Санкт-Петербургский политехнический университет Петра Великого

Институт машиностроения, материалов и транспорта

Высшая школа автоматизации и робототехники

**ОТЧЁТ ПО НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ**

**Исследование алгоритма автоматизации определения класса изделия по его 3D модели**

Выполнил

студент гр.3331506/70401 <*подпись*> И.А. Ляпцев

Руководитель

ст. преподаватель <*подпись*> А.С. Габриель

Научный консультант

м.н.с <*подпись*> В.В. Варлашин

«\_\_\_» \_\_\_\_\_\_\_\_\_\_ 2020 г.

Санкт-Петербург

2020

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 4](#_Toc61452652)

[1 Форматы 3D моделей 5](#_Toc61452653)

[1.1 Содержимое 3D моделей 5](#_Toc61452654)

[1.1.1 Кодирование геометрии моделей 5](#_Toc61452655)

[1.2 Хранение 3D моделей 8](#_Toc61452656)

[1.3 Основные форматы 11](#_Toc61452657)

[1.3.1 3ds…. 12](#_Toc61452658)

[1.3.2 Obj… 13](#_Toc61452659)

[1.3.3 fbx…. 14](#_Toc61452660)

[1.3.4 stl…… 14](#_Toc61452661)

[1.3.5 dae…. 15](#_Toc61452662)

[1.4 Вывод по разделу 16](#_Toc61452663)

[2 Классификация изделий 17](#_Toc61452664)

[3 Алгоритмы классификации 20](#_Toc61452665)

[3.1 Классификация на основе пошагового построения 20](#_Toc61452666)

[3.2 Классификация, используя типовые методы компьютерного зрения 21](#_Toc61452667)

[3.3 Классификация, используя граф Риба 24](#_Toc61452668)

[3.4 Классификация, используя воксели 25](#_Toc61452669)

[4 Выбор нейросети и ее параметров 28](#_Toc61452670)

[4.1 Выбор входных и выходных данных 28](#_Toc61452671)

[4.2 Архитектуры нейросетей 28](#_Toc61452672)

[4.3 Вывод по разделу 36](#_Toc61452673)

[ЗАКЛЮЧЕНИЕ 37](#_Toc61452674)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 39](#_Toc61452675)

# ВВЕДЕНИЕ

Классификация и архивирование инженерных работ облегчает повторное использование деталей в дальнейшем. Было подсчитано, что 75% проектных работ состоит из базового дизайна корпуса, который является вариациями и усовершенствованиями каких-либо предыдущих проектов.

Текущие каталоги описывают детали по различным атрибутам, таким как производитель, серия продуктов и просто номера моделей. Эти атрибуты требуют ручной записи и уточнения, и никак не описывают строение детали. Именно поэтому работы в областях компьютерного зрения, графики и др. стремятся автоматизировать процесс классификации деталей.

В этой работе исследуются методы распознавания и классификации паттернов для последующей идентификации и определения в базах данных САПР. Основная задача – улучшить текущие методы классификации и облегчить для инженеров поиск и повторное использование, тем самым улучшая управление жизненным циклом продукта, а также ускоряет разработку и проектирование продукта.

Для ее достижения определено несколько этапов – необходимо определить объект, над которым будут производиться операции, в данном случае – файл 3D модели определенного формата. Также необходимо определить, к каким классам будет привязан алгоритм, и в конце определить, на какой именно системе будет базироваться алгоритм. В результате необходимо получить решение, посредством которого можно разработать алгоритм, определяющий класс детали.

# Форматы 3D моделей

Изначально стоит определить, что именно подразумевается под понятием “3D модель”. 3D-модель – это объемная фигура в пространстве, создаваемая в специальной программе. За основу, как правило, принимаются чертежи, фотографии, рисунки и подробные описания, опираясь на которые, специалисты и создают виртуальную модель.

Прежде чем анализировать способы классификации и варианты алгоритмов, стоит определить, 3D модели какого именно формата мы будем использовать в дальнейшем. В этом разделе рассматриваются основные форматы представления 3D моделей, их достоинства и недостатки.

## Содержимое 3D моделей

Прежде чем говорит о кодировании форматов и их принципиальных достоинствах, стоит определить особенности 3D форматов. В этом разделе описываются общие особенности и содержимое форматов.

### Кодирование геометрии моделей

Основное, что существует у 3D модели любого формата – ее геометрия, т. е. общее строение фигуры, положение ее граней и вершин [1]. Это является базисом, который в дальнейшем уточняется с помощью различных элементов.

Существует три принципиальных метода кодирования информации о геометрии поверхности, каждый из которых имеет свои преимущества и недостатки. Это аппроксимирующая сетка (англ. approximate mesh), точная сетка (англ. precise mesh) и конструктивная блочная геометрия (англ. constructive solid geometry (CSG)).

В случае кодировки аппроксимирующей сеткой поверхность модели представляется сеткой многоугольников (или полигонов), чаще всего это треугольники, так как это наименьшая 2D фигура, с которой можно смоделировать форму (рисунок 1.1).

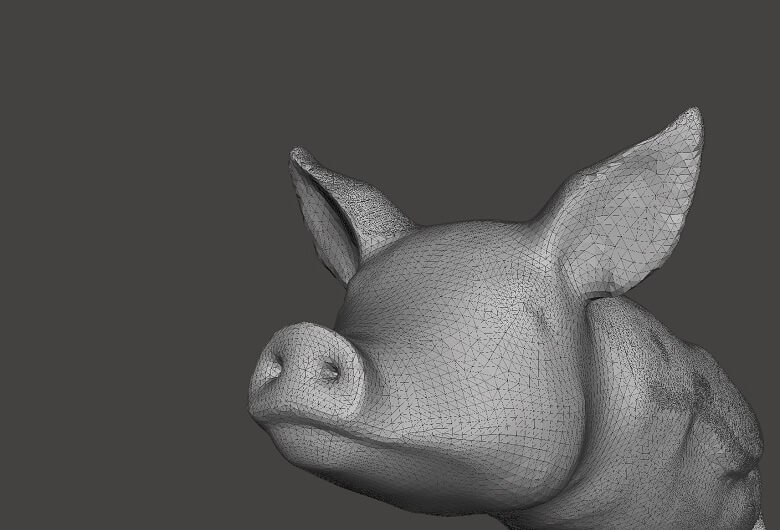


Рисунок 1.1 – Пример представления 3D модели

Вершины покрывающих многоугольников и вектор нормали сохраняются в файле, это и есть аппроксимация геометрии поверхности модели (рисунок 1.2)

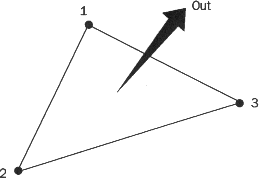
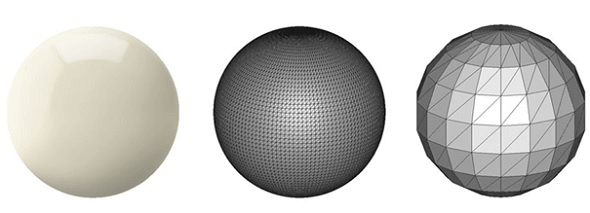


Рисунок 1.2 – Единица аппроксимирующей сетки

Процесс покрытия поверхности многоугольниками называется тесселяция, а форматы файлов – тесселированными. Чем больше многоугольников используется для представления фигуры, тем точнее получается аппроксимированная модель (рисунок 1.3).

