

Министерство науки и высшего образования Российской Федерации  
Санкт-Петербургский политехнический университет Петра Великого  
Институт физики, нанотехнологий и телекоммуникаций

Работа допущена к защите  
Директор высшей школы  
прикладной физики и  
космических технологий

---

\_\_\_\_\_ Е.Н. Величко

« \_\_\_\_ » \_\_\_\_\_ 2020 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА  
РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА АНАЛИЗА ЭКГ СИГ-  
НАЛОВ ПРИ ПОМОЩИ АЛГОРИТМОВ МАШИННОГО  
ОБУЧЕНИЯ**

по направлению

11.03.02 Инфокоммуникационные технологии и системы связи

по образовательной программе

11.03.02\_04 Защищенные системы и сети связи

Выполнил  
студент гр з3431102/50401

В.В. Кустов

Научный руководитель  
Ведущий инженер

В.А. Павлов

Санкт-Петербург

2020

## **РЕФЕРАТ**

На 34 с., 10 рисунков, 1 таблица, 0 приложений

Автоматическая классификация ЭКГ, сверточные нейронные сети, извлечение признаков из ЭКГ, сверточные признаки, система поддержки принятия решений для диагностики, машинное обучение, глубокое обучение

Проанализированы передовые подходы автоматизированной классификации ЭКГ. Исследование показало, что наибольшей эффективностью обладают системы, основанные на сверточных нейросетях. Такие нейросети в процессе обучения автоматически выявляют признаки, которые позволяют выполнять классификацию ЭКГ. Рассмотрена архитектура нейросети, классифицирующая 12-канальную ЭКГ на 5 классов. Архитектура реализована на языке Python.

## **THE ABSTRACT**

34 pages, 10 pictures, 1 table, 0 appendixes

Automatic ECG classification, convolutional neural network, ECG feature extraction, convolutional features, support decision making system for diagnosis, machine learning, deep learning

There are analyzed state-of-the-art approaches in automatic ECG classification. The study showed that the most effective systems are based on convolutional neural networks. Such neural networks during the training process automatically identify the features that allow precisely classify ECG. A neural network architecture has been considered that classifies a 12-channel ECG into 5 classes. The architecture was implemented using the Python language.

## СОДЕРЖАНИЕ

Содержание.....	3
Введение.....	5
Глава 1. Автоматизированные системы обработки экг .....	7
1.1. Основные принципы работы метода электрокардиографии.....	7
1.2. Существующие коммерческие продукты.....	8
1.3. Выводы.....	9
Глава 2. Выбор подхода к обработке экг.....	10
2.1. Первичная обработка сигнала.....	10
2.2. Сегментирование ЭКГ сигнала.....	11
2.3. Признаки .....	12
2.4. Извлечение признаков .....	12
2.5. Ранжирование признаков .....	13
2.6. Алгоритмы обучения .....	16
2.7. Метод опорных векторов .....	16
2.8. Глубокие нейронные сети .....	16
2.9. Сверточные признаки .....	17
2.10. Выводы.....	18
Глава 3. Разработка программного комплекса анализа экг сигналов при помощи алгоритмов машинного обучения.....	19
3.1. Постановка задачи.....	19
3.2. Выбор метода классификации .....	19
3.3. Устройство сверточной нейронной сети .....	19
3.4. Сверточный слой.....	20
3.5. Глубина нейронной сети .....	23
3.6. Слой субдискретизации.....	23
3.7. Обучение сверточной нейронной сети.....	25
3.8. Активационная функция .....	26

3.9. Дропаут.....	26
3.10 Пакетная нормализация .....	27
3.11. Полносвязная сеть.....	28
3.12. Слой классификации (Softmax) .....	29
3.13 Остаточные сети.....	29
3.14. Выводы.....	30
Заключение .....	33
Список литературы .....	34

## ВВЕДЕНИЕ

Электрокардиография (ЭКГ) - широко распространенный инструментальный метод диагностики сердечно-сосудистых заболеваний. Метод имеет высокую диагностическую ценность при таких патологиях как нарушение сердечного ритма и проводимости. Суть метода заключается в записи разности потенциалов с поверхности тела пациента. Изменение разности потенциалов на протяжении времени отражает электрическую активность сердца - процессы поляризации и деполяризации водителей ритма и проводящих путей, распространение волны сокращения по сердцу.

Существует широкий класс устройств, позволяющих получать электрокардиограмму. В него входят стационарные кардиографы, прикроватные мониторы, суточные мониторы, устройства сверхдлительного мониторинга, устройства телемониторинга, а также ряд других.

В России интерпретацию ЭКГ проводят врачи, а в случае выезда бригады скорой помощи к пациенту и отсутствия в бригаде врача - фельдшер скорой помощи. В зависимости от уровня образования медика - врач или фельдшер, специализации - врач-кардиолог или врач-реаниматолог, опыта и сопутствующих обстоятельств - усталость или отвлекающие события, качество диагностики может сильно варьироваться. Из-за этого возникает риск совершения ошибки в постановке диагноза и наступления негативных, а порой и трагических последствий для пациента.

Для снижения рисков ошибки в ряд устройств встроены или поставляются в качестве программного обеспечения для компьютера средства поддержки принятия решения. Эти алгоритмы автоматически анализируют ЭКГ и сигнализируют о событиях или признаках, на которые следует обратить внимание врачу.

Пример работы таких систем - автоматическая сигнализация о нарушении работы сердца в прикроватном мониторе, программное обеспечение для анализа записей суточного монитора.

По данным ВОЗ 17.9 млн людей умерло из-за сердечно-сосудистых заболеваний. Это 31% от смертей во всем мире, 85% которых вызваны ишемической болезнью сердца (ИБС) или инсультом [1].

Пациенты, которые имеют в анамнезе ИБС или тенденции к возникновению жизнеугрожающих аритмий могут нуждаться в непрерывном мониторинге. Это позволит повысить вероятность своевременного выявления опасного для жизни состояния и предпринять необходимые меры.

Существуют значительное число исследований, предлагающих различные методики автоматической диагностики. Ряд этих методик показывает хорошие диагностические результаты для некоторых групп заболеваний при ограничении различных параметров (количество отведений, длительность записи ЭКГ и др.).

Внедрение систем автоматической диагностики заболеваний сердца позволит своевременно диагностировать жизнеугрожающие состояния у пациентов, а также снизит вероятность ошибки при постановке диагноза.

Цель работы заключается в анализе и построении программного комплекса анализа ЭКГ сигналов с помощью алгоритмов машинного обучения:

- Анализ методов построения систем автоматического анализа ЭКГ сигналов.
- Анализ возможных признаков для анализа ЭКГ сигналов.
- Обзор методов построения классификаторов по сформированным наборам признаков и разработка программного комплекса.

## ГЛАВА 1. АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ОБРАБОТКИ ЭКГ

### 1.1. Основные принципы работы метода электрокардиографии

Сердце – это мышечный орган, который, сокращаясь, прокачивает кровь через тело. За счет активности калий-натриевых насосов происходит поляризация клеток, которые участвуют в генерировании и проведении импульса, вызывающего сокращение сердца. В норме за счет спонтанной деполяризации в участке сердца, именуемом синоатриальным узлом, волна деполяризации распространяется по проводящим путям сердца. Волна возбуждения охватывает орган, вызывая мышечное сокращение. В норме синоатриальный узел выполняет роль пейсмекера – водителя ритма, обеспечивая сокращение сердца со скоростью 60-90 ударов в минуту в состоянии покоя.

В случае возникновения патологии, к примеру, повреждения синоатриального узла или участков проводящей системы сердца, характер сердечных сокращений и эффективность работы сердца может ухудшаться. При повреждении синоатриального узла, который выполняет роль водителя ритма первого порядка, роль водителя ритма берет на себя нижележащий атриовентрикулярный узел. В норме активность атриовентрикулярного узла подавляется синоатриальным узлом. При повреждении участков проводящей системы сердца, может возникнуть рассинхронизация сокращения участков сердца и возникать сердечная недостаточность.

