

Előre betanított mélytanuló modellek holografikus mikroszkópos kép-osztályozásra

Áttekintés

Holografikus mikroszkópos adatok osztályozásához számos **előre betanított CNN (konvolúciós neurális háló)** használható, melyek kombinálni tudják az amplitűd, fázis és maszk csatornákat egy **többrétegű bemenetként**. A szakirodalomban bevált megoldás a **transzfer tanulás**: egy nagy adathalmazon (pl. ImageNet) betanított mélyháló finomhangolása a specifikus mikroszkópos adatkészletre ¹. Így a háló megtartja az általános mintafelismerő képességeit, de adaptálódik a holografikus képek sajátosságaihoz. Például MacNeil és munkatársai plankton-hologramok osztályozására négy különböző mélyhálót is kipróbáltak (VGG16, InceptionV3, ResNet50V2, Xception) előtanult súlyokkal ¹. Mindegyik modell gyorsan konvergált és magas pontosságot ért el néhány epoch alatt ², ami azt mutatja, hogy a transzfer tanulás jól működik ebben a doménben.

Fig. 1: Különböző plankton osztályok amplitűd képei egy holografikus mikroszkópos adathalmazból ³. A képek 128×128 pixeles ROI-k, melyeket a rekonstruált hologramokból vágtak ki. Az automatikus osztályozó modelleknek az ilyen morfológiai különbségeket kell megtanulniuk felismerni.

A fenti példában a legjobb teljesítményt az **Xception** modell érte el: ~89% precíziót és F1 értéket produkált, felülmúlva a többi hálót ⁴. Az **InceptionV3** ezzel szemben némileg gyengébben (kb. 83% precízió) szerepelt ⁴. Mindez azt sugallja, hogy **a korszerű, mély architektúrák (különösen az Xception) nagyon hatékonyak lehetnek** a holografikus mikroszkópos képek osztályozásában, így érdemes ilyen modellekkel kezdeni. Az alábbiakban bemutatunk néhány releváns modellt és jellemzőiket, valamint hogy hogyan integrálhatók az amplitűd-fázis-maszk hármas bemenet feldolgozására.

Javasolt modellek és jellemzőik

- **ResNet50** – Klasszikus, **50 réteges** mély CNN residuális blokkokkal (He et al. 2015). Széles körben bevált képosztályozó, **ImageNet előtanulással** könnyen elérhető PyTorch-ban (a `torchvision.models` részeként). MacNeil és mtsai. hologram-kísérletében a ResNet50 (v2 variáns) ~88–89% precíziót ért el finomhangolás után ⁴, megközelítve az Xception teljesítményét. Implementációs nehézség: *alacsony* – a modell könnyen betölthető és a bemeneti rétege 3 csatornás, így közvetlenül etethető az amplitűd-fázis-maszk képekkel. Finomhangolással várhatóan jó recall érhető el a *Chlorella* osztályra is.
- **Xception** – **Mély, depthwise-szeparált konvolúciós háló** (Chollet, 2017), amely MacNeil-ék tesztjeiben a legjobb eredményt adta holografikus plankton adatokon ⁴. Előnye a nagy tanulókapacitás és hatékony paraméter-kihasználás. Az Xception ~89% precíziót és F1-score-t produkált a több-osztályos plankton felismerésben ⁴, jelezve hogy jól megkülönbözteti a vizuálisan hasonló objektumokat is. Előtanított súlyai elérhetők (pl. Keras alkalmazások között, vagy PyTorch `timm` könyvtáron keresztül). Implementációs nehézség: *közepes* – PyTorch-ban nincs beépítve, de a **timm** csomaggal egyetlen sorral betölthető (`timm.create_model('xception', pretrained=True)`). A bemenet itt is 3-csatornásra

alakítható, így az amplitude, fázis, maszk kombinációját a modell első konvolúciója együttesen feldolgozza. Az Xception előnye lehet a magas pontosság (kevés fals pozitív), ami kulcs a *Chlorella* magas precíziójú felismeréséhez.

