

# Deep learning alapú kozmológiai vöröseltolódás-becslés fotometriai képek alapján

Bolyai Szakszeminárium

Mocskonyi Mirkó György

Fizikus MSc. I.

2020. november 4.

# Téma bemutatása

## Problémafelvetés

- Csillagászati felmérések: galaxisok égi koordinátái
- Távolság meghatározása nehéz
- Kozmológiai modell: **kozmológiai vöröseltolódás**  $\Rightarrow$  galaxis távolsága
- **Vöröseltolódás-becslés**

## Vöröseltolódás

- Doppler-effektus (zaj)
- Táguló világegyetem
- Gravitációs potenciál (elhanyagolható)
- $z = \frac{\lambda_o - \lambda_e}{\lambda_e}$
- Kinematikai:  $1 + z = \frac{\lambda_o}{\lambda_e} = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}$
- Tágulás:  $1 + z = \frac{\lambda_o}{\lambda_e} = \frac{a(t_o)}{a(t_e)}$  (Lemaître)
- Távolodási sebesség:  $v_r \Rightarrow$  Kozmológiai vöröseltolódás

## Vöröseltolódás

- Edwin Hubble:  $cz = H_0 \cdot D \Rightarrow$  csak kis  $z$ -kre
- Hubble-Lemaître-törvény:  $v_r = H_0 \cdot D$
- Sebesség-vöröseltolódás-reláció:  $v_r(z) = cH_0 \int_0^z \frac{dz'}{H(z')}; \left( H = \frac{\dot{a}}{a} \right)$

## Vöröseltolódás-becslés

- Sloan Digital Sky Survey (SDSS): Data Release 12
- Galaxisok spektruma: fotometriai és spektroszkópai felvételek
- Spektroszkópia  $\Rightarrow$  Vöröseltolódás; **Túlságosan időigényes**
- Fotometriai vöröseltolódás

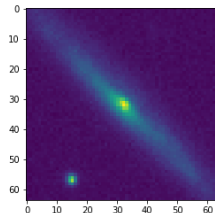
# Fotometriai vöröseltolódás-becslés

## Módszerek

- Színszűrők  $\Rightarrow$  **Magnitúdók**
- Feltevés: hasonló magnitúdók  $\Rightarrow$  hasonló vöröseltolódások
- Spektrális energiaeloszlás illesztése (SED-fitting):
  - Spektrum sablon: elméletből, vagy mérésekből
  - Nincs szükség spektroszkópai mintákra
- Gépi tanulás
  - Felügyelt tanuló algoritmusok
  - Spektroszkópai tanító halmaz szükséges
- Intergalaktikus por  $\Rightarrow$  **Vörösödés**

## Deep learning

- Színszűrők  $\Rightarrow$  **Fotometriai képek**
- Részletesebb galaxismorfológia
- Konvolúciós Neurális Háló (CNN)



# A konvolúciós hálókig

## Neurális háló

- Mesterséges intelligencia  $\Rightarrow$  Gépi tanulás
- $y = f(\text{objektum}, \mathbf{w})$ ; a  $\mathbf{w}$  súlyok optimalizálása
- Mesterséges neurális hálók  $\Rightarrow$  Mély tanulás (*deep learning*)

## Metódus

- Interakciók, aktivációk
- Veszteségfüggvény  $\Rightarrow$  minimalizálás
- *Backpropagation*  $\Rightarrow$  gradiens optimalizáció
- Batch, epoch

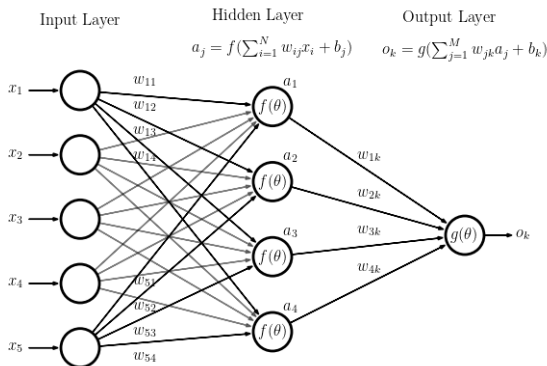


Figure: Teljesen kapcsolt neurális háló

# A konvolúciós hálóig

## Konvolúciós háló (CNN)

- Képek esetén környező pixelek számítanak  $\Rightarrow$  kevesebb paraméter
- Filterek/kernelek  $\Rightarrow$  konvolúció
- Stride, padding
- Összevonó réteg
- Paraméterek száma független a képek méretétől (magasság, szélesség)