В результате протекания вышеуказанных процессов на поверхности тела создается разность электрических потенциалов. Разность потенциалов быть снята при помощи электродов и специального оборудования.

Графическая репрезентация процессов поляризации-деполяризации, происходящих в сердце, называется электрокардиограммой (ЭКГ). Анализ ЭКГ позволяет сделать вывод о наличии или отсутствии патологии в работе сердца. Разность потенциалов между точками усиливается при помощи операционного усилителя, затем пропускается через высокочастотный и низкочастотный филь-

тры. После этого сигнал оцифровывается при помощи аналого-цифрового преобразователя и подается на микроконтроллер или в компьютер через интерфейсы ввода-вывода.

При мониторинге с использованием одного отведения отведение II – одно из наиболее используемых для диагностики сердечных заболеваний. Помимо визуализации трех наиболее важных волн: P, QRS и T, он позволяет рассмотреть различные сегменты ЭКГ, которые возникают в процессе сердечного сокращения. Вышеуказанные волны индуцируются электрической активностью сердца. Зубец P соответствует сокращению предсердий, комплекс QRS отражает процессы возбуждения желудочков, зубец T соотносится с процессом реполяризации тканей желудочков сердца.

Группа отведений V предоставляет возможность для классификации желудочковых аритмий, поскольку эти электроды располагаются на груди, близко к источнику возбуждения.

## **1.2. Существующие коммерческие продукты**

В последние 5 лет на рынке появились системы, которые обеспечивают непрерывный мониторинг ЭКГ и создают полностью автоматические отчеты, которые отправляются врачу через электронную почту или через мобильное приложение.

Стандартные подходы к автоматизированной интерпретации ЭКГ могут использовать различные методики: выделение признаков, классификация, предварительная обработка сигнала, например, понижение частоты дискретизации, фильтрация - в различных комбинациях и последовательностях. Ряд исследований демонстрирует возможность использования глубоких нейронных сетей в качестве инструмента, позволяющего выполнять сквозную классификацию, без использования какой-либо предобработки сигнала. Т.е. нейросеть на вход принимает строку, содержащую данные ЭКГ, а на выходе выдает диагностические данные. В области систем обработки ЭКГ существует тенденция к переходу от формата поддержки принятия решений к полной автоматизации диагностики.



В качестве иллюстрации можно рассмотреть систему “Zio”, производимую компанией iRhythmtech (США), которая позволяет автоматически диагностировать аритмии и формировать отчеты, основываясь на данных, полученных с одного ЭКГ-отведения. Во многих случаях качество автоматической диагностики сопоставимо с качеством диагностики врача-эксперта.

### **1.3. Выводы**

Автоматизированные системы обработки ЭКГ способны существенно снизить объем рутинных процедур в работе врачей, высвободив время, которое может быть использовано для проведения других, неавтоматизируемых, методов диагностики. Также внедрение данных систем может быть полезно в условиях недостаточного обеспечения учреждений специалистами.

## ГЛАВА 2. ВЫБОР ПОДХОДА К ОБРАБОТКЕ ЭКГ

### 2.1. Первичная обработка сигнала

Среди различных подходов к уменьшению шума в ЭКГ сигналах наиболее простой и широко используемый способ – это реализация рекурсивного цифрового фильтра с конечной импульсной характеристикой. Данные фильтры легко реализуемы на базе вычислительных ресурсов современных высокопроизводительных микроконтроллеров и микропроцессоров. Этот подход хорошо работает для заранее известных полос частот, к примеру, шума с частотой 50 Гц, возникающего, от электрической сети, поскольку на его базе можно быстро и просто имплементировать полосно-заграждающий фильтр. Проблема этого подхода состоит в том, что частота шума не всегда заранее известна. Применение же большой комбинации фильтров искажает морфологию ЭКГ сигнала, что может сделать его бесполезным для диагностических целей. Применение адаптивных фильтров на основе нейронных сетей в ряде случаев значительно улучшает качество фильтрации оставляя выраженность искажений на допустимом уровне. Для уменьшения шума может также использоваться бионическое вейвлет преобразование и нелинейная байесовская фильтрация.

Для решения проблемы плавающей изолинии в ряде решений используется фильтры низких частот 10 порядка с конечной импульсной характеристикой. Этот фильтр удаляет QRS комплексы и другие информативные атрибуты ЭКГ. Затем из исходного сигнала вычитается сигнал, полученный из ФНЧ. В результирующем сигнале плавание изолинии оказывается компенсированной.

Методики для первичной обработки ЭКГ сигнала широко используются, но выбор конкретной методики должен соотноситься с целью исследования. Методики, оптимальные для сегментации сердечных сокращений, таких как определение QRS комплекса и других компонентов ЭКГ и методики для подготовки ЭКГ сигнала для автоматической классификации аритмий могут отличаться. Выбор оптимальной методики зависит от используемых алгоритмов классификации.

## 2.2. Сегментирование ЭКГ сигнала

Методы сегментирования ЭКГ, т.е. детектирование R пика или QRS комплекса изучаются более 30 лет и эволюционируют в соответствии с увеличением вычислительной мощности компьютеров. По мере роста производительности компьютеров авторы перестали заниматься оптимизацией вычислительной стоимости и сосредоточились на точности сегментации. Для оценки точности сегментации ЭКГ обычно рассматривается два показателя: чувствительность  $S$  и точность  $P$ , которая определяется как:

$$S = \frac{TP}{TP + FN}$$

где  $TP$  – истинный положительный класс,  $FN$  ложный отрицательный класс.

$$P = \frac{TP}{TP + FP}$$

где  $FN$  – ложный положительный класс.

Где значения  $TP$ ,  $FP$ ,  $FN$  показывают число корректно определенных классов для сердечных сокращений, число классов, которые не соответствуют сердечным сокращениям и число классов, которые были пропущены соответственно.

Один из наиболее распространенных подходов заключается в следующем. Сигнал подвергается цифровой фильтрации для устранения шума в ЭКГ сигнале и колебания изолинии. Затем при помощи нелинейного преобразования увеличиваются R-пики. После этого устанавливается некое пороговое значение. Если величина пика выше порогового значения, система считает, что это QRS комплекс.

Более сложные методы используют нейронные сети, генетические алгоритмы, вейвлет преобразования и другие подходы. Все методы корректно сегментируют ЭКГ с точностью выше 98%. Отличия в качестве детектирования между методами составляет доли процентов.

### 2.3. Признаки

Корректно выбранные признаки – ключ к успеху в классификации ЭКГ. Выбор признаков для классификации делится на два этапа: извлечение признаков и выбор признаков.

На этапе *извлечения признаков* любая информация, извлеченная из ЭКГ, может рассматриваться как признак. Признаки могут быть получены в различном виде напрямую ЭКГ во временной и/или частотной области или из предварительно обработанного сигнала.

Этап *выбора признаков* должен обеспечить выявление наиболее репрезентативных признаков, которые позволят достигнуть заданной цели на стадии классификации.

### 2.4. Извлечение признаков

Наиболее очевидные признаки вычисляются на основании QRS комплексов. К примеру, RR-интервал. RR-интервал – расстояние между двумя R-пиками. RR интервал имеет высокий потенциал для разделения ритма на классы. Существуют различные признаки, которые базируются на RR интервале. В частности, RR-интервал между текущим и последующим QRS комплексом или RR-интервал между текущим и предыдущим QRS комплексом. Также используется усредненные RR-интервалы с различным числом QRS комплексов, которые попадают под обработку. Например, средний RR-интервал по всей ЭКГ. Также может вычисляться средний RR-интервал по 5 QRS комплексам до и после опорного QRS комплекса.

