

Technická univerzita v Košiciach  
Fakulta elektrotechniky a informatiky

Využitie hlbokého učenia pre  
modelovanie airglowu

Diplomová práca

**Technická univerzita v Košiciach**  
**Fakulta elektrotechniky a informatiky**

**Využitie hlbokého učenia pre  
modelovanie airglowu**

**Diplomová práca**

Študijný program: Hospodárska informatika  
Študijný odbor: Informatika  
Školiace pracovisko: Katedra kybernetiky a umelej inteligencie (KKUI)  
Školiteľ: doc. Ing. Peter Butka, PhD.  
Konzultant: RNDr. Šimon Mackovjak, PhD.  
Konzultant: Ing. Lenka Kališková

**Košice 2024**

**Bc. Michal Bencúr**

## **Abstrakt v SJ**

Skúmanie javov v blízkom vesmíre bola vždy dôležitá oblasť pre pochopenie vplyvu javov na našu Zem. Jedným z týchto javov je Airglow, ktorý je pozorovateľný ako žiarenie na nočnej oblohe. Cieľom práce je pomocou metód hlbokého učenia predikovať hodnoty tohto javu. Práca poskytuje teoretický prehľad javu Airglow, ako aj metód strojového učenia so zameraním na použitý model Pix2Pix. V práci sú použité dáta z celooblohových snímačov Bostonskej Univerzity, ktoré sú pomocou vybraného modelu využívané na predikciu častí snímanej oblasti dát NASA misie GOLD. Výsledky sú na záver porovnané s predošlými dosiahnutými výsledkami.

## **Kľúčové slová**

Airglow, hlboké učenie, predikcia, atmosféra, Pix2Pix

## **Abstrakt v AJ**

The study of phenomena in near-Earth space has always been an important area for understanding their influence on our Earth. One of these phenomena is the Airglow which is observable as a glow in the night sky. The goal of this work is to use deep learning methods to predict the values of this phenomenon. The thesis provides a theoretical overview of the Airglow phenomenon as well as machine learning methods, focusing on the Pix2Pix model used. The thesis uses data from Boston University's all-sky sensors, which are used to predict portions of the NASA GOLD mission's sensed data domain using the selected model. Finally, the results are compared with previous results obtained.

## **Kľúčové slová v AJ**

Airglow, deep learning, prediction, atmosphere, Pix2Pix

**TECHNICKÁ UNIVERZITA V KOŠICIACH**  
**FAKULTA ELEKTROTECHNIKY A INFORMATIKY**  
Katedra kybernetiky a umelej inteligencie

**Z A D A N I E**  
**D I P L O M O V E J P R Á C E**

Študijný odbor: **Informatika**  
Študijný program: **Hospodárska informatika**

Názov práce:

**Využitie hlbokého učenia pre modelovanie airglowu**  
Use of deep learning for airglow modeling

Študent: **Bc. Michal Bencúr**  
Školiteľ: **doc. Ing. Peter Butka, PhD.**  
Školiace pracovisko: **Katedra kybernetiky a umelej inteligencie**  
Konzultant práce: **Ing. Lenka Kališková**  
Pracovisko konzultanta: **Katedra kybernetiky a umelej inteligencie**

Pokyny na vypracovanie diplomovej práce:

1. Podat' teoretický prehľad problematiky metód hlbokého učenia, s dôrazom na ich použitie pre modelovanie airglowu.
2. Získať a predspracovať dáta pre modelovanie airglowu.
3. Navrhnuť, realizovať a vyhodnotiť experimenty s modelovaním airglowu pomocou metód hlbokého učenia.
4. Vypracovať dokumentáciu podľa pokynov katedry a vedúceho práce.

Jazyk, v ktorom sa práca vypracuje: **slovenský**  
Termín pre odovzdanie práce: **19.04.2024**  
Dátum zadania diplomovej práce: **31.10.2023**



*N. E. Vokorokos*  
prof. Ing. Liberios Vokorokos, PhD.  
dekan fakulty

### Čestné vyhlásenie

Vyhlasujem, že som diplomovú prácu vypracoval samostatne s použitím uvedenej odbornej literatúry.

Košice 19. 4. 2024

.....

*Vlastnoručný podpis*

## **Podakovanie**

Týmto by som chcel poďakovať doc. Ing. Petrovi Butkovi, PhD., RNDr. Šimonovi Mackovjakovi, Phd. a Ing. Lenke Kališkovovej za odbornú pomoc na tejto diplomovej práci a za ich čas, ktorý mi venovali pri tvorbe tejto práce.

# Obsah

Úvod	1
<b>1 Doména skúmania</b>	<b>2</b>
1.1 Airglow . . . . .	2
1.2 Celooblohové snímače Bostonskej Univerzity . . . . .	6
1.3 NASA misia GOLD . . . . .	8
<b>2 Strojové učenie</b>	<b>13</b>
2.1 GAN . . . . .	17
2.2 Pix2Pix . . . . .	19
2.3 Metriky . . . . .	21
<b>3 Modelovanie javu Airglow pomocou hlbokého učenia</b>	<b>23</b>
3.1 Analýza súčasného stavu . . . . .	23
3.2 Pochopenie cieľa . . . . .	27
3.3 Príprava dát . . . . .	29
3.4 Klasifikácia BU dát . . . . .	30
3.5 Príprava dát pre predikciu . . . . .	33
3.6 Explorácia modelu Pix2Pix . . . . .	37
3.7 Výsledný model a vyhodnotenie . . . . .	48
<b>4 Záver</b>	<b>53</b>
<b>Zoznam príloh</b>	<b>60</b>

## Zoznam obrázkov

1–1	Atmosféra Zeme a výšky javu Airglow s ich zafarbeniami. Zdroj: (Slovinský and Mackovjak, 2020) . . . . .	4
1–2	Rozmiestnenie staníc celooblohových staníc. Zdroj: Bostonská univerzita . . . . .	9
1–3	Snímanie pomocou jedného z kanálov prístroja misie GOLD. Zdroj: NASA . . . . .	10
1–4	Nákres snímača misie GOLD a všetky jeho časti. Zdroj: NASA . . . . .	11
2–1	Diagram fungovania GAN pri jeho učení. Zdroj: (Serag et al., 2019)	19
2–2	Princíp generovania obrazov pomocou Pix2Pix. Zdroj: (Isola et al., 2016) . . . . .	21
3–1	Orezaná snímka nočnej oblohy z 15. Mája 2021 o 7:48:48 . . . . .	31
3–2	Rozdelenie množstva jasných nocí za rok 2021 nad stanicou Arecibo	33
3–3	Znázornenie oblasti snímania stanice Arecibo na snímke dát misie GOLD . . . . .	34
3–4	Rozdelenie snímok vybraného časového úseku s priemerami pixelov snímok . . . . .	35
3–5	Znázornenie dostupnosti snímok v hodinových intervaloch. Červená farba sú dostupnosť z oboch zdrojov v danú hodinu, žltá iba z GOLD dát. . . . .	36
3–6	Porovnanie 3 po sebe idúcich dní predikcie stanice Mercedes voči originálnym výrezom. Horný rad sú predikované snímky zatiaľ čo spodný rad sú originálne výrezy z dát misie NASA GOLD . . . . .	49
3–7	Ukážka doplnenia 2 snímok z NASA misie GOLD . . . . .	51
3–8	Porovnanie predikovaných dát Airglow pomocou našej metódy kombinácie zdrojov dát . . . . .	52



## Zoznam tabuliek

3–1	Klasifikačný report použitého modelu . . . . .	32
3–2	Výsledky presnosti modelu na náhodne vybraných snímkach pri použití základných nastavení. . . . .	40
3–3	Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia batchsize na hodnotu 4. . . . .	42
3–4	Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavení ngf a ndf na hodnotu 128. . . . .	43
3–5	Výsledky presnosti modelu na náhodne vybraných snímkach počas zmeneného nastavenia lr na hodnotu 0.002. . . . .	45
3–6	Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia — lr_policy na spôsob kosínus. . . . .	46
3–7	Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia — lr_policy na spôsob step. . . . .	46
3–8	Výsledky presnosti finálneho modelu so zvolenými základnými nastaveniami. Snímky pre predikciu sú zvolené na základe ich postupnosti v dňoch . . . . .	50
3–9	Výsledky metrík 6 po sebe idúcich snímkov pre Severnú a Južnú pologu vytvorených ako výsledok bakalárskej práce. . . . .	51

## Slovník skratiek

**NM** - Nanometer, fyzikálna jednotka používaná pre vyjadrenie vlnovej dĺžky

**NASA** - National Aeronautics and Space Administration

**GOLD** - The Global-scale Observations of the Limb and Disk

**MSE** - Mean Squared Error

**PSNR** - Peak Signal-to-Noise Ratio

**GAN** - Generative Adversarial Networks

**CCD** - Charge-Coupled Devices

**CNN** - Konvolučné Neurónové Siete

**RNN** - Rekurentné Neurónové Siete

**GAN** - Generatívne Adverzné Siete

**AI** - Umelá Inteligencia

**GDX** - Gradient descent with momentum and adaptive learning rate

**SCG** - Scaled conjugate gradient

**LM** - Levenberg-Marguardt

**TCN** - Temporal Convolutional Network

## Úvod

Využitie strojového učenia sa v posledných rokoch výrazne rozšírilo. Začíname sa s ním stretávať v našom každodennom živote čoraz častejšie. Vďaka pokrokom vo vývoji výpočtovej techniky je možné aplikovať tieto princípy do programov a zariadení, ktoré používame každý deň. Strojové učenie má však využitie aj pri výskume mnohých fyzikálnych javov. Využitie týchto technológií umožňuje prácu s omnoho väčším množstvom dát. Vďaka modelom strojového učenia sa výskum mnohých oblastí urýchľuje.

Jednou z oblastí využitia je aj skúmanie javu v našej atmosfére označovaného ako Airglow ktorý predstavuje vznikajúce ako dôsledok ionizácie a excitácie atómov a molekúl v hornej atmosfére Zeme, slnečnými vysoko-energetickými fotónmi. Výskum tohto javu prináša mnoho odpovedí na otázky týkajúcich sa rozhrania medzi Zemou a blízkym vesmírom. Získavanie údajov o tomto jave môže prebiehať dvoma hlavnými spôsobmi. Prvým je pomocou snímačov z povrchu Zeme, ktoré snímajú nočnú oblohu. Ďalším spôsobom je pomocou prístrojov orbitujúcich Zem na satelitoch, ktoré snímajú smerom na našu planétu. Oba prístupy majú svoje výhody a nevýhody, a preto pre dosiahnutie najlepších výsledkov je potrebné ich kombinovať.

Práve spojenie týchto dvoch zdrojov s použitím hlbokého učenia je cieľom našej práce. Úlohou bude za použitia neurónovej siete predikovať hodnoty Airglow. Model bude aplikovaný na dáta misie NASA GOLD, ktoré snímajú iba určitú oblasť našej planéty pomocou snímača na satelite. Dopĺňanie týchto údajov o meraní bude možné pomocou využitia snímok z celooblohových snímačov Bostonskej Univerzity, ktoré budú na vstupe použitej neurónovej siete. Zameriavame sa na predikciu pomocou snímok, pričom ide o preklad obrazu na obraz. Vďaka týmto doplneným meraniam Airglow pomôžeme rozšíriť skúmanie o oblasti, ktoré by inak nebolo možné získať. V prvej kapitole sa budeme venovať teoretickému prehľadu problematiky.

# 1 Doména skúmania

## 1.1 Airglow

Vesmír a atmosféra našej planéty bola vždy zaujímavou oblasťou skúmania pre ľudí, či už išlo o fascináciu Mesiacom alebo Slnkom. Na nočnej oblohe je možné pozorovať množstvo svetelných objektov. V dnešnej dobe sa na oblohe nachádza aj svetelný smog, ktorý znižuje viditeľnosť týchto objektov. Ak by sme sa mu vyhli mimo veľkých miest stále by na oblohe bolo vidieť mierne žiarenie. Takýto jav sa označuje ako Airglow. Pozorovanie tohto javu je možné aj voľným okom, avšak je omnoho lepšie pozorovateľný pomocou špeciálnych staníc a teleskopov. Ide o slabú luminiscenciu v horných vrstvách zemskej atmosféry, ktoré fungujú ako rozhranie medzi procesmi vo vesmíre a na Zemi. Je to veľmi dynamické prostredie, ktoré je neustále ovplyvňované slnečným žiarením, vesmírnym počasím, ale aj atmosférickým počasím (Pfaff, 2012). Väčšina žiarenia sa nachádza v oblasti približne 80 km až 300 km nad povrchom Zeme, pričom najjasnejšia oblasť vo viditeľnej časti spektra sa nachádza vo výške približne 97 km. Airglow je netermálna emisia svetla pochádzajúca z excitovaných atómových alebo molekulových stavov. Zdrojom excitácie je priamo alebo nepriamo slnečné elektromagnetické žiarenie. Podobným žiarením, pozorovateľným voľným okom je aj polárna žiara, Airglow sa však na rozdiel od nej dá pozorovať na oblohe vo všetkých zemepisných šírkach nezávisle na čase. Pri použití snímačov je možné pozorovať tento jav nezávisle od svetelných podmienok, a teda aj počas dňa. Fotochemická luminiscencia, ktorá sa nazýva aj chemiluminiscencia, je spôsobená chemickou reakciou medzi slnečným žiarením a časticami prítomnými v horných vrstvách atmosféry. Slnečné svetlo dodáva potrebnú energiu, pomocou ktorej materiály zmenia svoj stav a následne začnú produkovať emisie s určitou vlnovou dĺžkou. (Khomich and et al., 2008) Medzi najvýznamnejšie emisie patria:

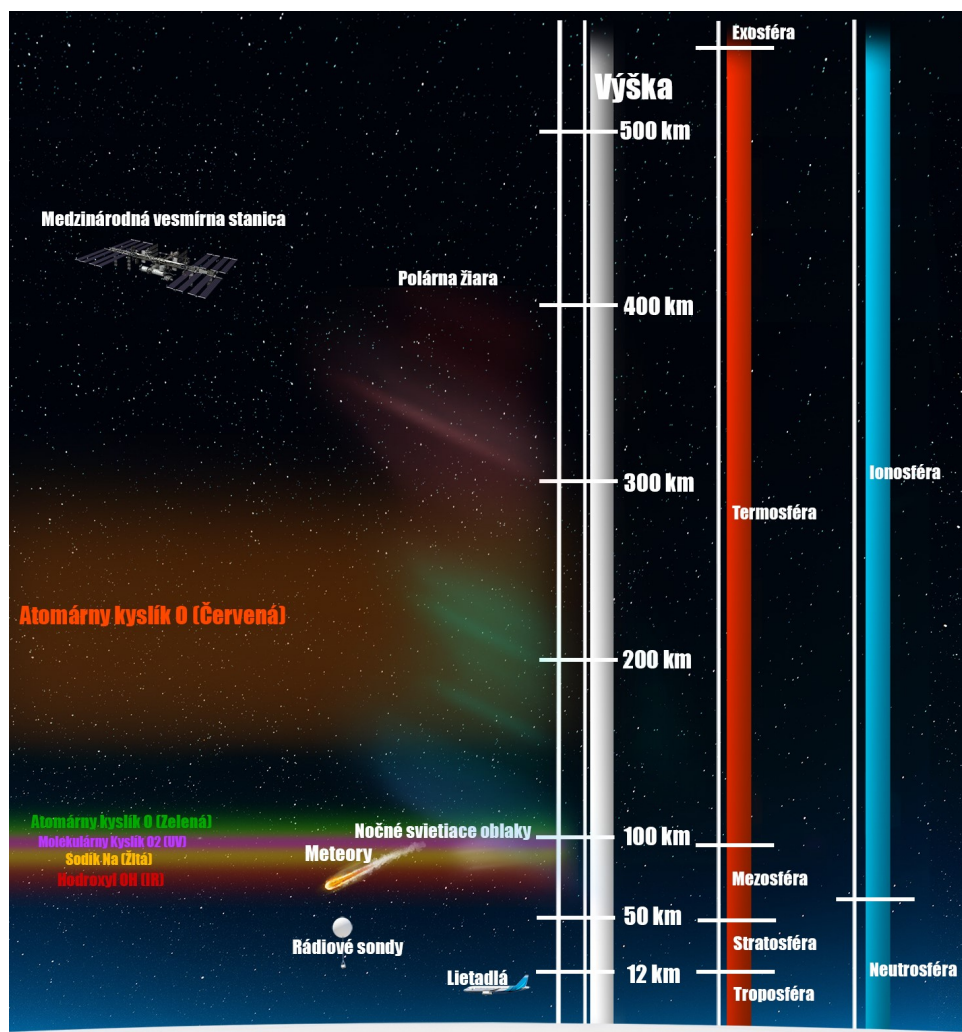
- **Emisie Sodíka** - Vyskytujú sa približne vo výške 80 km až 100 km nad zemským povrchom a pozorovať ich môžeme vo vlnovej dĺžke približne 590 nm, čo

predstavuje žltotooranžové sfarbenie.