Рисунок 1.3 – Аппроксимированная модель при различных уровнях точности

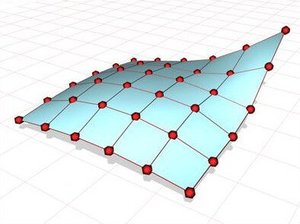
Вариант представления моделей в виде мелкой сетки позволяет добиться максимальной точности, используя функции Non-Uniform Rational B-Spline (или NURBS) вместо полигонов (рисунок 1.4). Эти параметрические поверхности состоят из небольшого количества взвешенных контрольных точек и набора параметров, называемых узлами. На основе узлов поверхность может быть вычислена математически путем плавной интерполяции по контрольным точкам.

Рисунок 1.4 – Пример патчей NURBS

Эти поверхности выглядят гладкими в любом масштабе и могут воспроизводить геометрию поверхности небольшой части 3D модели. Однако всегда есть компромисс. Хотя мелкая сетка точна при любом разрешении, они отрисовываются медленнее, и их следует избегать в приложениях, где важна быстрая визуализация.

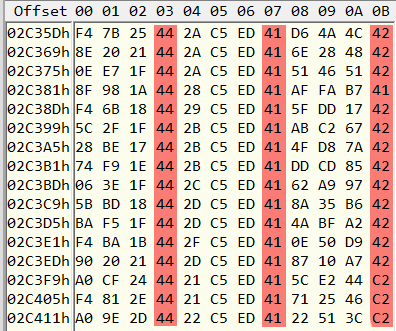
Последний вариант представления моделей – конструктивная блочная геометрия. Для этого варианта не используются сетки, а формы создаются путем простейших операций (сложение и вычитание) примитивных форм (кубы, сферы и т. д.). Большинство САПР используют этот принцип.

Конструктивная твердотельная геометрия подходит для описания проектируемых 3D моделей и очень удобна для пользователя. Еще одним большим преимуществом является то, что каждый отдельный шаг редактирования (сложение, вычитание, преобразования примитивных форм) сохраняется в этом формате файла 3D. Таким образом, любой шаг можно отменить и повторить в любой момент.

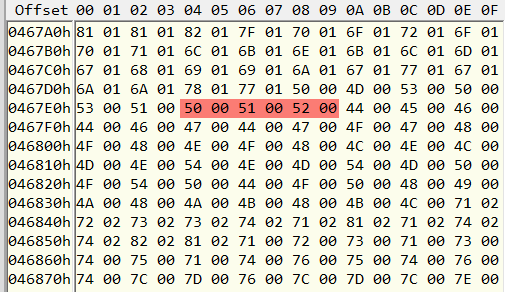
Очевидно, что при конвертации этого формата в формат на основе сетки теряется информация об отдельных этапах редактирования.

## Хранение 3D моделей

Если взять за основу наиболее часто встречающийся вариант с аппроксимирующей сеткой, то основной объём любого файла с моделью составляют несколько больших таблиц с данными о вершинах, о том, как они соединяются и как на них натягиваются текстуры.

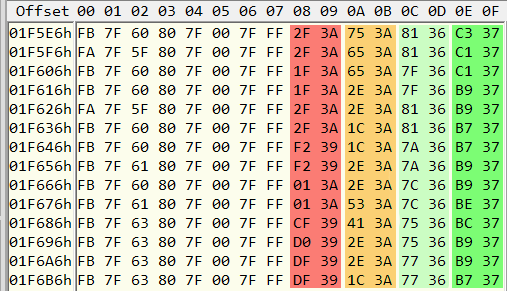
Начать стоит с вершин. Простой список с координатами x, y, z в шестнадцатеричном формате показан на рисунке 1.7.  
Рис 1.7 – Таблица координат x, y, z вершин

Так как чаще всего координаты хранятся в виде 32-битных плавающих чисел, их легко распознать внутри файла по повторяющимся через 4 байта цифрам в диапазоне 40-45, или для отрицательных чисел C0-C5. Конечно, могут встречаться и иные байты, но вероятность такого невелика. Так происходит потому, что диапазон координат 3D модели небольшой с точки зрения порядка, а порядок как раз хранится в старшем байте. Далее, нужна таблица, где указано, в каком порядке вершины соединяются в треугольники. Пример такой таблицы показан на рисунке 1.8.

Рисунок 1.8 – Пример хранения треугольников, моделирующих форму

Это 16-битные номера вершин, группами по три. Так как в моделях обычно не более нескольких сотен вершин, числа эти маленькие, и такую таблицу тоже легко видно визуально. В данном примере выделен один из треугольников, состоящий из вершин с номерами 50, 51 и 52.

И третья – таблица текстурных координат, чтобы привязать вершины к плоской текстуре, которую нужно на них натянуть, показана на рисунке 1.9.

Рисунок 1.9 – Таблица соответствия текстуры и вершин

Координаты x и y в пределах текстуры приводятся к диапазону от 0 до 1, и, если текстура имеет размер 2048x2048 или 4096x4096, нет смысла в высокой точности. Поэтому они чаще всего хранятся как плавающие числа с половинной точностью, 16 битные. Старший байт у них получается чуть больше 0x30, изредка доходя до 0x40 или чуть больше. Здесь красным и оранжевым выделены текстурные координаты, зеленым и светло-зеленым – координаты на карте освещения.

## Основные форматы

Форматы можно разделить по применимости на 2 группы – универсальные, которые используются во всех САПР, и собственные форматы, которые используются только в определенных САПР (к примеру, DWG, использующийся программой Autodesk).

Универсальные форматы служат «промежуточным звеном», позволяющим различным САПР использовать одну и ту же модель. Тогда, после создания модели, она преобразовывается в универсальный формат, возможно, как будет рассмотрено далее, с другим методом хранения/кодирования, который позволяет читать и редактировать модель в различных САПР.

В этом разделе рассматриваются наиболее часто встречающиеся форматы и также определяется, насколько хорошо они подходят для задачи классификации моделей. Основные критерии для определения подходящих форматов можно определить так: во-первых, они должны кодировать значения в удобном виде (в идеале – текстовом). Это позволит удобно их читать, что может понадобиться при разработке алгоритма. Также они должны читаться, а главное, создаваться в различных САПР, и хранить минимально необходимую информацию – нет нужды в информации о текстуре или о среде, так как для определения класса эти параметры значения не имеют.

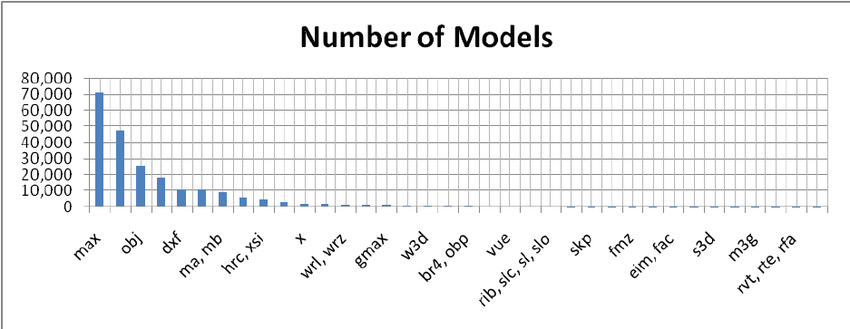
Сначала необходимо определить, какие именно форматы надо рассматривать. В этой работе выборка будет основана на общедоступных статистических данных загрузок для каждого формата. На рисунке 1.10 показана статистика количества моделей для основных форматов.

