

# Replikácia článku "Ore Image Classification Based on Improved CNN"

Dmytro Skrypchenko, Anatolii Porplytsia

## Abstract

Tento projekt je zameraný na replikáciu vybraného článku. Cieľom je na základe popisu uvedeného v článku implementovať modely a príslušné tréningové pipeline a následne porovnať získané výsledky s výsledkami uvedenými v pôvodnom článku.

## Index Terms

Ore classification, Deep learning, Convolution neural network, Transfer learning, Data augmentation, Squeeze-and-excitation networks

## I. ZHRNUTIE PÔVODNÉHO ČLÁNKU

V článku "Ore Image Classification Based on Improved CNN" autori riešia problém automatickej klasifikácie obrazov rúd, čo je významná úloha v ťažobnom priemysle. Tradičné metódy založené na vizuálnych vlastnostiach rúd, ako sú farba, textúra a lesk, sú neefektívne, časovo náročné a vyžadujú odbornú znalosť. Zároveň sú moderné metódy hlbokého učenia limitované nedostatkom dostatočne veľkých a rozmanitých množín označených dát [1].

Na prekonanie týchto problémov autori navrhli hybridný prístup, ktorý kombinuje transfer learning (prenos učenia), data augmentation (rozšírenie dát) a modul SENet (Squeeze-and-Excitation Networks). Ako dátový základ bol použitý verejne dostupný dataset z Kaggle s 957 obrázkami siedmich typov rúd (biotit, bornit, chryzokol, malachit, muskovit, pyrit, kremeň). Dátová množina bola rozdelená na tréningovú, validačnú a testovaciu časť v pomere 60:20:20 [2].

Testovaných bolo päť konvolučných neurónových sietí: AlexNet, VGG16, ResNet50, InceptionV3 a MobileNet. Najlepšie výsledky dosiahol model MobileNet, ktorý je efektívny a má nízku výpočtovú náročnosť vďaka použitiu separovateľných konvulúcií. Na zlepšenie presnosti bol na výstup MobileNetu pridaný modul SENet, ktorý zvyšuje dôležitosť relevantných kanálov a potláča irelevantné.

Hyperparameter	Value
Epochs	50
Learning Rate	0.0001
Batch Size	32
Optimizer	Adam
Loss function	Cross-entropy
Dropout	0.5

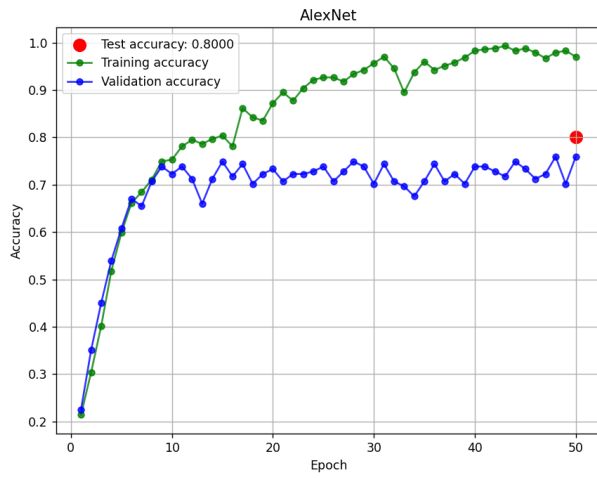
TABLE I: Hyperparametre použité pri tréňovaní modelov

Predspracovanie obrazu zahŕňalo normalizáciu obrázkov. Dropout bol pridaný ku každému modelu pred poslednú plne prepojenú vrstvu.

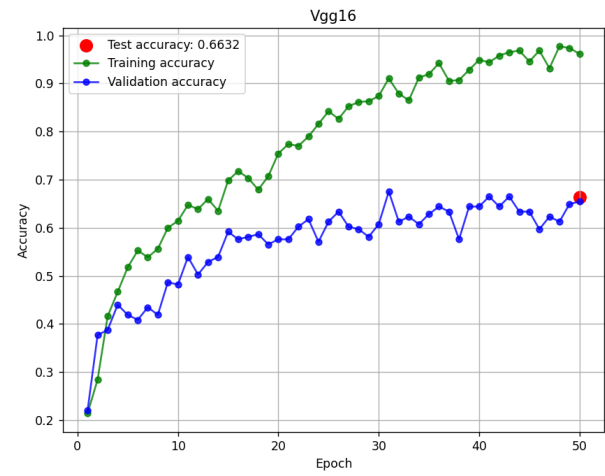
Výsledky ukazujú, že použitie transfer learningu a dátovej augmentácie výrazne zvýšilo presnosť všetkých modelov. Najvyššiu presnosť 96,89 % dosiahla kombinácia MobileNet + TL + DA + SENet. Navyše, testovanie na novom nezávislom datasete (243 obrázky) potvrdilo robustnosť riešenia s presnosťou 92,18 %.

