

Pénzérme-felismerés VGG16-tal

Készítették: Kovács Vanda Viktória és Jászfalusi Borka

Github link: <https://github.com/Borkalion/minta-kep-beadando/tree/main>

A beadandónkat saját fotóállományból álló képanyag alapján dolgoztuk fel magyar forint érmékkel (5-ös, 20-as, 10-es, 50-es, 100-as és 200-as). A projekt célja hogy ezeket az érmekeket a konvolúciós neurális háló automatikusan felismerje és bemutassa.

VGG16 modellt használtunk fel ImageNet-en előképezve Transfer Learning és finomhangolás segítségével. A rendelkezésre álló adatmennyiség viszonylag kisebb volt, így data augmentation technikákkal növeltük a varianciát, hogy a modell robusztusabban generalizáljon különböző fényviszonyokra és érme-orientációkra.

A végső háló egy új érmefotó alapján képes előre jelezni, milyen címlet szerepel a képen.

Felhasznált technológiák

- Python, PyTorch, torchvision
- VGG16 előtanított konvolúciós neurális háló
- Transfer learning + fine-tuning
- Data augmentation: forgatás, tükrözés, fényerő/kontraszt módosítás, véletlen crop
- train_test_split és DataLoader-ek
- Adam optimalizáló, CrossEntropyLoss
- Kiértékelés: pontosság, classification report, konfúziós mátrix

Adatkészlet bemutatása

A következő címleteket tartalmazza a feltöltött képanyag:

- **5 Ft**
- **10 Ft**
- **20 Ft**
- **50 Ft**
- **100 Ft**
- **200 Ft**

A fájlnevekben található az osztálycímke ("100_forint_23_min.jpg" → címke: 100).

A képek összegyűjtése után a datasetet véletlenszerűen, de stratifikáltan felosztottuk:

- **Train:** 70%
- **Val:** 15%
- **Test:** 15%

Ez biztosítja, hogy mindegyik címlet reprezentált marad minden részhalmazban.

Modell és tanítási folyamat

A modell alapja a `torchvision.models.vgg16`, ImageNet súlyokkal. A tanítást két szakaszra bontottuk:

1. Fej tanítása (head-only stage)

- A konvolúciós rétegek befagyasztva
- Tanuló réteg: a teljesen összekötött classifier blokk
- Tanulási ráta: $3e-4$
- Epochok száma: 8

A cél, hogy a modell a magas szintű jellemzőkből megtanulja a címletekhez tartozó döntési határokat.

2. Finomhangolás (fine-tuning)

- Az utolsó konvolúciós blokk feloldása
- Tanulási ráta: $3e-5$
- További epochok finomhangolásra: 15

Ez a szakasz az érmek lassan változó mintázatait (fényesség, textúra, domborulatok) segít megtanítani.

CrossEntropyLoss veszteségfüggvényt használtunk Adam féle optimalizálóval és veszteség alapján LR schedulerrel.

A tanítás során history-ban tároltuk el a *train loss*, *train accuracy*, *val loss* és *val accuracy* értékeket. A legjobb validációs eredményt adó modell külön mentésre került.

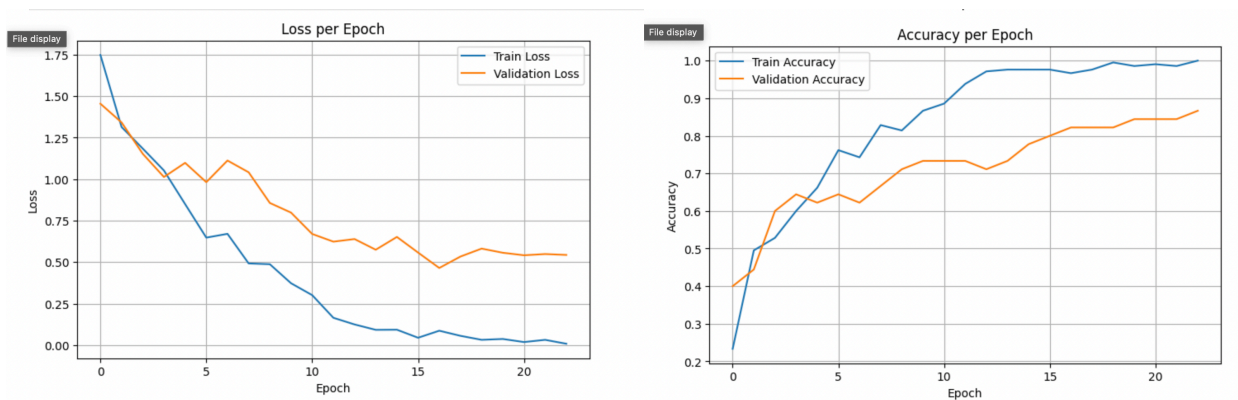
Data augmentation

A tanító halmazra véletlenszerű transzformációkat alkalmaztunk:

