

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky
Katedra kybernetiky a umelej inteligencie

Modelovanie airglowu pomocou hlbokého učenia

Bakalárska práca

AIRGLOW MODEL

Používateľská príručka

Vedúci bakalárskej práce:

Doc. Ing. Peter Butka, PhD.

Bakalár:

Michal Bencúr

Konzultant bakalárskej práce:

Ing. Viera Maslej Krešňáková, PhD.

Košice 2022

Obsah

1	Funkcia programu	1
2	Súpis obsahu dodávky	2
3	Inštalácia programu	3
3.1	Požiadavky na programové prostriedky	3
4	Použitie programu	4
4.1	Stiahnutie dát misie GOLD	4
4.2	Explorácia dát	4
4.3	Príprava dát	5
4.4	Vytvorenie modelov a predikcií	5
4.5	Vyhodnotenie pomocou metrík	5
5	Popis vstupných, výstupných a pracovných súborov	6

1 Funkcia programu

Úlohou našich skriptov je stiahnuť dáta NASA misie GOLD, ktoré sú vo formáte `.nc`, a previesť ich do grafickej formy vo formáte `.png`. Následne ich roztriediť medzi množiny Severnej a Južnej pologule. Natrénovať model pomocou algoritmu Pix2Pix, orezať výsledné predikcie do formátu `.png` a porovnať ich s očakávanými obrázkami. Jednotlivé úlohy sú rozdelené do viacerých skriptov:

- Stiahnutie potrebných dát zo stránky misie, skrátiť ich názov a uložiť ich vo formáte `.nc`: `download.ipynb`
- Vytvorenie a uloženie grafických verzií dát pre rôzne zobrazenia vo formáte `.png`: `plots.ipynb`
- Grafické zobrazenie ukážka pokrytia dát za jeden deň:
`all_plots_at_one.ipynb`
- Grafické zobrazenie ukážka prekrytia dát za jeden deň: `overlap_plots.ipynb`
- Vytvorenie a uloženie grafických verzií dát určených ako množina dát pre učenie modelu a roztriedenie podľa pologule vo formáte `.png`:
`prediction_plots.ipynb`
- Natrénovanie modelu z pripravených dát pomocou algoritmu Pix2PixHD a vytvorenie predikcií z tohto modelu: `pix2pix.ipynb`
- Orezanie vytvorených predikcií aj pôvodných obrázkov aby ich bolo možné porovnať: `crop_images.ipynb`
- Porovnanie vytvorených predikcií a pôvodných snímok pomocou metrík:
`compare.ipynb`

2 Súpis obsahu dodávky

Priložené CD obsahuje kópiu github repozitáru `github.com/MichalBencur03/Airglow-modeling-by-deep-learning` tejto práce:

- priečinok obsahujúci Jupyter notebooky s kódom tejto práce: `code`,
- ukážku jedného dátového netCDF súboru z dôvodu miesta: `code/data`,
- priečinok obsahujúci algoritmus Pix2PixHD: `code/pix2pixHD`,
- ukážku exploatačných dát pre jeden deň roku 2021: `code/plots`
- základné inštrukcie: `README.md`,
- systémovú príručku v elektronickej forme: `SP_Bencur_Michal.pdf`,
- používateľskú príručku v elektronickej forme: `PP_Bencur_Michal.pdf`,
- bakalárska práca v elektronickej forme: `BP_Bencur_Michal.pdf`.

3 Inštalácia programu

3.1 Požiadavky na programové prostriedky

- Python 3.8.6
- Anaconda Navigator
- Jupyter Notebook
- Knižnice: `requests`, `bs4`, `urllib`, `os`, `glob`, `pathlib`, `netCDF4`, `matplotlib`, `numpy`, `datetime`, `dominate`, `opencv-python`, `Pillow`, `sewar`, `scipy`, `pandas`, `mpl_toolkits`

4 Použitie programu

V tejto kapitole sa budeme zaoberať použitím skriptov ktoré sme uviedli v kapitole 1.

4.1 Stiahnutie dát misie GOLD

Stiahnutie dát vo formáte NetCDF s príponou `.nc` je uskutočnené pomocou skriptu `download.ipynb`. V tomto súbore je potrebné zadať začiatočný deň v roku do premennej `day`, rok do premennej `year` a počet dní ktoré chceme stiahnuť do premennej `how_many`. Hlavnou úlohou tohto skriptu je stiahnuť dáta a uložiť ich do priečinku `data/`.

4.2 Explorácia dát

Celková explorácia dát je uskutočnená pomocou nasledujúcich skriptov:

- `plots.ipynb`
- `all_plots_at_one.ipynb`
- `overlap_plots.ipynb`

Tieto súbory pracujú so súbormi s príponou `.nc` a vytvárajú ich grafické zobrazenia. Súbor `plots.ipynb` vytvára grafy pre zobrazenia hodnôt 1356 žiarenia, LBH žiarenia, SZA, a samotný graf bez preloženia na povrch Zeme. Tieto zobrazenia pripravuje pre každý deň. Súbory `all_plots_at_one.ipynb` a `overlap_plots.ipynb` vytvárajú grafické zobrazenia pre pokrytie a prekrytie dát. Hlavnou úlohou týchto skriptov je vytvorenie dát pre exploráciu vo formáte `.png` a ich uloženie do priečinku `plots/`. Použitie týchto skriptov je dobrovoľné keďže vytvoriť predikciu je možné aj bez ich použitia.

