

МФТИ

**Классификация 3D-объектов на основе облака точек
с использованием архитектуры PointNet и датасета
ModelNet**

Филак Александр

Долгопрудный
2025

1 Отчёт по практическому занятию №5

1.1. Введение

Целью данной работы было практическое знакомство с задачей классификации трёхмерных объектов, представленных в виде неупорядоченных облаков точек. Для этого была использована архитектура нейронной сети PointNet, предложенная в 2017 году и ставшая классическим подходом для работы с точечными данными. Эксперименты проводились на датасете ModelNet10, содержащем 10 категорий бытовых объектов. Работа включала в себя подготовку данных, реализацию модели, её обучение, оценку и визуализацию результатов.

1.2. Подготовка данных и настройка модели

В качестве исходных данных использовался датасет ModelNet10. Каждый 3D-меш конвертировался в облако точек с помощью библиотеки Open3D. Для унификации входа из каждого облака случайным образом выбиралось 1024 точки, а их координаты нормализовались в диапазон $[-1, 1]$. Полученные данные были сохранены для ускорения последующей загрузки.

Основные параметры эксперимента:

- **Датасет:** ModelNet10 (10 классов: bathtub, bed, chair, desk, dresser, monitor, night-stand, sofa, table, toilet).
- **Количество точек на объект:** 1024.
- **Архитектура:** PointNet с включённым модулем Feature Transform (T-Net) для повышения инвариантности к преобразованиям.
- **Разделение данных:** 80% (320 объектов) на обучение, 20% (80 объектов) на тестирование.
- **Размер батча:** 32.
- **Оптимизатор:** Adam с learning rate = 0.001.
- **Функция потерь:** CrossEntropyLoss.
- **Количество эпох:** 50.

1.3. Ход обучения и результаты

Обучение модели проводилось в течение 50 эпох. Наблюдалась стабильная сходимость: функция потерь на тренировочном наборе уменьшилась с 2.46 до 0.45, а точность возросла с 19% до 85%. На тестовой выборке наилучшая точность (**accuracy**) составила **80%**, достигнутая на 42-й эпохе.

Ключевые моменты процесса обучения:

- Первые признаки обучения появились уже на 5-й эпохе (точность 30% на тесте).
- Значительный скачок качества произошёл на 22-й эпохе (точность ~64%).
- Модель достигла плато качества около 75-80% после 35-й эпохи, что указывает на достаточную сложность задачи для выбранного объёма данных и архитектуры.

График изменения функции потерь и точности в процессе обучения представлен ниже.



Рис. 1: Динамика функции потерь и точности в ходе обучения

1.4. Анализ качества модели

Для детальной оценки качества модели на тестовой выборке была построена матрица ошибок (confusion matrix).

Анализ матрицы позволяет сделать следующие выводы:

1. **Лучше всего** классифицируются объекты классов **chair** и **table**. Это может быть связано с их хорошо выраженной и уникальной геометрической структурой.
2. **Наибольшие трудности** возникли с классами **night-stand** и **dresser**. Вероятно, это обусловлено их визуальной и геометрической схожестью (оба являются видами мебели для хранения с горизонтальными поверхностями и ящиками). Модель часто путает их между собой.
3. Класс **bed** также иногда ошибочно классифицируется как **sofa**, что логично из-за схожести формы в упрощённом представлении облаком точек.

Визуализация нескольких примеров с предсказаниями модели подтверждает эти наблюдения.