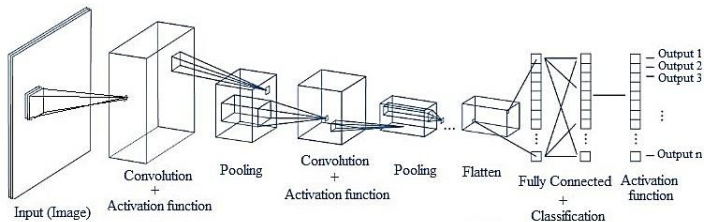


Figure: Tipikus tanítás CNN-nel

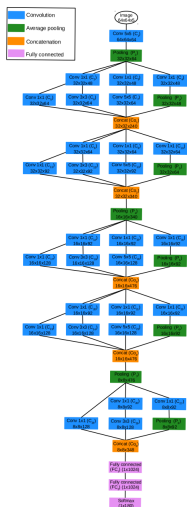
# Adatok, modell

## Adatok

- SDSS spektroszkópia  $\Rightarrow z < 0.4$  tartományban jó lefedettség
- Main Galaxy Sample (SQL)
- $64 \times 64$  pixeles  $(u, g, r, i, z)$  képek  $\Rightarrow$  **515 610 galaxis**
- Vörösödési értékek
- Képek augmentációja

## Modell

- *Keras, TensorFlow 1.4.1.*
- Pasquet et al, 2019 (P19)  $\Rightarrow$  **Inception Module CNN**  $\rightarrow$  **Dense**
- **ADAM** optimizer, *Cross-Entropy* veszteségfüggvény
- 25 824 344 paraméter
- Bemenet: kép + vörösödés
- Kimenet: **PDF** (180 bin)



## Aktivációk

- CNN aktiváció:

$$\text{PReLU}(x) = \begin{cases} x & \text{ha } x \geq 0 \\ \alpha x & \text{ha } x < 0. \end{cases}$$

- Kimenet aktiváció:

$$\text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

## Veszteségfüggvény

- *sparse categorical cross-entropy*
- *Cross-Entropy*:  $CE = - \sum_{i=1}^{180} z_{\text{true},i} \ln(z_{\text{pred},i})$

## Optimalizáló

- Sztochasztikus gradiens ereszkedés  $\Rightarrow$  Adam
- Változtatott paraméter: Tanulási ráta
- Batch méret: 64



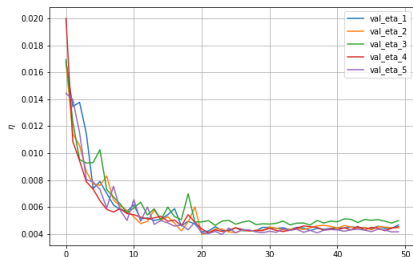
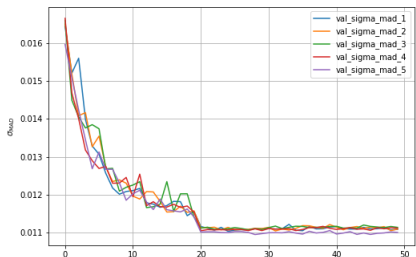
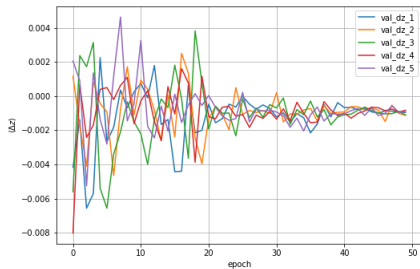
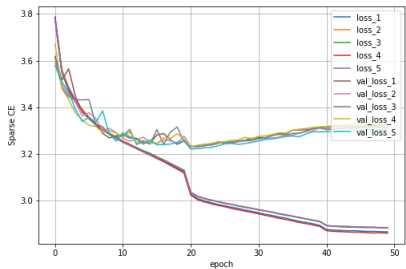
## Metrikák

- $z_{phot} = \sum_{k=1}^{180} PDF(z_k)z_k$
- Reziduális:  $\Delta z = \frac{z_{phot} - z_{spec}}{1 + z_{spec}}$
- $\langle \Delta z \rangle$
- $\sigma_{MAD} = 1.4826 \cdot \text{Median}(|\Delta z - \text{Median}(\Delta z)|)$
- *Outlier*-ek hányada ( $\eta$  [%]):  $|\Delta z| > 0.05$

## Futtatás

- ADAM tanulási ráta átállításával:  
20 + 20 + 10 epoch
- *Cross-Validation*: 5 tanítás:  
80%-20%
- Metrikák kiértékelése  $\Rightarrow$  legjobb  
súlyok kiválasztása
- Futtatások átlagolása
- 90 óra

# Metrikák



## Metrikák

- $z_{phot} = \sum_{k=1}^{180} PDF(z_k)z_k$
- Reziduális:  $\Delta z = \frac{z_{phot} - z_{spec}}{1 + z_{spec}}$
- $\langle \Delta z \rangle$
- $\sigma_{MAD} = 1.4826 \cdot \text{Median}(|\Delta z - \text{Median}(\Delta z)|)$
- *Outlier*-ek hányada ( $\eta$  [%]):  $|\Delta z| > 0.05$