Рисунок 1.10 – Статистика количества моделей [2]

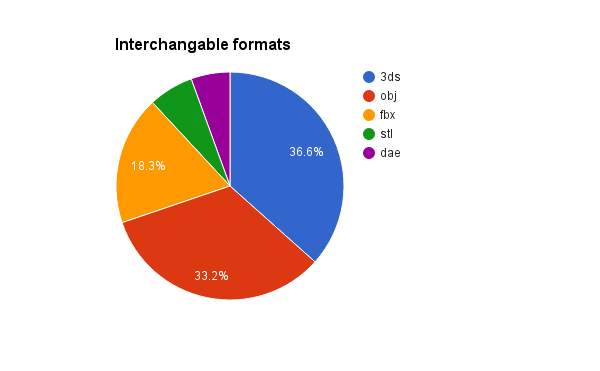
Далее необходимо исключить редко встречающиеся и не-универсальные форматы, получим диаграмму, представленную на рисунке 1.11.

Рисунок 1.11 – Статистика скачиваний для универсальных форматов

Далее рассмотрим каждый формат и определим, насколько он подходит для решения поставленной задачи, исходя из ранее обозначенных критериев.

### 3ds

Формат файла 3ds является основным форматом программного обеспечения AutoDesk 3ds Max. Это двоичный формат, состоящий из блоков, содержащих различную информацию. Блоки содержат идентификацию, указывающую, какая информация там сохраняется и смещение до следующего фрагмента [3]. Таким образом, программное обеспечение, которое не поддерживает определенные свойства рендеринга могут просто игнорировать их.

Формат файла 3ds поддерживает: геометрию в форме вершин / граней и параметрических поверхностей, текстуры, физические свойства материала, преобразования, информацию о камере и источники света.

### Obj

Формат файла obj – это текстовый открытый формат файла, разработанный Wavefront Technologies (сейчас Alias ​| Wavefront). Формат был принят другими поставщиками приложений для 3D графики, и некоторые из них могут импортировать / экспортировать. Формат файла obj состоит из ряда строк, каждая из которых содержит ключ и различные значения. Ключ в каждой строке указывает информацию о типе, которой следует следовать. По этой причине формат файла obj не требует заголовка. В таблице 1.1 показаны основные ключи, используемые при кодировании.

Таблица 1.1. Файловые ключи obj

|  |  |
| --- | --- |
| Ключ | Описание |
| # | Комментарий |
| v | Вершина |
| l | Линия |
| f | Грань |
| vt | Координата текстуры |
| vn | Нормаль |
| g | Группа |
| … | … |

Простой пример, содержащий одно лицо, будет выглядеть так:

v 0.0 0.0 0.0

v 0.0 1.0 0.0

v 1.0 0.0 0.0

f 1 2 3

Формат файла obj поддерживает: геометрию в форме вершин / ребер / граней и параметрические поверхности, нормали вершин, текстуры, свойства материалов и группы.

### fbx

FBX может храниться на диске в виде бинарных или ASCII данных, FBX SDK поддерживает оба формата. Оба формата не задокументированы, однако формат ASCII имеет древовидную структуру с чёткими обозначениями идентификаторов, в то время как бинарный формат не задокументирован, но есть неофициальная спецификация, представленная Blender Foundation.

Формат файла fbx поддерживает: геометрию в форме вершин / граней и параметрических поверхностей, текстуры, физические свойства материала, преобразования, информацию о камере и источники света.

### stl

STL (от англ. stereolithography) – формат файла, широко используемый для хранения трёхмерных моделей объектов для использования в аддитивных технологиях. Информация об объекте хранится как список треугольных граней, которые описывают его поверхность, и их нормалей. STL-файл может быть текстовым (ASCII) или двоичным [4]. Свое название получил от сокращения термина «Stereolithography», поскольку изначально применялся именно в этой технологии трехмерной печати.

Файл типа ASCII STL начинается со строки:

*solid name*

где name – необязательная строка (но, если name опущено, всё равно должен быть пробел после solid). Файл продолжается произвольным числом треугольников, описываемых следующим способом:

*facet normal ni nj nk*

*outer loop*

*vertex v1x v1y v1z*

*vertex v2x v2y v2z*

*vertex v3x v3y v3z*

*endloop*

*endfacet*

где каждое n и v – число с плавающей точкой в формате: знак, мантисса, «e», знак, экспонента, например «-2.648000e-002». Файл завершается строкой:

*endsolid name*

Недостатки формата:

* Невысокая точность геометрии (в версии ASCII можно задавать произвольную точность).
* Большой объём файла для сложных моделей.

### dae

COLLADA – это формат, разработанный для обмена между 3D приложениями. Использует открытый стандарт XML для обмена форматами, которые в противном случае были бы несовместимы.

COLLADA был задуман как промежуточный формат для переноса файлов. Реализована поддержка таких программ, как Maya (используя ColladaMaya); 3ds Max (при помощи ColladaMax); Poser (v.7.0); Lightwave 3D (version 9.5); Cinema 4D (MAXON); Softimage|XSI; Houdini; MeshLab; CityScape, CityEngine, SketchUp, Blender, modo и Strata 3D. Adobe Photoshop с версии CS3.

Игровые движки, такие как Unreal engine, Unity и Torque 3D, также поддерживают этот формат. Библиотека Assimp также поддерживает формат COLLADA.

Формат по принципам использования и возможностям близок к fbx, который уже был рассмотрен ранее.

## Вывод по разделу

Исходя из поставленных задач и критериев, можно выделить несколько форматов, которые подходят для дальнейшей работы. Во-первых, это 3ds, который является универсальным и достаточно удобным, но хранит излишнюю информацию, в том числе преобразования, информацию о камере и источниках света.

Формат fbx и dae, в свою очередь, также хранят излишнюю информацию, но имеют идентификаторы, которые достаточно удобны при использовании. Главный недостаток форматов – закрытый исходный код.

Формат stl крайне удобен тем, что имеется возможность задавать точность моделей, что позволяет уменьшить размер файла, но для алгоритма метод хранения – набор треугольников – может быть неудобен.

Формат obj имеет достаточно удобное строение файла, идентификаторы для вершин, граней и т. д., а также хранит минимально необходимую информацию об объекте.

В итоге можно сказать, что лучшими форматами для работы являются obj, как основной, и stl, как альтернативный.

# Классификация изделий

Далее после того, как определен формат хранения модели, необходимо определить, чем руководствоваться при их классификации. В этой работе была выбрана Единая система конструкторской документации (далее ЕСКД), а именно «Общероссийский классификатор изделий и конструкторских документов ОК 012-93» [5].

ЕСКД, согласно ГОСТ 2.001-2013 – комплекс государственных стандартов, устанавливающих взаимосвязанные правила, требования и нормы по разработке, оформлению и обращению конструкторской документации, разрабатываемой и применяемой на всех стадиях жизненного цикла изделия (при проектировании, разработке, изготовлении, контроле, приёмке, эксплуатации, ремонте, утилизации) [6]. Основное назначение ЕСКД – установление единых опциональных правил, требований и норм выполнения, оформления и обращения конструкторской документации, именно поэтому для задачи классификации была выбрана именно эта система.

Классификатор ЕСКД представляет собой систематизированный свод наименований классификационных группировок объектов классификации – изделий основного и вспомогательного производства всех отраслей народного хозяйства, общетехнических документов и их кодов и является составной частью Единой системы классификации и кодирования технико-экономической информации (ЕСКК ТЭИ).