- **Emisie hydroxylových radikálov** - Vyskytujú sa vo výškach od 87 km do 95 km. Pozorovateľné sú vo vlnovej dĺžke od 650 nm do 1800 nm, ich sfarbenie je teda červenej farby a tiež výrazne zasahujú do infračervenej časti spektra.
- **Emisie molekulárneho kyslíka** - Taktiež sa vyskytujú vo výške 87 km do 95 km a pozorovať sa dajú vo vlnovej dĺžke 120 nm až 490 nm, čo predstavuje ultrafialové, fialové až modré sfarbenie.
- **Emisie atómového kyslíka** - Nachádzajú sa vo výške 100 km a vyššie pričom sú pozorovateľné v 3 rôznych vlnových dĺžkach. Pri 558 nm, čo predstavuje zelenú farbu, 630 nm a 632 nm, čo je červené sfarbenie.

Spomenuté rozdelenie je možné vidieť aj na Obrázku 1–1. Pozorovanie tohto javu v noci je síce možné, ale stále je jeho viditeľnosť slabá. Žiara, ktorou svieti na povrch Zeme je približne rovnaká ako keby sme sa snažili osvetľovať povrch Zeme sviečkou vo výške 91 metrov. V prípade pozorovania v infračervenej oblasti ide o osvetlenie približne 1000-krát silnejšie vďaka čomu je jednoduchšie ho pozorovať (Gattinger, 1968).

Meranie javu Airglow je potrebné z dôvodu skúmania a predpovede zmien v ionosfére a strednej atmosfére. Tieto merania a pozorovania môžu poskytnúť informácie o vplyve globálnych faktorov, ako je napríklad zvýšený obsah oxidu uhličitého či metánu na dynamiku atmosféry. Analýzou emisií Airglow môžu výskumníci posúdiť citlivosť žiarenia na rôzne globálne zmeny a študovať mechanizmy zmien tohto javu. Jedným z kľúčových potenciálnych využití skúmania a predpovedania Airglow je zisťovanie trendov súvisiacich so zmenami teploty a zloženia v mezosfére ako aj spodnej termosfére. Tieto merania môžu pomôcť sledovať zmeny koncentrácie atómového kyslíka a rotačných teplôt, čo ponúka cenné informácie o dynamike atmosféry a dlhodobých trendoch. Pozorovanie a skúmanie je možné hlavne dvoma



**Obrázok 1 – 1** Atmosféra Zeme a výšky javu Airglow s ich zafarbeniami. Zdroj: (Slovinský and Mackovjak, 2020)

spôsobmi, a to vďaka pozemnému meraniu alebo meraniu pomocou satelitných prístrojov (Toshinori Maihara and et al., 1993).

Airglow sa podľa uhlu dopadu slnečných lúčov na povrch našej planéty a času merania delí na:

- **Dayglow** – Pozorovaný počas dňa v dôsledku interakcie medzi slnečným žiarením a zložkami atmosféry. Slnečné lúče dopadajú na povrch zeme priamo a je osvetlená celá pozorovaná časť atmosféry našej planéty. Emisie denného Airglow sa často pozorujú v blízkom infračervenom spektre a zohrávajú klú-

čovú úlohu pri pochopení energetickej dynamiky a zloženia horných vrstiev atmosféry (Chakrabarti et al., 1983).

- **Nightglow** – Emisie pretrvávajúce v horných vrstvách atmosféry aj po západe Slnka. Žiarenie je primárne spôsobené chemickými reakciami, ktoré pokračujú aj v neprítomnosti priameho dopadu slnečného svetla, poprípade len malého množstva slnečných lúčov. Toto vedie k emisii svetla so špecifickými vlnovými dĺžkami. Slnečné lúče pri tomto jave dopadajú na povrch Zeme, pozorovanú časť atmosféry, pod uhlom menším ako  $-18^\circ$ , čo znamená že jav sa dá pozorovať len počas takzvanej astronomickej noci. K emisiám tohto javu prispievajú procesy ako rekombinácia iónov, chemiluminiscencia a radiačná rekombinácia elektrónov. Nočné pozorovanie môže poskytnúť cenné informácie o zložení a dynamike horných vrstiev atmosféry počas nočných hodín, keď chýba priame slnečné žiarenie (Silverman, 1970).

História výskumu javu Airglow sa začína začiatkom 19. storočia, pričom poznatky o existencii tohto javu boli pravdepodobne známe už pred rokom 1800. Jedným z prvých míľnikov v pozorovaní tohto javu bolo Yntemovo fotometrické stanovenie vzdušného svitu, kedy sa však tento jav nazýval Earthlight (Yntema, 1909). Prvotné výskumy sa zameriavali na určenie vplyvu rôznych svetelných zdrojov k svietivosti nočnej oblohy a zaoberali sa problémami izolácie pozemských atmosférických emisií od svetelných zdrojov z vesmíru, ako sú napríklad hviezdy. Významný pokrok bol spôsobený zlepšením pozorovacích nástrojov, ktoré umožnili podrobnú analýzu emisií Airglow v rôznych vrstvách atmosféry omnoho detailnejšie. Následný výskum bol rozšírený o oblasti ionosféry ovplyvňujúce Airglow, úlohu ozónovej chémie na tento jav a zmeny intenzity spôsobené Slnečnými a Mesačnými cyklami. Vďaka rôznym misiám a kampaniam sa podarilo uskutočniť aj výskum Airglow mimo hraníc Zeme a tento jav bolo možné pozorovať aj na iných planétach, ako napríklad Marse. V súčasnej dobe je skúmanie možné vďaka rôznym pozemným staniciam ako aj satelitom zameraných na túto oblasť. Novou oblasťou tohto skúmania je taktiež využitie

strojového učenia pre predikciu výskytu Airglow, čo sa snaží priniesť nové poznatky a doplniť hodnoty tohto javu kedy skúmanie nie je možné (Savigny, 2017).

Na zachytávanie a sledovanie Airglow sa používajú 2 hlavné metódy získavania snímok. Prvou metódou sú stanice nachádzajúce sa na Zemi snímajúce nočnú oblohu. Zachytenie a pozorovanie tohto javu si vzhľadom na jeho jemnosť vyžaduje špecializované techniky a vybavenie. Hlavným nástrojom na tento účel sú celooblohové snímače (angl. *All-sky Airglow imagers*) vybavené špeciálnymi objektívmi, ktoré umožňujú širokohlé pozorovanie javu Airglow. Tieto snímače pracujú tak, že zhotovujú snímky nočnej oblohy s vysokým rozlíšením, pričom vďaka ich rozmiestneniu sa vyhýbajú svetelnému znečisteniu, aby sa zabezpečila viditeľnosť žiarenia. Ďalšou metódou sledovania a zaznamenávania hodnôt Airglow je snímanie zo satelitov na orbite Zeme. Satelity sú vybavené špeciálnymi snímačmi určenými pre zachytávanie hodnôt Airglow. Výhodou týchto snímačov je, že dokážu snímať podobu a hodnoty žiarenia aj počas zamračených nocí kedy by z povrchu našej planéty nebolo možné Airglow pozorovať. V dátach z týchto snímačov sa však často môžu vyskytovať oblasti, ktoré nie sú zamerané z dôvodu rotácie Zeme, ako to je napríklad pri snímači misie NASA GOLD (angl. *The Global-scale Observations of the Limb and Disk*, skr. GOLD) (Otsuka et al., 2004).

## 1.2 Celooblohové snímače Bostonskej Univerzity

Jeden z hlavných spôsobov snímania a skúmania Airglow je použitie pozemných snímačov oblohy. Využitie týchto prístrojov je jednoduchšie a menej finančne náročné ako použitie snímačov v satelitoch orbitujúcich našu planétu. Nevýhodou však je, že snímanie potrebuje jasnú nočnú oblohu pre získavanie použiteľných snímok. Bostonská univerzita prostredníctvom Centra pre vesmírnu fyziku využíva celooblohové snímacie zariadenia na štúdium rôznych javov v horných vrstvách atmosféry vrátane snímania podoby a hodnôt Airglow. Tieto snímače sú dôležité pri pozorovaní mezoféry, termosféry a ionosféry, pričom poskytujú kľúčový pohľad na dynamiku



a procesy prebiehajúce v atmosfére našej planéty. Snímače sú navrhnuté tak, aby zachytávali celú oblohu pomocou širokouhlého objektívu a poskytovali 180° zorné pole pohľadu. Snímky zachytené týmito prístrojmi sú obzvlášť užitočné na štúdium žiarenia v Zemskej atmosfére. Jednou z hlavných technických súčastí týchto prístrojov je zariadenie s nábojovou väzbou označované ako CCD (angl. *Charge-Coupled Devices*, skr. CCD), ktoré premieňa prichádzajúce fotónové signály na elektrické signály, vhodné na spracovanie a analýzu. Tieto technológie umožňujú vysokokvalitné zobrazovanie žiarenia v rôznych vlnových dĺžkach, od viditeľného svetla až po infračervenú oblasť. Konštrukcia snímača zvyčajne zahŕňa systém objektívov, optický systém pozostávajúci zo šošoviek a filtrov na výber konkrétnych vlnových dĺžok a detektor CCD. Filtre sú dôležité pre izoláciu špecifických javov, ktoré sú predmetom zájmu, ako napríklad rôzne hodnoty emisií Airglow. Údaje získané z celooblohových snímačov sa podrobujú rôznym technikám spracovania obrazu pre lepšiu presnosť snímok, ako napríklad odstránenie hviezd a iného šumu z pozadia vesmíru. Dáta z týchto snímačov sú dostupné na stiahnutie v rôznych filtroch a sú v podobe snímok nočnej oblohy (Smith et al., 2006).

Celooblohové snímacie zariadenia používané Bostonskou univerzitou sú umiestnené na niekoľkých kľúčových staniciach po celom svete, hlavne v oblastiach s menším výskytom svetelného smogu. Pre túto prácu sú dôležité stanice nachádzajúce sa v oblasti Južnej Ameriky z dôvodu, že ide o oblasť, nad ktorou je aj najviac dostupných snímok z dát NASA misie GOLD. Stanice, nachádzajúce sa v tejto oblasti použité v tejto práci sú:

- **Arecibo** - Observatórium je napriek zrúteniu svojho ikonického rádioteleskopu v roku 2020 významným miestom pre výskum atmosféry. Nachádza sa v Portoriku a táto celooblohová zobrazovacia jednotka sa využíva na nočný zber údajov o ionosférickom Airglow, čo prispieva k štúdiu dynamiky ionosféry v tejto oblasti.
- **Villa de Leyva** - Nachádza sa blízko mesta Boyacá v Kolumbii a zohráva kľú-

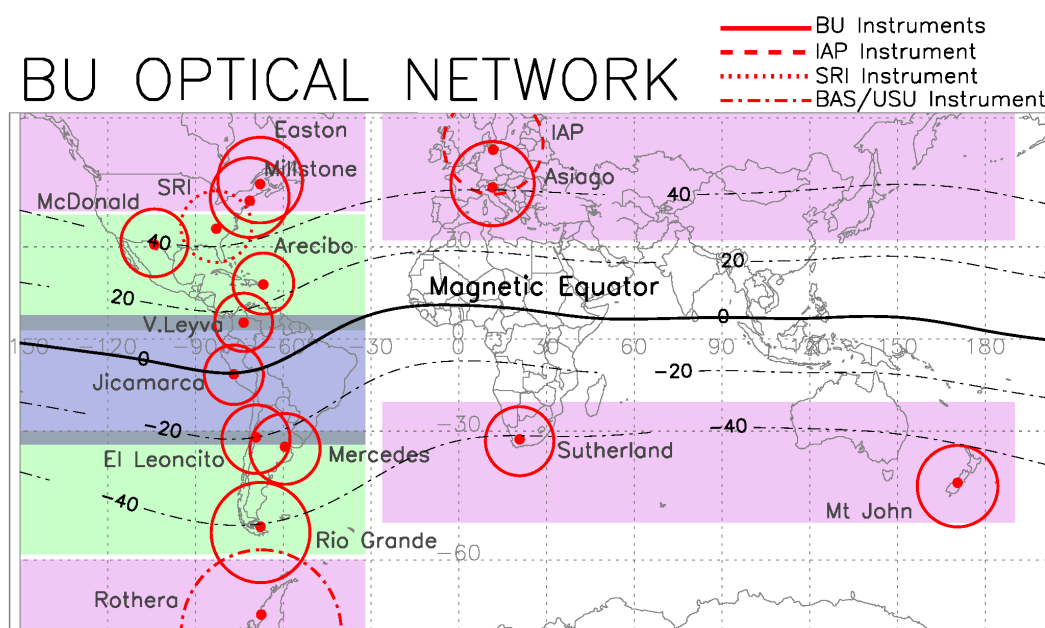
čovú úlohu pri pochopení šírkového rozsahu a pohybu ionosférických porúch.

- **El Leoncito** - Snímač je umiestnený v provincii San Juan v Argentíne a je významný pre pozorovanie horných vrstiev atmosféry na južnej pologuli. Táto stanica spolu s ďalšími v regióne poskytuje údaje o ionosférickej dynamike ovplyvňovanej faktormi, ako je napríklad juhoatlantická magnetická anomália.
- **Jicamarca** - Rádiové observatórium Jicamarca sa nachádza neďaleko Limy v Peru a je známe svojimi radarovými pozorovaniami ionosféry.
- **Mercedes** - Stanica Mercedes sa nachádza v provincii Buenos Aires v Argentíne a spolu so stanicou El Leoncito pracuje na štúdiu ionosférických porúch v geomagneticky konjugovaných bodoch voči staniciam Arecibo a Villa de Leyva.
- **Rio Grande** - Táto stanica sa nachádza v najjužnejšej oblasti Argentíny. Má ideálnu polohu na pozorovanie horných vrstiev atmosféry a ionosférických javov nad južným pólom vrátane Airglow a iných javov vesmírneho počasia.

Tieto stanice sú súčasťou globálnej siete observatórií, ktoré umožňujú výskumníkom z Bostonskej univerzity monitorovať a študovať atmosférické javy v rôznych zemepisných šírkach a dĺžkach. Rozmanitosť týchto miest je kľúčová pre pochopenie globálnej dynamiky horných vrstiev atmosféry vrátane zmien v jave Airglow, ionosférických porúch a vplyvu vesmírneho počasia na životné prostredie Zeme (*Aeronomy*, n.d.). Rozmiestnenie týchto staníc je možné vidieť na Obrázku 1 – 2.