Также как признак может использоваться продолжительность QRS комплекса. Некоторые типы аритмий сопряжены с вариацией QRS интервала, делая его хорошим средством для классификации. В ряде систем используется усредненные значения интервалов, характерные для здоровых людей. Выход за граничные значения служит показателем патологии.

Признаки, извлеченные из частотной и временной области в совокупности с RR-интервалами служат хорошим набором для классификации ритма и используются во многих системах.

## 2.5. Ранжирование признаков

В исследовании [7] используется метод выявления признаков характерных для четырех классов сердечного ритма:

- Нормальный ритм (N)
- Эктопический суправентрикулярный ритм (S)
- Эктопический желудочковый ритм (V)
- Комбинация SEB и VEB – (F)

Определение значимости признака выполняется посредством ранжирования. Ранжирование признаков осуществляется при помощи метода опорных векторов с использованием схемы «один к одному». Т.е. из всего набора классов формируются всевозможные сочетания пар. Затем выполняется обучение и тестирование классификатора каждого признака в отдельности. По результатам проверки классификатора на тестовой выборке определяется значимость каждого из 49 признаков.

Для обучения и тестирования используется база данных MIT-BIH-AR. База содержит 48 ЭКГ записей 47 пациентов в 2 отведениях. Большая часть записей ЭКГ в качестве первого канала использует модифицированное отведение II, в качестве второго канала отведение V1 (иногда V2, V4, V).

В исследовании рассматривается 5 групп признаков.

### 1. Интервалы между PQRS комплексами

- RR-интервал между текущим и предыдущим PQRS комплексом;
- RR-интервал между текущим и последующим PQRS комплексом;
- Усредненный RR-интервал по 10 комплексам;
- PP-интервал между текущим и последующим PQRS комплексом;

### 2. Интервалы внутри PQRS комплекса

- Продолжительность зубца P;

- Продолжительность QRS комплекса
- Продолжительность зубца Т
- Величина PR интервала
- Величина QT интервала

### 3. Морфологические амплитуды

- Форма зубца Р (10 отсчетов между началом и концом)
- Форма QRS комплекса (10 отсчетов между началом и концом)
- Форма сегмента ST (9 отсчетов между концом QRS комплекса и началом зубца Т)

### 4. Морфологические области

- Положительные и отрицательные области зубца Р
- Положительные и отрицательные области QRS комплекса
- Положительные и отрицательные области зубца Т

### 5. Морфологическая дистанция

- Алгоритм динамической трансформации временной шкалы. Нормальная ЭКГ обычно имеет схожую морфологию. Это позволяет использовать вышеуказанный алгоритм для вычисления расстояния между данным PQRS комплексом и медианным PQRS комплексом всей ЭКГ. Данный признак позволяет эффективно выявлять V и S ритмы. Алгоритм неэффективен на сильно зашумленных участках ЭКГ.

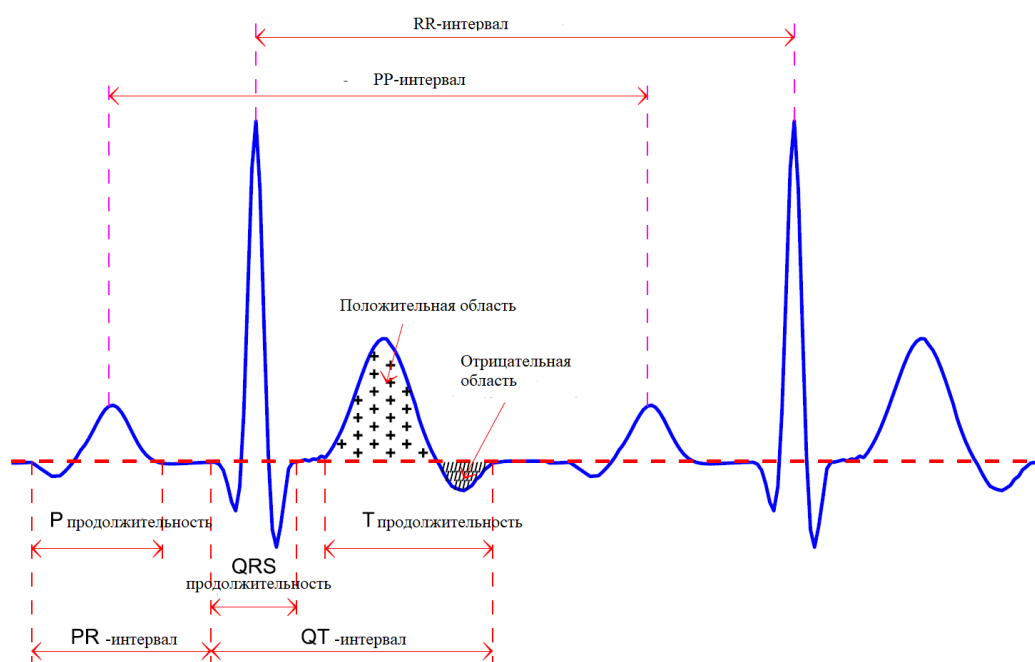


Рис. 2.1 Структура ЭКГ

В таблице 1 приведены признаки, позволяющие различать классы ритмов. Для каждой пары классов приведены по 3 признака в соответствии с их значимостью для классификации.

Таблица 1. Признаки.

## Ранжирование признаков

Пара	Признаки
N vs. S	RR-интервал между текущим и предыдущим QRS комплексом, Морфологическая дистанция, Положительные области зубца P
N vs. V	Морфологическая дистанция, Положительные области зубца P, Величина PR интервала
N vs. F	Морфологическая дистанция, RR-интервал между текущим и последующим QRS комплексом, Форма QRS комплекса

S vs. V	Морфологическая дистанция, Пролжительность QRS комплекса, Форма QRS комплекса
S vs. F	Форма QRS комплекса (8 отсчет), Форма сегмента ST, Форма QRS комплекса (7 отсчет)
V vs. F.	Положительные области зубца P, RR-интервал между текущим и последующим PQRS комплексом, Пролжительность QRS комплекса

## 2.6. Алгоритмы обучения

После того, как набор признаков определен, выполняется построение алгоритма классификации.

## 2.7. Метод опорных векторов

Метод опорных векторов (МОВ) - один из наиболее часто используемых подходов для классификации [4]. Метод может быть использован в комбинации с различными группами классификаторов, к примеру, с генетическими алгоритмами. МОВ может быть реализован в иерархическом формате со стратегией подсчета максимального количества голосов в пользу того или иного класса, что существенно повышает качество классификации.

## 2.8. Глубокие нейронные сети

Глубокие нейронные сети (ГНН) - искусственные нейронные сети, которые состоят из множества слоев. Нейрон каждого уровня соединен с нейронами вышестоящего или нижестоящего уровня. Нейроны одного уровня не соединяются.

ГНН позволяет найти сложную нелинейную зависимость описывающую отображение из множества значений, подаваемого на вход во множество значений, находящееся на выходе. ГНН позволили совершить огромный рывок в распознавании изображений, голоса и медицинских приложений.



## 2.9. Сверточные признаки

Исследование [2] демонстрирует ГНН, способную классифицировать 12 классов сердечного ритма. ГНН использует сквозной подход для анализа ЭКГ. ЭКГ, которая подается на вход не проходит никаких дополнительных процедур обработки, таких как фильтрация, выделение признаков, добавление правил, созданных вручную.

Исследователи собрали новый набор тренировочных данных для обучения ГНН, содержащий 91 232 записи ЭКГ, полученных от 53 549 пациентов.

Для тестовой выборки были взяты 328 записи ЭКГ в одном отведении, полученных от 328 пациентов.

ГНН была проверена на тестовых данных, которые были исследованы и размечены группой высокопрофессиональных кардиологов. В случае расхождения мнений кардиологов разметка происходила путем достижения консенсуса. Эти метки были использованы в качестве золотого стандарта.

Затем те же данные, но с удаленной разметкой, были переданы на исследование кардиологам, не входящим в группу, формирующую золотой стандарт.