В Классификатор ЕСКД включены классификационные характеристики изделий – деталей, сборочных единиц, комплектов, комплексов (ГОСТ 2.101-68 ""ЕСКД. Виды изделий""), на которые разработана и разрабатывается конструкторская документация по ЕСКД, в том числе стандартные изделия, а также общетехнические документы (нормы, правила, требования, методы и т. д.) на изделия, входящие в Классификатор ЕСКД.

Обозначение изделий и конструкторских документов устанавливается по ГОСТ 2.201-80 "ЕСКД. Обозначение изделий и конструкторских документов". Обозначение основного конструкторского документа (чертежа детали или спецификации) включает: код организации-разработчика (четыре знака), код классификационной характеристики (шесть знаков), код порядкового регистрационного номера (три знака) [7].

Классификационная характеристика является основной частью обозначения изделия и его конструкторского документа. Код классификационной характеристики изделия присваивается по Классификатору ЕСКД и представляет собой шестизначное число, последовательно обозначающее класс (первые два знака), подкласс, группу, подгруппу, вид (по одному знаку). Структура обозначения кода классификационной характеристики имеет вид:

XX X X X X

1,2 цифры – Класс

3 цифра – Подкласс

4 цифра – Группа

5 цифра – Подгруппа

6 цифра – Вид

Для данной работы стоит заострить внимание на классах 71-75, и их производных. Эти классы представлены в таблице 2.1.

Таблица 2.1. Используемые классы изделий.

|  |  |
| --- | --- |
| Обозначение класса | Подробная информация |
| 71 | Детали – тела вращения типа колец, дисков, шкивов, блоков, стержней, втулок, стаканов, колонок, валов, осей, штоков, шпинделей и др. |
| 72 | Детали – тела вращения с элементами зубчатого зацепления; трубы, шланги, проволочки, разрезные, сектора, сегменты; изогнутые из листов, полос и лент; аэрогидродинамические; корпусные, опорные; емкостные; подшипников. |
| 73 | Детали – не тела вращения корпусные, опорные, емкостные. |
| 74 | Детали – не тела вращения плоскостные; рычажные, грузовые, тяговые; аэрогидродинамические; изогнутые из листов, полос и лент; профильные; трубы. |
| 75 | Детали – тела вращения и (или) не тела вращения, кулачковые, карданные, с элементами зацепления, арматуры, санитарно-технические, разветвленные, пружинные, ручки, уплотнительные, отсчетные, пояснительные, маркировочные, защитные, посуды, оптические, электрорадиоэлектронные, крепежные. |

Очевидно, что чем точнее необходимо классифицировать изделие, тем сложнее и точнее необходимо сделать любой алгоритм классификации, и тем больше вероятность ошибки. Поэтому было принято решения для первых версий алгоритма определять только класс изделия.

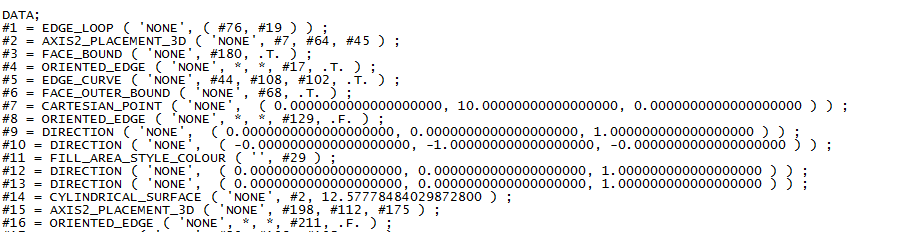
# Алгоритмы классификации

Для классификации моделей применяется большое количество алгоритмов, и далеко не все подходят для задачи классификации инженерных изделий по ГОСТ. Ниже рассмотрены возможные алгоритмы и их применимость.

## Классификация на основе пошагового построения

Как уже было сказано ранее, в большинстве САПР для создания изделий используются конструктивная твердотельная геометрия, которая позволяет построить модель сложной формы за счет действий с моделями простейших форм. Первый вариант классификации моделей – определять шаги построения моделей, а затем на их основе выделять характеризующие для каждого класса черты (к примеру, для деталей вращения – построение окружности и выдавливание по ней).

Этот подход выделяется тем, что не требует никаких особенных вычислительных мощностей, а сами пошаговые операции вполне возможно получить, к примеру, из САПР SolidWorks с помощь файла STEP (Рисунок 3.1)

Рисунок 3.1 – Представление файла STEP в текстовом редакторе

Но этот подход имеет несколько существенных минусов. Во-первых, он не является универсальным, так как для большинства классов невозможно определить какие-либо характерные шаги при построении. Во-вторых, высока вероятность ошибок, так как при большом количестве шагов необходимо учитывать не просто какой-либо один шаг, а их набор.

В итоге было принято решения разрабатывать алгоритм, не используя этот подход из-за предположительно невысокой точности и крайне высокой сложности разработки.

## Классификация, используя типовые методы компьютерного зрения

Для извлечения информации об объекте можно воспользоваться методами, активно использующимися в компьютерном зрении. Это позволит уменьшить количество информации об объекте до минимально необходимого, и соотносить только информацию, полученную после обработки 3D моделей.

Сначала стоит выделить основные понятия, и далее перейти к алгоритму. В компьютерном зрении существует несколько методов извлечения информации из модели. Первый – использование контуров, по которым в дальнейшем будут определяться параметры. Для тех же целей используются шаблоны.

В первом случае контур составляется как место быстрого изменения функции интенсивности изображения. Но у этого метода есть существенные ограничения, которые не позволяют использовать их в этом случае. Из ограничений контурного анализа стоит выделить два – неоднозначность определения контура, к примеру, при наличии помех или одинаковой яркости объекта и фона.

Для шаблонов также стоит выделить ограничения, которые, ввиду поставленной задачи, являются критическими: шаблон зависит от масштаба, углов обзора, поворота картинки (для 2D изображений) и наличия физических помех. Также возможны ложные срабатывания алгоритма, когда искомого объекта на нет, но имеются общие детали у шаблона и области на входном изображении.

Тогда выделим третий метод, которым технически можно воспользоваться для решения поставленной задачи – дескрипторы и особые точки.

Особая точка – это точка изображения, окрестность которой *o(m)* можно отличить от окрестности любой другой точки изображения *o(n)* в некоторой другой окрестности особой точки *o2(m)*. В качестве окрестности точки изображения для большинства алгоритмов берётся прямоугольное окно, составляющее размер *KxK* пикселей. Процесс определения особых точек достигается путем использования детектора и дескриптора.

В 1992 *Haralick* и *Shapiro* выделили следующие требования к особым точкам в виде следующих свойств:

•        *Отличимость (distinctness)* – особая точка должна явно выделяться на фоне и быть отличимой (уникальной) в своей окрестности.

•        *Инвариантность (invariance)* – определение особой точки должно быть независимо к аффинным преобразованиям.

•        *Стабильность (stability)* – определение особой точки должно быть устойчиво к шумам и ошибкам.

•        *Уникальность (uniqueness)* – кроме локальной отличимости, особая точка должна обладать глобальной уникальностью для улучшения различимости повторяющихся паттернов.

•        *Интерпретируемость (interpretability)* – особые точки должны определяться так, чтобы их можно было использовать для анализа соответствий и выявления интерпретируемой информации из изображения.

Детектор – это метод извлечения особых точек из изображения. Детектор обеспечивает инвариантность нахождения одних и тех же особых точек относительно преобразований изображений.

Дескриптор – производит описание найденных ключевых точек, оценивая их позиции через описание окружающих областей. В свою очередь, дескрипторы должны обеспечивать инвариантность нахождения соответствия между особыми точками относительно преобразований изображений. Ищутся особенности, которые более устойчивые к повороту, масштабированию и т. д.