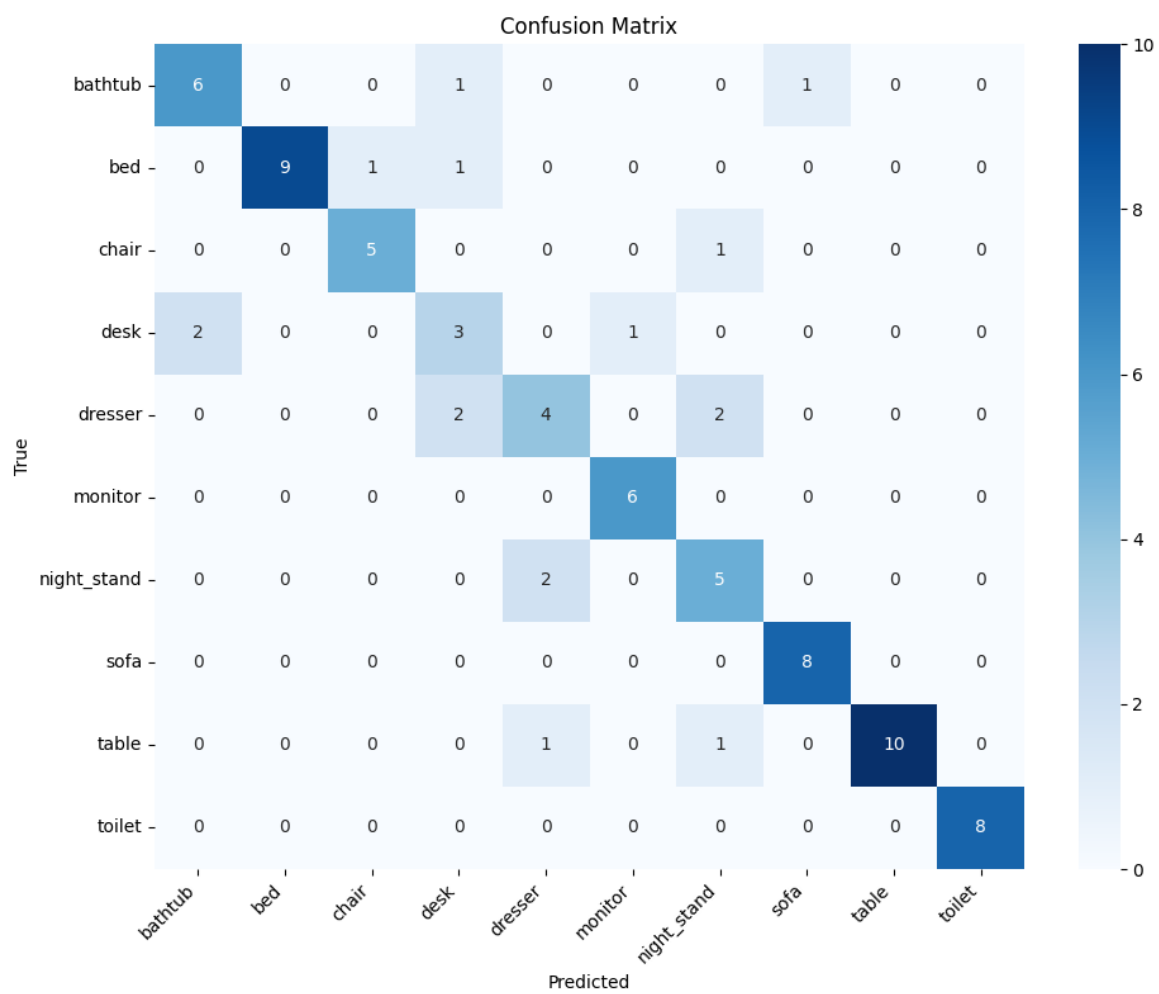


Рис. 2: Матрица ошибок для 10 классов ModelNet10

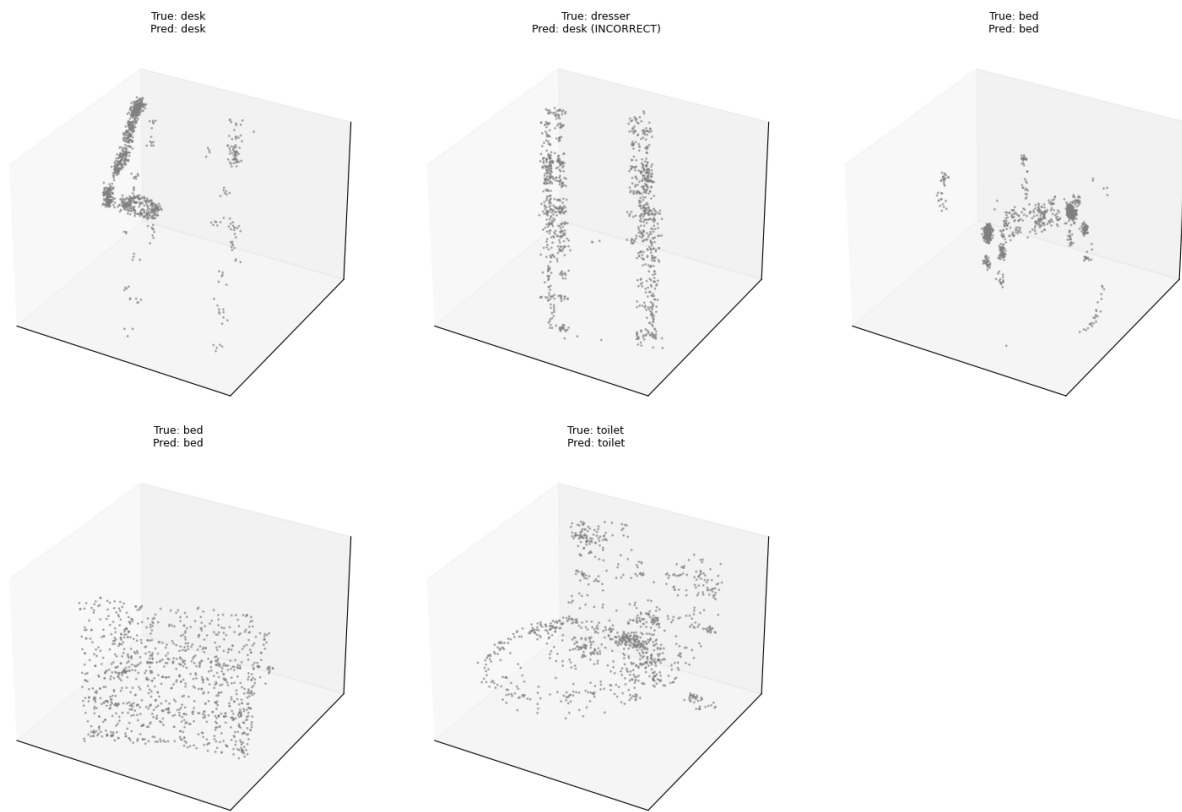


Рис. 3: Примеры классификации объектов моделью PointNet

1.5. Ответы на контрольные вопросы

1. **Какие категории объектов классифицируются хуже всего и почему?** Хуже всего классифицируются **night_{stand}** (тумбочка) и **dresser** (комод). Основная причина — их высокая семантическая и геометрическая схожесть в контексте облака точек из 1024 вершин. Оба объекта представляют собой параллелепипеды с выдвижными элементами, что затрудняет их различение для модели без более детального контекста или текстур.
2. **Влияет ли количество точек (512, 1024, 2048) на качество модели?** В рамках данной работы использовалось фиксированное количество точек (1024). Однако, основываясь на теории и результатах оригинальной статьи PointNet, можно утверждать, что увеличение количества точек (до определённого предела) обычно улучшает качество классификации, так как позволяет точнее описать геометрию объекта. Уменьшение же количества точек (например, до 512) может привести к потере важных деталей и снижению ассурасу.
