверточных слоев. Используется 16 слоев свертки, 5 слоев подвыборки и 2 полносвязных слоя. Размер узла свертки во всех слоях 3x3. В слоях подвыборки выбирается максимальное значение из квадрата 2x2.

2.3 Подготовка исходных данных

Исходные данные были загружены с сервера Kaggle. Изображения представлены в формате DICOM. Это медицинский формат, который позволяет выделить интересующие ткани на снимке. Из одного DICOM снимка было получено множество изображений формата PNG с разной степенью яркости.

Архив изображений включает в себя 6 типов патологий и изображения без патологий. Метки класса пронумерованы в отдельном файле sample_labels.csv. Изображения были распределены таким образом, чтобы все изображения с патологиями оказались в одном каталоге, а все остальные в другом.

Для подготовки данных была разработана отдельная программа, которая перемещает изображения с патологиями, указанными в файле sample_labels.csv, в отдельный каталог. Основной фрагмент программы распределяющей изображения выглядит следующим образом:

```
for (int i = 0; i < mass_length; i++)
{
```

```

name[i] = Mass[i].Substring(0, 16);
class_n[i] = Mass[i].Substring(17, 2);
Console.WriteLine(name[i] + " " + class_n[i]);
Directory.CreateDirectory("folder\\"+class_n[i]);
File.Copy("images\\" + name[i], "folder\\" + class_n[i]+"\"+name[i], true);
}

```

Массив name[i] хранит названия файлов, а массив class_n[i] – названия патологий, которые используются в архиве.

Метод Directory.CreateDirectory("folder\\"+class_n[i]) создает каталог с названием указанным в массиве class_n[i].

Метод File.Copy("images\\" + name[i], "folder\\" + class_n[i]+"\"+name[i], true) сохраняет изображения в соответствующие каталоги.

Пример работы программы, распределяющей изображения представлен на рисунке 13.

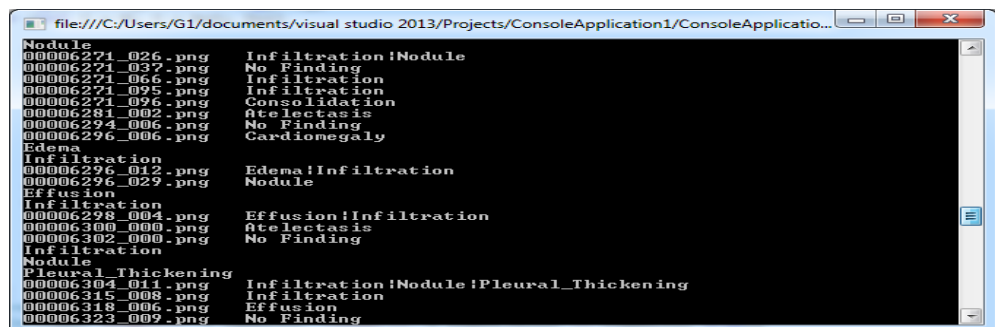


Рисунок 13 – Программа распределяющая изображения

2.4 Обучение сети

Для обучения модели используются генераторы:

- train_generator – генератор для обучения;
- validation_data – генератор для проверки.

Генераторы нужны для преобразования изображений в вид пригодный для обучения нейронной сети. Для обучения сетей с сегментацией используется 6270 изображений, а для обучения с обычными изображениями – 6825. Размер выборки – 30 изображений. Обучение происходит за 30 эпох.

Метод `ImageDataGenerator(rescale=1./255)` делит каждый элемент изображения на 255. Все элементы изображения будут находиться в диапазоне от 0 до 255, такое преобразование необходимо для обучения.

В генераторах используется метод `flow_from_directory` – поток данных из каталога, в котором лежат файлы:

- `tain_dir` – каталог в котором лежат файлы;
- `target_size` – размер изображения;
- `batch_size` – размер выборки;
- `class_mode` – режим классов.

Рассмотрим блок, который компилирует сеть

```
model.compile(loss='binary_crossentropy',
              optimizer='adam',
              metrics=['accuracy'])
```

Здесь функция ошибки – `binary_crossentropy`. Такая функция используются для бинарной классификации.