Дескриптор – вектор, описывающий структуру окрестностей особых точек. Как правило, эти вектора формируются на основе набора значений первых и вторых производных изображения в точке.

Дескрипторы должны быть: специфичны (отличаем разные точки), локальны (зависеть только от небольшой окрестности), инвариантны (к искажениям\изменениям освещенности), просты в вычислении.

Дескрипторы, основанные на градиенте яркости, инвариантны к сдвигу яркости.

При описании дескриптора в простой окрестности имеются следующие недостатки: небольшой сдвиг приводит к существенному изменению, не инвариантен к повороту.

Для представления изображения можно использовать гистограммы. Это распределение каких-то признаков (цвет, текстура, глубина). Чаще всего используются гистограммы ориентированных градиентов.

*SIFT:* для детекции используется детектор *DOG* (разница гауссиана). Ориентация идет по градиентам (что позволяет построить гистограмму). Устойчив к изменению освещенности и небольшим сдвигам.

*SURF* – быстрая аппроксимация идей *SIFT*. Решает две задачи – поиск особых точек и создание дескриптора. Обнаружение особых точек основано на вычислении детерминанта матрицы Гессе. Особые точки представляют из себя круги (диаметр показывает масштаб точки). Показаны линии градиента яркости. Дескриптор имеет 64 или 128 чисел описания ключевой точки.

Для нахождения особой точки *SURF* пробегает по пикселям изображения и ищет максимум гессиана (матрица Гессе). Гессиан зависит только от перепада яркости и инвариантен относительно вращения, но не инвариантен к масштабу. Поэтому нужно перебирать разные масштабы фильтров и поочередно применять их к пикселю.

*SURF* разбивает все множество масштабов на октавы. Каждая октава содержит в себе 4 фильтра. Первая октава применяется для каждого второго пикселя, вторая – для каждого четвертого и т. д. (чтобы вычисления производились быстрее).

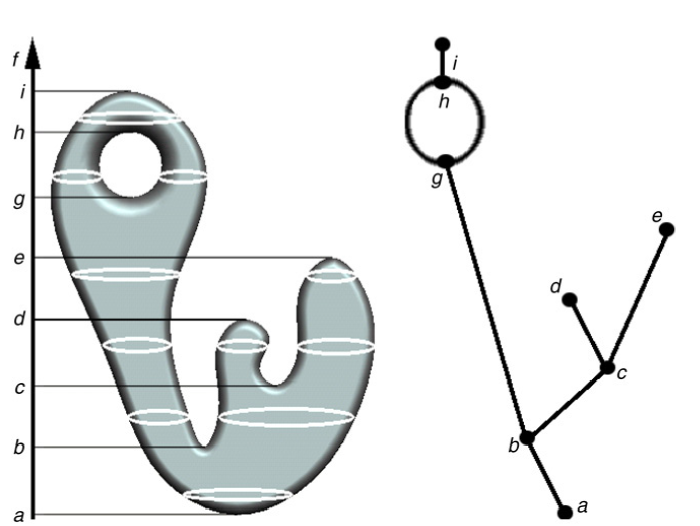
Алгоритм *HOG:* рассматриваем каждый пиксель и его окрестность. Находим направление, в котором цвет становится темнее (градиент), таким образом изображение заменяется на набор векторов, показывающих поток от света к темноте. Далее выделяются области и в них находятся средние направления.

К сожалению, применение типовых методов к 3D моделям достаточно сложно относительно других, специальных методов, описанных далее, из-за чего от их применения решено было отказаться.

## Классификация, используя граф Риба

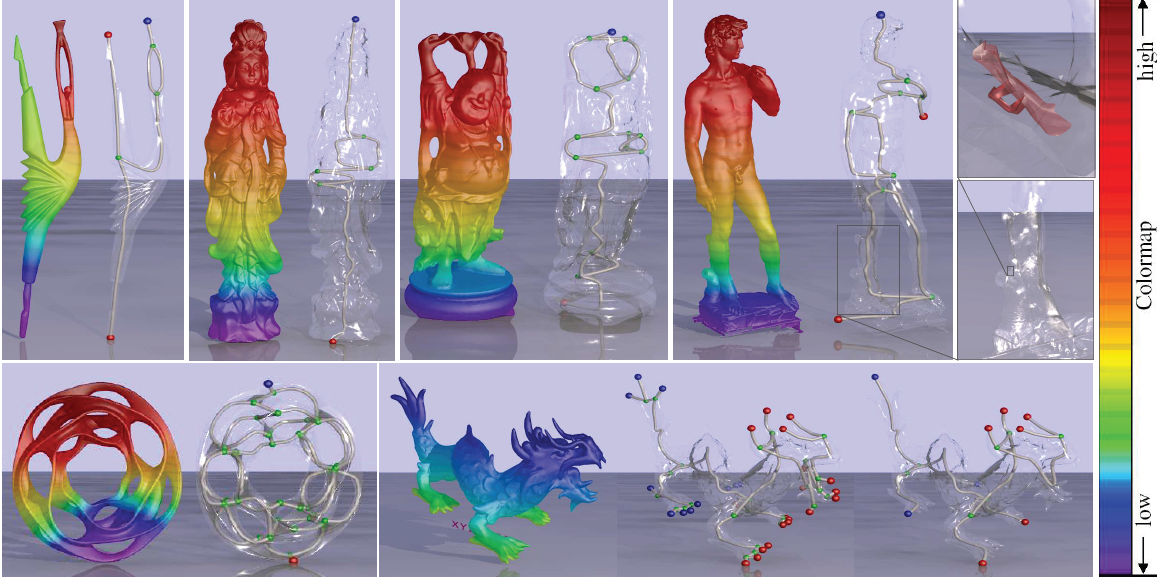
Для анализа форм и получения основной топологической информации об объекте также можно использовать дескриптор формы, который был изобретен Жоржем Рибом в его работе, датированной 1946 годом.

Граф Риба передает топологическую информацию, относящуюся к наборам уровней функции, определенной на форме. Пример представления графа Риба показан на рисунке 3.2. Буквы от a до f обозначают соответствие между критическими значениями и узлами графа [8].

Рисунок 3.2 – Представление Графа Риба относительно высоты.

Вершиной графа Риба называют точку, отвечающую особому слою, содержащему критическую точку (т. е. точку функции, дифференциал в которой равен нулю). Вершину графа Риба назовём концевой, если она является концом ровно одного ребра графа. Все остальные вершины назовём внутренними.

Для построения графа Риба существуют эффективные алгоритмы [9], [10], и данный метод позволяет достаточно точно описывать различные сложные модели (пример показан в работе Valerio Pascucci и др., рисунок 3.3).

Рисунок 3.3 – Пример использования графа Риба для представления объектов сложной формы

Но без методов соотношения таких графов данный способ классификации для решения данной задачи неприменим. И, как будет сказано далее, использование графов Риба для представления модели является избыточным, поэтому от этого метода решено было отказаться.

## Классификация, используя воксели

Исходя из описанной выше проблемы – необходимости соотносить объекты, можно подобрать другой метод представления объектов. Для этого объекты можно поделить на участки, каждый из которых далее будет сравниваться с исходным и будет определяться, к какому классу принадлежит объект с данной 3D моделью. Для этой задачи лучше всего подойдет вокселизация.

Воксель (объёмный (англ. volumetric) и пиксель (англ. pixel)) – элемент объёмного изображения, содержащий значение элемента растра в трёхмерном пространстве [11]. Воксели являются аналогами двумерных пикселей для трёхмерного пространства.