### 1.3 NASA misia GOLD

Ďalším spôsobom sledovania Airglow je snímanie pomocou prístrojov na satelitoch orbitujúcich našu planétu. Na rozdiel od pozemných snímačov získavajú informácie o hodnotách toho javu z opačnej strany. Pri ich pozorovaní smerom na Zem nemajú problém s oblačnými nocami, pretože sa Airglow nachádza až nad hranicou oblakov. Misia GOLD je významným pokrokom v snahe pochopiť vrstvy atmosféry našej



**Obrázok 1 – 2** Rozmiestnenie staníc celoblokových staníc. Zdroj: Bostonská univerzita

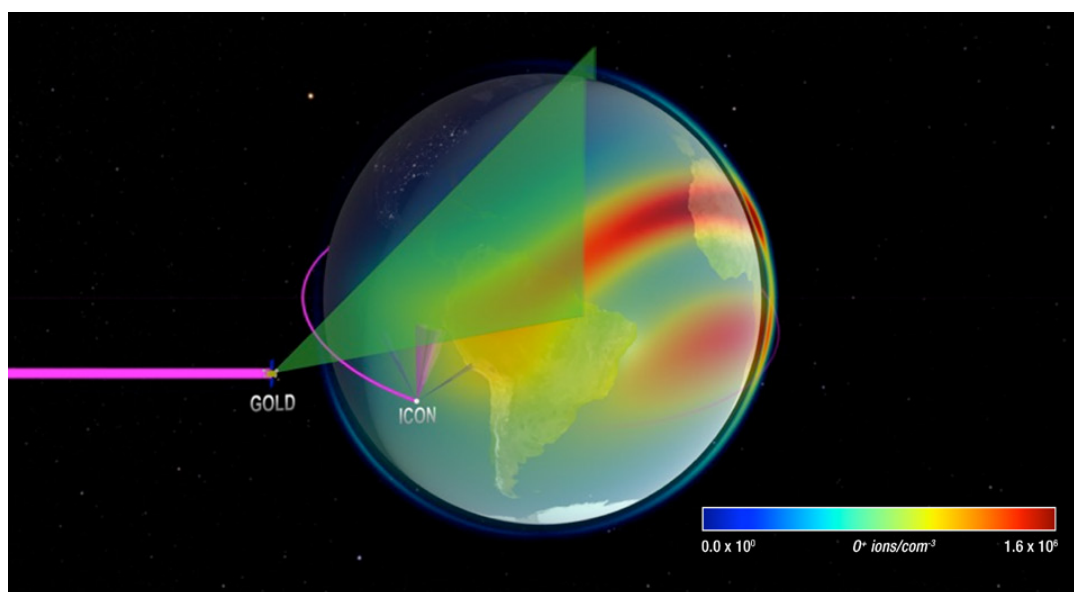
planéty ako aj okraj, kde sa stretáva vesmír s našou planétou. Misia GOLD, ktorú iniciovala NASA ako misiu príležitosti, bola vypustená na geostacionárnu obežnú dráhu na palube komerčnej komunikačnej družice SES-14 začiatkom roka 2018. Ide o komerčnú družicu, čo predstavuje strategickú spoluprácu medzi vedeckými cieľmi a komerčnými podnikmi. Cieľom tejto misie je preskúmať termosféru a ionosféru, ktoré sú kritickými zložkami zemskej atmosféry a zohrávajú kľúčovú úlohu v rádiovkej komunikácii, satelitnej navigácii a celkovom stave vesmírneho prostredia našej planéty. Keďže sa GOLD nachádza vo výške viac ako 35 000 kilometrov nad zemou, má jedinečnú pozíciu na pozorovanie dynamického vzájomného pôsobenia medzi hornými vrstvami atmosféry a prívodom slnečného žiarenia ako aj vplyvu počasia v nižších vrstvách atmosféry. Umiestenie a spôsob snímania misie je možné vidieť na Obrázku 1 – 3. Misia GOLD má niekoľko hlavných cieľov:

- Skúmať, ako horné vrstvy atmosféry reagujú na slnečnú aktivitu a geomagnetické búrky.
- Pochopiť vplyv atmosférických veľkoškálových vln a vln vychádzajúcich z niž-

šieho prostredia.

- Skúmať premenlivosť systému termosféra-ionosféra za rôznych podmienok.
- Zlepšiť našu schopnosť predpovedať udalosti vesmírneho počasia s priamymi dôsledkami pre technológie a komunikáciu na Zemi.

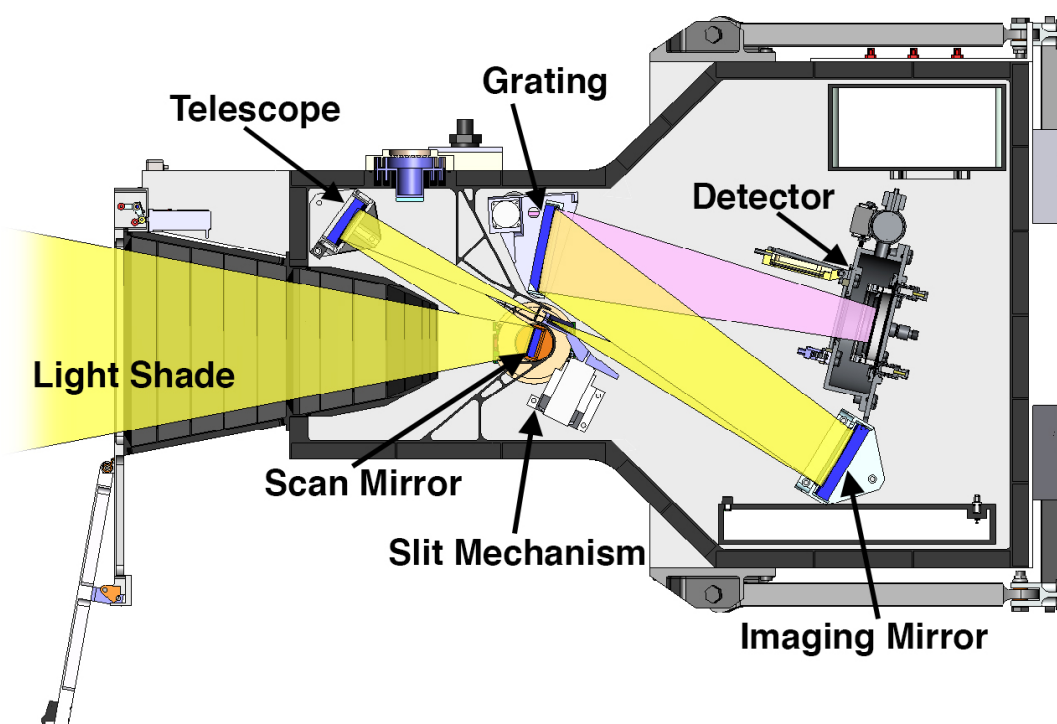
Prostredníctvom komplexných pozorovaní misia pomôže objasniť zložité procesy na pozadí riadenia správania termosféry a ionosféry, a tým prispieť k nášmu širšiemu chápaniu systému a vzájomného pôsobenia medzi Zemou a vesmírom. Táto misia sa zaoberá nielen základnými vedeckými otázkami, ale usiluje sa aj o poskytnutie využiteľných poznatkov pre zlepšenie odolnosti moderných technologických systémov voči výkyvom vesmírneho počasia (Eastes et al., 2017).



**Obrázok 1 – 3** Snímanie pomocou jedného z kanálov prístroja misie GOLD. Zdroj: NASA

Snímač misie GOLD je navrhnutý na pozorovanie zemskej termosféry a ionosféry s vysokou presnosťou aj vďaka svojej schopnosti pracovať v rôznych kanáloch v ultrafialovom spektre. Tento prístroj využíva viacero kanálov, ktoré mu umožňujú dôkladne analyzovať zloženie, teplotu a hustotu plynov v horných vrstvách atmosféry zachytením ich jedinečných ultrafialových signatúr. Nákres snímača je možné vidieť

na Obrázku 1 – 4. Dvojkanálová konfigurácia spektrografu je kľúčovou vlastnosťou, ktorá zvyšuje jeho pozorovacie schopnosti. Každý kanál je vybavený vlastnou sadou detektorov a optiky pre nezávisle zachytávanie snímok a spektier. Táto konfigurácia nielen zdvojnásobuje rýchlosť získavania údajov, ale zabezpečuje aj redundanciu a prevádzkovú spoľahlivosť, keďže každý kanál môže slúžiť ako záloha pre druhý. Okrem toho sa tieto kanály dajú nastaviť na špecifické vlnové dĺžky, čo umožňuje spektrografu selektívne sa zamerať na emisie, ako napríklad atómového kyslíka a molekulárneho dusíka. Schopnosť spektrografu prepínať medzi pozorovacími režimami a kanálmi mu taktiež umožňuje prispôbiť sa rôznym vedeckým cieľom a meniacim sa podmienkam prostredia. Táto flexibilita je kľúčová pre ciele tejto misie (*Gold Instrument*, n.d.).



**Obrázok 1 – 4** Nákras snímača misie GOLD a všetky jeho časti. Zdroj: NASA

Prínos tejto misie za jej zatiaľ krátke pôsobenie mal veľký dopad na výskum našej atmosféry. Jedným z najdôležitejších prínosov misie GOLD bola schopnosť sledovať každodenné zmeny v atmosfére s 30-minútovou frekvenciou pozorovania. Toto vysoké

časové rozlíšenie umožnilo vedcom pozorovať následky slnečných erupcií, geomagnetických búrok a iných vplyvov vesmírneho počasia na horné vrstvy atmosféry v reálnom čase. Údaje odhalili, ako tieto udalosti môžu viesť k dramatickým zmenám teploty, hustoty a zloženia v termosfére, čo môže následne ovplyvniť štruktúru a správanie ionosféry. Pozorovania misie GOLD objasnili napríklad aj mechanizmy vzniku a vývoja rovníkových ionizačných anomálií a rovníkových plazmových bublín, tieto javy môžu mať vážne následky, pretože dokážu narušiť signály GPS a rádiovú komunikáciu. Ďalším prínosom tejto misie bolo objasnenie interakcie medzi počasím u nás na Zemi a tým, ktoré sa deje vo vesmíre. Koreláciou atmosférických vln a prílivu z nižších vrstiev atmosféry so zmenami v termosfére-ionosfére GOLD ukázal, ako môže pozemské počasie ovplyvňovať javy vesmírneho počasia. Toto prepojenie medzi jednotlivými vrstvami je dôležité pri pochopení celého spektra faktorov spôsobujúcich zmeny vo vyšších vrstvách atmosféry. Globálna perspektíva systému misie taktiež poskytla cenné poznatky o pozdĺžnych zmenách v atmosfére. Pozorovaním nad západnou pologulou GOLD identifikoval významné lokálne rozdiely v reakcii na slnečnú a geomagnetickú aktivitu. Údaje z misie GOLD nielenže rozšírili vedecké poznatky, ale majú aj praktický význam pre zlepšenie modelov predpovedania vesmírneho počasia. Začlenením pozorovaní z misie GOLD do týchto modelov chcú vedci zvýšiť presnosť predpovedí rôznych javov. Takéto zlepšenie predpovedí je nevyhnutné na ochranu satelitov, komunikačných sietí a navigačných systémov pred nepriaznivými podmienkami vesmírneho počasia. Dáta misie sú taktiež dostupné na stiahnutie na stránke projektu a sú často aktualizované keďže misia stále prebieha. Od spustenia misie boli tieto dáta taktiež niekoľkokrát aktualizované o zlepšenia vytvorené po spustení tejto misie Eastes et al. (2020).

## 2 Strojové učenie

Strojové učenie je metóda spracovanie dát pomocou výpočtovej techniky, ktorá sa snaží nájsť spojitosť a učí sa z poskytnutých dát. Existuje veľké množstvo princípov takeéhoto učenia určeného na rôzne situácie. S postupom času je tento spôsob spracovania údajov a dát čoraz lepší a častejšie využiteľný pre rôzne domény. Základom strojového učenia je vývoj algoritmov, ktoré analyzujú údaje, učia sa z nich a robia predpovede alebo rozhodnutia na základe dát, na ktorých boli tieto algoritmy naučené. Celý princíp trénovania spočíva v spätnej kontrole dát a následného upravovania modelu pomocou váh. Algoritmus, teda časom dokáže vytvárať presnejšie výsledky, čo však závisí od mnohých parametrov, ako napríklad množstva vstupných údajov alebo rýchlosti učenia. Avšak príliš dlhé učenie môže spôsobiť preučenie modelu. V takom prípade model dosahuje dobrú presnosť pri učení, no pri použití na nových dátach je presnosť omnoho menšia. Modely strojového učenia sa líšia od tradičných programovacích paradigiem, teda logika a rozhodnutia nie sú priamo dané tvorcom modelu. Pri strojovom učení má výpočtová technika schopnosť učiť sa vzory a odvodzovať poznatky bez toho, aby boli explicitne naprogramované na každú konkrétnu úlohu. Podstatou učenia je využívanie štatistických modelov a algoritmov na vytváranie matematických modelov založených na vzorových údajoch označovaných ako trénovacia množina dát. Tieto matematické modely vytvárajú predpovede alebo rozhodnutia bez toho, aby boli explicitne naprogramované na vykonávanie danej úlohy. Tento proces umožňuje strojom zlepšovať svoj výkon a výsledky vzhľadom na rozsiahlosť a množstvo dostupných údajov. Účinnosť týchto modelov závisí od kvality a množstva údajov, na ktorých sa môžu naučiť ako aj od sofistikovanosti algoritmov. Vstupné dáta, na ktorých sa daný model učí sú veľmi dôležitou zložkou, pretože malé množstvo dát môže spôsobiť, že model nebude veľmi presný, no zároveň aj príliš veľké množstvo dát môže spôsobovať pre model problémy. Najdôležitejšia je tým pádom kvalita dát, ktorá sa docieľi ich predspracovaním (Alpaydin, 2021).

Hlboké učenie je jeden z princípov strojového učenia a predstavuje významný

skok vpred v schopnosti algoritmov analyzovať a interpretovať komplexné údaje. Tento princíp využíva štruktúry známe ako neurónové siete, ktoré čerpajú inšpiráciu zo zložitých biologických sietí nachádzajúcich sa v ľudských mozgoch. Skladajú sa z viacerých vrstiev neurónov, ktoré sú navzájom prepojené. Neuróny v každej vrstve vykonávajú špecifické výpočty pomocou vstupných údajov a svoje výsledky odovzdávajú ďalšej vrstve. Spojenia medzi neurónmi sú vážené pričom váhy sa upravujú počas fáz tréningu s cieľom zlepšiť presnosť pri vytváraní predpovedí alebo klasifikácií. Funkčnosť neurónových sietí je hierarchická. Počiatočné vrstvy zvyčajne zvládajú základné rozpoznávanie vzorov, ako sú hrany alebo textúry. Ako údaje postupujú sieťou, ďalšie vrstvy interpretujú čoraz zložitejšie aspekty údajov. Napríklad stredné vrstvy môžu tieto hrany a textúry skladať do rozpoznateľnejších foriem, ako sú časti objektov. Nakoniec hlbšie vrstvy dosiahnu komplexnejšie chápanie, čo umožní sieti rozpoznať celé objekty alebo komplexné scenáre. Táto schopnosť hierarchického spracovania modelov hlbokého učenia im umožňuje riešiť úlohy zložitosti a rozsahu aké boli predtým nedosiahnuteľné, ako napríklad rozbor identifikácia objektov na obrázkoch s pozoruhodnou presnosťou. Aj pri tejto metóde je však dôležitá príprava údajov na učenie, pretože modely sú zložité a nie vždy sa dá jednoducho povedať podľa čoho sa model rozhodol o výsledku. Príkladom môže byť určovanie objektu na obrázku. Model sa môže príliš zamerať na pozadie namiesto samotného objektu, čo môže viesť k omylnej klasifikácii. Výpočtové modely sú základom strojového učenia a hlbokého učenia. Sú schopné identifikovať vzory v údajoch, ktoré sú príliš zložité alebo rozsiahle na to, aby ich ľudia dokázali rozoznať. Táto schopnosť robí zo strojového a hlbokého učenia výkonné nástroje pre širokú škálu aplikácií, od vylepšovania spotrebiteľských produktov až po vedecký výskum. V dnešnej dobe s narastajúcim množstvom dostupných údajov, ako aj schopností výpočtovej techniky potenciál strojového učenia a hlbokého učenia poskytovať poznatky a zlepšovať rozhodovacie procesy sa len zvyšuje. Tieto technológie sú síce revolučné, ale nie bez výziev. Vývoj presných modelov strojového učenia si vyžaduje nielen značné množstvo vopred dostupných údajov, ale aj značné odborné znalosti pri výbere vhodných



algoritmov, ladení ich parametrov a overovaní výkonnosti modelu. Výsledné modeli sú často krát veľmi matematicky zložité. Je teda problém povedať podľa čoho sa daný model rozhodol bez zdĺhavého procesu spätného skúmania. Tento problém môže vyvolávať etické otázky týkajúce sa zodpovednosti za výsledky modelov a ich využítí v oblastiach, kde je vysvetliteľnosť dôležitá, ako napríklad zdravotníctvo (Goodfellow et al., 2016).

História strojového učenia siaha až do 50. rokov 20. storočia kedy bola predstavená prvá neurónová sieť. Postupom rokov však postup tejto technológie napredoval veľmi pomaly. Dôvodom bol nedostatočný výkon výpočtovej techniky. Skutočný prelom nastal až začiatkom 21. storočia s nárastom výkonu vďaka grafickým procesorom a množstva dostupných dát. V tomto období sa hlboké učenie stalo omnoho viac rozšíreným spôsobom pre riešenie problémov, pričom neurónové siete dosiahli prelomové výsledky v úlohách, ako napríklad rozpoznávanie obrazu a podobne. V súčasnosti sú strojové učenie a hlboké učenie neoddeliteľnou súčasťou širokého spektra aplikácií v rôznych odvetviach a disciplínach. Vo vedách o atmosfére a vesmíre prináša strojové učenie revolúciu v tom, ako výskumníci analyzujú a interpretujú rozsiahle súbory údajov zo satelitov, klimatických modelov a iných zdrojov. Algoritmy strojového učenia sú v súčasnosti rozhodujúcimi nástrojmi na predpovedanie vývoja počasia, pochopenie zmeny klímy, analýzu povrchu planét a zlepšenie nášho chápania vesmíru. Vďaka identifikácii zložitých vzorcov v údajoch, ktoré by človek nedokázal rozoznať, umožňuje strojové učenie nové poznatky a objavy v tak rozmanitých oblastiach, ako je astrofyzika alebo monitorovanie životného prostredia. Pomocou prediktívnych algoritmov je taktiež možné získavať nové dáta, vďaka ktorým je možné získať širší prehľad o rôznych javoch v atmosfére našej Zeme (Fradkov, 2020).