## II. VÝSLEDKY NATRÉNOVANÝCH MODELOV

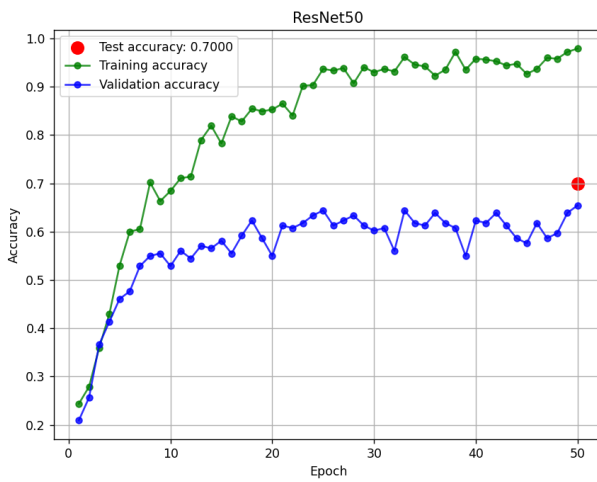
Na napísanie kódu bola použitá verzia Python 3.9.0 a používal sa framework PyTorch verzie 2.5.1.



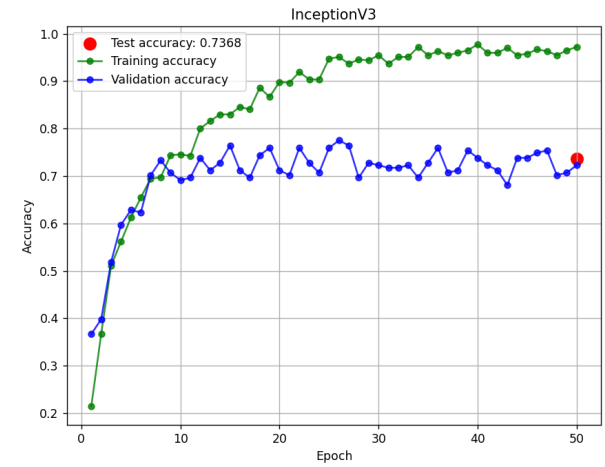
(a)



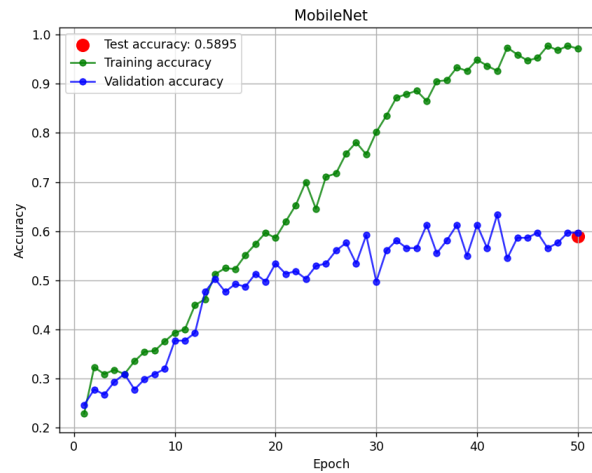
(b)



(c)



(d)



(e)

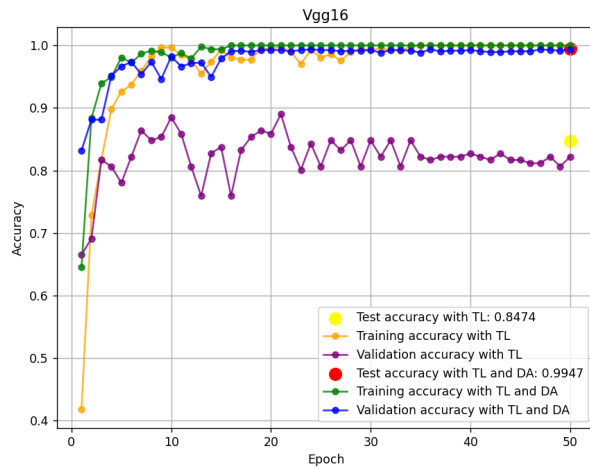
Fig. 1: (a) Výsledky pre AlexNet. (b) Výsledky pre VGG16. (c) Výsledky pre ResNet50. (d) Výsledky pre InceptionV3. (e) Výsledky pre MobileNet.