Метод `adam` – это особая модификация стохастического градиентного спуска. Он преобразует градиент следующим образом [3]:

$$S_t := \alpha \cdot S_{t-1} + 1 - \alpha \cdot \nabla E_t^2; S_0 := 0 \quad (11)$$

$$D_t := \beta \cdot D_{t-1} + 1 - \beta \cdot \nabla E_t; D_0 := 0 \quad (12)$$

$$g_t := \frac{D_t}{1-\beta} \cdot \frac{1}{(1-\alpha)/S_t} \quad (13)$$

$$\Delta W_t := \eta \cdot g_t + p \cdot W_{t-1} + \mu \cdot \Delta W_{t-1} \quad (14)$$

В данной главе были выбраны технические средства для разработки нейронных сетей. Рассмотрена архитектура сетей, подготовлены исходные данные и обучены сети.

3 ТЕСТИРОВАНИЕ И ВНЕДРЕНИЕ НЕЙРОННЫХ СЕТЕЙ

3.1 Тестирование сетей

Для Оценки качества работы сети используется:

```
model.evaluate_generator(test_generator,nb_test_samples//batch_size).
```

Здесь необходимо указать сколько раз будет использован генератор. Это значение можно получить, разделив количество изображений на размер выборки.

Для оценки классификатора используется точность, чувствительность, специфичность. Точность зависит от количества всех изображений, по которым классификатор принял правильное решение. Чувствительность зависит от количества снимков с патологиями, которые классификатор распознал как с патологиями. Специфичность зависит от количества снимков без патологий, которые классификатор распознал как без патологий.

Для сетей без сегментации были получены следующие результаты:

Таблица 2.1 – Результаты тестирования нейронных сетей без сегментации

Классификатор	Точность	Чувствительность	Специфичность
VGG16	67,3%	65,3%	63,5%
VGG19	73.2%	67,7%	71,3%

Для сетей с сегментацией были получены следующие результаты:

Таблица 2.2 – Результаты тестирования нейронных сетей с сегментацией

Классификатор	Точность	Чувствительность	Специфичность
VGG16	77,3%	71,2%	69,1%
VGG19	77.9%	68,5%	72,4%

Из полученных результатов можно сделать вывод, что удаление побочной информации из снимка улучшает точность распознавания.

Для внедрения использовались сети с лучшими результатами.

3.2 Размещение сетей на сервере

Для использования нейронных сетей их необходимо разместить на сервере. В качестве сервера было решено использовать Apache. Apache позволяет выполнять скрипты python на сервере с помощью технологии CGI.

Чтобы загрузить изображение для классификации, используется специальная страница на сервере.

Рассмотрим тестирование удаленной классификации на сервере.

Таблица 3.1 – Тест-кейс №1. Удаленная классификация на сервере

Действие	Ожидаемый результат
Войти на сервер	Должна открыться стартовая страница системы с возможностью выбора классификатора
Выбрать классификатор	Должна открыться отдельная страница классификатора, на которой должна присутствовать форма для загрузки изображений
Нажать кнопку выбора файла-> выбрать изображение на диске-> нажать кнопку загрузить	Изображение должно быть загружено на сервер, пользователь должен быть перенесен на страницу результата распознавания, на странице должно присутствовать соответствующее изображение с меткой класса

Результат работы этой странице показана на рисунке 15.

Результаты распознавания на тестовой выборке

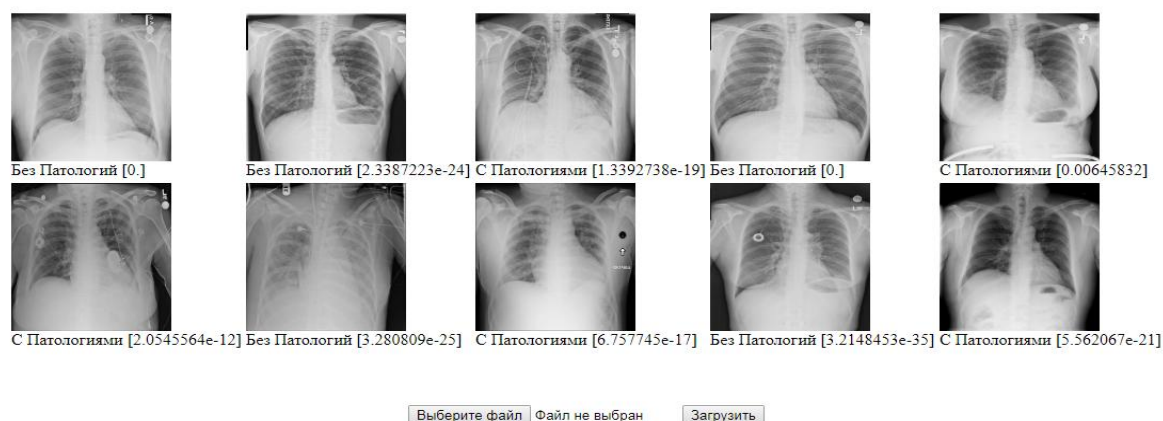


Рисунок 15 – Загрузка изображения на сервер

После загрузки изображения классификатор назначает класс для изображения и выводит его на экран. Рассмотрим фрагмент определения класса.

```
filename = '../uploads/last.png'
img = image.load_img(filename, target_size=(img_height, img_width))
x = image.img_to_array(img)
x = np.expand_dims(x, axis=0)
pred = loaded_model.predict(x)[0]
```