Spôsoby strojového učenia možno rozdeliť do niekoľkých typov, pričom každý z nich má svoj jedinečný prístup a oblasť použitia. Medzi základné typy strojového učenia patria:

- **Kontrolované učenie** - Najrozšírenejšia forma strojového učenia, pri ktorej

sa algoritmus učí z označeného súboru údajov. Počas tréovania algoritmu vytvára predpovede na základe vstupných údajov, ktoré sú následne kontrolované so správnymi výsledkami. Model podľa úspešností predpovedí upraví váhy a tento proces opakuje. Medzi bežné aplikácie patria regresné a klasifikačné úlohy, ako napríklad detekcia spamu alebo predpovedanie cien akcií (Nasteski, 2017).

- **Nekontrolované učenie** - V prípade nekontrolovaného učenia sa algoritmus trénuje na údajoch bez označení o aké dáta ide a čo v nich má hľadať. Model sa snaží určitým spôsobom usporiadať údaje tak, aby opísal ich štruktúru. Model sa snaží v tréovacích dátach hľadať súvislosti bez toho, aby vedel či tam nejaké súvislosti existujú. Môže to znamenať zhlukovanie, identifikácia skupín údajov alebo odhad hustoty. Medzi aplikácie patrí segmentácia zákazníkov, detekcia anomálií a organizácia veľkých knižníc obsahu do rôznych kategórií (Usama et al., 2019).
- **Učenie odmenou a trestom** - Spôsob, pri ktorom sa agent učí robiť rozhodnutia vykonávaním určitých akcií v prostredí na dosiahnutie cieľov. Agent sa učí na základe výsledkov svojich akcií a nie na základe explicitného učenia. Často sa používa v robotike, hrách, navigácii a pri optimalizácii prideľovania zdrojov. Princíp spočíva v hodnotení výsledkov modelu. Každý výstup je hodnotený oproti tomu predošlému či išlo o zlepšenie alebo nie. Algoritmus snaží pri učení vytvoriť optimálne riešenie problému (François-Lavet et al., 2018).

Medzi kľúčové architektúry v hlbokom učení patria:

- **Konvolučné neurónové siete** - Špecializujú sa na analýzu obrazových údajov a stoja za pokrokom v oblasti rozpoznávania a spracovania obrazu. Využívajú vrstvy filtrov, ktoré automaticky zisťujú vlastnosti, ako sú hrany v počiatočných vrstvách a zložitejšie vlastnosti, ako sú tvary alebo objekty v hlbších vrstvách. Tieto siete počas konvolučného procesu aplikujú na vstupný

obraz rôzne filtre, čím zmenšujú veľkosť obrazu a zameriavajú sa na najdôležitejšie detaily. Potom nasledujú vrstvy združovania, ktoré ďalej komprimujú obrazové údaje a zjednodušujú informácie bez straty kritických vlastností. Sú schopné identifikovať tváre a objekty s vysokou presnosťou, vďaka čomu sú nevyhnutné pre aplikácie v bezpečnostných systémoch alebo autonómnych vozidlách (Albawi et al., 2017).

- **Rekurentné neurónové siete** - Na rozdiel od iných neurónových sietí sa vyznačujú schopnosťou prepojiť predchádzajúce informácie s aktuálnou úlohou, čo ich robí ideálnymi pre aplikácie, kde je kontext kľúčový. Tento pamäťový efekt sa dosahuje prostredníctvom slučiek v sieti, ktoré umožňujú zachovanie informácií. V každom kroku sekvencie RNN (angl. *Recurrent Neural Network*, skr. RNN) berie do úvahy nielen aktuálny vstup, ale aj informácie, ktoré si uchovala z predchádzajúcich vstupov. Táto vlastnosť umožňuje RNN vykazovať časovo dynamické správanie. V dôsledku toho sa využívajú pri preklade jazyka a v prediktívnom písaní. Tréningom na veľkých súboroch údajov sa naučia predpovedať ďalšiu pravdepodobnú položku v sekvencii, vďaka čomu sú efektívne pri úlohách, ako je navrhovanie ďalšieho slova vo vete (Sherstinsky, 2020).
- **Generatívne adverzné siete** - O tomto type neurónových sietí píšeme viac v kapitole 2.1.

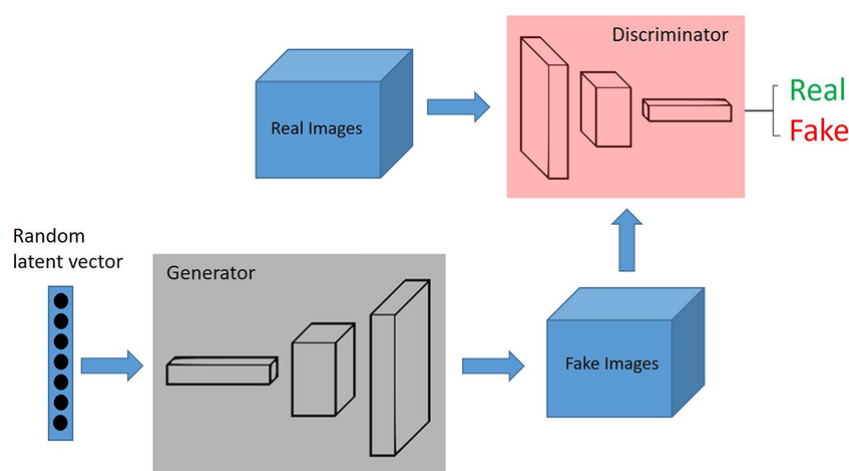
## 2.1 GAN

Generatívne adverzné siete (angl. *generative adversarial network*) sú typom hlbokého učenia navrhnutého na generovanie čo najrealistickejších údajov, a preto sa často používajú práve na generovanie snímok. Tento prístup využíva nekontrolované učenie pomocou neurónových sietí. Koncept GAN, ktorý prvýkrát navrhol Ian Goodfellow a jeho tím v roku 2014 (Goodfellow et al., 2014). Tento rámec má korene v teórii hier s nulovým súčtom, kde sa dva modely neurónových sietí zapájajú do nepretržitého

konkurenčného procesu. Táto súťaž poháňa oba modely k iteratívnemu zlepšovaniu ich metód generovania a vyhodnocovania údajov bez toho, aby vyžadovali označené súbory údajov. Každý z týchto modelov sa teda snaží byť lepším ako ten druhý. Na rozdiel od kontrolovaného učenia, ktoré sa spolieha na údaje so známymi výsledkami, GAN sa učia napodobňovať a inovovať na základe komplexných súborov údajov autonómne. Tento rámec otvoril nové cesty v oblasti učenia bez dohľadu, keďže na rozdiel od tradičných metód učenia s dohľadom nevyžadoval na trénovanie označené súbory údajov. V priebehu rokov sa GAN dočkali mnohých vylepšení a variácií, ako napríklad podmienené GAN alebo hlboké konvolučné GAN, pričom každá z nich rieši špecifické výzvy, ako je stabilita tréningu, kolaps režimu a generovanie obrazov s vysokým rozlíšením. Všestrannosť GAN viedla k ich uplatneniu v širokej škále oblastí. V oblasti spracovania obrazu a počítačového videnia sa GAN používajú napríklad na úlohy, ako je preklad obrazu na obraz. Okrem toho zohrávajú GAN kľúčovú úlohu pri rozširovaní údajov, čím zvyšujú rozmanitosť súborov údajov používaných na trénovanie iných modelov strojového učenia a tým zlepšujú ich presnosť a robustnosť. V oblasti skúmania vesmíru sa dajú využiť na zvýšenie kvality snímok zachytených rôznymi prístrojmi alebo na generovanie nových údajov na miestach, kde ich je menší počet a doplniť tak chýbajúce časti (Goodfellow et al., 2020).

GAN sa skladá z dvoch neurónových sietí, a to generátora a diskriminátora. Fungovanie GAN závisí od ich dynamickej interakcie medzi sebou. Generátor vytvára syntetické údaje z náhodného vstupného šumu s cieľom kopírovať podobu skutočných údajov. Diskriminátor, ktorý je často binárny klasifikátor, sa snaží klasifikovať vstupné údaje buď ako skutočné alebo ako generované. Počas procesu trénovania sa generátor a diskriminátor iteratívne zlepšujú na základe ich vzájomného výkonu, pričom sa na aktualizáciu ich váh používajú metódy spätného šírenia a gradientného zostupu. Cieľom je dosiahnuť bod, v ktorom diskriminátor už nedokáže spoľahlivo rozlíšiť generované údaje od skutočných údajov, čo znamená, že generátor sa účinne naučil distribúciu skutočných údajov. Celý proces sa dá vysvetliť ako falzifikácia

umenia, kde sa generátor snaží vytvoriť čo najlepší falzifikát originálneho diela a diskriminátor sa snaží určiť či ide o falzifikát alebo nie. Tento adverzný proces učenia je charakterizovaný stratovou funkciou, ktorá kvantifikuje výkonnosť generátora aj diskriminátora. Tréning zahŕňa striedavú optimalizáciu diskriminátora s cieľom maximalizovať jeho presnosť pri identifikácii skutočných a falošných údajov a aktualizáciu generátora s cieľom minimalizovať schopnosť diskriminátora rozlišovať generované údaje od skutočných údajov. Rovnováha tohto procesu, keď generátor vytvára dokonalé repliky skutočných údajov a diskriminátor má šancu 50/50 správne identifikovať skutočné a falošné údaje, znamená úspešné trénovanie GAN (Mogren, 2016). Tento proces medzi generátorom a diskriminátorom je možné vidieť na Obrázku 2 – 1.



**Obrázok 2 – 1** Diagram fungovania GAN pri jeho učení. Zdroj: (Serag et al., 2019)

## 2.2 Pix2Pix

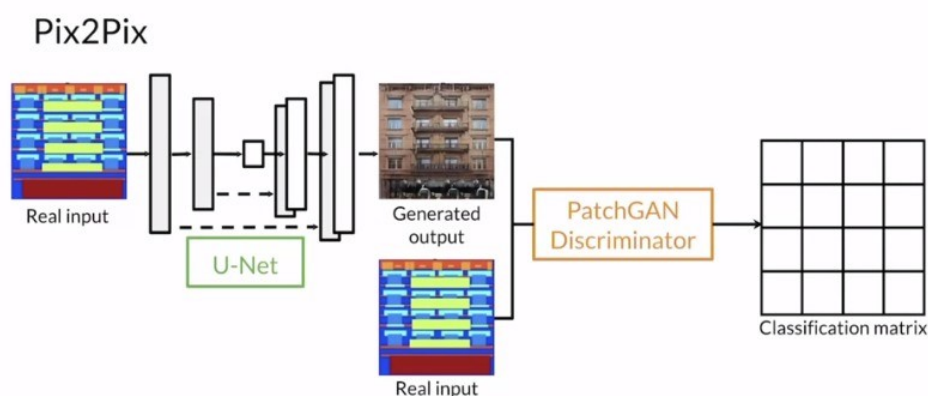
Pix2Pix je generatívna adverzná sieť navrhnutá na úlohy prekladu obrazu na obraz, teda jej cieľom je generovať obraz podľa podoby iného obrazu. Tento model umožňuje transformáciu obrázkov z jedného štýlu alebo domény do inej, napríklad premenu náčrtov na fotorealistické obrázky, alebo konverziu satelitných snímok na mapy. Všestrannosť Pix2Pix otvára nespočetné množstvo aplikácií v rôznych oblas-

tiach. Ako je možné vidieť na Obrázku 2–2, základný mechanizmus Pix2Pix zahŕňa rovnako ako iné GAN modely dva hlavné komponenty:

- **Generátor** - Pix2Pix využíva architektúru založenú na U-Net, ktorá je charakteristická svojou štruktúrou kodér-dekodér s vzájomnými spojeniami. Tieto spojenia spájajú zodpovedajúce vrstvy v kodéri a dekodéri, čo umožňuje priamy tok informácií v sieti. Táto konštrukcia pomáha zachovať dôležité detaily z vstupného obrazu a efektívne ich preniesť do generovaného obrazu, čo umožňuje vytvárať vysokokvalitné preklady. Zahrnutie centrálnej kompresnej vrstvy pomáha sieti pri kompresii a následnej dekompresii informácií, čo podporuje vytváranie nového obsahu pri zachovaní základných štrukturálnych detailov vstupného obrazu.
- **Diskriminátor** - Využíva PatchGAN, ktorý sa zameriava na posudzovanie pravosti malých častí obrazu a nie celého obrazu. Tento prístup umožňuje diskriminátoru efektívne zachytiť a analyzovať štruktúru a jemné detaily obrazu, čo z neho robí výkonný nástroj na rozlišovanie medzi skutočnými a falošnými obrazmi. Penalizovaním generátora za nezrovnalosti v mierke týchto políčok model zabezpečuje, že vygenerované obrázky nielenže vyzerajú celkovo realisticky, ale majú aj presvedčivé textúry a detaily.

Model Pix2Pix je možné použiť v rôznych odvetviach pracujúcich s obrazom. Priamočiarym príkladom je umenie kedy sa model dá využiť na transformáciu obrazov do iných štýlov alebo v prípade vygenerovaného videa pridanie väčšej reálnosti, prípadne generovanie pokračovania pohybu. Okrem sveta umenia zohráva Pix2Pix kľúčovú úlohu pri vylepšovaní fotografií. Ponúka možnosť pridať farbu pre čierno-biele fotografie alebo oživiť historické snímky. Okrem toho dokáže zvýšiť rozlíšenie obrázkov, čím sa stanú jasnejšími a detailnejšími alebo previesť denné scény na nočné scény. V oblasti medicíny je prínos Pix2Pix neoceniteľný. Pomáha pri konverzii 2D obrázkov na 3D modely, čím poskytuje komplexnejší pohľad na anatomické štruktúry, ktorý je kľúčový pre diagnostiku, vzdelávacie účely a plánovanie liečby.

Okrem toho môže model vylepšiť lekárske skeny, čím sa zviditeľnia detaily a pomôže zdravotníckym pracovníkom pri analýze a interpretácii zložitých snímok. Aplikácia Pix2Pix sa rozširuje aj na geografické informačné systémy. Urbanisti a pracovníci monitorujúci životné prostredie môžu model použiť na konverziu leteckých fotografií na podrobné mapy. Táto schopnosť je nevyhnutná na plánovanie mestskej infraštruktúry, monitorovanie zmien životného prostredia a efektívnejšie riadenie prírodných zdrojov (Isola et al., 2016).



**Obrázok 2 – 2** Princíp generovania obrazov pomocou Pix2Pix. Zdroj: (Isola et al., 2016)

## 2.3 Metriky

Porovnávanie podobnosti dvoch obrazov môže byť veľmi subjektívne v prípade porovnávania len pomocou pohľadu. Preto je dôležité pre zistenie podobnosti predikovaných obrazov voči pôvodným číselné vyjadrenie. Takéto objektívne porovnanie je možné vďaka použitiu metrík. V prípade obrazu je najbežnejším spôsobom porovnávanie každého pixelu medzi dvoma snímkami. Pre porovnávanie výsledkov nášho modelu sme vybrali nasledujúce metriky:

- **MSE (angl. *Mean Square Error*)** - Bežne používaným meradlom kvality obrazu. Je to plne referenčná metrika, čo znamená, že na výpočet sú potrebné

dva obrázky, a to pôvodný obrázok a testovací obrázok, ktorý môže byť skreslenou verziou pôvodného obrázku. MSE predstavuje priemer štvorcov rozdielov medzi príslušnými pixelmi týchto dvoch obrazov. Čím bližšie je hodnota MSE k nule, tým menšia je chyba medzi týmito dvoma obrazmi, čo znamená lepšiu kvalitu alebo menšie skreslenie (Willmott and Matsuura, 2005). Vzorec pre výpočet metriky je:

$$MSE(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (2.1)$$

kde  $x$  a  $y$  sú dva obrazy, pri ktorých  $i = 1, 2, \dots, N$ , kde  $N$  je počet pixelov a  $x_i$  a  $y_i$  sú hodnoty  $i$ -tých pixelov v  $x$  a  $y$ , v tomto poradí.