Model	Training Accuracy	Validation Accuracy	Test Accuracy
AlexNet	0.9702	0.7592	0.8
VGG16	0.9614	0.6545	0.6632
ResNet50	0.9789	0.6545	0.7
InceptionV3	0.9719	0.7225	0.7368
MobileNet	0.972	0.5969	0.5895

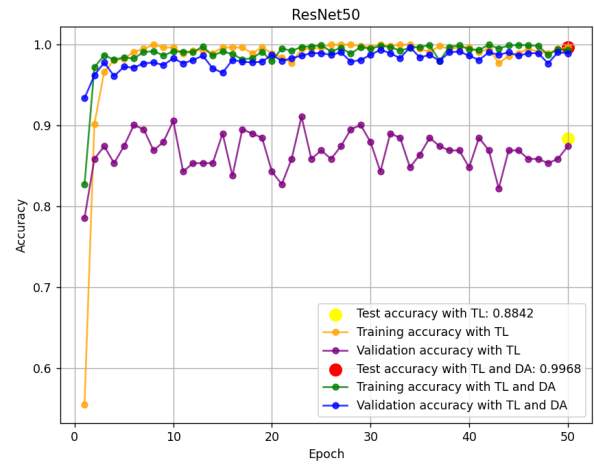
TABLE II: Výsledky natrénovaných modelov

V našej replikácii dosiahol najvyššiu testovaciu presnosť model AlexNet. Relatívne horšie výsledky dosiahli modely VGG16, ResNet50 a InceptionV3, pričom najnižšiu testovaciu presnosť mal model MobileNet. V porovnaní s pôvodným článkom vidíme, že aj tam bol MobileNet najhorším modelom (0,4404), avšak najlepší výkon tam dosiahli modely VGG16 a InceptionV3 (približne 0,75). Čo sa týka trénovacej presnosti, v našej replikácii ju všetky modely dosiahli nad úrovňou 0,95. V pôvodnom článku mali všetky modely taktiež pomerne vysokú trénovaciu presnosť, s výnimkou MobileNetu, ktorý dosiahol len úroveň porovnateľnú s testovacou presnosťou. Validačná presnosť bola v oboch prípadoch – u nás aj v pôvodnom článku – približne rovnaká ako testovacia.

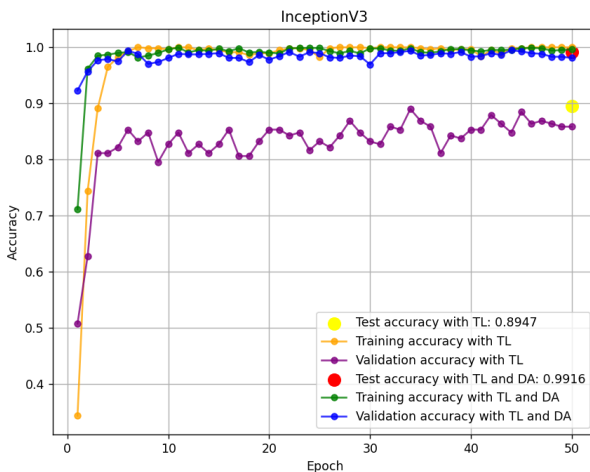
### III. VÝSLEDKY POUŽITIA TRANSFER LEARNINGU A DÁTOVEJ AUGMENTÁCIE



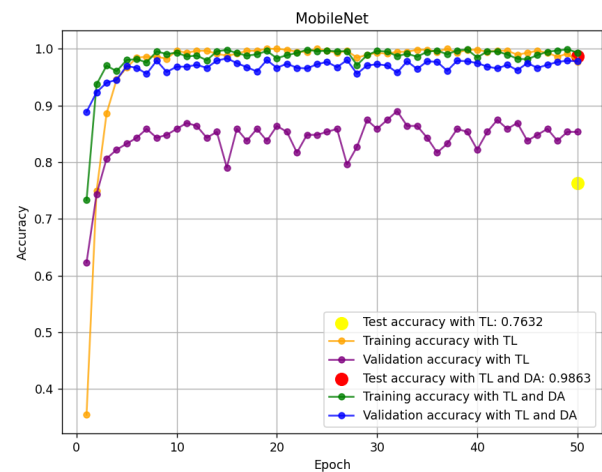
(a)



(b)



(c)



(d)

Fig. 2: (a) Výsledky pre VGG16. (b) Výsledky pre ResNet50. (c) Výsledky pre InceptionV3. (d) Výsledky pre MobileNet.

