Mélytanulás Nagy Házi Feladat - Beadandó Dokumentáció
Point Cloud Segmentation with graph neural networks

Készítette:

Lovácsi Kristóf - N3EEWB

Dremák Gergely - KSHSLY

https://github.com/kristof39612/pointcloud_v3

https://github.com/WingSMC/point_cloud_client

Mélytanulás Nagy Házi Feladat - Beadandó Dokumentáció
Point Cloud Segmentation with graph neural networks 0

Bevezető

Előzetes feladatkiírás:

The goal of this project is to delve into point cloud segmentation using Graph Neural Networks (GNNs). You have to work with the ShapeNet dataset, which contains multiple object categories (from airplanes to wine bottles). By choosing this project, you'll gain expertise in deep learning and spatial data analysis.

Miért fontos a választott probléma megoldása?

A pontfelhők szegmentációja alapvető jelentőségű számos modern technológiai alkalmazásban, mint például az autonóm járművek érzékelésében, a 3D modellezésben, az orvosi képalkotásban és a virtuális valóságban. Ezekben a kontextusokban a pontfelhők megfelelő szegmentálása lehetővé teszi az egyes objektumok vagy elemek pontos meghatározását és elkülönítését a környezettől.

A hagyományos gépi tanulási technikák korlátai miatt a Graph Neural Network (GNN) alapú megoldások egyre nagyobb népszerűségnek örvendenek. Ezek a megközelítések hatékonyan kezelik a pontfelhők nem-euklideszi struktúráját, valamint a bonyolult térbeli mintázatok azonosítását permutációs invariancia mellett. A ShapeNet adatbázis felhasználása különösen fontos volt a probléma megoldásában, mivel változatos adathalmazokat kínál, amelyek lehetővé teszik a modellek általánosítási képességeinek értékelését.

Ez a projekt nem csupán a pontfelhők feldolgozásában nyújt tapasztalatokat, hanem a GNN-ek hatékony alkalmazásában is, amelyek az adattudomány és a gépi tanulás más területein is hasznosíthatók.

Megvalósított munka

A feladatkiírásból kiindulva elsőként irodalomkutatást végeztünk a témában és a PointNet neurális hálózatban. Előkészítettük a PointNet PyTorch-ban is működő modelljét, amit az ajánlott Shapenet adatkészlettel együtt be lehet tanítani és klasszifikációs, valamint szegmentációs feladatokat lehet segítségével végrehajtani pontfelhőkön. Feladatunkban továbbá, hogy a végfelhasználó számára is érthetőbb, kézenfoghatóbb lehessen a megoldásunk, úgy létrehoztunk egy frontend szolgáltatást is, melyben a felhasználó a már betanított modell segítségével tudja ezeket a feladatokat végrehajtani egy grafikus felületen.

Kapcsolódó munkák

Típus	Cím	Szerzők	Link
Cikk	PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation	Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas	https://arxiv.org/abs/1612.00593
Cikk	PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space	Charles R. Qi, Li Yi, Hao Su, Leonidas J. Guibas	https://arxiv.org/abs/1706.02413
Cikk	Dynamic Graph CNN for Learning on Point Clouds	Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, Justin M. Solomon	https://arxiv.org/abs/1801.07829
Cikk	PointCNN: Convolution On X-Transformed Points	Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, Baoquan Chen	https://arxiv.org/abs/1801.07791
Repository	PointNet2 Repository	Charles R. Qi	https://github.com/charlesq34/pointnet2
Adathalmaz	Shapenet	Stanford	https://shapenet.org/

Tanítás

Adatok előkészítése

A Shapenet Partanno V0 adatkészletet használtuk, melyben 16 különböző, mindennapi objektumról találhatóak felcímkézett adatok, mint például lámpa, vagy laptop. Ez az adatkészlet minden osztályhoz kapcsolódóan tartalmaz több “.pts” kiterjesztésű fájlt, mely magát a pontfelhőt tartalmazza az adott objektumról, valamint a szegmentálási feladatokhoz kapcsolódóan tartalmaz további “.seg” kiterjesztésű fájlokat, mely az objektum pontfelhőjében található pontokhoz hozzárendeli az elvárt szegmentálást. Például egy repülőnél akár a szárnyak és a hajtóművek pontjait felcímkézi egy plusz adattal (szegmentációs osztály egész számú azonosítója), hogy szétválasztásra tudjon kerülni az adott része az objektumnak szegmentációnál.