- **PSNR (angl. *Peak Signal-to-Noise Ratio*)** - Široko používaná metrika na meranie kvality obrazu, najmä pri stratovej kompresii obrazu. Hodnotí pomer medzi maximálnou možnou silou signálu a silou šumu, ktorý ovplyvňuje kvalitu jeho reprezentácie. PSNR sa zvyčajne vyjadruje v logaritmickú stupnici v decibeloch. Vyššie hodnoty PSNR naznačujú lepšiu kvalitu obrazu (Hore and Ziou, 2010). Vzorec pre túto metriku je:

$$PSNR = 20 \log_{10} \left( \frac{MAX_f}{\sqrt{MSE}} \right), \quad (2.2)$$

kde  $MAX_f$  je maximálna hodnota signálu, ktorá existuje v originálnom obraze, ktorý nebol predikovaný.



## 3 Modelovanie javu Airglow pomocou hlbokého učenia

### 3.1 Analýza súčasného stavu

S pokrokom vo vývoji výpočtovej techniky sa využitie a zakomponovanie metód strojového učenia stalo čoraz častejším. V posledných rokoch často označované ako AI (angl. *Artificial Intelligence*, skr. AI) sa so strojovým učením stretávame častejšie, ako si uvedomujeme. AI je súčasťou mnohých programov a systémov, s ktorými sa stretávame každý deň. Pokrok tejto technológie spôsobil všestranné využitie metód strojového učenia a inak tomu nie je ani pri využití pre vedecké účely. Pri výskume našej atmosféry a vesmíru je AI veľmi často používané napríklad na detekovanie objektov na snímkach. Pri pozemných teleskopoch môže ísť o detekovanie hviezd, planét alebo bleskov. Pri snímkach zo satelitov môže ísť o detekovanie rôznych vesmírnych objektov, ako sú napríklad čierne diery. Využitie metód strojového učenia pre výskum javu Airglow je menej rozšírené, avšak v posledných rokoch sa táto oblasť posunula výrazne a jeho využitie sa stalo o niečo častejším. Pre bližšiu analýzu využitia sme vybrali práce:

- **Metodika získavania miery objemovej emisie z intenzity emisie Airglow žiarenia pozorovaného kombináciou techník Abelovej inverzie a hlbokého učenia** (Duann et al., 2022)

Zameraním tejto práce je zlepšenie procesu získavania množstva emisie vyprodukovanej z daného objemu za stanovený čas z pozorovaní intenzity airglowu v blízkosti okraja Zeme, čo je rozhodujúce pre štúdium ionosféry. Na riešenie tohto problému autori navrhujú metodiku, ktorá kombinuje techniky hlbokého učenia s Abelovou inverziou. Táto kombinácia umožňuje optimalizovať proces získavania dát vyplnením nepozorovaných alebo nespojitých pozorovaní v profiloch intenzity Airglow žiarenia, čím sa uľahčí ich použitie s Abelovou inverziou na určenie rýchlosti produkcie emisie. V štúdiu využili 3 algoritmy hlbokého

učenia GDX (angl. *Gradient descent with momentum and adaptive learning rate*, skr. GDX), SCG (angl. *Scaled conjugate gradient*, skr. SCG) a LM (angl. *Levenberg-Marquardt*, skr. LM) na presnú extrapoláciu neúplných profilov intenzity. Tieto modely boli použité a overované pomocou dát o intenzite emisie javu Airglow z misie NASA GOLD. Autori porovnávajú výkonnosť modelov hlbokého učenia s výsledkami získanými z Abelových transformovaných pozorovaní v reálnom čase. Najlepšie výsledky v práci mal GDX pri počte neurónov od 20 do 45. Najmenej úspešný bol algoritmus LM. V závere autori konštatovali, že by bol vhodný ďalší výskum a vyskúšanie aj iných modelov hlbokého učenia. Avšak použitie metód hlbokého učenia pri práci s javom Airglow a aj jeho využitie pri dopĺňaní chýbajúcich častí merania intenzít žiarenia je hodnotné a môže v budúcnosti pomôcť pri ďalšom výskume oblastí atmosféry našej planéty.

- **Analýza údajov 2D Airglow snímača s ohľadom na dynamiku pomocou strojového učenia** (Sedlak et al., 2023)

V tejto štúdii sa tvorcovia zamerali na použitie strojového učenia pri analýze veľkého množstva údajov zo snímača OH Airglow v ovzduší s ohľadom na dynamiku tohto javu. Bol tu použitý klasifikačný algoritmus TCN (angl. *Temporal Convolutional Network*, skr. TCN) na automatické triedenie snímok do troch tried: dynamické (snímky, ktoré obsahujú pohyby v malom meradle), pokojné (snímky s jasnou oblohou so slabými odchýlkami Airglow) a oblačné (snímky, na ktorých oblačnosť bráni analýze Airglow). Algoritmus dosiahol priemernú presnosť 0.82 pri klasifikácii obrázkov. Automatickou kategorizáciou snímok sa výskumníci zamerali na identifikáciu dynamických epizód, ktoré by potenciálne mohli vykazovať javy turbulencie. Ďalšou analýzou v rámci dynamickej triedy identifikovali podmnožinu sérií snímok vykazujúcich turbulenciu, čo umožnilo odhadnúť parametre turbulencie, ako napríklad rýchlosť šírenia energie. Vďaka výskumu poukázali na využiteľnosť neurónových sietí pre identifikáciu turbu-

lentných epizód a prospech využitia pri skúmaní javu Airglow.

- **Analýza klasifikácie Airglow obrazu na základe vizualizácie mapy prvkov** (Lin et al., 2023)

Podobne ako v predošlej práci, výskumníci využívajú hlboké učenie pre klasifikovanie snímok nočnej oblohy pri výskume Airglow. Dáta sú zachytené pomocou celooblohových snímačov a klasifikujú ich na jasné a nejasné noci. Použitá metóda hlbokého učenia je založená na CNN. Výskumný tím navrhol architektúru hlbkej CNN pozostávajúcej z desiatich vrstiev. Vstupná vrstva spracuje snímky v sivej škále do matice s rozmermi 128x128x1. Konvolučné vrstvy s aktivačnými funkciami ReLU sa učia extrahovať vlastnosti zo snímok. Výsledky štúdie ukazujú, že model CNN, ktorý vytvorili dosiahol vysokú presnosť klasifikácie 99%. Model bol natrénovaný na súbore údajov obsahujúcom 1 688 obrázkov. Proces klasifikácie bol vylepšený vizualizáciou máp príznakov, ktoré pomohli pri sledovaní procesu učenia a optimalizácii modelu identifikáciou a opravou problémov, ako je zameranie modelu na umelé svetlá na okrajoch snímok. Tento model mal lepšie výsledky zatiaľ čo využíval menej výpočtovej techniky, tým pádom sa im podarilo vylepšiť klasifikáciu snímok javu Airglow pomocou hlbokého učenia.

- **Klasifikácia bublín rovníkovej plazmy pomocou konvolučnej neurónovej siete a techniky podporného vektora** (Thanakulketsarat et al., 2023)

Predmetom tohto článku nie je priamo jav Airglow, ale ide o využitie metód strojového učenia na jav, ktorý je možné pozorovať v podobnej oblasti a veľa krát je zachytávaný rovnakými snímačmi. Práca je zameraná na klasifikáciu rovníkových plazmových bublín pomocou konvolučných neurónových sietí a techník podporných vektorov. Cieľom štúdie bolo vyvinúť systémy detekcie a klasifikácie bublín v reálnom čase na pomoc komunitě zaoberajúcej sa vesmírnym počasím. Modely sú natrénované s použitím snímok z radarového systému

na stanici Chumphon v Thajsku. Zistenia tejto štúdie ukázali, že kombinovaný model konvolučných neurónových sietí a techník podporných vektorov dosiahol v najlepšej verzii presnosť 93.08%. Touto prácou preukázali, že aplikácia metód hlbokého učenia by bola vhodná aj na iné oblasti výskumu Airglow.

- **Štatistické charakteristiky nočných stredne veľkých pohyblivých ionosférických porúch z 10-ročného pozorovania Airglow metódou strojového učenia** (Lai et al., 2023)

Výskum tejto práce sa zameriava na analýzu štatistického výskytu a charakteristík šírenia ionosférických porúch pomocou strojového učenia. Konkrétne sa v štúdiu analyzovali údaje o Airglow zozbierané od októbra 2011 do decembra 2021 v Xinglongu v Číne. Použité metódy hlbokého učenia využívali algoritmy na automatickú identifikáciu a extrakciu charakteristík šírenia a morfológických charakteristík ionosférických porúch v obrazoch Airglow s vlnovou dĺžkou 630 nm. V štúdiu použili konvolučné neurónové siete a algoritmy regionálnych konvolučných neurónových sietí na klasifikáciu a detekciu udalostí ionosférických porúch z veľkého súboru údajov. Klasifikačný model dosiahol presnosť 96.9% a detekčný model mal presnosť medzi 70% a 85%. Zistenia prispievajú k pochopeniu mechanizmov, ktoré stoja za vznikom a šírením ionosférických porúch a zdôrazňujú účinnosť metód strojového učenia pri analýze veľkých súborov údajov z atmosférických pozorovaní.

Po analýze článkov je jasné, že využitie strojového učenia pri skúmaní javu Airglow alebo javov v oblasti jeho výskumu je veľmi užitočné. Výskumníci využívajú tieto metódy pri rôznych oblastiach vrátane dopĺňania chýbajúcich častí dát. Pri práci s dátami Airglow žiadna z prác nepoužívala prepojenie pozemného snímania Airglow spolu so satelitným snímaním. Práve z tohto dôvodu je naša práca prínosom kedy pomocou kombinácie týchto dvoch snímaní a modelu hlbokého učenia Pix2Pix doplníme snímky hodnôt o dáta, ktoré by inak nebolo možné získať.

### 3.2 Pochopenie cieľa

Cieľom tejto práce je predikovanie hodnôt javu Airglow pomocou hlbokého učenia, konkrétne neurónových sietí. Ďalším dôležitým aspektom tejto práce je kombinácia dát o Airglow z 2 rôznych zdrojov. Ide o kombináciu snímok vytvorených pomocou pozemných celooblohových kamier a snímok vyhotovených misiou NASA GOLD, ktoré pochádzajú zo satelitných snímačov. Oblasť výskytu Airglow je často skúmaná oblasť kvôli jej dôležitosti pre výskum rôznych oblastí Atmosféry. Dopĺňanie chýbajúcich hodnôt nameraných emisií javu pomocou strojového učenia sa v posledných rokoch začalo využívať častejšie, avšak kombinácia dvoch rôznych zdrojov nie je bežná, a preto je prístup, ktorý vytvárame v tejto práci veľmi prospešný pre výskum Airglow. Táto práca nadväzuje na poznatky a zistenia počas bakalárskej práce. Cieľom tejto práce bolo taktiež využitie metód strojového učenia pre predikciu javu Airglow. Využité boli dáta misie GOLD, na ktoré bola aplikovaná neurónová sieť, pričom výsledkom bolo predikcia dát ako pokračovanie sekvenvie snímok zachytených touto misiou. Úlohou tejto práce je teda rozšíriť nadobudnuté poznatky a otestovať prístup predikcie kedy sú na vstupe dáta z celooblohových snímačov. Myšlienka tohto spojenia dát je doplniť snímky javu Airglow z misie GOLD o dáta získané z iného zdroja, čím dokážeme doplniť hodnoty na miesta, ktoré nebolo možné predikovať predošlým spôsobom. Pre dosiahnutie cieľa bolo potrebné prejsť nasledujúcimi krokmi:

- **Príprava dát misie GOLD** - Dáta misie GOLD sme museli získať zo stránky misie a následne ich upraviť do grafickej podoby. Zo získaných snímok bolo následne potrebné získať aj ich hodnoty v oblasti snímačov Bostonskej Univerzity. Celý proces získavania a vytvárania grafickej podoby bol vytvorený ako súčasť bakalárskej práce, na ktorú nadväzujeme. Práca s týmito dátami už berie do úvahy uskutočnenie procesu, ktorý bol súčasťou bakalárskej práce.
- **Stiahnutie dát zo snímačov Bostonskej Univerzity** - Snímky z celooblohových snímačov sú dostupné na stránke projektu Bostonskej Univerzity.

Stránka ale neobsahuje priamu možnosť sťahovania veľkého množstva údajov. Našou úlohou bolo prísť na spôsob, ako ich stiahnuť za požadované obdobia.

- **Príprava dát z celooblohových snímachov** - Stiahnuté snímky zo snímachov bolo potrebné ďalej upraviť pre ich klasifikáciu ako aj pre vstup do modelu pre dopĺňanie dát misie GOLD. Tento proces zahŕňal hlavne odstránenie nežiadanych častí dát, ako sú ich okraje a odstraňovanie hviezd na snímkach.
- **Klasifikácia snímok z celooblohových snímachov** - Pre prácu s dátami sme potrebovali získať snímky, ktoré neobsahujú žiadne prekážky, ako sú oblačky alebo mesiac. Samotná klasifikácia by zabrala dlhú dobu, a preto sme sa rozhodli natrénovať jednoduchý klasifikačný model, ktorý by nám mohol pomôcť. Klasifikácia však nie je cieľom tejto práce, preto sme sa nesnažili získať, čo najlepší model. Išlo len o pomoc pri získavaní vhodných snímok pre vstup do predikčného modelu.
- **Výber časového úseku pre predikciu** - Pri vstupe do neurónovej siete pre predikciu boli potrebné len snímky s jasnou oblohou. Po získaní klasifikovaných dát bolo potrebné zistiť časový úsek, v ktorom sa nachádzalo najviac nocí s jasnou oblohou. Takýmto spôsobom bolo potrebné skontrolovať niekoľko rokov dát.
- **Výber a explorácia modelu pre predikciu** - Kvôli práci s grafickými dátami bolo potrebné vybrať správny model. Tieto dáta sa nachádzajú ako na vstupe tak aj na výstupe. Ide teda o preklad obrazu na obraz. Neurónové siete a obzvlášť GAN sú v tejto oblasti úspešné, preto sme sa zmerali hlavne na túto oblasť strojového učenia. Po výbere modelu bolo potrebné zistiť jeho optimálne nastavenia. Natrénovali sme niekoľko rôznych verzií algoritmu, aby sme získali čo najlepšie výsledky predikcie dát misie GOLD.
- **Vyhodnotenie výsledkov** - Pri porovnávaní podobnosti predikovaných a pôvodných snímok bolo potrebné nájsť vhodné metriky, ktorými ich môžeme

porovnať na úrovni pixelov. Výber bol taktiež závislý na porovnávaní s výsledkami z bakalárskej práce. Bolo potrebné vyhodnotiť získané výsledky s predošlými získanými predikciami ako na úrovni presnosti, tak zároveň aj podľa toho či je s nimi možné predikovať oblasti, ktoré neboli pôvodne dostupné.