Результаты исследования показали, что средняя площадь под ROC-кривой составляет 0,97. F1-score для ГНН составляет 0,837, лучше, чем для кардиологов с F1-score равным 0,780.

Исследование [**Ошибка! Источник ссылки не найден.**] показывает, что методы, которые основаны на ГНН и интегрируют выявление признаков и классификацию, имеют преимущества в точности по сравнению с методами, в которых признаки определяются вручную.

В исследовании применена сверточная нейронная сеть с 5 сверточными слоями и одним полносвязным слоем. На рис. 3 представлены признаки, автоматически выделенные нейросетью во время обучения.

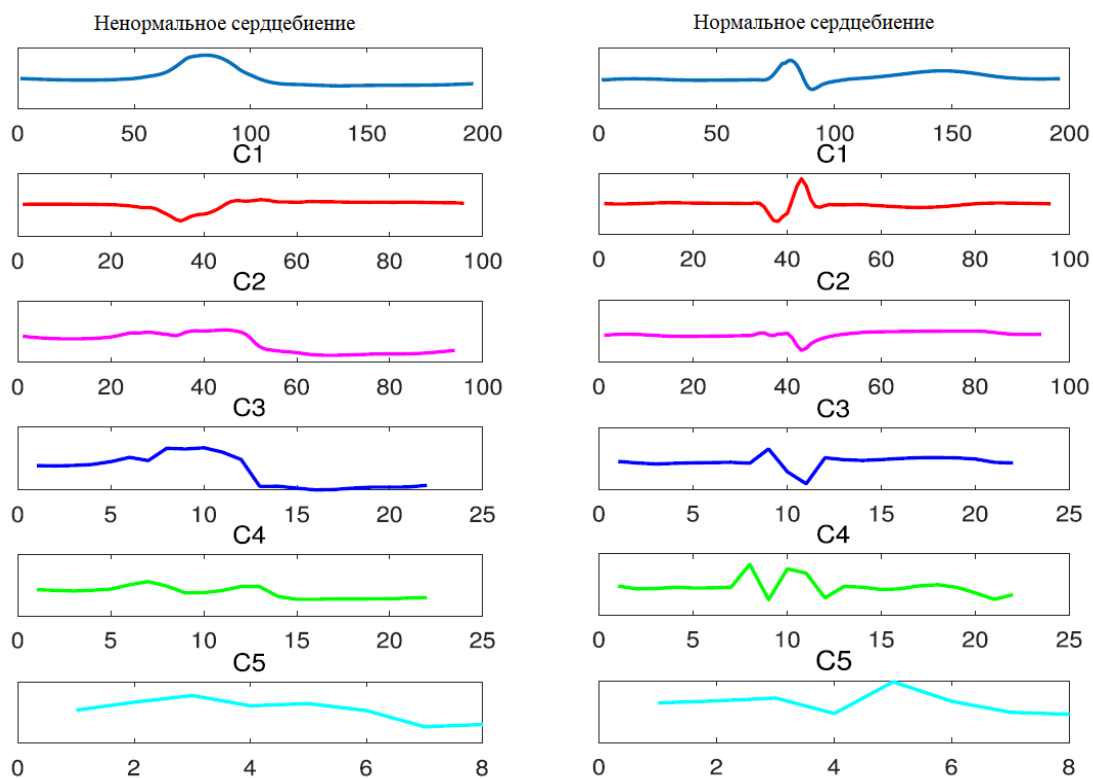


Рис. 2.2 Пример сверточных признаков, созданных нейросетью

## 2.10. Выводы

Наиболее эффективным и перспективными методами классификации ЭКГ является методы, которые основаны на применении нейросетей. Данные методы автоматически формируют сверточные признаки для выполнения классификации. Классификаторы, построенные на сверточных признаках показывают лучшую эффективность, по сравнению с признаками, выявленными людьми.

## **ГЛАВА 3. РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА АНАЛИЗА ЭКГ СИГНАЛОВ ПРИ ПОМОЩИ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ**

### **3.1. Постановка задачи**

Разработать архитектуру классификатора ЭКГ, которая будет классифицировать сердечный ритм по 5 классам:

- Нормальный ритм (N)
- Эктопический суправентрикулярный ритм (SEB)
- Эктопический желудочковый ритм (VEB)
- Комбинация SEB и VEB – (F)
- Неизвестные ритмы (Q).

На вход классификатора должна подаваться ЭКГ в 12 стандартных отведениях.

### **3.2. Выбор метода классификации**

В качестве алгоритма классификации выбрана сверточная нейронная сеть. Рассмотрим этот алгоритм подробнее.

Сверточная нейронная сеть – это алгоритм глубокого обучения, который принимает на вход обучающую выборку – векторы с каким-либо типом объектов (картинками, сигналами), присваивает веса различным признакам объектов и позволяет различить найденные признаки у объектов тестовой выборки.

### **3.3. Устройство сверточной нейронной сети**

Сверточная нейронная сеть имеет некоторое сходство с визуальной корой головного мозга. Одиночные нейроны отвечают на стимулы определенного ограниченного участка рецептивного поля. Набор таких чувствительных участков покрывает все рецептивное поле, при этом границы участков накладываются друг на друга (рис. 3.1).

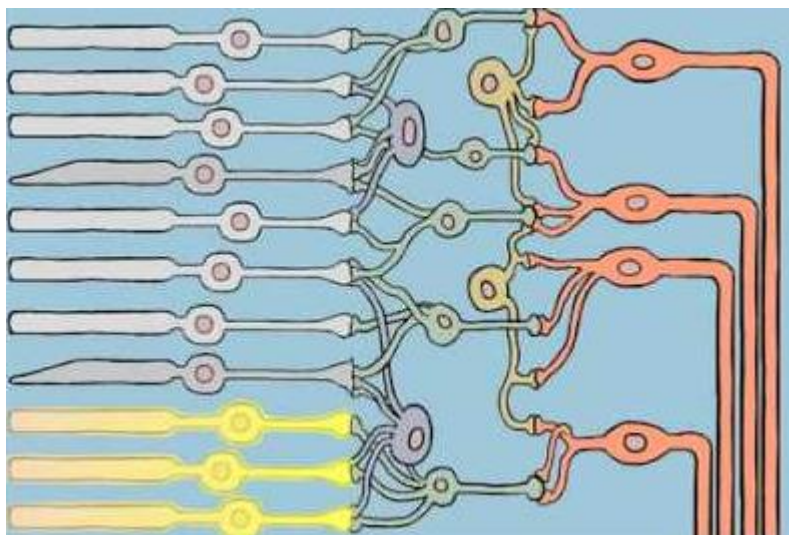


Рис. 3.1 Схема рецептивного слоя

В случае работы с объектами, требующими высоких вычислительных ресурсов, например, с картинками в высоком разрешении, Сверточная нейронная сеть выполняет роль инструмента, упрощающего процесс обработки. Она снижает количество обрабатываемых данных без потери признаков, которые важны для классификации.

В современных нейросетях помимо сверточного слоя используется целый ряд компонентов, которые обеспечивают важные оптимизации. Эти оптимизации снижают вычислительную стоимость, ускоряют обучение нейросети, а в случае с большими нейросетями делают обучение возможным в принципе.

### 3.4. Сверточный слой

Сверточный слой (ядро свертки) – представляет собой матрицу, имеющую набор весов. Сверточный слой имеет меньший размер, по сравнению с матрицей, содержащей исходное изображение.

Цель операции свертки – это извлечение признаков таких, например, как углы или ориентация элементов изображения. Сверточный слой накладывается на изображение, после чего осуществляется последовательное перемножение матрицы сверточного слоя и тех участков матрицы изображения, которые покрыты сверточным слоем. Затем полученные элементы суммируются. Результатом проведения операции свертки является карта признаков (рис. 3.2).