## Futtatás

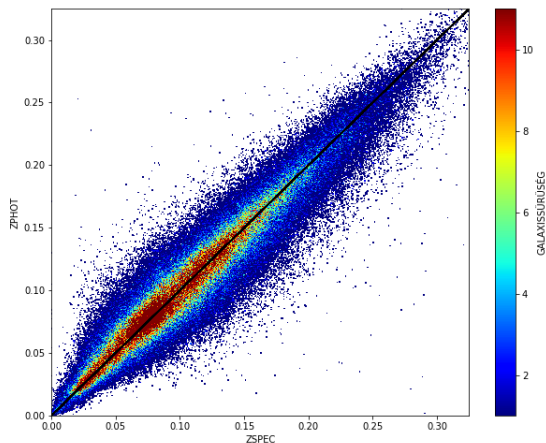
- ADAM tanulási ráta átállításával:  
20 + 20 + 10 epoch
- 5 tanítás: 80%-20%
- Metrikák kiértékelése  $\Rightarrow$  legjobb súlyok kiválasztása
- Futtatások átlagolása

### Átlagolt metrikák

|  | CNN     | P19     | B16     |
|--|---------|---------|---------|
| $\langle  \langle \Delta z \rangle  \rangle$ | 0.00015 | 0.00010 | 0.00062 |
| $\langle \sigma_{MAD} \rangle$               | 0.01116 | 0.00912 | 0.01350 |
| $\langle \eta \rangle$ [%]                   | 0.44    | 0.31    | 1.34    |

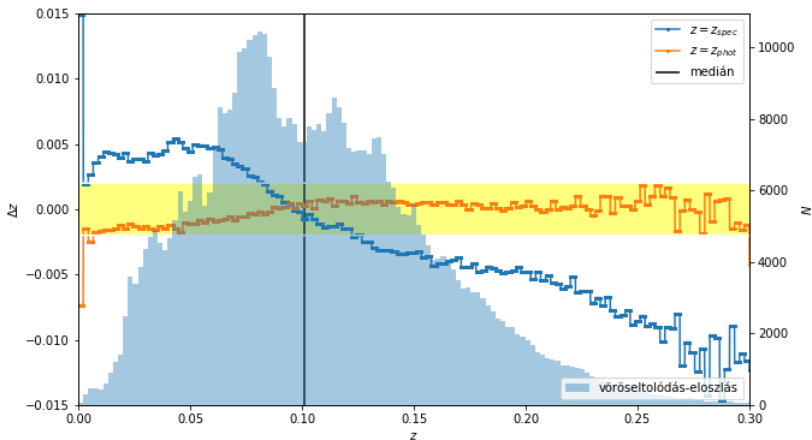
# Fotometriai vöröseltolódások

A CNN becsült fotometriai vöröseltolódásainak összehasonlítása a spektroszkópai vöröseltolódásokkal



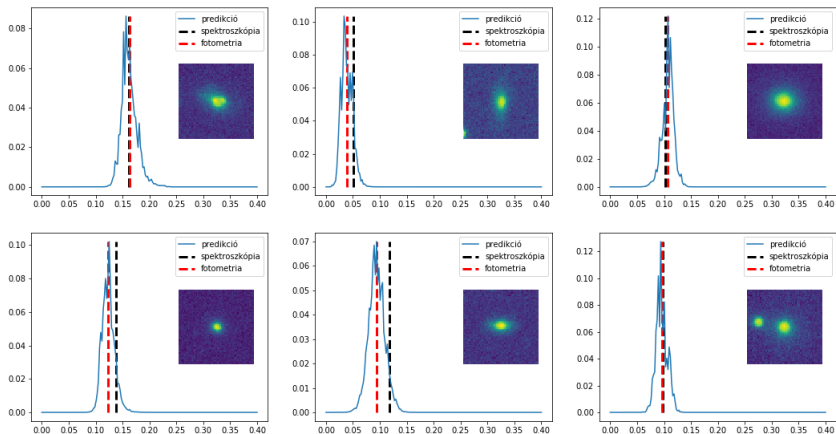
# Fotometriai vöröseltolódások

A  $\Delta z$  eltérések a spektroszkópai, valamint a fotometriai vöröseltolódások függvényében  $\Rightarrow$  *Euclid* misszióknak megfelelő pontosság



# Fotometrikus vöröseltolódások

Becsült fotometriai vöröseltolódás-eloszlások és a hozzájuk tartozó fotometriai vöröseltolódások, a spektroszkópai vöröseltolódások, valamint a  $64 \times 64$  pixeles fotók a megfelelő galaxisról



## További elemzés

- Vörösödés értékek segítenek-e?  $\Rightarrow \Delta z$  értékek
- Inklináció befolyásoló hatása  $\Rightarrow$  **augmentáció**
- Zavaró jelenségek a képeken: szomszéd, változó méret, fényesség  $\Rightarrow$  szomszéd segít

## Kitekintés

- SWARP eljárás
- *ensemble learning*
- Képekből közvetlenül nem kinyerhető információk modellbe foglalása:
  - SDSS kamerahibák
  - Égi és galaktikus koordináták
  - Földi atmoszféra zavaró hatása
- Képekhez tartozó magnitúdók

## Összefoglalás

- A teljesítmény megfelel az elvárásoknak, a különbségek megmagyarázhatóak
- *Euclid* misszió követelményeinek eleget tesz  $\Rightarrow \Delta z < 0.002$
- Releváns információk kinyerhetőek a képekből
- Sikeres vöröseltolódás-becslés

Köszönöm a figyelmet!