### 3.3 Príprava dát

Jedným z dvoch zdrojov dát pre túto prácu sú snímky nočnej oblohy z celooblohových kamier Bostonskej Univerzity. Dostupné sú pomocou archívu dát, ktorý obsahuje roztriedené snímky podľa miesta, kde boli snímky vytvorené, pozorovanej vlnovej dĺžky, roku, mesiaca a následne dňa, kedy boli vyhotovené. Pre našu prácu boli dôležité dáta vlnovej dĺžky 6300. Ide o vlnovú dĺžku, pri ktorej je jav Airglow najlepšie pozorovateľný v ionosfére a zároveň emisia prichádza z rovnakých výšok ako emisia, ktorú pozoruje NASA GOLD. Stanice, ktorých snímky nočnej oblohy budeme v práci využívať sme vybrali observatória v Južnej Amerike. Tento výber bol dôležitý keďže dáta z druhého zdroja, misie NASA GOLD, sa zameriavajú práve na oblasť Južnej Ameriky a Atlantického oceánu. Z tohto dôvodu bolo potrebné stiahnuť dáta z lokácií Arecibo, Villa de Leyva, El Leoncito, Jicamarca, Mercedes a Rio Grande. Všetky parametre rozdelenia snímok v archíve sú súčasťou URL. Týmto spôsobom je aj pri chýbajúcej možnosti hromadného sťahovania možné stiahnuť väčšie množstvo snímok pomocou kódu. Vytvorili sme skript, pomocou ktorého je možné postupne prejsť URL každého dňa v roku. Po vytvorení URL smerujúceho na konkrétny deň je potrebné jednotlivé snímky stiahnuť pomocou metódy web scrapingu. Ide o metódu prechádzania celej štruktúry stránky za účelom hľadania konkrétnych informácií. Celý kód sme vytvorili v programovacom jazyku Python a sťahovanie snímok bolo uskutočnené cez knižnicu `BeautifulSoup`. Naprogramovaná funkcia sa pozrie na dané URL stránky a pozbiera všetky URL pre obrázky na stránke. To však zahŕňa aj iné obrázky ako len snímky nočnej oblohy, ako napríklad logo archívu. Snímky nočnej oblohy sú vo formáte `.gif`, vďaka čomu výber správnych obrázkov

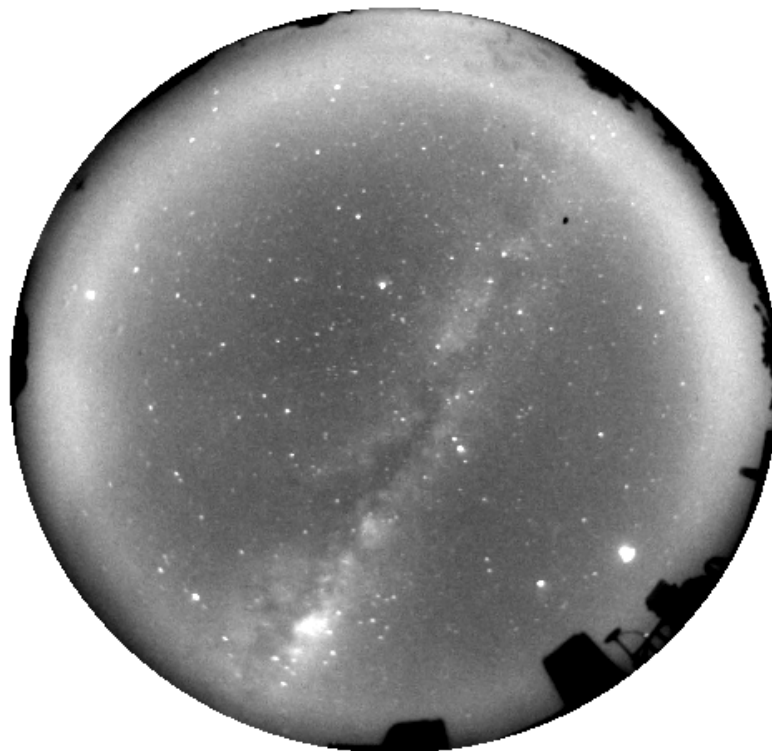
nebol náročný. Nami vytvorený kód podľa zadaného mesiaca prechádza každý deň v mesiaci a stiahne každý dostupný obraz pre dané obdobie. Snímky boli následne uložené do priečinku s názvom podľa stiahnutej stanice a roku. V tomto priečinku sa nachádzajú podpriečinky pre mesiace a dni, do ktorých sú jednotlivé snímky roztriedené podľa ich zhotovenia. Informácie o zhotovení snímok sa nachádzajú na snímkach samotných, no sú aj súčasťou názvu. Pomenovanie snímok sa skladá z čísla dňa v roku a času, kedy bol snímok vyhotovený. Stiahnuté boli dáta z rokov 2020, 2021 a 2022. Tento proces bol najprv uskutočnený iba pre stanicu umiestnenú na stanici Arecibo. Dôvodom bolo vytvorenie klasifikačného modelu na týchto dátach. Dáta z ostatných staníc boli stiahnuté rovnakým kódom po zistení obdobia, z ktorého dáta budú použité pri predikcii. Celá práca s kódom sa uskutočnila na servery určenom pre dátovú analytiku, ktorý nám poskytla Katedra kybernetiky a umelej inteligencie FEI TUKE. Pre prácu na tomto servery sme používali rozhranie notebooku Jupyterlab.

Pred samotným vytváraním modelu klasifikácie bolo potrebné pripraviť tréningovú množinu. Kvôli triedeniu na jednotlivé triedy bolo potrebné vytvoriť jeden priečinok obsahujúci všetky snímky za daný rok. Snímky sme taktiež pomocou knižnice PIL kruhovo orezali, aby obsahovali iba samotnú oblohu a žiadne ďalšie informácie, ktoré by mohli následne skresliť model. Dáta z roku 2020 sme ručne roztriedili do troch priečinkov: jasná, zamračená a polojasná obloha. Tento proces nám zabral dlhšiu dobu, keďže išlo o viac ako 17 000 snímok. Tieto triedy sú neskôr označované pri klasifikácii ako čísla 0, 1 a 2. Dôležité pre našu prácu sú snímky s jasnou oblohou. Orezaný snímok jasnej nočnej oblohy je možné vidieť na Obrázku 3–1.

### 3.4 Klasifikácia BU dát

Pre vytvorenie klasifikačného modelu sme použili konvolučnú neurónovú sieť. Cieľom nebolo vytvoriť model s čo najlepšou presnosťou, ale model, ktorý by nám aspoň čiastočne pomohol pri triedení dát z ostatných staníc Bostonskej Univerzity. Dáta,





**Obrázok 3 – 1** Orezaná snímka nočnej oblohy z 15. Mája 2021 o 7:48:48

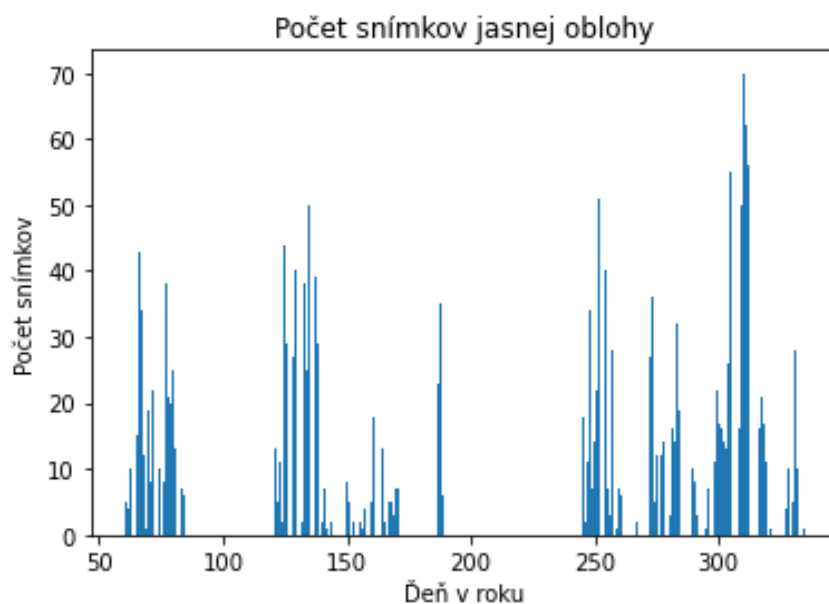
ktorých klasifikáciu sme urobili sami mali nerovnomerný počet snímok v triedach. Prvým krokom bolo zistiť, ktorý priečnik má najmenej snímok. Tento krok je dôležitý z dôvodu, že pri trénovaní modelu potrebujeme pracovať s rovnakým množstvom snímok v každej triede, aby sme sa vyhli skresleniu modelu. Pri každej triede sme náhodne zamiešali snímky a následne vybrali také množstvo, aké bolo pri najmenej početnej najmenšej triede. Výsledný počet snímok v každej triede bol 2107. Model pre klasifikáciu sme vytvorili pomocou knižnice *Keras*. Dáta sme rozdelili na trénovaciu a testovaciu množinu v pomere 80 ku 20. Tento proces vytvárania modelu sme opakovali viackrát s použitím rôznych nastavení až do momentu, kedy sme nedosiahli uspokojivé výsledky. Metriky finálového modelu je možné vidieť v Tabuľke 3 – 1. Pri zvolenom modeli bola taktiež použitá augmentácia prostredníctvom horizontálneho otočenia snímok. Model používa 3 konvolučné vrstvy s aktiváciou ReLU

a na optimalizáciu je použitá metóda Adam.

	precision	recall	f1-score	support
Jasná obloha	0.85	0.85	0.85	439
Oblačná obloha	0.88	0.89	0.88	421
Polooblačná obloha	0.72	0.71	0.71	405

**Tabuľka 3 – 1** Klasifikačný report použitého modelu

Natrénovaný klasifikačný model sme použili na dátach z rokov 2021 a 2022, aby sme získali poznatky o oblačnosti. Dáta z roku 2020 sme už mali ručne klasifikované. Išlo iba o dáta zo stanice Arecibo, dôvodom pre rozhodnutie je fakt, že nad touto oblasťou sníma NASA misia GOLD najčastejšie. Spôsobené to je vďaka jej skenovaniu dvoma kanálmi, pričom prvý sa zameriava na severnú a druhý na južnú pologuľu. Tieto kanály sa však prekrývajú v oblasti rovníka, vďaka čomu máme pri oblasti Arecibo dvojnásobný počet snímok z misie GOLD. Práve kvôli tomuto faktoru sme sa rozhodli vybrať obdobie bez oblačnosti podľa tejto stanice. Z klasifikovaných snímok sme zistili, že najviac snímok jasnej oblohy a zároveň dostatočný počet po sebe idúcich dní s jasnou oblohou za určité obdobie bolo v období koncu roka 2021, konkrétne okolo 27. Októbra 2021. Z tohto dôvodu sme sa rozhodli pre ďalší postup pracovať s dátami medzi 7. Septembrom 2021 a 16. Decembrom 2021. Rozdelenie dní je možné vidieť na Obrázku 3–2. Ide o úsek 100 dní, podľa ktorého sme následne znovu použili nami vytvorený kód na sťahovanie dát pre získanie obdobia aj z ostatných vybraných staníc celooblohových kamier. Na dáta z ostatných staníc sme aplikovali natrénovaný model, ktorý nám našiel snímky s jasnou oblohou aj pri ostatných staniach. Snímky, ktoré označil ako nevhodné sme však ešte dodatočne roztriedili aj ručne. Dôvodom boli niektoré stanice, pri ktorých tento model nenašiel všetky snímky s jasnou oblohou. Aj napriek tomu bol však tento klasifikačný model veľkou pomocou pri výbere vhodných snímok. Klasifikáciou dát sme získali 14038 snímok, ktoré boli pripravené pre ďalší postup.

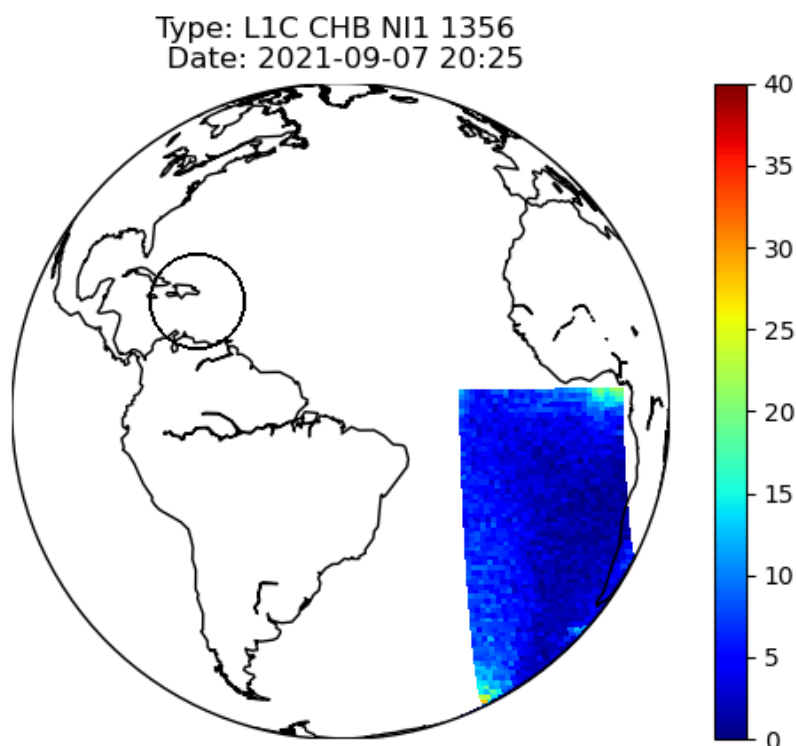


Obrázok 3–2 Rozdelenie množstva jasných nocí za rok 2021 nad stanicou Arecibo

### 3.5 Príprava dát pre predikciu

Dáta z celooblohových snímačov obsahujú na okrajoch rôzne objekty, ako napríklad stromy, elektrické stĺpy alebo budovy. Z tohto dôvodu bolo potrebné snímky zo staníc orezať na menšiu vnútornú časť, aby sme odstránili okraje nesnímanej oblohy. Zostali nám kruhové výrezy nočnej oblohy nad každou z vybraných staníc. Pre tieto stanice sme následne museli upraviť aj snímky z misie GOLD. Táto misia zachytáva veľkú plochu, pričom snímky zo staníc zachytávajú len malú časť oblasti ionosféry ak uvažíme výšku 250 km. Na snímkach dát misie GOLD sa taktiež nachádzajú obrysy kontinentov a najväčších riek. Tieto vo výslednej predikcii mať nepotrebujeme, a preto sme ich museli odstrániť. Pre snímky GOLD sme nevytvárali celý proces ich sťahovania a úprav. Túto časť už sme mali dostupnú vďaka kódom bakalárskej práce, v ktorej sme pracovali práve s dátami GOLD. Celý proces zahŕňal stiahnutie dát a ich transformáciu do grafickej podoby. Odstránenie obrysov na nami vytvorených snímkoch bolo jednoduché kedy sme museli urobiť iba miernu zmenu v týchto kódach. Po úprave sme vytvorili snímky nanovo pre nami vybrané obdobie. Podľa polohy

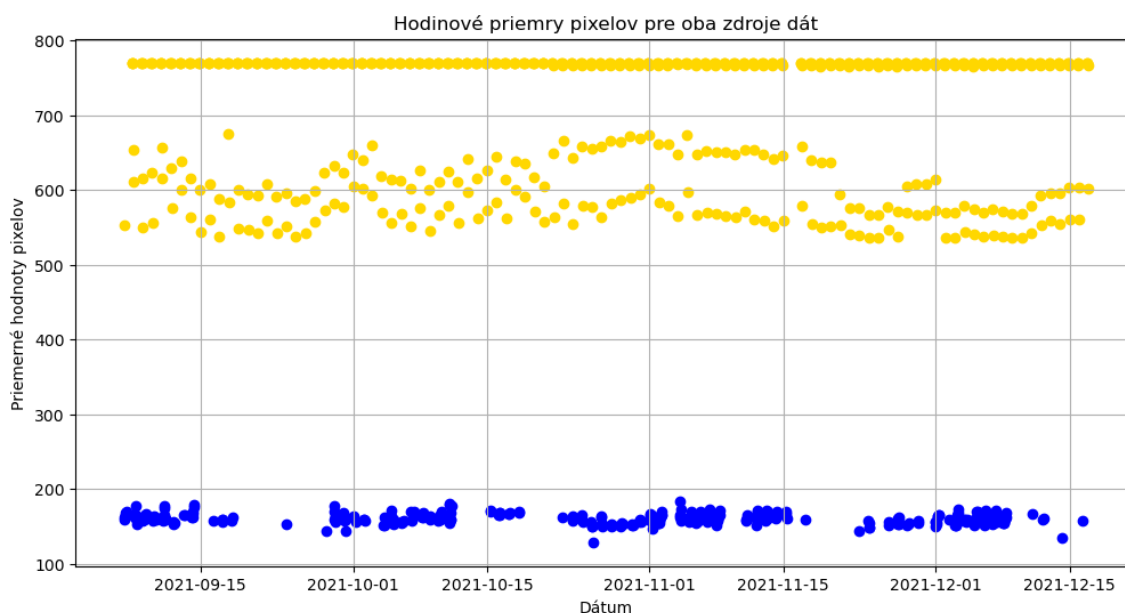
staníc sme vytvorili kruhový výrez zo snímok GOLD v okolí staníc. Príklad takéhoto výrezu pre stanicu Arecibo je možné vidieť na Obrázku 3–3. Umiestenie týchto výrezov sme získali zo stránky, na ktorej sú dostupné aj samotné dáta Bostonskej Univerzity.