- RandomResizedCrop
- RandomRotation
- Horizontal/Vertical Flip
- ColorJitter (fényerő, kontraszt, szaturáció)
- Opcionális GaussianBlur
- Normalizálás ImageNet statisztikákkal

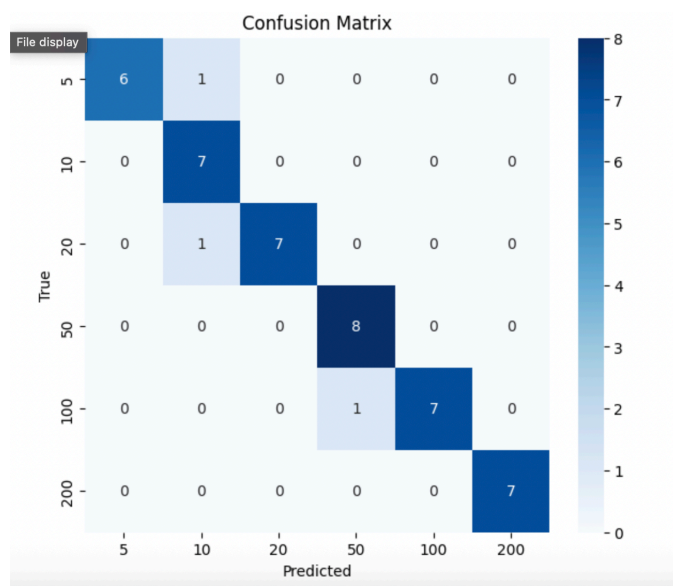
A validation és test halmazra csak determinisztikus transzformációk kerültek (átméretezés, center crop, normalizálás), hogy az értékelés konzisztens legyen.

Train-validation görbék



A veszteség- és pontossággörbék alapján a modell hatékonyan tanul: a tanulási veszteség folyamatosan csökken, miközben a tanulási pontosság közel 100%-ig javul. A validációs veszteség és pontosság szintén jelentős javulást mutat az első néhány epoch során, azonban később lassabban fejlődik, illetve ingadozóvá válik. A két görbe szétválása arra utal, hogy a modell a tanulási folyamat második felében elkezd túlilleszkedni: a tréningadatokon tovább javul, miközben a validációs teljesítmény már csak kisebb mértékben fejlődik. Összességében a modell jó tanulóképességet mutat, de a későbbi epochokban megjelenő overfitting miatt érdemes lenne korai megállást vagy valamilyen regularizációs módszert alkalmazni.

Konfúziós mátrix



A konfúziós mátrix azt mutatja, hogy a modell a legtöbb osztályt helyesen kategorizálja. A főátlóban látható értékek magasak, ami azt jelzi, hogy az egyes kategóriák többségét pontosan felismeri (pl. a 50-es és 200-as osztályokat hibátlanul). A néhány tévesztés elsősorban a 5–10 és 20–10 osztálypárok között jelenik meg, de ezek száma alacsony.

Összességében a modell jó osztályozási teljesítményt mutat, kevés téves besorolással és erős osztályonkénti pontossággal.

Classification Report

A klasszifikációs riport alapján a modell minden osztályon jó teljesítményt nyújt, magas precision, recall és f1-score értékekkel. Az összesített pontosság 93%, ami azt jelzi, hogy a modell megbízhatóan és következetesen osztályozza a mintákat.

CLASSIFICATION REPORT					
File display	precision	recall	f1-score	support	
5	1.00	0.86	0.92	7	
10	0.78	1.00	0.88	7	
20	1.00	0.88	0.93	8	
50	0.89	1.00	0.94	8	
100	1.00	0.88	0.93	8	
200	1.00	1.00	1.00	7	
accuracy			0.93	45	
macro avg	0.94	0.93	0.93	45	
weighted avg	0.95	0.93	0.93	45	

Helyes-hibás predikciók