- **16 különböző osztály**
- Tartalmazza a pontfelhő adatokat (.pts fájlok) és a hozzájuk tartozó szegmentációs címkéket (.seg fájlok).
- A train-test-validation split az adathalmazban előre definiált, amelyet változtatás nélkül használtunk:
    - Train: 75%
    - Validation: 15%
    - Test: 10%

Modell tanítása

- **Hyperparaméterek**:
    - Epochs: 25
    - Learning rate: 0.001
    - Batch size: 32
- **Hardver**: NVIDIA RTX 3090 GPU.
- **Eredmény**: ~98.6% train accuracy és ~97.5% validation accuracy és magas általánosítási teljesítmény.
- **Docker-támogatás**: A modell tanításához Dockerfile készült, amely segíti a gyors környezettelepítést.

Kiértékelés

Mérőszámokat a teljesítmény értékelésére:

Classification Task

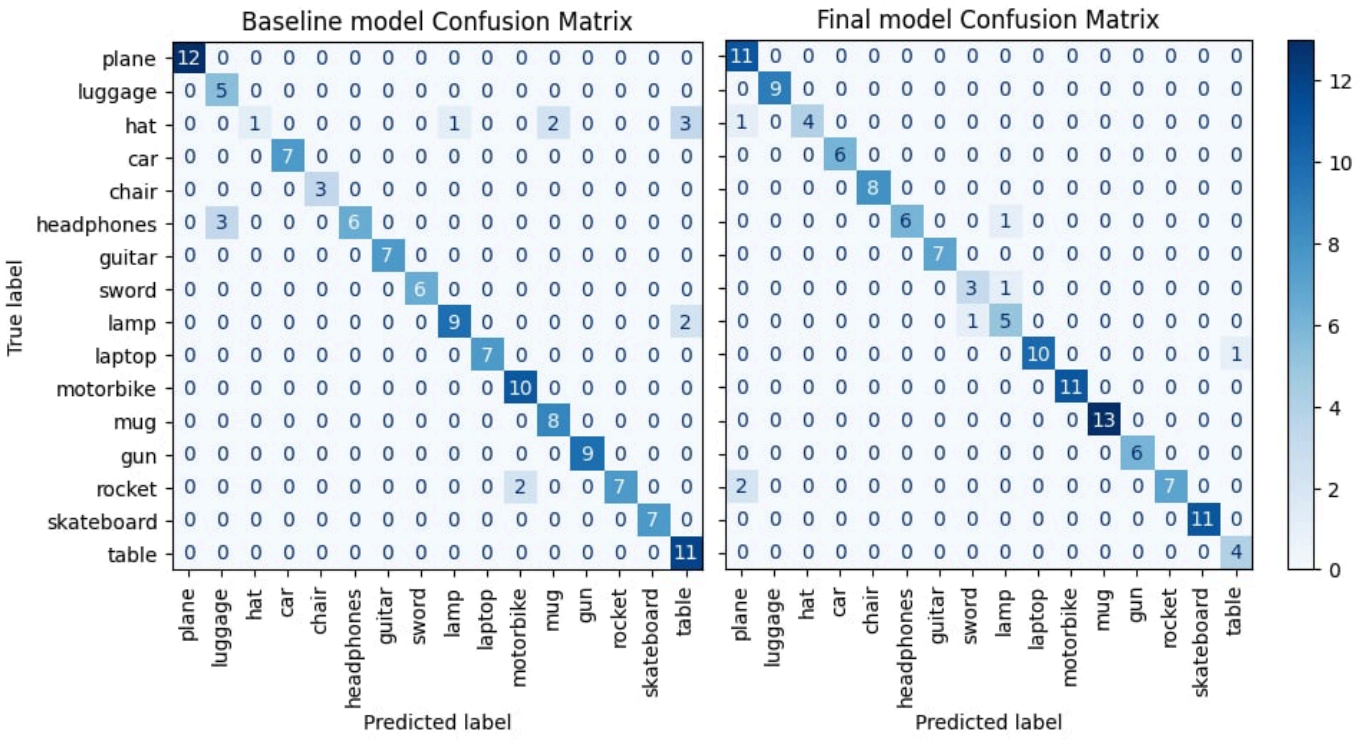
1. **Classification Accuracy**: A helyesen besorolt pontfelhők aránya az összeshez képest.
2. **Precision**: Az összes pozitív osztályozás közül hány volt valóban helyes.
3. **Recall**: Az összes valódi pozitív közül hányat talált meg a modell.
4. **F1 Score**: A Precision és Recall harmonikus átlaga, amely kiegyensúlyozott mérőszám a pontosság és a lefedettség között.