**Obrázok 3–3** Znázornenie oblasti snímania stanice Arecibo na snímke dát misie GOLD

Ďalším krokom bolo vytvorenie číselných priemerov snímok BU dát a GOLD dát. Cieľom bolo zistiť aké množstvo snímok BU dát máme v momente kedy snímky GOLD dát nemajú nameranú hodnotu Airglow. Tieto priemery boli vytvorené ako priemerná hodnota pixelov na snímkach. Výpočet bol uskutočnený pomocou knižnice PIL a grafy pomocou knižnice Matplotlib. Sledovali sme dáta na hodinovej báze pre každý deň v intervale vybraných 100 dní. Ako je možné vidieť na Obrázku 3–4, dáta z misie GOLD majú veľké množstvo snímok, ktoré majú vysokú priemernú hodnotu farby pixelov. Dôvodom sú prípady, kedy sa na snímke nenachádza

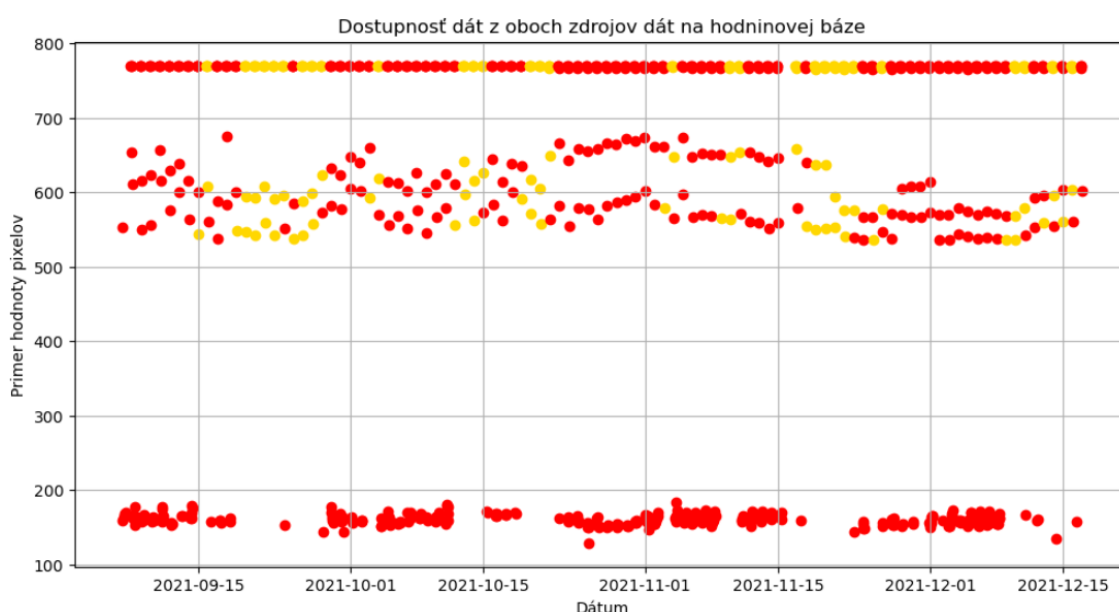
nameraná hodnota Airglow, a teda celý kruhový výrez je bielej farby, čo zapríčiní aj vysokú priemernú hodnotu farby pixelov. Ostatné body označujúce hodinové priemery označujú miesta, kedy máme namerané hodnoty Airglow. Na dátach Airglow označených žltou farbou je taktiež vidno rozdelenie na dva rady. Horný rad má vyššie hodnoty keďže na snímkach sa nachádza aj biela farba. Dáta misie GOLD sa s časom posúvajú smerom nad Atlantický oceán. Tento posun spôsobuje, že v určité hodiny je nad stanicami Bostonskej Univerzity nameraná hodnota iba na polovici výrezu.



**Obrázok 3–4** Rozdelenie snímok vybraného časového úseku s priemermi pixelov snímok

Dôležitou informáciou, ktorú sme potrebovali taktiež zistiť boli časy kedy máme dostupné snímky z celoblokových snímačov a misie GOLD naraz. Takéto snímky sú potrebné pre tréovanie nami vybraného predikčného modelu. Ako je možné vidieť na Obrázku 3–5 prienik týchto snímok je častý, bude však potrebné odfiltrovať snímky misie GOLD s najväčšími priemernými hodnotami pixelov. V intervaloch kedy máme snímky z oboch zdrojov dát bolo potrebné vytvoriť dvojice. Takýchto dvojíc v hodinových intervaloch existuje niekoľko. Prepočet na priemerné pixely je len obrazný a slúžil iba na pochopenie výskytu dát oboch zdrojov, ich rozdelenia

a ktoré časti dát sme museli odfiltrovať. Prvým krokom pri tejto filtrácii bolo odstrániť GOLD dáta bez hodnôt Airglow. Vytvorili sme kód, ktorý pomocou knižnice PIL získal priemernú farbu snímku a následne pomocou knižnice Shutil presunul snímky, ktoré neobsahovali iba bielu farbu do nového priečinka. Rozhodli sme sa pre tento krok z dôvodu, aby sme stále mali dostupné aj všetky dáta spolu v prípade zmeny v procese alebo neskoršiemu testovaniu.



**Obrázok 3 – 5** Znážornenie dostupnosti snímok v hodinových intervaloch. Červená farba sú dostupnosť z oboch zdrojov v danú hodinu, žltá iba z GOLD dát.

Pri dátach z celooblohových snímačov sme taktiež potrebovali zmierniť dopad hviezd na učenie modelu. Algoritmus by sa mohol zamerať na rozmiestnenie hviezd na oblohe miesto štruktúr Airglow. Ďalším krokom predspracovania bolo odstrániť alebo aspoň zmierniť ich dopad na snímkach. Otestovali sme niekoľko riešení vrátane hľadania hviezd a nahradenia ich priemernou farbou pozadia. Tento proces však nebol vhodný kvôli veľkému množstvu hviezd, ktoré boli príliš malé a vybraný spôsob mal problém ich nájsť. V prípadoch kedy ich aj našiel, nahradil veľkú časť snímky rovnakou farbou, pretože išlo o veľa malých hviezd na jednom mieste. Takéto nahradenie je problémové, pretože strácame informácie o výskyte Airglow. Z tohto

dôvodu sme sa rozhodli využiť Gaussovske rozostrenie. Tento spôsob je pomerne častým spôsobom pre úpravu snímok nočnej oblohy. Aplikovanie sme docielili pomocou kódu, kde bolo rozostrenie použité na všetky snímky nočnej oblohy. Hviezdy na snímkach neboli úplne odstránené týmto procesom no bol zmiernený ich dopad pri tréňovaní modelu.

Spárované snímky by sa ešte nedali použiť priamo pri predikcii. Vo vybraných intervaloch bolo potrebné vyriešiť problém kedy v jednom hodinovom intervale existovalo niekoľko snímok z celooblohových snímačov ako aj z misie GOLD. Práve z tohto dôvodu sme sa vo finálovej predikcii rozhodli použiť polhodinové intervaly namiesto hodinových. Týmto spôsobom sme sa ale stále úplne nezbavili viacerých snímok vrámci intervalu. Pripravili sme kód, pomocou ktorého sme spojili všetky snímky do jedného využitím knižníc PIL a Numpy. Takúto úpravu sme aplikovali na všetky dostupné dáta z celooblohových snímačov. Pri aplikovaní tejto úpravy na dáta misie GOLD sme kód pozmenili z dôvodu, že aj po odfiltrovaní sa stále vyskytujú snímky, kedy len polovica plochy obsahuje dáta o hodnotách Airglow. Na takýchto snímkach je teda časť snímku bielej farby. V prípade kombinácie takéhoto snímku s tým, kde sú hodnoty Airglow zachytené na celej jeho ploche, sa biela časť snímku nebrala do úvahy. Časť snímku bez hodnôt Airglow bola nahradená hodnotami zo spájaného snímku.

Vytvorené snímky sme následne pomenovali rovnakým spôsobom, aby mohol vybraný model rozpoznať, ktoré snímky z celooblohových snímačov majú pár v dátach z misie GOLD. Bolo však potrebné v ich názve zachovať všetky dôležité údaje. Pomenovanie týchto snímok pozostáva z dátumu, času a stanice, kde boli snímky vyhotovené, respektíve pri dátach GOLD, nad ktorou stanicou sa dáta nachádzajú.

### 3.6 Explorácia modelu Pix2Pix

Pri výbere modelu bolo potrebné uvažovať nad faktom, že naše vstupné dáta sú snímky, a teda sme potrebovali algoritmus, ktorý dokáže na základe podoby jedného

obrazu predpovedať podobu druhého. Pre túto úlohu sme zvolili model Pix2Pix, ktorý je dostatočne robustný, aby nám poskytol všetky potrebné funkcionality, no zároveň nie je príliš zložitý. Tento model sme už taktiež mali vyskúšaný na dátach misie GOLD vďaka predošlej práci, kde bola využitá jeho upravená verzia pre predikovanie snímok pomocou sekvencie dát. Upravená verzia bola pôvodne vytvorená pre prácu s videom a pracovala so sekvenciou snímok. Model dokázal predikovať niekoľko snímok po ich sekvencii na vstupe. Použitím podobného modelu taktiež máme možnosť priameho porovnania princípu použitia nových vstupných dát oproti predošlému využitiu.

Súbory modelu sú dostupné na stiahnutie cez GitHub a obsahujú python kódy ako aj priečinkov s testovacím datasetom, priečinkov pre výslednú predikciu a priečinkov pre ukladanie priebežných verzií natrénovaného modelu pomocou epoch. Pre lepšiu exploráciu a kontrolu výsledkov sme upravili kód tak, aby sa model ukladal po každej desiatej epoche. V priečinku s datasetom sme museli vytvoriť náš dataset. Vytvorili sme kód, v ktorom sme pomocou knižníc `Os` a `Shutil` presúvali snímky do predom pripravených priečinkov v súboroch modelu. Oba zdroje dát obsahovali kombinované snímky na polhodinových intervaloch a mali rovnaký názov. Vytvorený kód presúval iba snímky, ktoré mali rovnaký názov v oboch zdrojoch dát. Výsledkom bol dataset s dvojicami snímok v priečinkoch pomenovaných `A` a `B`. Model má následne pri tréovaní priečinkov `B` na vstupe a priečinkov `A` na výstupe. Teda podľa snímok z celooblohových kamier v priečinku `B` predikuje, ako by mali vyzeráť snímky z dát GOLD v priečinku `A`. Pre prácu s modelom sme vytvorili Jupyter notebook. Pri spúšťaní tréovania modelu Pix2Pix sme využívali:

- **`!python`**

Časť príkazu, ktorý je na začiatku. Algoritmus je natrénovaný v jazyku python v niekoľkých súboroch. Pri spustení je potrebné definovať interpreter pre tento jazyk. Pri našej práci využívame prostredie Jupyter notebookov, v ktorom je potrebné pred celý tento príkaz pridať výkričník.



- **train.py**

Názov samotného súboru, ktorý obsahuje spustenie tréovania. Obsahuje všetky dostupné nastavenia pre model ako aj ich základné hodnoty. Všetky ostatné súbory potrebné pre natréovanie modelu sú taktiež odkazované v tomto skripte.

- **-- dataset**

Spustenie modelu vyžaduje definovanie názvu datasetu. Ide o povinný parameter pri každom spustení tréovania. Zdroj dát sa nachádza v priečinku **Dataset**, kde už môže byť ľubovoľne pomenovaný priečinok s datasetom, v našom prípade pomenovaný **Airglow**.

- **-- epoch\_count**

Model Pix2Pix umožňuje pokračovať v tréovaní od zadanej epochy v prípade ak má v priečinku **Checkpoint** uložený model. Predvolená hodnota 1 znamená začiatok tréovania od začiatku, pričom sa predpokladá, že nie je uložená žiadna predchádzajúca verzia natréovaného modelu. Nastavenie bolo veľmi užitočné kvôli dĺžke tréovania modelu. Ďalším využitím je aj prípad kedy sa tréovanie nečakane skončilo a mohli sme v ňom pokračovať.

- **-- seed**

Parameter zabezpečuje reprodukovateľné náhodné procesy pri každom tréovaní modelu. Pri testovaní rôznych nastavení modelu zaručí, že rozdielnosť výsledkov nie je zapríčinená náhodne, ale spôsobili ich zmeny v nastaveniach.

- **test.py**

Používaný v prípade tvorby predikcie už natréovaného modelu. Aj napriek názvu nejde iba o testovanie, ale aj samotné vytváranie predikcie. V takom prípade je potrebné zdrojové snímky, podľa ktorých budeme predikcie vytvárať vložiť do pripraveného priečinku, ktorý sa nachádza v priečinku **Dataset**.

Model sme na začiatku natrénovali na jeho základných nastaveniach, aby sme získali predstavu o jeho presnosti. Pri vytváraní testovacích výsledkov bol model natrénovaný po epochu 200 a výstupné snímky z modelu boli porovnané s pôvodnými snímkami pomocou metrík MSE a PSNR. Výsledky modelu so základnými nastaveniami je možné vidieť v Tabuľke 3–2. Ako je možné vidieť jeden zo snímok zo stanice Mercedes mal omnoho horšie výsledky ako ostatné vybrané snímky. Dôvodom tejto odchýlky je existencia len čiastočne nameraných emisií Airglow v originálnom snímku, a teda polovica snímku je biela. Niekoľko takýchto snímok zostalo aj po úprave. Dôvodom bolo, že pri pol hodinovom intervale neexistovali žiadne ďalšie snímky, ktoré by túto časť doplniť mohli. Model túto časť napriek tomu doplnil, a preto sa nezhodujú. Ostatné snímky dosiahli o niečo lepšie výsledky, avšak stále nie uspokojivé. Ďalším krokom bolo otestovať nastavenia modelu a zistiť optimálne nastavenia pre predikciu na našich dátach.

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	1657.21	15.93
3.10.	23:30	riogrande	1250.73	17.15
23.9.	00:00	elleoncito	1918.55	15.30
24.11.	00:00	colombia	716.47	19.57
<b>1.11.</b>	<b>22:30</b>	<b>mercedes</b>	<b>18551.67</b>	<b>5.44</b>
24.9.	23:00	elleoncito	1051.83	17.91
28.9.	23:00	riogrande	2680.02	13.84
10.9.	00:00	elleoncito	1219.39	17.26
27.9.	23:30	colombia	1282.9	17.04
28.10.	23:30	mercedes	2282.86	14.54

**Tabuľka 3–2** Výsledky presnosti modelu na náhodne vybraných snímkach pri použití základných nastavení.

Model Pix2Pix má pri trénovaní niekoľko nepovinných nastavení. Rozhodli sme sa preto otestovať vplyv zmeny niektorých nastavení na výsledky modelu. Na otes-

tovanie boli použité rovnaké snímky, ako pri testovaní základných metrík pre porovnanie. Taktiež sme nespúšťali tréning dvoch rôznych zmenených parametrov naraz z dôvodu, že sme sa snažili zistiť dopad jednotlivých nastavení samostatne. Jedinou výnimkou sú nastavenia, kedy je rovnaký parameter dostupný pre generátor aj diskriminátor, v takom prípade boli zmenené oba parametre. Dostupné parametre pre model Pix2Pix sú:

- **-- batch\_size**

Určenie množstva snímok trénovacej množiny, ktoré sa majú spracovať pred aktualizáciou vnútorných parametrov modelu. Základné nastavenie je na hodnote 1. Teda model bude aktualizovaný po každej jednotlivej tréningovej vzorke. Väčšie veľkosti dávok môžu zvýšiť rýchlosť a stabilitu tréningu, ale vyžadujú viac pamäte. Pri testovaní parametra sme hodnotu na 4. Teda model sa aktualizuje po každej štvrtej snímke. Ako je možné vidieť v Tabuľke 3–3, nastavenie nepomohlo zlepšiť výsledky pri tréningu na našich dátach. Metriky dosahujú horšie hodnoty, a to vo veľa prípadoch až dvojnásobne. Parameter by bol viac užitočný v prípade väčšieho množstva dát, ako máme dostupné.

- **-- test\_batch\_size**

Plní rovnakú úlohu ako **-- batch\_size**, ale počas fázy testovania. Určuje počet testovacích vzoriek, ktoré model spracuje pred aktualizáciou váh. Pri veľkosti 1 model spracuje každú testovaciu vzorku samostatne. Výber veľkosti dávky počas testovania môže ovplyvniť rýchlosť testovania a využitie pamäte, ale na rozdiel od tréningu neovplyvňuje učenie modelu.