Воксельные модели часто используются для визуализации и анализа медицинской и научной информации, так как они позволяют получить упрощенной представление модели.

Для поставленной задачи обрабатываемые модели можно поделить на равное число вокселей, что позволить избавиться от проблемы с масштабом. Далее необходимо с помощью алгоритмов соотнести каждый воксель и по совокупному результату определить класс модели. Для поставленной задачи лучше всего подойдет нейросеть, так как входные данные заданы четко, известно точное количество объектов на входе (т. к. мы сами делим модель на воксели), и известен необходимый результат – вероятность принадлежности объекта к тому или иному классу деталей. Как пример можно воспользоваться нейросетью, представленной пользователем *EJ Shim* (см. рисунок 3.3, [12]). Данная нейросеть использует библиотеку моделей, по которым определяется предмет.

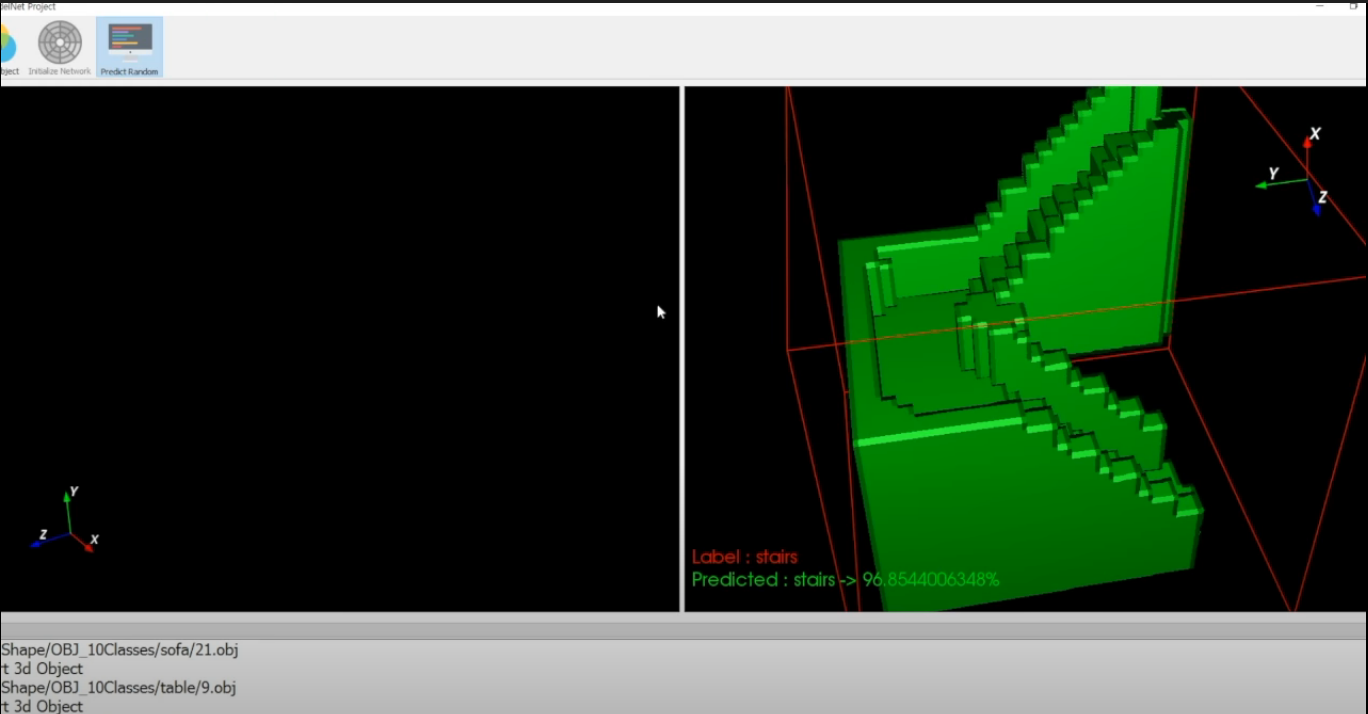


Рисунок 3.3 – Классификация моделей с использованием вокселей

Следующим шагом будет определение алгоритма создания воксель-моделей. Для этого также найдены существуют алгоритмы [13], более подробно создание воксель-моделей в этой работе рассматриваться не будет.

# Выбор нейросети и ее параметров

Далее необходим еще один шаг – нужно определить структуру нейросети. Для этого определим параметры и задачу нейросети, и исходя из них выберем лучший вариант реализации.

## Выбор входных и выходных данных

Первое, от чего отталкиваются при создании нейросети – выбор и определение количества входных данных, а также необходимый результат. Как уже было сказано ранее, в результате необходимо получить вероятность принадлежности объекта к некоторому определенному классу (к которому вероятность максимальна). Определить это необходимо исходя из некоторого набора вокселей, их количество будет неизменно, так как разделение будет производиться автоматически на нужное число по каждой из осей.

Проблема, которую можно встретить при разработке нейросети – необходимо ли определять ориентацию детали в пространстве, и без экспериментальной проверки узнать это невозможно.

## Архитектуры нейросетей

Существует несколько элементарных конфигураций нейросетей, каждая из который выполняет свои задачи.

1. **Нейронные сети прямого распространения** (feed forward neural networks, FF или FFNN) **и перцептроны** (perceptrons, P) прямолинейны, они передают информацию от входа к выходу (рисунок 4.1) [14].

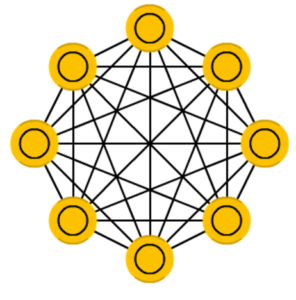
Рисунок 4.1 – *FFNN*

Нейронные сети часто описываются в виде модели из нескольких слоев, где каждый слой состоит из входных, скрытых или выходных нейронов. Нейроны одного слоя не связаны между собой, а соседние слои обычно полностью связаны. Самая простая нейронная сеть имеет две входных клетки и одну выходную, и может использоваться в качестве модели логических вентилей. FFNN обычно обучается по методу обратного распространения ошибки, в котором сеть получает множества входных и выходных данных. Этот процесс называется обучением с учителем, и он отличается от обучения без учителя тем, что во втором случае множество выходных данных сеть составляет самостоятельно. Вышеупомянутая ошибка является разницей между вводом и выводом. Если у сети есть достаточное количество скрытых нейронов, она теоретически способна смоделировать взаимодействие между входным и выходными данными. Практически такие сети используются редко, но их часто комбинируют с другими типами для получения новых.

1. Сети радиально-базисных функций (radial basis function, RBF) – схожи с **нейронными сетями прямого распространения**, но используют радиально-базисные функции как функции активации.

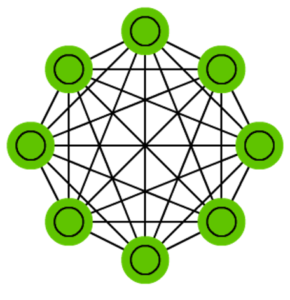
Радиальная функция – это любая вещественная функция, значение которой зависит только от расстояния до начала координат или от расстояния между некоторой другой точкой, называемой центром. В качестве нормы обычно выступает евклидово расстояние, хотя можно использовать и другие метрики.

1. **Нейронная сеть Хопфилда** (Hopfield network, HN) – это полносвязная нейронная сеть с симметричной матрицей связей. Во время получения входных данных каждый узел является входом, в процессе обучения он становится скрытым, а затем становится выходом (рисунок 4.2) [15].