3. **Какие преимущества PointNet по сравнению с CNN для 3D данных?** Ключевое преимущество PointNet — эффективная работа с неупорядоченным и инвариантным к перестановкам представлением данных (облаками точек). В отличие от CNN, которые требуют предварительного преобразования 3D-объектов в регулярные структуры (например, воксельные сетки или многовидовые проекции), что ведёт к потере информации

или значительному увеличению вычислительной сложности, PointNet обрабатывает "сырые" точки, сохраняя исходную геометрию. Модули T-Net обеспечивают пространственную инвариантность, а операция max-pooling агрегирует глобальные признаки.

1.6. Вывод

В ходе работы была успешно реализована и обучена модель PointNet для классификации 3D-объектов датасета ModelNet10. Достигнутая итоговая точность на тестовой выборке составила 80%, что является хорошим результатом для задачи классификации 10 классов. Были отработаны навыки подготовки облаков точек, реализации нейросетевой архитектуры, проведения цикла обучения и комплексной оценки результатов с помощью метрик и визуализации. Эксперимент подтвердил эффективность подхода PointNet для работы с неупорядоченными точечными данными и выявил классы объектов, требующие более детального подхода к различению.

1.7. Использованные источники

1. Qi et al., "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation", CVPR 2017.
2. Официальный сайт датасета ModelNet: <https://modelnet.cs.princeton.edu>
3. Официальная реализация PointNet: <https://github.com/charlesq34/pointnet>