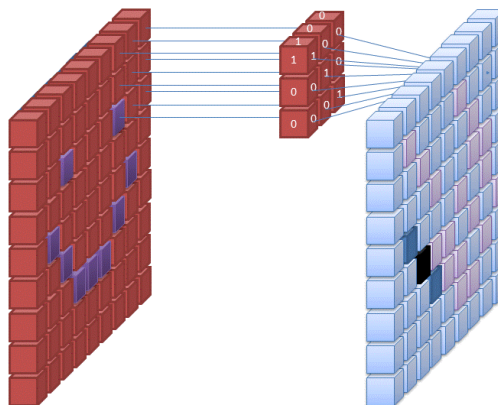


Рис. 3.2 Сверточная нейронная сеть: изображение, признак и карта признаков

Сверточные слои обеспечивают извлечение признаков не только из первоначального объекта, например, из изображения, но и из карты признака, которая была получена в результате выполнения предыдущих вычислений. В случае работы с изображением применительно к одному сверточному слою считаем, что  $\mathbf{X} \in \mathbb{R}^{H \times W \times D}$  представляет подаваемые на вход данные (изображение или карта признака) с тремя измерениями, где  $H$  и  $W$  представляют высоту и ширину, а  $D$  определяет число каналов в изображении или число слоев в карте признака.

Обозначим параметры фильтра (сверточное ядро) как  $\mathbf{W} \in \mathbb{R}^{h \times w \times D \times D'}$ , где  $h \times w$  представляют собой размер сверточного окна,  $D$  соответствует числу каналов, а  $D'$  - это число фильтров в данном сверточном слое. Для пространственного пути  $\mathbf{x} \in \mathbb{R}^{h \times w \times D}$  с тем же размером, что и ядро  $\mathbf{W}_k \in \mathbb{R}^{h \times w \times D}$  с помощью приведения путем векторизации  $\mathbf{x}$  и  $\mathbf{W}_k$  к  $\mathbf{x}_v \in \mathbb{R}^{(h \times w \times D) \times 1}$  и  $\mathbf{W}_k \in \mathbb{R}^{(h \times w \times D) \times 1}$  выходной нейрон  $s$ , который представляет собой число, может быть вычислен путем скалярного произведения  $\mathbf{W}_{kv}^T \mathbf{x}_v \in \mathbb{R}^1$ , где  $k = 1, 2, \dots, D'$ .

Затем соответствующая выходная карта признаков  $\mathbf{F}_k$  может быть представлена как  $\mathbf{X} * \mathbf{W}_k \in \mathbb{R}^{H' \times W' \times 1}$ , где  $k = 1, 2, \dots, D'$ .

Символ '\*' представляет собой операцию свертки,  $H' \times W'$  представляет размер получившейся карты признаков.

Для сверточного слоя, если сверточный шаг равен 1 и карта признаков, подаваемая на вход, дополнена до соответствующего размера

$((H + h - 1) \times (W + w - 1))$ , результирующая карта признаков будет иметь тот же размер, что и подаваемая на вход. То есть  $H' = H$  и  $W' = W$ .

Другими словами, если сверточный шаг больше, чем 1 или, если дополнение карты признаков не произведено, то выходная карта признаков будет меньше, чем та, которая подавалась на вход. То есть  $H' < H$  и  $W' < W$ .

Исходя из вышесказанного, операция свертки будет либо сохранять размер карты признаков, либо уменьшать его  $H' \leq H$  и  $W' \leq W$ .

Далее следует слой нелинейной активации и пулинговые слои. Для изображений первые несколько сверточных слоев отвечают за извлечение признаков с низким уровнем абстракции, такие как признаки границ объекта, а другие, более глубокие сверточные слои, на основе признаков с низким уровнем абстракции формируют более высокоуровневые признаки. Если рассматривать в качестве объекта лицо человека, то в этом случае могут быть сформированы такие абстрактные признаки как фрагменты элементов лица – контуры или текстуры глаза, носа, рта, волос и пр.

Однако существуют и определенные отличия в том случае, когда сверточная сеть применяется для классификации ЭКГ сигнала. На выходе промежуточного сверточного слоя из-за того, что признак имеет форму вектора в результате процедуры свертки создается не карта признаков, а вектор признаков.

Вход  $X' \in \mathbb{R}^{H \times 1 \times D}$  представляет векторы признаков  $D$ , каждый из которых имеет размер  $H \times 1$ . Таким образом, фильтр можно обозначить как  $W' \in \mathbb{R}^{h \times 1 \times D \times D'}$  для ЭКГ. На первом входном слое нейронной сети,  $D$  будет равняться 12. Это означает, что в случае с ЭКГ на нейронную сеть подается матрица из 12 строк.

Также в случае с обработкой ЭКГ выходная карта признаков  $F'_k$  может быть представлена как  $X' * W'_k \in \mathbb{R}^{H' \times 1 \times 12}$ , где  $k = 1, 2, \dots, D'$ .

Следовательно, чтобы модифицировать сверточную нейронную сеть для работы с ЭКГ сигналом, размер сверточного окна должен быть изменен на  $h \times 1$  (напр.  $3 \times 1$ ), вместо  $h \times w$

(напр.  $3 \times 3$ ). Таким образом, выполняется операция между векторами, а не между матрицами, что является упрощением.

Для эффективной классификации в нейросети сверточные фильтры должны иметь малый размер, по сравнению с размером вектора ЭКГ-сигнала, а также фильтры должны варьироваться по размеру.

Кроме того, малый размер фильтра обусловлен тем, что в ряде случаев патологическая ЭКГ отличается от нормальной весьма незначительно. Кроме того, размер фильтра влияет на вычислительную стоимость, что сокращает время обучения и ускоряет работу обученной нейросети. Это может быть актуально тогда, когда нейронная сеть используется на мобильных устройствах. Нейронная сеть может содержать набор фильтров различных масштабов для того, чтобы реагировать на признаки, имеющие различные размеры [5].

Архитектура разрабатываемой нейросети подразумевает наличие фильтров с шириной 16 семплов. Количество фильтров  $32 * 2^k$ , где  $k$  – гипер-параметр, который начинается с 0 и увеличивается на 1 в 2, 5, 8, 10 блоках.

### **3.5. Глубина нейронной сети**

С увеличением глубины нейронной сети информация, важная для классификации информация извлекается и сливается, увеличивается набор признаков с высокой абстракцией и сеть должна демонстрировать лучшую предсказательную способность. Однако вместе с увеличением глубины может возникнуть состояние переобучения, когда нейросеть будет показывать хорошие результаты только на объектах из обучающей выборки. Также может проявиться эффект затухания градиента. Нейросеть состоит из 24 слоев.

### **3.6. Слой субдискретизации**

Слой субдискретизации отвечает за уменьшение пространственного размера карты признаков, а также за поддержание инвариантности возмущения в данных, содержащихся в картах признаков (рис. 3.3). Под возмущением подразумевается большое значение элемента матрицы (по сравнению с другими элементами), которое связано с наличием какого-либо доминантного признака. Значение может находиться в любой части карты признаков, однако процедура пулинга выявит его. Также данный подход позволяет уменьшить вычислительную мощность, необходимую обработки данных.