Рисунок 4.2 – HN

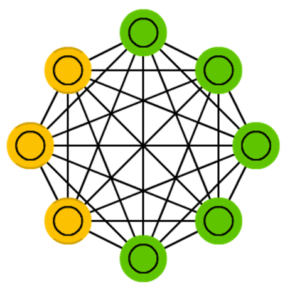
Сеть обучается так: значения нейронов устанавливаются в соответствии с желаемым шаблоном, после чего вычисляются веса, которые в дальнейшем не меняются. После того, как сеть обучилась на одном или нескольких шаблонах, она всегда будет сводиться к одному из них (но не всегда – к желаемому). Она стабилизируется в зависимости от общей «энергии» и «температуры» сети. У каждого нейрона есть свой порог активации, зависящий от температуры, при прохождении которого нейрон принимает одно из двух значений (обычно -1 или 1, иногда 0 или 1).  Такая сеть часто называется сетью с ассоциативной памятью; как человек, видя половину таблицы, может представить вторую половину таблицы, так и эта сеть, получая таблицу, наполовину зашумленную, восстанавливает её до полной.

1. **Цепи Маркова** (Markov chains, MC или discrete time Markov Chains, DTMC) – это предшественники машин Больцмана (*BM*) и сетей Хопфилда (*HN*) (рисунок 4.3) [16]. Их смысл можно объяснить так: каковы мои шансы попасть в один из следующих узлов, если я нахожусь в данном?

Рисунок 4.3 – DTMC

Каждое следующее состояние зависит только от предыдущего. Хотя на самом деле цепи Маркова не являются НС, они весьма похожи. Также цепи Маркова не обязательно полносвязны.

1. **Машина Больцмана** (Boltzmann machine, BM) очень похожа на сеть Хопфилда, но в ней некоторые нейроны помечены как входные, а некоторые – как скрытые (рисунок 4.4). Входные нейроны в дальнейшем становятся выходными. Машина Больцмана – это стохастическая сеть. Обучение проходит по методу обратного распространения ошибки или по алгоритму сравнительной расходимости.  В целом процесс обучения очень похож на таковой у сети Хопфилда [17].

Рисунок 4.4 – BM 

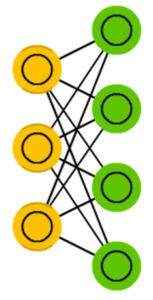
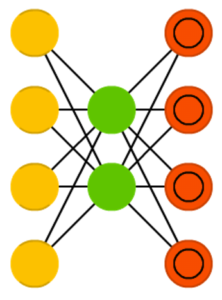
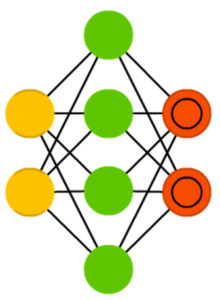
1. **Ограниченная машина Больцмана** (restricted Boltzmann machine, RBM) удивительно похожа на машину Больцмана и, следовательно, на сеть Хопфилда. Единственной разницей является её ограниченность. В ней нейроны одного типа не связаны между собой (рисунок 4.5). Ограниченную машину Больцмана можно обучать как FFNN, но с одним нюансом: вместо прямой передачи данных и обратного распространения ошибки нужно передавать данные сперва в прямом направлении, затем в обратном. После этого проходит обучение по методу прямого и обратного распространения ошибки.

Рисунок 4.5 – RBM

1. **Автокодировщик** (autoencoder, AE) чем-то похож на *FFNN*, так как это скорее другой способ использования *FFNN*, нежели фундаментально другая архитектура. Основной идеей является автоматическое кодирование (в смысле сжатия, не шифрования) информации [18]. Сама сеть по форме напоминает песочные часы, в ней скрытые слои меньше входного и выходного, причём она симметрична (рисунок 4.6). Сеть можно обучить методом обратного распространения ошибки, подавая входные данные и задавая ошибку равной разнице между входом и выходом.

Рисунок 4.6 – AE

1. **Разреженный автокодировщик** (sparse autoencoder, SAE) – в каком-то смысле противоположность обычного. Вместо того, чтобы обучать сеть отображать информацию в меньшем «объёме» узлов, мы увеличиваем их количество. Вместо того, чтобы сужаться к центру, сеть там раздувается (рисунок 4.7) [19]. Сети такого типа полезны для работы с большим количеством мелких свойств набора данных. Если обучать сеть как обычный автокодировщик, ничего полезного не выйдет. Поэтому кроме входных данных подаётся ещё и специальный фильтр разреженности, который пропускает только определённые ошибки.

Рисунок 4.7 – SAE

1. **Вариационные автокодировщики**(variational autoencoder, VAE) обладают схожей с *AE* архитектурой, но обучают их иному: приближению вероятностного распределения входных образцов. В этом они берут начало от машин Больцмана. Тем не менее, они опираются на байесовскую математику, когда речь идёт о вероятностных выводах и независимости, которые интуитивно понятны, но сложны в реализации [20]. Если обобщить, то можно сказать, что эта сеть принимает в расчёт влияния нейронов. Если что-то одно происходит в одном месте, а что-то другое – в другом, то эти события не обязательно связаны, и это должно учитываться (рисунок 4.8).

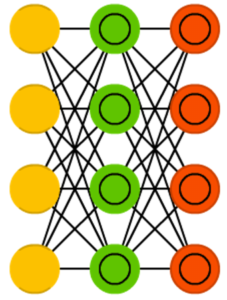
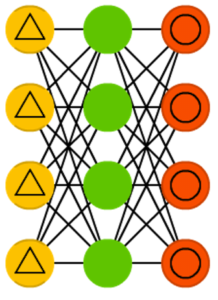


Рисунок 4.8 – *V*AE

1. Шумоподавляющие автокодировщики*(denoising autoencoder, DAE)* – это *AE*, в которые входные данные подаются в зашумленном состоянии. Ошибку мы вычисляем так же, и выходные данные сравниваются с зашумленными. Благодаря этому сеть учится обращать внимание на более широкие свойства, поскольку маленькие могут изменяться вместе с шумом [21].

 Рисунок 4.9 – *DAE*

1. **Сеть типа «deep belief»** (deep belief networks, DBN) – это название, которое получил тип архитектуры, в которой сеть состоит из нескольких соединённых RBM или VAE. Такие сети обучаются поблочно, причём каждому блоку требуется лишь уметь закодировать предыдущий (рисунок 4.10) [22]. Такая техника называется «жадным обучением», которая заключается в выборе локальных оптимальных решений, не гарантирующих оптимальный конечный результат. Также сеть можно обучить (методом обратного распространения ошибки) отображать данные в виде вероятностной модели. Если использовать обучение без учителя, стабилизированную модель можно использовать для генерации новых данных.

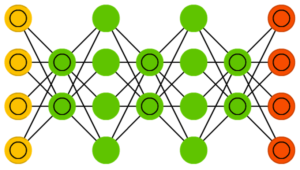


Рисунок 4.10 – DBN

1. **Свёрточные нейронные сети** (convolutional neural networks, CNN)**и глубинные свёрточные нейронные сети** (deep convolutional neural networks, DCNN) сильно отличаются от других видов сетей. Обычно они используются для обработки изображений, реже для аудио [23]. Типичным способом применения *CNN* является классификация изображений: если на изображении есть кошка, сеть выдаст «кошка», если есть собака – «собака». Такие сети обычно используют «сканер», не парсящий все данные за один раз. Например, если у вас есть изображение 200×200, вы не будете сразу обрабатывать все 40 тысяч пикселей. Вместо это сеть считает квадрат размера 20 x 20 (обычно из левого верхнего угла), затем сдвинется на 1 пиксель и считает новый квадрат, и т. д. Эти входные данные затем передаются через свёрточные слои, в которых не все узлы соединены между собой. Эти слои имеют свойство сжиматься с глубиной, причём часто используются степени двойки: 32, 16, 8, 4, 2, 1. На практике к концу *CNN* прикрепляют *FFNN* для дальнейшей обработки данных (рисунок 4.11). Такие сети называются глубинными (*DCNN*).