Model	Experiment	Training Accuracy	Validation Accuracy	Test Accuracy	Improvement
VGG16	TL	1.0	0.822	0.8474	0.1842
	TL + DA	1.0	0.9937	0.9947	0.3315
ResNet50	TL	0.9965	0.8743	0.8842	0.1842
	TL + DA	0.9937	0.9895	0.9968	0.2968
InceptionV3	TL	1.0	0.8586	0.8947	0.1579
	TL + DA	0.9965	0.9811	0.9916	0.2548
MobileNet	TL	0.9772	0.8534	0.7632	0.1737
	TL + DA	0.9933	0.9779	0.9863	0.3986

TABLE III: Výsledky použitia transfer learningu a dátovej augmentácie

Vo výsledkoch našej replikácie vidíme, že testovacie presnosti modelov trébovaných s použitím iba transfer learningu sú porovnateľné s výsledkami v pôvodnom článku a pohybujú sa v rozmedzí 0,76 – 0,89 (v pôvodnom článku od 0,78 do 0,86). Naopak, pri použití transfer learningu v kombinácii s dátovou augmentáciou dosiahli všetky modely v našej replikácii testovaciu presnosť nad 0,97, pričom najlepší výkon mal model VGG16. V pôvodnom článku použitie dátovej augmentácie takisto výrazne zvýšilo presnosť modelov, no najvyššiu presnosť (0,95) dosiahol model MobileNet, zatiaľ čo ostatné modely sa pohybovali v rozmedzí 0,82 – 0,87. Čo sa týka trébovacej presnosti, tá sa v oboch prípadoch (u nás aj v pôvodnom článku) pohybuje medzi 0,97 a 1. Validačná presnosť je vo všeobecnosti porovnateľná s testovacou.

#### IV. VÝSLEDKY POUŽITIA MODULU SE-NET

Autori pôvodného článku, aby ešte zvýšili výkonnosť modelu MobileNet, okrem použitia transfer learningu a dátovej augmentácie pridali pred plne prepojenú vrstvu aj modul SE-Net. Týmto spôsobom zvýšili presnosť modelu o 2 %, z 94 % na 96 %. V našej replikácii po pridaní modulu SE-Net sa presnosť zvýšila o 1 %.

Experiment	Test Accuracy
MobileNet	0.5895
MobileNet + TL	0.7632
MobileNet + TL + DA	0.9863
MobileNet + TL + DA + SENet	0.9947

TABLE IV: Výkonnosť modelu MobileNet po použití vybraných metód

#### V. CONCLUSION

Na testovanie autori vytvorili navyše nový nezávislý dataset, ktorý obsahoval 243 obrázkov. V našej replikácii sme vytvorili vlastný dataset so 92 obrázkami zozbieranými z internetu. Následne sme otestovali všetky vyššie uvedené modely na tomto novom datasete.

Model	Accuracy
MobileNet + TL + DA	0.8478
ResNet50 + TL + DA	0.8478
VGG16 + TL + DA	0.8478
MobileNet + TL + DA + senet	0.8261
ResNet50 + TL	0.7935
MobileNet + TL	0.7717
VGG16 + TL	0.7500
AlexNet	0.6522
InceptionV3	0.6413
ResNet50	0.5870
VGG16	0.5870
MobileNet	0.5326
InceptionV3 + TL	0.4891
InceptionV3 + TL + DA	0.3261

TABLE V: Presnosť všetkých modelov na novom datasete

Z tabuľky vyššie vidíme, že iba pri jednom modeli (InceptionV3) použitie transfer learningu a dátovej augmentácie znížilo presnosť klasifikácie. Pri všetkých ostatných modeloch tieto metódy výrazne zlepšili výkonnosť, čo potvrdzuje ich užitočnosť, najmä pri práci s malými datasetmi. Naopak, pridanie modulu SE-Net zhoršilo presnosť modelu MobileNet.

Pokiaľ ide o faktory, ktoré mohli ovplyvniť rozdiely vo výsledkoch replikácie, identifikovali sme niekoľko kľúčových problémov. Hlavným z nich je absencia uvedenia náhodného seedu použitého pri inicializácii modelov v pôvodnom článku, čo môže viesť k variabilite vo výsledkoch. Ďalej, popis dátovej augmentácie bol neúplný – nebolo špecifikované, do akej miery bol pôvodný dataset rozšírený, ani aké konkrétne transformácie boli aplikované na obrázky. Napokon, pre finálne testovanie bol v pôvodnej práci použitý nový, nezávislý dataset, avšak chýbal akýkoľvek odkaz alebo prístup k nemu, a preto sme boli nútení vytvoriť vlastný testovací dataset.

Celkovo možno skonštatovať, že aj napriek určitým rozdielom vo výsledkoch sa potvrdilo, že použitie transfer learningu a dátovej augmentácie vedie k zlepšeniu presnosti modelov. Naopak, prínos modulu SE-Net zostáva nejasný a bolo by vhodné jeho efektívnosť ďalej overiť v rozšírených experimentoch.

## REFERENCES

- [1] W. Zhou, Z. Chen, X. He, H. Song, H. Gu, and J. Liu, "Ore image classification based on improved CNN," *Computers & Structures*, vol. 263, p. 106750, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0045790622001173>
- [2] A. Brempong, "Minerals Identification Dataset," Kaggle, 2022. [Online]. Available: <https://www.kaggle.com/datasets/asiedubrempong/minerals-identification-dataset>