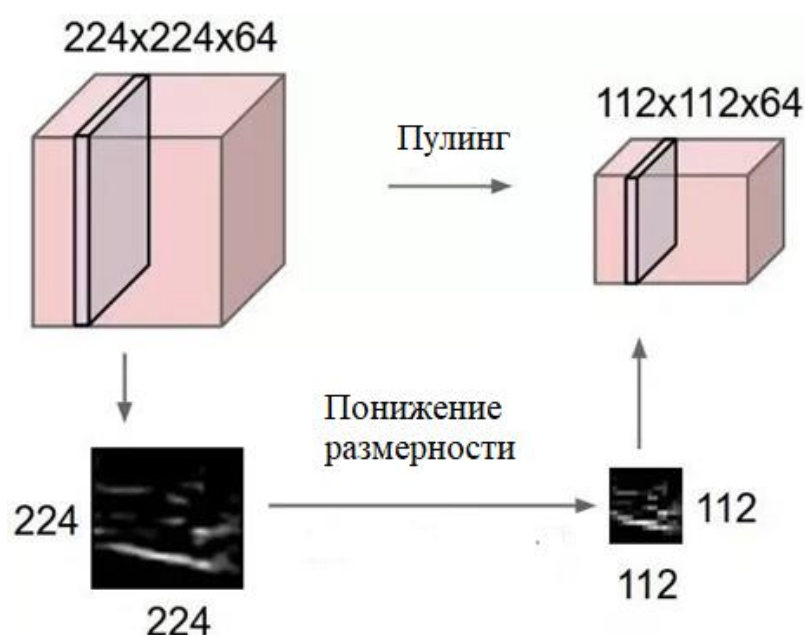


Рис. 3.3 Процедура пулинга

Существует два типа пулинга: пулинг с использованием функции максимума, когда возвращается максимальное значение участка, покрытого ядром, и пулинг с использованием функции среднего, когда возвращается среднее значение участка.

Пулинг с функцией максимума также может использоваться как средство для подавления шума. Таким образом, данные, подаваемые в сверточную нейронную сеть в ряде случаев могут не нуждаться в предварительной обработке.



Пулинг с функцией среднего подавляет шум с меньшей эффективностью. Так как фильтрация – важная часть обработки ЭКГ, в работе будет использован пулинг с функцией максимума.

Распространенный вариант пулинга – слой с фильтрами размера  $2 \times 2$ , величиной шага 2. При подаче на вход матрицы  $4 \times 4$  после прохождения пулинга отбрасывается 75% данных:

$$f_{X,Y}(S) = \max_{a,b=0}^1 f_{2X+a,2Y+b}$$

Значения, полученные с пулингового слоя передаются на полносвязную нейронную сеть с нелинейной активационной функцией.

### 3.7. Обучение сверточной нейронной сети

Для обучения сверточной нейронной сети используется метод прямого распространения, которая позволяет вычислить функцию потерь  $L(y, y') = [y' \neq y]$ , где  $y$  – фактическое значение из обучающей выборки, а  $y'$  – значение на выходе нейросети.

После вычисления функции потерь используется метод обратного распространения для вычисления параметров фильтра при помощи производной сложной функции. Затем параметры обновляются при помощи алгоритма градиентного спуска.

Градиентный спуск – это итеративный алгоритм оптимизации, который позволяет найти минимум функции (рис. 3.4).

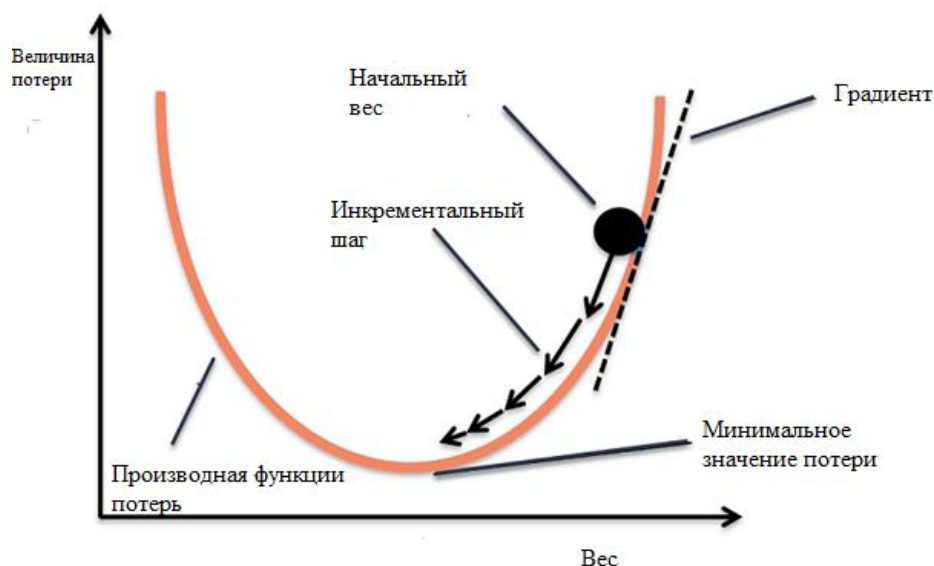


Рис. 3.4 Градиентный спуск

Цель алгоритма градиентного спуска – минимизация заданной функции, а именно, функции потерь обучаемой нейросети. Градиентный спуск выполняется в 2 шага. 1 шаг – вычисление наклона (градиента), который является производной первого порядка функции в заданной точке. 2 шаг – Движение в противоположную сторону относительно стороны, в которой увеличивается наклон.

Прямое распространение и обратное распространение выполняются итеративно до схождения функции потерь.

### 3.8. Активационная функция

Для сверточной нейронной сети затухание градиента может существенно замедлять схождения функции потерь. Такая проблема может быть вызвана неподходящей активационной функцией, которая используется для создания нелинейного отношения между входом и выходом. Для снижения затухания градиента в современных системах часто применяется активационная функцию ReLu:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

Если значение положительно, она возвращает  $x$ , если равно 0 или отрицательно, возвращается 0. Данная функция обеспечивает более быструю сходимость по сравнению с сигмоидальной функцией.

### 3.9. Дропаут

Дропаут (англ. dropout) – метод, который в процессе обучения на каждой итерации исключает из процесса обучения узлы нейросети с вероятностью  $p$  [6]. Исключенный нейрон всегда возвращает значение 0.

Дропаут повышает обобщающую способность нейросети. В переобученной сети нейроны в нижележащих слоях могут требовать консенсуса (единогласного решения) от нейронов из вышележащих слоев для того, чтобы выдать положительное решение относительно наличия признака. Нейросеть с такими свойствами показывает хороший результат на обучающей выборке, но плохой на тестовой. Так как объекты в тестовой выборке отличаются, нейроны в промежуточных слоях не находятся в состоянии консенсуса. За счет дропаута нейроны обучаются принимать решение, не исходя из наличия консенсуса, а по большинству голосов, таким образом, решается проблема переобучения.

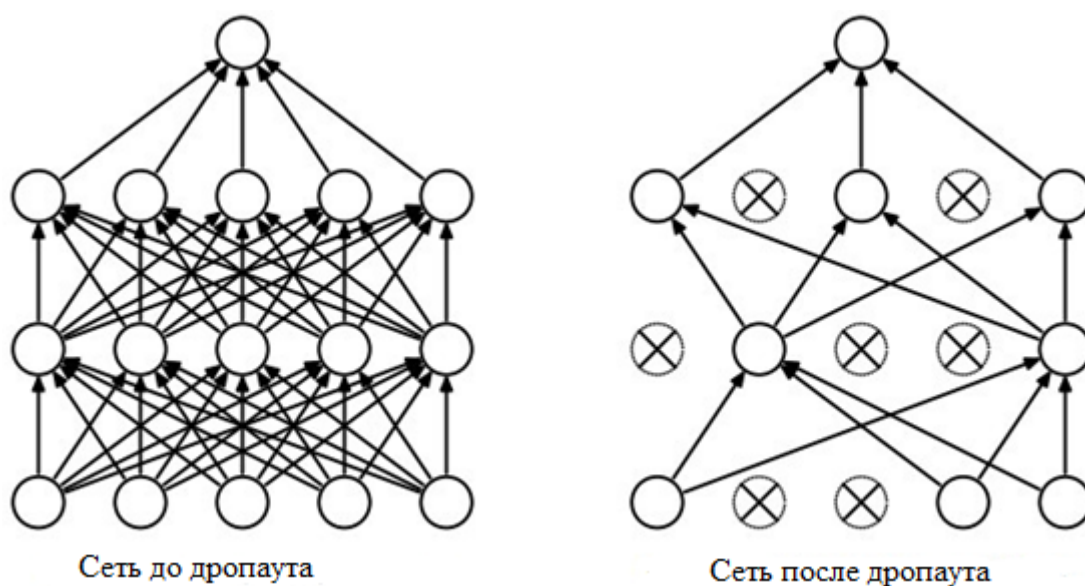


Рис. 3.5 Дропаут

Для разрабатываемой сети параметр дропаута имеет величину 0.2.