- **ResNeXt-50 (32x4d)** – Egy **ResNet-variáns** (Xie et al. 2017) csoportos konvolúciókkal, amely kiváló eredményeket mutatott mikroszkópos algaképeken. Pant és munkatársai módosított ResNeXt architektúrát alkalmaztak több algafaj (*Pediastrum* genus) automatikus azonosítására, és **98,45%** osztályozási pontosságot értek el, 0,98 feletti F1-score mellett ⁵ ⁶. Ez kiemelkedően magas teljesítmény, ami jelzi a modell robusztusságát. A ResNeXt-50 előre betanított verziója elérhető PyTorch-ban (pl. `resnext50_32x4d` a Torchvision-ben). Implementációs nehézség: *alacsony/közepes* – hasonló a ResNet-hez, beépítetten használható. Várhatóan jól finomhangolható a 5-osztályos (*Chlorella*, *debris*, *Haematococcus* stb.) feladatra is. Előnye, hogy a **cardinality** (párhuzamos szűrőcsoportok) révén gazdagabb jellemzőket tanulhat, ami segíthet a *Chlorella* vs. hasonló kis részecskék megkülönböztetésében.
- **Egyéb modern modellek** – Alternatívaként szóba jöhetnek újabb, nagy teljesítményű modellek is, mint pl. az **EfficientNet** család vagy a Vision Transformer alapú classifier-ek. Bár ezekre konkrét hivatkozást a holografikus mikroszkópiában kevésbé találtunk, általánosságban ismertek kiváló pontosságukról a képosztályozási feladatokban. Az EfficientNet például méretarányosan mélyül és szélesedik, így kis adathalmazokra is hatékony lehet; PyTorch-ban a `timm` könyvtár biztosít előtanított EfficientNet-B0...B7 modelleket. Mivel a futási sebesség nem elsődleges szempont, akár egy nagyobb EfficientNet-B4/B7 is alkalmazható a magasabb pontosság érdekében – azonban érdemes megjegyezni, hogy az ilyen modellek finomhangolása több adatot és gondos hiperparaméter-beállítást igényelhet. **Vision Transformer (ViT)** alapú modellek szintén elérhetők előtanítva (pl. ImageNet21k-n betanítva), és bizonyos orvosi képalkotási feladatokban jól teljesítenek; implementációjuk PyTorch-ban szintén a `timm` csomaggal lehetséges. Ezek a modernebb modellek potenciálisan javíthatják az osztályozás pontosságát, de az implementáció komplexitása és az adaptáláshoz szükséges adat mennyisége kissé nagyobb lehet, mint a hagyományos CNN-eknél.

Az amplitude, fázis és maszk csatornák kombinálása

Fontos, hogy a választott modell **egyidejűleg felhasználja mindhárom bemeneti csatorna információját**. A legegyszerűbb megoldás erre, ha **háromcsatornás képként** kezeljük az adatot: pl. az amplitude képet, a fázisképet és a maszkot összeillesztjük a színcsatornák mentén. Így a standard CNN első rétege automatikusan feldolgozza mindhárom inputot. Ezt a módszert alkalmazta például Guo és munkatársai is, akik a hologramok rekonstruált amplitude- és fázisképét kétcsatornás inputként adták egy 3D CNN-nek ⁷. Hasonlóan, a harmadik csatornaként megadott bináris maszk biztosítja, hogy a háló **az objektum alakjára és helyére** is figyeljen. Ez praktikusán fókuszálja a modellt a releváns régióra, csökkentve a háttér zaj hatását. Alternatív megközelítés lehet külön alhálózatot használni az amplitude/fázis és a maszk feldolgozására, majd a magas szintű jellemzőket egyesíteni – ám általában nincs erre szükség, mert egy **egyetlen háló elegendő kapacitással** bír, hogy a maszk-információt is integrálja (pl. a maszk éleiből a tárgy körvonala kinyerhető az első konvolúciós rétegben).

Gyakorlati tanács: ügyeljünk az egyes csatornák normalizálására. A fázis-képek értékei és a maszk bináris értékei eltérő skálán mozoghatnak; célszerű pl. minden csatornát 0–1 közé skálázni vagy standardizálni, hogy a háló egyformán kezelje őket. Továbbá, ha az amplitude és fázis képek zajosak, a maszk segítségével nullázhatók a háttér-pixelek (maszkolással), ami még a hálózat bemenete előtt javíthatja a jel/zaj arányt. Összességében a háromféle információ együttes bevitele növelheti az

osztályozás **robosztusságát** és pontosságát, különösen a hasonló méretű és formájú részecskék elkülönítésénél.