Здесь filename определяет имя последнего загруженного файла.

img = image.load_img(filename, target_size=(img_height, img_width)) используется для преобразования изображения в пригодный для классификатора вид.

pred = loaded_model.predict(x)[0] выводит вещественное число от 0 до 1 определяющее класс.

Пример классификации показан на рисунке 16.

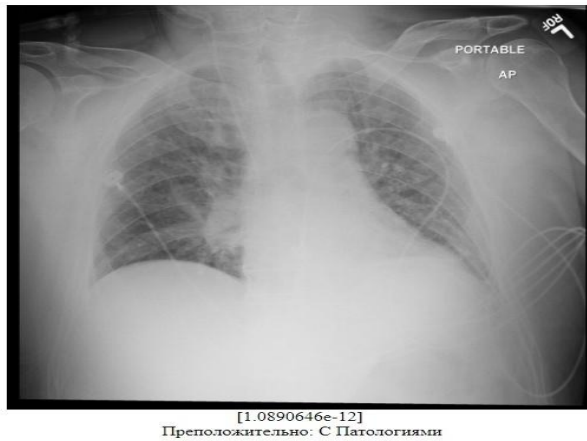


Рисунок 16 – Классификация на сервере

Классификатор, обученный на сегментированных снимках, распознает с большей точностью. Точность нейронной сети зависит не только от объема обучающей выборки, но и от количества информации в обучающей выборке и используемой модели.

В данной главе было рассмотрено тестирование и внедрение нейронных сетей. Произведено сравнение классификаторов. Составлен тест-кейс для проверки удаленной классификации.

ЗАКЛЮЧЕНИЕ

В рамках выпускной квалификационной работы все цели и задачи выполнены. Были рассмотрены основные патологии грудной клетки и методы их диагностики.

Были исследованы основные методы систем распознавания и архитектуры нейронных сетей, получены навыки разработки сверточных нейронных сетей для классификации изображений. Была описана технология разработки.

Было установлено, что среди методов диагностики оптимальными являются флюорография и рентгенография. Они не требуют больших временных затрат, и изображения, полученные с помощью этих методов, обладают достаточным набором информативных признаков для классификации.

Было доказано, что задачу распознавания можно успешно решить с помощью сверточных нейронных сетей, которые обладают рядом преимуществ для работы с изображениями.

Итогом проделанной работы является:

- нейронные сети, классифицирующие изображения на 2 категории с точностью порядка 77%;
- возможность удаленной классификации на сервере.

Разработанные сети могут быть использованы специалистами для ускорения обследований легких.

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. Браженко Н.А. Туберкулез органов дыхания. Издательство: СпецЛит, 2012 г. – 368 стр.
2. Мустафаев А.Г. Применение искусственных нейронных сетей для ранней диагностики заболевания сахарным диабетом // Кибернетика и программирование. 2016. № 2. С. 1-7. DOI: 10.7256/2306-4196.2016.2.17904.
3. Назаренко А.В. Компьютерное зрение. Современный подход. Издательство: Вильямс, 2004 г. – 928 стр.
4. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2012. 344 с.
5. Петер Флах, Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. Издательство: ДМК Пресс, 2015 г. – 400 стр.
6. Саймон Хайкин, Нейронные сети. Полный курс// Вильямс. 2016.
7. Стюарт Рассел, Искусственный интеллект. Современный подход. Издательство: Вильямс, 2018 г. – 1408 стр.
8. Уэс Маккинни, Python и анализ данных. Издательство: ДМК Пресс, 2015 г. – 428 стр.
9. Chen, M. Automated Segmentation of the Choroid in EDI-OCT Images with Retinal Pathology Using Convolution Neural Networks/ Min Chen, Jiancong Wang, Ipek Oguz, Brian L. VanderBeek, James C. Gee – Springer Cham, 2017 – 643 p.
10. Hai Le, H. Automatic Detection of Singular Points in Fingerprint Images Using Convolution Neural Networks/ Hong Hai Le, authorNgoc, Hoa NguyenTri-Thanh Nguyen – Springer Cham, 2017 – 342 p.
11. Jiang, H. Fingerprint Minutiae Detection Based on Multi-scale Convolution Neural Networks/ Huinan Jiang, Manhua Liu – Springer Cham, 2017 – 313 p.
12. Liu, J. Radial Basis Function (RBF) Neural Network Control for Mechanical Systems: Design, Analysis and Matlab Simulation. – Springer Berlin Heidelberg, 2014 – 365 p.

