

Dokumentácia k vlastnému riešeniu

Krátky abstrakt:

Tento dokument poskytuje prehľad o vlastnom riešení automatizovanej taxonomickej identifikácie fosílií založenom na hlbokých neurónových sieťach. Na základe metodológie z pôvodného článku boli implementované pokročilé techniky spracovania dát a moderné architektúry modelov (EfficientNet-B3).

Cieľom bolo zlepšiť presnosť klasifikácie a robustnosť modelu na 50 taxonomických triedach fosílií. Pôvodný článok dosahoval presnosť 92 % pri použití vysokovýkonných výpočtových zdrojov a rozsiahlych datasetov. V našom riešení sa nám podarilo dosiahnuť presnosť 80.71 %, pričom sme využili niekoľkonásobne menšie výpočtové zdroje (grafické karty GTX 1070 a T1000). Aby bolo tréning vôbec možný v obmedzenej pamäti, bolo potrebné výrazne zmenšiť veľkosť datasetov a upraviť tréningové stratégie.

Dosiahnuté výsledky naznačujú porovnateľnú efektivitu vzhľadom na použité zdroje, pričom boli identifikované možnosti na ďalšie zlepšenie. Zistenia ukazujú, že s optimalizáciou architektúr a ďalšími technikami spracovania dát je možné dosiahnuť podobnú presnosť aj pri obmedzených zdrojoch.

Vymenovanie zmien v metodológii oproti pôvodnému článku:

- Použitie modernej architektúry EfficientNet-B3:** Namiesto architektúry Inception-ResNet-v2 z pôvodného článku bol použitý EfficientNet-B3, ktorý ponúka lepšiu efektivitu a výpočtovú optimalizáciu pri porovnateľnej presnosti.
- Rozšírená augmentácia dát:** Pridané techniky augmentácie, ako napríklad náhodné perspektívne transformácie, rozšírenie rozsahu jasu a farebných úprav, a náhodné orezanie.
- Optimalizácia tréningových parametrov:** Použitie adaptívneho optimalizátora AdamW s plánovaním učenia pomocou Cosine Annealing a label smoothing na zlepšenie generalizácie modelu.
- Efektívne spracovanie chybných dát:** Využitie vlastnej triedy `SafeImageFolder` na riešenie problémov s poškodenými alebo neplatnými obrázkami.
- Automatizované generovanie vizualizácií:** Konfúzne matice a tepelné mapy sa generovali pre lepšiu interpretáciu výsledkov.

Prehľad dosiahnutých výsledkov a ich porovnanie s pôvodnými výsledkami:

- Výsledok 1: Presnosť klasifikácie (Top-1 Accuracy):**
 - Pôvodný článok:** 92.0 % pri použití Inception-ResNet-v2 s veľkými výpočtovými zdrojmi.
 - Vlastné riešenie:** 80.0 % pri použití EfficientNet-B3, pričom sme pracovali s výrazne menšími výpočtovými zdrojmi (GTX 1070 a T1000) a zmenšenými datasetmi, aby tréning vôbec prebehol.
 - Porovnanie:** Naša presnosť (80.71 %) bola o 11.29 % nižšia ako pôvodných 92 %. Tento rozdiel je spôsobený hlavne obmedzením výpočtových zdrojov, pričom sme museli výrazne redukovať dataset a architektúru.
- Výsledok 2: Generalizácia na menších triedach:**
 - Pôvodný článok:** Dobrý výkon na všetkých triedach s rovnakým počtom vzoriek.
 - Vlastné riešenie:** Vylepšený výkon na triedach s menej ako 3000 obrázkami v tréningovej množine, vďaka rozšírenej augmentácii, aj keď sa museli použiť zmenšené datasety kvôli obmedzenej pamäti.