Modellösszehasonlítás és implementáció

Az alábbiakban összefoglaljuk a szóba jöhető modellek főbb jellemzőit, várható teljesítményét és az implementáció ráfordítását, különös tekintettel a *Chlorella* osztály detektálására:

- **Pontosság és visszahívás (recall):** A legtöbb említett modell megfelelően finomhangolva képes legalább *0.5 recall* értéket elérni a *Chlorella* kategóriára, miközben a precízió magas marad. Például a **ResNeXt** alapú algaklaszter-azonosító 0.98 feletti *F1-score*-ja arra utal, hogy mind a recall, mind a precision kiváló volt minden osztályban ⁶. A **plankton-hologram transzfer tanulás** kísérletben pedig minden modellnél 0.8 feletti *F1-score* adódott ⁴, tehát a recall értékek is ehhez közeli tartományban voltak. Külön kiemelendő, hogy a *Chlorella* (vagy hasonló gömb alakú algák) felismerése tipikusan könnyebb a mélyhálóknak, mint a struktúrált törmelék vagy amorf debris kategóriáké, mert a *Chlorella* jellegzetes homogén, kerek sejtmorfológiával bír. Ennek megfelelően **várhatóan teljesül az ≥ 0.5 recall** követelmény, sőt magasabb is lehet, ha elegendő tanító példa áll rendelkezésre. A magas precision biztosítása érdekében érdemes kiegyensúlyozottan tanítani a modellt, vagy szükség esetén *cost-sensitive* módszerekkel a *Chlorella*-ra vonatkozó fals pozitívokat büntetni – de a tapasztalat azt mutatja, hogy egy jól generalizáló mélyháló önmagában is kevés téves riasztást ad (MacNeil-éknél pl. az Xception modell osztályonkénti precíziója is közel 0.9 volt ⁴).
- **Implementációs nehézség:** PyTorch környezetben a legegyszerűbben bevezethető modellek a **ResNet/ResNeXt család** (már a standard könyvtár részei, egy-két soros inicializálással), illetve a **DenseNet** vagy **VGG** is hasonlóan elérhető. Ezeknél csak a *klasszifikációs réteget* kell módosítani 5 kimeneti osztályra, majd finomhangolni a hálót. Az **Xception** és EfficientNet modellek PyTorch implementációja a `timm` könyvtár használatát igényli (amely azonban iparági standard, jól dokumentált). Kód szintjén tehát egyik javasolt modell sem jelent komoly akadályt. **TensorFlow 2** esetén hasonlóan elérhetők ezek az architektúrák alacsony szinten: pl. a Keras nélküli implementációhoz a `tf.keras.applications` modul kikerülendő, de használhatók nyílt forrású példakódok. Mivel a felhasználó explicit módon kizárta a Kerast, érdemes PyTorch-ban maradni, vagy TF-ben saját definícióval felépíteni a modellt (pl. online elérhetők nyers TF implementációk a ResNet, Xception stb. számára).
- **Elérhető források és dokumentáció:** Az ajánlott modellek mind **nyilvánosan dokumentáltak és hozzáférhetőek**. Például a **PyTorch hivatalos dokumentációja** tartalmaz példakódot a ResNet finomhangolására és bemutatja a szükséges lépéseket (adatbetöltés, train/val loop, stb.). Konkrét alkalmazási esettanulmányként MacNeil és munkatársai a teljes plankton osztályozó pipeline-juk forráskódját megosztották GitHubon ⁸ – ebben ugyan Keras környezetben valósították meg a transzfer tanulást (az Xception és más modellek felhasználásával), de a kód és a leírás hasznos támpont lehet a saját implementációhoz is. Hasonlóképpen, Roy *et al.* (2024) közreadta egy **mikroalga képosztályozó** tanulmány adatait és kódját ⁹ ¹⁰, melyben *Chlorella vulgaris* is szerepel. Ők ugyan hagyományos gépi tanulási módszereket (SVM, k-NN) is összehasonlítottak a deep learning megközelítésekkel, de eredményeik szerint egy felhő alapú CNN (Azure Custom Vision) érte el a legjobb, ~97.7%-os pontosságot a három algafaj (*Chlorella*, *Chlamydomonas*, *Spirulina*) szétválogatásában ¹¹. Ez is alátámasztja, hogy egy jól tanított mélyháló kiválóan működik a mikroalgák felismerésére. Az ő GitHub repójuk (RoyWeiiii/Scope_3_Microalgae...) tartalmazza az adat-előkészítés lépéseit, ami segítség lehet a saját holografikus dataset feldolgozásában is.