Segmentation Task

1. **Train Loss és Validation Loss**: A tanítási és validációs adathalmazokon számított veszteségértékek, amelyek a modell konvergenciáját jelzik.
2. **Train Accuracy és Validation Accuracy**: Az egyes pontok helyes szegmentálásának aránya a tanító és validációs adathalmazon.
3. **Segmentation Accuracy**: Az összes pont helyes szegmentálásának aránya az egész adathalmazon.
4. **Mean IoU (Intersection over Union)**: Az átlagos IoU érték az összes kategóriára, amely az átfedés és egyesítés arányát méri minden szegmentációs osztályra.
5. **Dice Coefficient**: Az IoU-hoz hasonló átfedési mérőszám, amely az osztályok pontosságát méri, kiemelve a hibák minimalizálását.

Az eredmények részletezése

Classification	BASELINE	FINAL
Classification Accuracy	0.8984	0.9766
Precision	0.9244	0.9768
Recall	0.8984	0.9766
F1 Score	0.8849	0.9758



Segmentation	BASELINE	FINAL
Train loss	0.2430	0.2298
Val loss	0.3158	0.3251
Train accuracy	0.9121	0.9168
Val accuracy	0.8956	0.8964
Segmentation Accuracy	0.8679	0.8874
Mean IoU	0.7816	0.8077
Dice Coefficient	0.8665	0.8861

API és Front-end

A projekt során létrehozott **Fastify API** és **Svelte alapú frontend** lehetővé teszi a pontfelhők osztályozási és szegmentálási feladatainak egyszerű elérését. Az API két (hivatalos) endpointot biztosít:

- **/octetclassify**: Bináris adatként (octet-stream) küldött pontfelhőt fogad, és az osztályazonosítót adja vissza JSON formátumban.
- **/octetsegment**: Bináris adatként küldött pontfelhőt fogad, és JSON formátumban adja vissza a szegmentációhoz tartozó színezési adatokat.

A frontend egy **THREE.js** alapú 3D vizualizációs felületet kínál, amely saját shaderekkel jeleníti meg a pontfelhőket. A frontend repójában CI/CD pipeline működik, amely minden új cimkézésnél automatikusan készít egy új Docker image-et, és feltölti azt a Docker Hub-ra.

Az API és a frontend futtatásához részletes útmutató található a **GitHub repókban**, amely bemutatja, hogyan lehet fejlesztői környezetben elindítani az alkalmazást, vagy Docker segítségével futtatni azt. A dokumentációban külön kitérünk a különböző futtatási módokra és az API használatára.

Konklúzió

A projekt feladat zárásaként elmondható, hogy a feladatkiírásban leírt előzetes követelményeket teljesítettük, továbbá a felhasználói felület biztosításával színesebbé és egyedivé tettük a megoldásunkat. Megtanultuk a pontfelhők osztályozásához és szegmentálásához szükséges alapvető fogalmakat és elmélyülhettünk a témában a félév alatt. Továbbfejlesztési lehetőség lehet még, hogy a modell a Shapenet újabb verziójára is feltanítható lehessen, de az újabb verziójú adatkészletekben a formátum különbözik a V0-ban megadott (“.pts”,”.seg”) formátumtól, tehát a kódbázist át kellene alakítani, hogy ezeket befogadja. A modell teljesítményének javítása érdekében kísérletezést érdemelhetnek más, korszerűbb hálózati architektúrák is.

Appendix

LLM-et használtunk

- Kód kommentálásra (GitHub copilot (GPT-4o és o1 modell))
- Felhasznált könyvtárak és keretrendszerek dokumentációja helyett
- Ennek a dokumentumnak az átnézésére