### 3.10 Пакетная нормализация

Перед подачей на вход (на первый сверточный слой) нейросети данные должны быть нормализованы.

К примеру, в случае работы с изображениями в формате RGB, каждый цветовой канал кодируется значениями от 0 до 255. В результате на вход подаются значения от 0 до 1.

Нормализация позволяет приводить данные к одному масштабу в результате чего увеличивается скорость схождения функции потерь и уменьшается время обучения.

Аналогично тому, как процедура приносит пользу при применении к данным, подаваемым на первый слой, она может быть применена к данным в скрытых слоях. Этот подход называется пакетной нормализацией.

Пакетная нормализация делает слои более независимыми друг от друга, позволяет использовать более высокий коэффициент скорости обучения, уменьшает эффект переобучения имеет эффект регуляризации [**Ошибка! Источник ссылки не найден.**].

Использование пакетной нормализации позволяет в меньшей степени использовать дропаут.

### 3.11. Полносвязная сеть

Сверточные сети могут быть существенно более глубокими, точными и иметь более высокую скорость тренировки, если они содержат короткие связи между слоями, близкими к входу и к выходу сети. Сети, которые соответствуют данному условию называются полносвязными сетями.

В то время как обычные сверточные сети с  $L$ -слоями имеют по одной связи между каждым слоем, полносвязные сети имеют  $\frac{L(L+1)}{2}$  связей. Для слоя  $L$  слоя используются все карты признаков из предыдущих слоев. Также карта признаков слоя  $L$  используется для всех нижележащих слоев.

### 3.12. Слой классификации (Softmax)

Функция используется в качестве последнего слоя нейронной сети. На вход Softmax подается вектор из последнего линейного слоя нейронной сети многоклассовой классификации, а на выходе получается вектор той же размерности с распределением вероятностей потенциальных исходов.

Функция описывается формулой  $\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}}$ , где  $z$  – вектор, который имеет размерность  $K$ .

### 3.13 Остаточные сети

Остаточные сети увеличивают скорость обучения за счет снижения влияния затухания градиента, вследствие того, что прямое и обратное распространение проходит через меньшее количество слоев (рис. 3.6).

Имеем матрицу весов  $W^{l-1,l}$  для связи весов между слоями  $l-1$  и  $l$  и матрицу весов  $W^{l-2,l}$  для связи весов между слоями.

Операция прямого распространения может выглядеть следующим образом:

$$a^l := g(W^{l-1,l} \cdot a^{l-1} + b^l + W^{l-2,l} \cdot a^{l-2})$$

$$a^l := g(Z^l + W^{l-2,l} \cdot a^{l-2}),$$

где  $a^l$  – выходное значение нейронов из слоя  $l$ ,  $g$  – *активационная функция слоя  $l$* ;  $W^{l-1,l}$  – матрица весов для нейронов между слоями  $l-1$  и  $l$  и  $Z^l = W^{l-1,l} \cdot a^{l-1} + b^l$ .

В общем виде это может быть выражено как

$$a^l := g(Z^l + \sum_{k=2}^l W^{l-k,l} \cdot a^{l-k}),$$

В последние годы остаточные сети позволили преодолеть ранее недостижимый барьер в 100 слоев в архитектуре нейросетей.

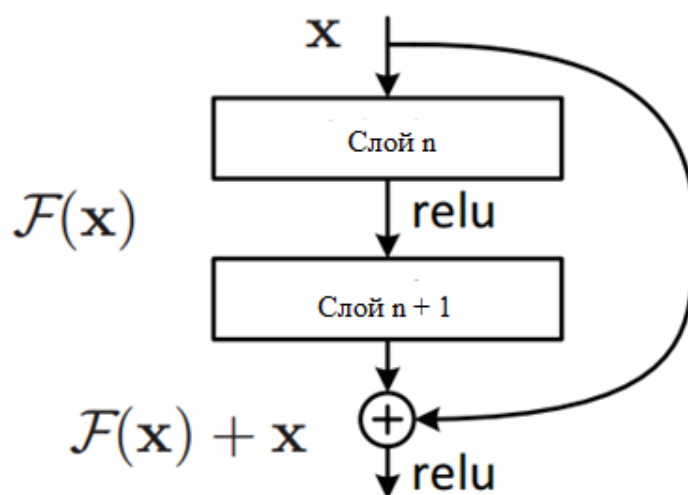


Рис. 3.6 Блок с остаточной сетью

Проектируемая нейросеть имеет 11 блоков остаточных сетей. По сверточные сети на блок.

### 3.14. Выводы

По результатам исследования разработана архитектура нейросети (рис. 3.7), которая состоит из следующих компонентов:

- сверточная сеть (23 слоя)
- пакетная нормализация
- активационная функция (ReLU)
- дропаут
- полносвязная сеть (1 слой)
- функция Softmax