Összefoglalva: a feladatra több megfelelő előre betanított modell is elérhető. **ResNet50** egy biztos kiindulópont egyszerű implementációval; a **ResNeXt** architektúrával várhatóan tovább javítható a pontosság, míg az **Xception** mélyháló bizonyítottan kiváló a holografikus bio-képek osztályozásában (bár PyTorch-ban kissé több konfigurációt igényel). Mindezek a modellek kombinálni tudják az amplitűd, fázis és maszk információt – jellemzően a csatornák egyesítésével a bemenetben – és **finomhangolással** elérhető velük a kitűzött ≥ 0.5 *Chlorella recall* magas precízió mellett. A választás során figyelembe vehető még az adatmérték és a számítási erőforrás: ha elegendő tanító adat van, a komplexebb modellek (Xception, ResNeXt-101, EfficientNet-B5 stb.) nagyobb pontosságot hozhatnak. Kevesebb adatnál előnyösebb lehet egy egyszerűbb háló visszafogottabb paraméterszámmal (pl. ResNet18/34 vagy MobileNet-féle könnyített modellek), hogy elkerüljük a túltanulást. Mivel azonban a kérdésben a pontosság az elsődleges, érdemes a nagyobb kapacitású modellek felé orientálódni, és szükség esetén **adathalmaz-bővítéssel (data augmentation)** javítani a háló generalizációját. Az említett forráskódok és publikációk jó kiindulási alapot nyújtanak a megvalósításhoz, és tartalmaznak linkeket a **modell implementációkhoz illetve repókhoz** is (pl. MacNeil-ék GitHubja a holografikus plankton osztályozáshoz ⁸). A kiválasztott modellt ezek alapján testre szabva és a saját holografikus mikroszkópos adatokon betanítva várhatóan teljesíthetők lesznek a kitűzött mérőszámok (különös tekintettel a *Chlorella* detektálás recall ≥ 0.5 céljára).

Források:

1. MacNeil et al., “Plankton classification with high-throughput submersible holographic microscopy and transfer learning,” BMC Ecol. Evol. (2021) – különböző előtanított CNN-ek (VGG16, InceptionV3, ResNet50V2, Xception) alkalmazása holografikus plankton képekre, Xception modellel ~89% precízió elérése ¹ ⁴. Kód: GitHub – Holographic-Transfer-Learning repo ⁸.
2. Pant et al., “ResNeXt CNN topology-based deep learning model for Pediastrum identification,” Algal Research 48 (2020) – módosított ResNeXt modell 42 000 mikroszkópos algakép alapján, 98% feletti osztályozási pontossággal ⁵ ⁶.
3. Roy et al., “AI-driven microalgae cultivation: comparing ML and DL image classifiers,” Algal Research 72 (2024) – *Chlorella*, *Chlamydomonas*, *Spirulina* képosztályozása; Azure DL modell ~97.7% pontosság ¹¹. Adat és kód elérhető: Scope_3_Microalgae GitHub repó ¹² ¹⁰.
4. Guo et al., “Automated plankton classification from holographic imagery with deep CNNs,” Limnol. Oceanogr.: Methods 19(1):21-36 (2021) – holografikus inline mikroszkópos planktonképek osztályozása mélyhálózattal; amplitűd és fázis információ együttes feldolgozása a bemenetben ⁷. (Lásd még MacNeil et al. munkáját, amely e tanulmány továbbfejlesztett transzfer tanulási változatát mutatja be.)
5. Frontiers in Marine Science (2023) – “Vision Meets Algae: microalgae detection and classification” – összefoglaló cikk az algák képi azonosításáról, objektum-detektálási megoldásokkal. Leírja, hogy modern CNN-alapú detektorok (pl. YOLOv3 Darknet53 backbone-nal) is sikeresen alkalmazhatók algák felismerésére ¹³ ¹⁴. Megjegyzés: ha a teljes képen több objektum detektálása lenne cél, a YOLO-szerű modellek integrálhatják a lokalizációt és osztályozást; jelen esetben azonban a maszkok révén adottak az ROI-k, így elegendő a fenti klasszifikációs modelleket használni az egyes kivágott objektumokra.

¹ ² ³ ⁴ ⁸ Plankton classification with high-throughput submersible holographic microscopy and transfer learning | BMC Ecology and Evolution | Full Text
<https://bmcecol-evol.biomedcentral.com/articles/10.1186/s12862-021-01839-0>

⁵ ⁶ ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum (2020) | Gaurav Pant | 77 Citations
<https://scispace.com/papers/resnext-convolution-neural-network-topology-based-deep-4p8j2zjx5i>

7 Illustration of the 3D-model architecture. The input of the 3D network...

https://www.researchgate.net/figure/Illustration-of-the-3D-model-architecture-The-input-of-the-3D-network-is-the-amplitude_fig2_365087403

9 10 11 12 GitHub - RoyWeiiii/Scope_3_Microalgae_shape_texture_convolution_classification: The goal of this study is to classify microalgae of different species such as *Chlorella vulgaris* FSP-E, *Chlamydomonas reinhardtii*, and *Spirulina platensis*, using machine learning (ML) and deep learning (DL) methods

https://github.com/RoyWeiiii/Scope_3_Microalgae_shape_texture_convolution_classification

13 14 Frontiers | Vision meets algae: A novel way for microalgae recognition and health monitor

<https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2023.1105545/full>