4.3 Príprava dát

Príprava dát určených pre prácu s algoritmom Pix2PixHD je uskutočnená pomocou skriptu `prediction_plots.ipynb`. Tento súbor pracuje s dátami vo formáte `.nc` a vtvára grafické zobrazenie dát ktoré následne roztriedi podľa pologule. Hlavnou úlohou tohto skriptu je vytvoriť a uložiť dáta podľa príslušnej pologule do priečinku `pix2pixHD/datasets/nazov_` vo formáte `.png`.

4.4 Vytvorenie modelov a predikcií

Pred použitím skriptu `pix2pix.ipynb` je potrebné v priečinku `pix2pixHD/datasets/` vytvoriť priečinky `test_frames` a `train_frames` a premiestniť do nich dáta na učenie a testovanie. Hlavnou úlohou tohto skriptu je vytvorenie modelu algoritmu Pix2PixHD a vytvorenie predikovaných obrázkov z tohto modelu vo formáte `.png`.

4.5 Vyhodnotenie pomocou metrík

Pred vyhodnotením predikcií pomocou metrík je potrebné orezať predikované aj pôvodné obrázky na rovnakú veľkosť. Pre tento krok je potrebné použiť skript `crop_images.ipynb`. Hlavnou úlohou skriptu `compare.ipynb` je porovnať prebie s pôvodnými obrázkami pomocou metrík MSE a PSNR.

5 Popis vstupných, výstupných a pracovných súborov

Vstupné dáta pre vytvorenie množín na tréovanie a učenie modelov vo formáte `.png`:

- Stiahnuté dáta misie GOLD vo formáte `.nc`

Výstupné obrázky predikcií z množiny dát vo formáte `.png`:

- Predikované snímky misie vo formáte `.png`

Výstupnými dátami Jupyter notebooku `download.ipynb` sú:

- Množina dát zo stránky misie GOLD vo formáte `.nc`

Vstupnými dátami Jupyter notebooku `plots.ipynb` sú:

- Dáta misie GOLD vo formáte `.nc`

Výstupnými dátami Jupyter notebooku `plots.ipynb` sú:

- Vizualizácia dát pre žiarenie označené ako 1356 vo formáte `.png`
- Vizualizácia dát pre žiarenie označené ako LBH vo formáte `.png`
- Vizualizácia dát pre SZA vo formáte `.png`
- Vizualizácia dát pre pre žiarenie označené ako 1356 bez preloženia na povrch Zeme vo formáte `.png`

Vstupnými dátami Jupyter notebooku `all_plots_at_one.ipynb` sú:

- Dáta misie GOLD vo formáte `.nc`

Výstupnými dátami Jupyter notebooku `all_plots_at_one.ipynb` sú:

- Vizualizácia pokrytia Zemského povrchu dátami misie GOLD vo formáte `.png`

Vstupnými dátami Jupyter notebooku `overlap_plots.ipynb` sú:

- Dáta misie GOLD vo formáte `.nc`

Výstupnými dátami Jupyter notebooku `overlap_plots.ipynb` sú:

- Vizualizácia prekrytia medzi dátami misie GOLD vo formáte `.png`

Vstupnými dátami Jupyter notebooku `prediction_plots.ipynb` sú:

- Dáta misie GOLD vo formáte `.nc`

Výstupnými dátami Jupyter notebooku `prediction_plots.ipynb` sú:

- Vizualizácia dát pre žiarenie označené ako 1356 z dôvodu vytvorenia množín pre predikčné modely vo formáte `.png`

Vstupnými dátami Jupyter notebooku `pix2pix.ipynb` sú:

- Množina dát pre tréovanie modelu Severnej pologule vo formáte `.png`
- Množina dát pre testovanie modelu Severnej pologule vo formáte `.png`
- Množina dát pre tréovanie modelu Južnej pologule vo formáte `.png`
- Množina dát pre testovanie modelu Južnej pologule vo formáte `.png`

Výstupnými dátami Jupyter notebooku `pix2pix.ipynb` sú:

- Natrénovaný model pre predikciu Airglow Severnej pologule `checkpoints/year_north`
- Natrénovaný model pre predikciu Airglow Južnej pologule `checkpoints/year_south`
- Výsledné predikcie modelu Severnej pologule vo formáte `.jpg`
- Výsledné predikcie modelu Južnej pologule vo formáte `.jpg`

Vstupnými dátami Jupyter notebooku `crop_images.ipynb` sú:

- Predikcie natrénovaných modelov vo formáte `.jpg`
- Originálne porovnávacie dáta vo formáte `.png`

Výstupnými dátami Jupyter notebooku `crop_images.ipynb` sú:

- Upravené predikcie a originálne dáta vhodné pre porovnanie vo formáte `.png`

Vstupnými dátami Jupyter notebooku `compare.ipynb` sú:

- Upravené predikované snímky vo formáte `.png`
- Upravené originálne snímky vo formáte `.png`