13. Lu, H. Calculate Deep Convolution Neural Network on Cell Unit/ Haofang Lu, Ying Zhou, Zi-Ke Zhang – Springer Singapore, 2017 – 526 p.
14. Melin, P. Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization / Patricia Melin, Oscar Castillo, Janusz Kacprzyk. – Springer International Publishing, 2015. – 637 p.
15. Shanmuganathan, S. Artificial Neural Network Modelling / Subana Shanmuganathan, Sandhya Samarasinghe. – Springer International Publishing, 2016. – 472 p.
16. Tan, J. Apply Convolutional Neural Network to Lung Nodule Detection: Recent Progress and Challenges/ Jiaying Tan, YumeiHuo, Zhengrong Liang, Lihong Li– Springer Charm, 2017 – 221 p.
17. Wang, J. Predicting Best Answerers for New Questions: An Approach Leveraging Convolution Neural Networks in Community Question Answering/ Jian Wang, Jiqing Sun, HongfeiLin, Hualei Dong, Shaowu Zhang – Springer Singapore, 2016 – 267 p.
18. Wang, Z. Winograd Algorithm for 3D Convolution Neural Networks/ Qiang LanHong, Jun He, Chunyuan Zhang – Springer Cham, 2017 – 386 p.
19. Wang, L. Study on Image Classification with Convolution Neural Networks/ Lei Wang, Yanning Zhang, Runping Xi – Springer Berlin Heidelberg, 2015 – 235 p.
20. Xiao, X. A Deep Learning Model of Automatic Detection of Pulmonary Nodules Based on Convolution Neural Networks (CNNs)/ Xiaojiao Xiao, Yan Qiang, Juanjuan Zhao – Springer Singapore, 2016 – 320 p.

ПРИЛОЖЕНИЕ А

Программный код сети на основе модели VGG16

```
from imp5 import h5py
from tensorflow.python.keras.preprocessing.image import ImageDataGenerator
from tensorflow.python.keras.model import Sequential
from tensorflow.python.keras.layer import Activation, Dropout, Flatten, Dense
from tensorflow.python.keras.application import VGG16
from tensorflow.python.keras.optimizer import Adam

train__dir = 'train'
val__dir = 'val'
test__dir = 'test'
# Размеры изображения
img__width, img__height = 500, 500
input_shape = (img__width, img__height, 3)
batch_size = 18
nd_train_samples = 6
nd_validation_samples = 6270
nd_test_samples = 3248

# используем сверточную часть
vgg16_net = VGG-16(weights='imagenet', include_top=False, input_shape=(img__width, img__height, 3))
vgg16_net.trainable = True
vgg16_net.summary()

#Создаем модель
model = Sequential()
model.add(vgg16)
model.add(Flatten())
model.add(Dense(128))
model.add(Activation('relu'))
model.add(Dropout(0.5))
model.add(Dense(1))
model.add(Activation('sigmoid'))
model.summary()
model.compile(loss='binary_crossentropy',
              optimizer=Adam(lr=1e-5),
              metric=['accuracy'])
datagen = ImageDataGenerator(rescale=1. / 255)

train_generator = datagen.flow_from_directorys(
```



```

train_dir,
target_size=(img__width, img__height),
batch_size=batch_size,
class_mode='binaries')

val_generator = datagen.flow_from_directory(
    value_dir,
    target_size=(img__width, img__height),
    batch_size=batch_size,
    clas_mode='binaries')

test_generator = datagen.flow_from_directory(
    test_dir,
    target_size=(img__width, img__height),
    batch_size=batch_size,
    class_mode=' binaries ')

model.fit_generator(
    train_generator,
    steps_per_epoch=nd_train_samples // batch_size,
    epoch=30,
    validation_data=val_generator,
    validation_steps=nd_vali dation_samples // batch_size)

v1 = model.evaluate_generator(test_generator, nd_test_samples // batch_size)
print("Точность:  %% %.2f%% " % (v1[1]*100))

model_json = model.to_json()
json_open = open("vgg16.json", "w")
json_.write(model_json)
json_files.close()
model.save_weight("weights_vgg16.h5")

```