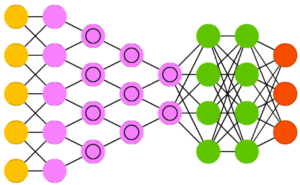


Рисунок 4.11 – *CNN* + *FFNN*

1. **Развёртывающие нейронные сети** (deconvolutional networks, DN), также называемые обратными графическими сетями, являются обратным к свёрточным нейронным сетям (рисунок 4.12). Представьте, что вы передаёте сети слово «кошка», а она генерирует картинки с кошками, похожие на реальные изображения котов. *DNN* тоже можно объединять с *FFNN*. Стоит заметить, что в большинстве случаев сети передаётся не строка, а какой бинарный вектор: например, <0, 1> – это кошка, <1, 0> – собака, а <1, 1> – и кошка, и собака.

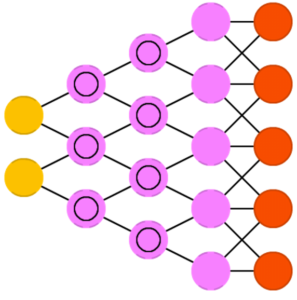


Рисунок 4.12 – DN

## Вывод по разделу

Достаточно очевидно, что для поставленной задачи лучше всего выбрать CNN, или сверточную нейросеть. Действительно, задача определения и классификации изображения очень схожа с задачей классификации 3D моделей, тем более, когда входные данные схожи – вместо пикселей используются воксели.

# ЗАКЛЮЧЕНИЕ

Классификация изделий облегчает повторное использование деталей в дальнейшем, а также позволяет систематизировать изделия, которые в целом являются слабоизмененными вариациями одной модели. Именно поэтому задача классификации изделий является востребованной и достаточно актуальной.

Стоит отметить основные выведенные аспекты алгоритма классификации изделий. Во-первых, был определен лучший формат, в котором стоит подавать на вход модели, является obj, и, как альтернатива, fbx. Эти форматы предоставляют минимально необходимый набор данных, а именно необходимое в дальнейшем положения вершин и граней, и также удобный доступ к этим данным посредством ключей, что облегчит дальнейшую обработку. Также эти форматы являются универсальными, что позволит при необходимости преобразовывать другие форматы и получать универсальный алгоритм для изделий, созданных в различных САПР.

Также были определены классы изделий, с классификации которых достаточно начать тестировать возможный алгоритм. Это классы 71-75 по классификатору, а именно различные детали вращения, трубы и корпуса. У данных классов есть подклассы, классификация которые является дополнительной задачей, точность результатов которой еще предстоит определить.

В качестве алгоритма классификации, как наиболее удобный и быстрый, был выбран следующий алгоритм: сначала необходимо разделить модель на фиксированное количество вокселей, что позволит четко обозначить входные данные, которые далее подаются в CN-нейросеть, которая на основе предоставленных базисных, или шаблонных, объектов определит вероятность принадлежности данного изделия к конкретному классу.

Так как в данной работе предоставлены только исследовательские наблюдения, в качестве продолжения можно выполнить непосредственную разработку и тестирование алгоритма для различных основных классов, и, как было описано ранее, выбор возможных для определения подклассов, для которых точность работы будет больше необходимой.

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Разбор форматов: 3d-модели изнутри // [habr.com]. URL: habr.com/ru/post/263009/ (дата обращения: 15.11.2020).
2. An overview of 3D data content, file formats and viewers // [researchgate.net]. URL: researchgate.net/publication/228976876\_An\_overview\_of\_3D\_data\_content\_file\_formats\_and\_viewers (дата обращения: 29.11.2020).
3. Исследование 3d форматов хранения данных в интеллектуальных системах виртуальной реальности // [Международный журнал экспериментального образования]. URL: expeducation.ru/ru/article/view?id=7186 (дата обращения: 02.12.2020).
4. The Most Common 3D File Formats // [all3dp.com]. / URL: all3dp.com/3d-file-format-3d-files-3d-printer-3d-cad-vrml-stl-obj/ (дата обращения: 02.12.2020).
5. Классификатор ЕСКД // [КлассИнформ]. URL: classinform.ru/ok-eskd/kod.html (дата обращения: 03.12.2020).
6. Единая система конструкторской документации (ЕСКД). Общие положения [Текст]: ГОСТ 2.001-2013. Введ. 2013-11-22.  М.: Стандартинформ, 2011.
7. Единая система конструкторской документации (ЕСКД). Обозначение изделий и конструкторских документов [Текст]:   
   ГОСТ 2.201-80. Введ. 1980-03-21. М.: Стандартинформ, 2011.
8. S. Biasotti, Giorgi D., Spagnuolo M., Falcidieno B. Reeb graphs for shape analysis and applications // DOI 10.1016/j.tcs.2007.10.018 28.02.2008.
9. Doraiswamy H., Natarajan V. Efficient algorithms for computing Reeb graphs // DOI 10.1016/j.comgeo.2008.12.003 07.08.2009.
10. Pascucci V., Scorzelli G., Bremer P., Mascarenhas A. Robust On-line Computation of Reeb Graphs: Simplicity and Speed // DOI 10.1145/1276377.1276449 09.2007.
11. Cohen-Or D., Kaufman A. Fundamentals of Surface Voxelization // DOI 10.1006/gmip.1995.1039. 11.1995.
12. EJModelNet // [github.com]. URL: https://github.com/EJShim/EJModelNet (дата обращения: 15.12.2020).
13. Voxelizer and SDF // [github.com]. URL: https://github.com/rFalque/voxelization\_and\_sdf (дата обращения: 25.12.2020).
14. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain // DOI [10.1037/h0042519](https://psycnet.apa.org/doi/10.1037/h0042519). 1958.
15. Hopfield J. Neural networks and physical systems with emergent collective computational abilities // DOI [10.1073/pnas.79.8.2554](https://doi.org/10.1073/pnas.79.8.2554). 01.10.1982.
16. Hayes B. First Links in the Markov Chain // American Scientist. –2013. –№101. –С. 92.
17. Michael I., Terrence J. Learning and Relearning in Boltzmann Machines // Graphical Models: Foundations of Neural Computation. –2001.
18. Bourlard, H., Y. Kamp Auto-association by multilayer perceptrons and singular value decomposition // Biological Cybernetics. –2004. –№59. –С. 291-294.
19. M. Ranzato Efficient Learning of Sparse Representations with an Energy-Based Model // NIPS. –2006. –№19.
20. Diederik P., M. Welling Auto-Encoding Variational Bayes // CoRR. –2014.
21. P. Vinsent Extracting and composing robust features with denoising autoencoders // ICML. –2006. –№08.
22. Y. Bengio Greedy Layer-Wise Training of Deep Networks // NIPS. –2006.
23. LeCun, Y. Gradient-based learning applied to document recognition // Proceedings of the IEEE. –1998. –№86.