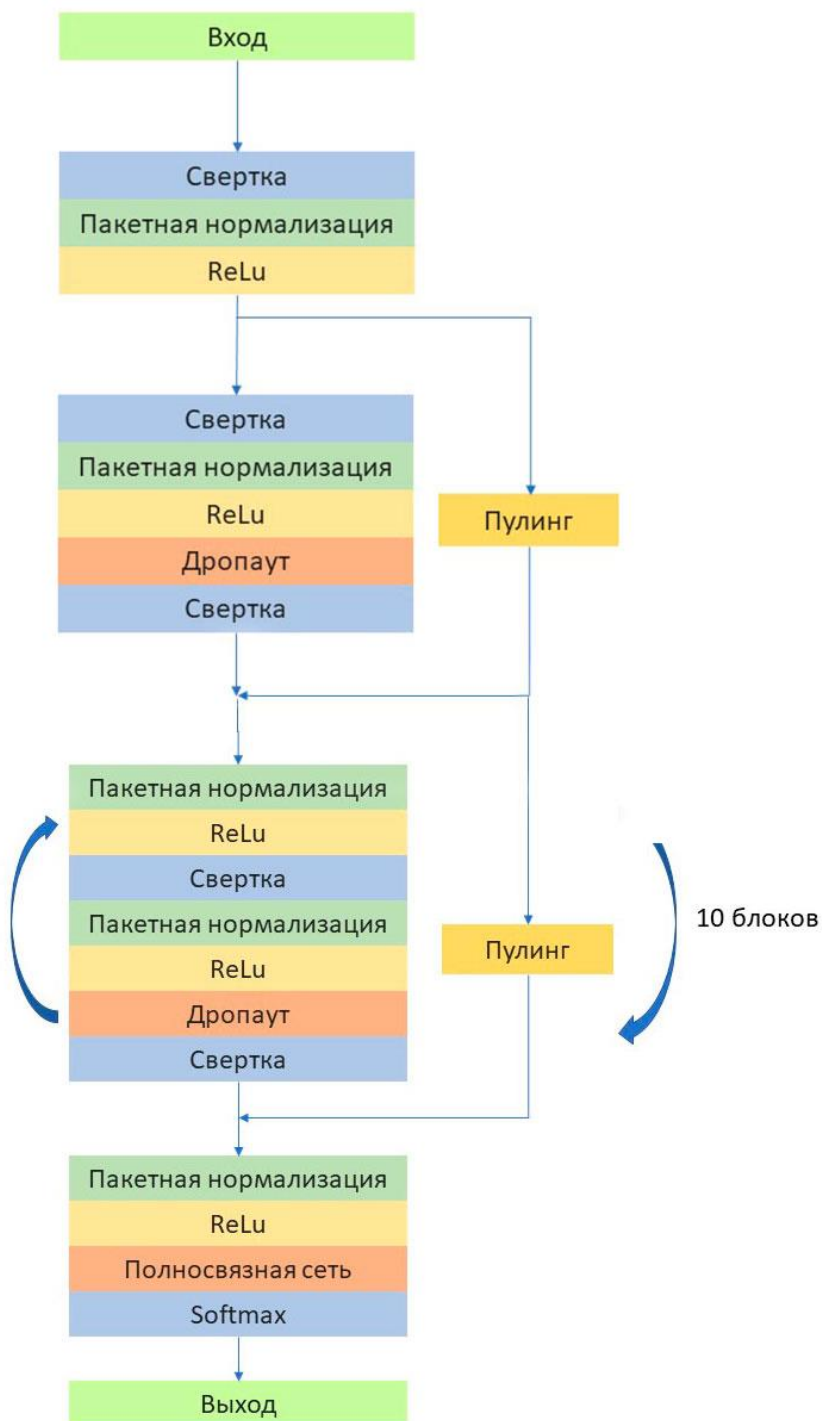


Рис. 3.7 Архитектура нейросети

Архитектура классификатора реализована на языке программирования Python.

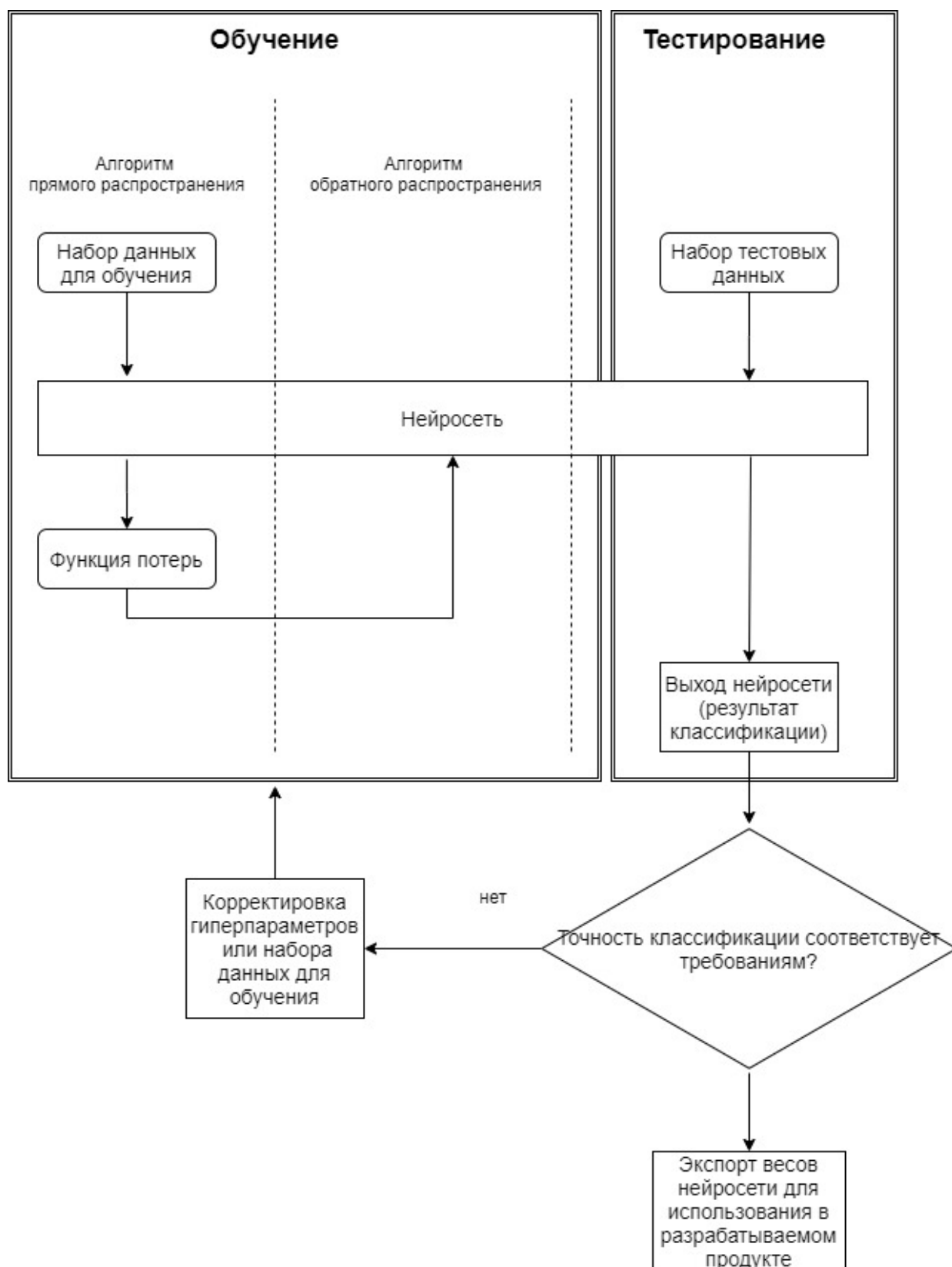


Рис. 3.8 Блок-схема алгоритма



## ЗАКЛЮЧЕНИЕ

В работе были проанализированы передовые подходы автоматизированной обработки ЭКГ и архитектуры систем классификации. Исследование показало, что наибольшей эффективностью обладают системы, основанные на сверточных нейросетях. Такие нейросети в процессе обучения автоматически выявляют признаки, которые позволяют выполнять классификацию ЭКГ.

Нейросеть способна выявить больше признаков, чем это может сделать человек, используя аналитические методы. При этом, эти признаки более значимы для классификации.

Классификаторы на основе нейросети способны классифицировать ЭКГ, как минимум, на уровне сопоставимом с человеком или даже лучше человека.

Так исследования показали, что средняя площадь под ROC-кривой составляет 0,97. F1-score для ГНН составляет 0,837, лучше, чем для кардиологов с F1-score равным 0,780.

В качестве научной основы работы были выбраны исследования последних 5 лет, описывающие лучшие в отрасли результаты. Классификаторы, описанные в этих исследованиях на вход принимают 1-канальную ЭКГ. На основании данных исследований была разработана архитектура нейросети, классифицирующая 12-канальную ЭКГ на 5 классов. Ожидается, что разработанная архитектура позволит точнее выполнять диагностику на основании более полного набора данных о работе сердца.

Архитектура классификатора реализована на языке программирования Python. Алгоритм может быть встроен в виде программных компонентов в системы, осуществляющие холтеровское мониторирование, а также в аппаратуру для снятия ЭКГ для экспресс-диагностики состояния пациента. Также на основе компонентов для классификации ЭКГ может быть построен онлайн-сервис, позволяющий осуществлять дистанционную постановку диагноза.

## СПИСОК ЛИТЕРАТУРЫ

1. Cardiovascular diseases (CVDs) // World Health Organization official site – 2017.  
– URL: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) — (дата обращения: 25.01.2020).
2. Hannun A., Rajpurkar P., Haghpanahi M. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network // J. Nature Medicine – 2019. – N 25 65–69
3. Huang G., Liu Z. Densely Connected Convolutional Networks // arXiv — 2017. — Single paper — 1-9
4. Li Y., Pang Y. Patient-specific ECG classification by deeper CNN from generic to dedicated // Neurocomputing — 2018. — 314 — 336-346
5. Luz E., Schwartz W. ECG-based Heartbeat Classification for Arrhythmia Detection: A Survey // Computer Methods and Programs in Biomedicine — 2016. — 127—144-164.
6. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv — 2015. — Separated conference paper — 1-14
7. Srivastava N., Hinton G. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. // Journal of Machine Learning Research — 2014. — 15 — 1929-1958
8. Zhang Z., Dong J. Heartbeat classification using disease-specific feature selection // Computers in biology and medicine — 2014. — 46C — 79-89.