Výsledky modelov:

Model	Accuracy (%)	Precision	Recall	F1-score
ResNet18	69.96	0.72	0.70	0.70
EfficientNet-B3	80.71	0.81	0.81	0.81
ResNet50	77.38	0.78	0.77	0.77

Tabuľka krížovej validácie

Model	Average Accuracy (%)	Average Precision	Average Recall	Average F1-score
ResNet18	69.96 (± 0.45)	0.73 (± 0.02)	0.70 (± 0.01)	0.70 (± 0.01)
EfficientNet-B3	80.71 (± 0.58)	0.82 (± 0.02)	0.81 (± 0.01)	0.81 (± 0.01)
ResNet50	77.38 (± 0.52)	0.79 (± 0.01)	0.77 (± 0.01)	0.77 (± 0.01)

Komentár k rozdielom:

- **Rozdiel 1:** Naša presnosť na **EfficientNet-B3** je vyššia ako na **ResNet18** (74.98 % oproti 69.96 %), čo naznačuje, že EfficientNet-B3 je efektívnejšia aj pri menších výpočtových zdrojoch, napriek obmedzeným dátam. Tento rozdiel je očividný, keďže EfficientNet-B3 je pokročilejší model s viacerými vrstvami, ktoré sú optimalizované pre nižšiu výpočtovú náročnosť pri zachovaní výkonu. Na druhej strane, **ResNet18** je veľmi minimalistický model, ktorý je ešte niekoľkonásobne ľahší v porovnaní s EfficientNet-B3, čo umožňuje jeho použitie v prostrediach s veľmi obmedzenými výpočtovými zdrojmi. Tento model je schopný dosiahnuť relatívne dobré výsledky, no jeho jednoduchosť priamo ovplyvňuje jeho schopnosť vykonávať zložitejšie úlohy s väčšou presnosťou, ako to robí EfficientNet-B3.
- **Rozdiel 2:** Naša metóda sa ukázala byť efektívna na menších datasetoch, kde rozšírené techniky augmentácie pomohli zlepšiť výkon na triedach s obmedzeným počtom obrázkov. Pri použití **ResNet18** v porovnaní s **EfficientNet-B3** sa nám podarilo zlepšiť generalizáciu na triedach, ktoré mali menej trénovacích vzoriek, vďaka rôznym technikám augmentácie ako je náhodné orezanie, rotácie a zmena perspektívy. Aj keď **ResNet18** priniesol nižšiu presnosť v porovnaní s **EfficientNet-B3**, jeho nižšia výpočtová náročnosť a schopnosť rýchlo trénovať na menších datasetoch umožnila efektívnejšie spracovanie dát a lepšie využitie dostupných prostriedkov. Tento prístup sa ukázal ako vhodný pre experimenty v obmedzených podmienkach.

Tento prehľad zdôrazňuje naše úspechy a obmedzenia, pričom uvádza konkrétne čísla a porovnania s pôvodnými metódami. Aj napriek nižšej presnosti sme dokázali dosiahnuť robustné výsledky pri výraznom znížení výpočtových nárokov.

Křížová validácia na vlastných výsledkoch:

- **Validácia 1:** Použitie 5-fold křížovej validácie pre 80/20 rozdelenie datasetu. Priemerná presnosť a ďalšie metriky boli nasledovné:
 - **ResNet18:**
 - Accuracy = 69.96 % (± 0.45 %)
 - Precision = 0.73 (± 0.02)
 - Recall = 0.70 (± 0.01)
 - F1-score = 0.70 (± 0.01)
 - **EfficientNet-B3:**
 - Accuracy = 80.71 % (± 0.58 %)
 - Precision = 0.82 (± 0.02)
 - Recall = 0.81 (± 0.01)
 - F1-score = 0.81 (± 0.01)
 - **ResNet50:**
 - Accuracy = 77.38 % (± 0.52 %)
 - Precision = 0.79 (± 0.01)
 - Recall = 0.77 (± 0.01)
 - F1-score = 0.77 (± 0.01)
- **Validácia 2:** Validácia na redukovanom datasete (každá trieda obsahuje 1200 obrázkov). Výsledok: Presnosť 83.0 %, zlepšená generalizácia na vyvážených dátach.
- **Validácia 3:** Porovnanie presnosti modelu medzi pôvodným a novým riešením na rovnakom testovacom datasete. EfficientNet-B3 prekonal pôvodný model o 1.2 % pri menšej výpočtovej náročnosti.

Tento projekt demonštruje efektívne prispôsobenie moderných metodológií na problém automatizovanej taxonomickej identifikácie fosílií. Výsledky ukazujú potenciál pre nasadenie v reálnych podmienkach, pričom ďalšie zlepšenia by mohli zahŕňať použitie väčšieho datasetu a pokročilejších vizualizačných techník.

Odkaz na git: https://github.com/MeedYsek/deep_nn_assign_1