- **-- input\_nc**

Nastavenie vstupného počtu kanálov. Pre obrázky vo farebnom priestore RGB sú to zvyčajne 3 (jeden kanál pre červenú, jeden pre zelenú a jeden pre modrú farbu). Ak by sme pracovali s obrázkami v odtieňoch sivej, potrebovali by sme zmeniť túto hodnotu. Nastavenie musí zodpovedať skutočnému počtu kanálov

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	6407.2	10.06
3.10.	23:30	riogrande	1506.58	16.35
23.9.	00:00	elleoncito	2900.95	13.51
24.11.	00:00	colombia	3160.54	13.13
1.11.	22:30	mercedes	21680.69	4.77
24.9.	23:00	elleoncito	1500.04	16.36
28.9.	23:00	riogrande	2248.69	14.61
10.9.	00:00	elleoncito	1931.63	15.27
27.9.	23:30	colombia	5797.13	10.49
28.10.	23:30	mercedes	3405.86	12.81

**Tabuľka 3–3** Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia batchsize na hodnotu 4.

vo vstupných údajoch, aby model údaje správne spracoval. Dáta na vstupe sú z celooblohých kamier Bostonskej Univerzity. Z dôvodu zhody farebnej škály medzi vstupnými a výstupnými dátami nebolo potrebné nastavenie nemeniť.

- **-- output\_nc**

Podobne, ako nastavenie `-- input_nc` určuje počet kanálov, ale na výstupných dátach. Rovnako musí byť nastavené na hodnotu 3. Snímky GOLD farebne odlišujú hodnoty Airglow, teda musí byť nastavenie na hodnote 3.

- **-- ngf**

Tento parameter riadi počet filtrov v konvolučnej vrstve generátora modelu. Počet filtrov priamo ovplyvňuje kapacitu a zložitosť siete, pričom viac filtrov umožňuje sieti naučiť sa zložitejšie funkcie za cenu zvýšených nárokov na výpočtovú techniku. Predvolená hodnota tohto nastavenia je 64. Nastavenie sme sa rozhodli otestovať zmenením parametru na hodnotu 128 filtrov. Išlo, teda o dvojnásobné zvýšenie.

- **-- ndf**

Podobne ako **-- ngf**, ide o určenie počtu filtrov v konvolučnej vrstve diskriminátora modelu. Nastavenie počtu filtrov môže ovplyvniť citlivosť a presnosť diskriminátora, čo má vplyv na celkový výkon algoritmu. Základné nastavenie je taktiež na hodnote 64. Kvôli testovaniu predošlého nastavenia bolo potrebné rovnako zmeniť aj tento parameter na hodnotu 128. Výsledok testovania je možné vidieť v Tabuľke 3–4. Hodnoty v niektorých prípadoch boli síce lepšie, ale pri porovnaní s ostatnými išlo o mierne zhoršenie, aj napriek väčšej zložitosti modelu. V prípade jedného snímku zo stanice Colombia išlo dokonca o takmer trojnásobné zhoršenie. Trénovanie samotného modelu taktiež trvalo výrazne dlhšie, a preto sme sa ani toto nastavenie nerozhodli využiť pri finálovom tréovaní.

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	1544.68	16.24
3.10.	23:30	riogrande	1316.41	16.93
23.9.	00:00	elleoncito	2052.68	15.01
24.11.	00:00	colombia	2142.88	14.82
1.11.	22:30	mercedes	23366.72	4.44
24.9.	23:00	elleoncito	1173.37	17.43
28.9.	23:00	riogrande	2911.29	13.48
10.9.	00:00	elleoncito	1415.97	16.62
27.9.	23:30	colombia	1648.27	15.96
28.10.	23:30	mercedes	2018.74	15.07

**Tabuľka 3–4** Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavení ngf a ndf na hodnotu 128.

- **-- niter**

Tento parameter určuje počet iterácií, pri ktorých sa model bude trénovať s

počiatočnou rýchlosťou učenia. Základná hodnota pre nastavenia je 100 iterácií. Pri našom testovaní sme parameter nemenili z dôvodu, že sme neočakávali výrazné zvýšenie presnosti s nárastom iterácii vzhľadom k času, ktorý by bol na takéto trénovanie potrebný.

- **-- niter\_decay**

Je to počet iterácií, počas ktorých sa rýchlosť učenia lineárne zníži na nulu po uplynutí počiatočných iterácií. Rovnako, ako nastavenie pre iterácie je základné nastavenie 100 iterácií. Spolu s nastavením **-- niter** ide o nastavenie počtu iterácií, počas ktorých sa bude model trénovať. Tieto parametre sme nechali na ich pôvodných hodnotách, ktoré sú preddefinované, a teda sme vždy trénovali na 200 epochách.

- **-- lr**

Určenie počiatočnej rýchlosti učenia pre optimalizátor Adam. Riadenie veľkosti krokov, ktoré vykoná počas trénovania. Základné nastavenie je na hodnote 0.0002. Testovali sme zmenu hodnoty na 0.002. Výsledky zmeny tohto parametru je možné vidieť v Tabuľke 3–5. Hodnoty porovnaných snímok sú výrazne horšie, ako pri pôvodnom nastavení. Tento parameter necháme na jeho základnej hodnote.

- **-- lr\_policy**

Tento argument definuje spôsob úpravy miery učenia v priebehu tréningu. Základným nastavením spôsobu úpravy miery učenia je lambda, ale dostupné sú aj step, plateau, a kosínus. Pre tento parameter sme sa rozhodli uskutočniť dva testy. Pri prvom sme spôsob zmenili na kosínus. Ako je možné vidieť v Tabuľke 3–6 hodnoty porovnaných snímok sú relatívne porovnateľné s tými pri použití spôsobu lambda. Pri niektorých išlo o zlepšenie, no pri väčšine išlo o mierne zhoršenie hodnôt metrík, a preto sme sa nerozhodli využiť tento spôsob. Ďalším testom bol parameter step, ktorého výsledky je možné vidieť

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	8315.55	8.93
3.10.	23:30	riogrande	8288.19	8.94
23.9.	00:00	elleoncito	6582	9.94
24.11.	00:00	colombia	3655.45	12.5
1.11.	22:30	mercedes	38502.36	2.27
24.9.	23:00	elleoncito	6647.3	9.91
28.9.	23:00	riogrande	10451.18	7.93
10.9.	00:00	elleoncito	5772.59	10.51
27.9.	23:30	colombia	7378.62	9.45
28.10.	23:30	mercedes	12132.02	7.29

**Tabuľka 3–5** Výsledky presnosti modelu na náhodne vybraných snímkach počas zmeneného nastavenia lr na hodnotu 0.002.

v Tabuľke 3–7. Výsledky porovnania v tomto prípade nie sú až tak podobné a ide o celkom výrazné zhoršenie. Počas tohto testovania sme taktiež chceli vyskúšať spôsob plateau, avšak pri trénoch s touto zmenou dochádzalo k častému zastaveniu trénoch. Tento problém sa nám bohužiaľ nepodarilo vyriešiť, a teda sme nezískali hodnoty, ktoré by sme mohli porovnávať. Vďaka otestovaniu vybraných parametrov sme zistili, že najlepšie bude použiť spôsob lambda.

- **-- lr\_decay\_iters**

Určuje, ako často sa má rýchlosť učenia znižovať. Nastavuje počet iterácií trénoch, po ktorých sa miera učenia vynásobí vopred určeným faktorom gama, aby sa postupne znižovala. Znižovanie pomáha pri jemnom doladovaní váh modelu, čo umožňuje presnejšie úpravy a potenciálne vyšší výkon. Predvolené nastavenie je každých 50 iterácií. Táto metóda pomáha zabrániť tomu, aby model pri postupujúcom trénoch nestratil optimálne váhy udržiavaním príliš

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	1591.23	16.11
3.10.	23:30	riogrande	1290.49	17.02
23.9.	00:00	elleoncito	2217.54	14.67
24.11.	00:00	colombia	1822.81	15.52
1.11.	22:30	mercedes	12332.26	7.22
24.9.	23:00	elleoncito	1124.82	17.61
28.9.	23:00	riogrande	2753.41	13.73
10.9.	00:00	elleoncito	1605.53	16.07
27.9.	23:30	colombia	1243.66	17.18
28.10.	23:30	mercedes	3053.76	13.28

**Tabuľka 3–6** Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia — lr\_policy na spôsob kosínus.

Dátum	Čas	Stanica	MSE	PSNR
10.9.	00:00	elleoncito	2055.80	15.01
3.10.	23:30	riogrande	1457.25	16.49
23.9.	00:00	elleoncito	2552.18	14.06
24.11.	00:00	colombia	4370.24	11.72
1.11.	22:30	mercedes	22678.76	4.57
24.9.	23:00	elleoncito	1502.53	16.36
28.9.	23:00	riogrande	2713.86	13.79
10.9.	00:00	elleoncito	3283.89	12.96
27.9.	23:30	colombia	3035.56	13.30
28.10.	23:30	mercedes	3952.15	12.16

**Tabuľka 3–7** Výsledky presnosti modelu na náhodne vybraných snímkach počas použitia nastavenia — lr\_policy na spôsob step.



vysokej rýchlosti učenia. Naše tréovanie bolo uskutočnené na 200 epochách, teda základné nastavenie bolo dostačujúce.

- **-- beta1**

Špecifický parameter pre optimalizátor Adam. Riadi exponenciálnu rýchlosť rozpadu pre odhady prvého momentu, ktoré sú kľúčové pre možnosti adaptívnej miery učenia optimalizátora. Predvolená hodnota parametra je nastavená na 0.5. Ide o spravodlivé vyváženie najnovšieho gradientu a minulých gradientov. Vyššia hodnota by dávala väčšiu váhu minulým gradientom, čo by viedlo k plynulejším, ale potenciálne pomalším aktualizáciám, zatiaľ čo nižšia hodnota zdôrazňuje nedávne gradienty, čo umožňuje citlivejšie, ale menej stabilné aktualizácie. Z tohto dôvodu sme sa rozhodli parameter nechať na jeho základnej hodnote, ktorá je vybalancovaná medzi týmito dvoma situáciami.

- **-- cuda**

Aktivácia tohto príkazu umožňuje modelu využívať na výpočty špeciálne prostriedky grafického procesora NVIDIA, čím sa v porovnaní s výpočtom založeným na CPU výrazne urýchľujú procesy tréovania a odvodzovania. Systém musí byť kompatibilný s touto funkciou kvôli čomu sa nám ju nepodarilo zahrnúť pri tréovaní modelov. Nevieme teda aký dopad by mal na rýchlosť tréovania.

- **-- threads**

Nastavenie počtu vlákien, ktoré sa majú použiť pre načítavanie údajov. Nastavenie tohto počtu môže pomôcť vyvážiť záťaž medzi predbežným spracovaním údajov a výpočtom modelu. Základné nastavenie je na 4 vlákna.

- **-- lamb**

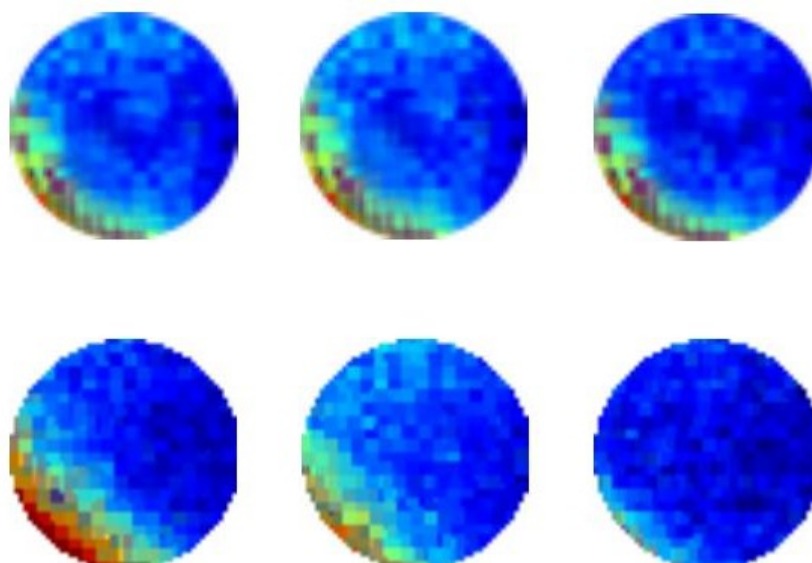
Určenie váhy člena L1 v stratovej funkcii modelu. Vyššia hodnota kladie väčší dôraz na presné zhody na úrovni pixelov, čo môže viesť k menej rozmazaným

obrázkom, ale môže tiež potlačiť schopnosť modelu generovať nové prvky. Základné nastavenie je na hodnote 10. V prípade predikovania hodnôt Airglow na snímkach rozmazanie nie je problémové no zároveň je potrebné generovať nové prvky. Rozhodli sme sa hodnotu parameteru nemeniť.

### 3.7 Výsledný model a vyhodnotenie

Po otestovaní všetkých nastavení modelu sme sa rozhodli využiť základné parametre modelu Pix2Pix. Trénovanie prebehlo na všetkých dostupných dátach vo vybranom intervale 100 dní. Model bol trénovaný po 200 epoch, pričom sme ukladali každú 10 epochu. Najlepšie výsledky boli dosiahnuté počas epochy 170. Pre porovnanie vytvorených snímok sme použili metriky MSE a PSNR. Výber snímok, ktoré sme vybrali pre testovanie bol uskutočnený na základe ich dennej postupnosti v rámci jednotlivých staníc kedy bola obloha na snímkach jasná. Takéto porovnanie je dôležité pre vývoj hodnôt Airglow. Porovnanie 3 po sebe idúcich dní pri stanici Mercedes s originálnymi výrezmi dát GOLD je možné vidieť na Obrázku 3–6. Ako je možné vidieť v Tabuľke 3–8, pri porovnaní s predošlými výsledkami dosahujú metriky horšie hodnoty. Pri metrike MSE platí, čím je číslo bližšie k 0, tým sú si obrázky podobné. Pri metrike PSNR je to zase opačne. Najhoršia hodnota je pri snímke z 24.9. zo stanice Jicamarca. V tomto prípade je to však očakávané. Pôvodná snímka obsahuje značnú časť plochy kedy nemáme údaje o hodnote Airglow. Presnosť je teda zlá kvôli doplneniu tejto oblasti. Ako je možné vidieť pri snímkach zo stanice Mercedes, hodnoty sú najlepšie z celej tabuľky. Dôvod týchto výsledkov je zapríčinený oblasťou s nízkou hodnotou nameraných emisií Airglow. Ide o dobrú predikciu, pretože náš model nepredpovedal výskyt veľkého množstva Airglow v oblasti, kde by ani nemal byť. Pri ostatných snímkach sú hodnoty už relatívne podobné.

V Tabuľke 3–9 je možné vidieť výsledky použitia predošlého postupu pri bakalárskej práci. Tento princíp spočíva vo vytváraní pokračovania sekvencie dát GOLD. V tomto prípade bola na vstupe sekvencia 5 snímok a predikovaných bolo 6 nasle-



**Obrázok 3–6** Porovnanie 3 po sebe idúcich dní predikcie stanice Mercedes voči originálnym výrezom. Horný rad sú predikované snímky zatiaľ čo spodný rad sú originálne výrezy z dát misie NASA GOLD

dujúcich snímok. V porovnaní s výsledkami nášho nového postupu sú tieto metriky omnoho lepšie. Je však potrebné zobrať do úvahy, že tieto snímky boli porovnávané ako celok, teda snímok celej Zeme, zatiaľ čo naše nové výsledky sú porovnávané na základe použitej oblasti nad jednotlivými stanicami, ako bolo možné vidieť na Obrázku 3–3. Predošlé výsledky majú lepšie hodnoty aj kvôli ploche porovnávania, ktorá mala veľké množstvo pixelov mimo predikcie. Takéto miesta sú totožné s pôvodnými obrazmi a prispievajú k lepším metrikám. Avšak, aj keď zoberieme tento fakt do úvahy, naše nové výsledky dosahujú v metrikách horšie výsledky.

Pri porovnávaní výsledkov je v tomto prípade potrebné pozrieť sa aj na oblasť a možnosti predikcie. Predošlý spôsob dokázal predikovať iba snímky, ktoré nasledovali po skutočných. Náš nový princíp nie je závislý na dátach z misie GOLD, a teda dokážeme predikovať hodnoty misie Airglow aj na miestach, ktoré predošlý spôsob nedokázal, ako je možné vidieť na Obrázku 3–7. Ďalším dôležitým aspektom je či dokázal model predikovať oblasti výskytu Airglow okolo rovníka. Práve v tejto oblasti sú emisie javu najvýraznejšie. Porovnanie predikcie je možné vidieť

Dátum	Čas	Stanica	MSE	PSNR
23.9.	23:30	mercedes	1294.99	17.01
24.9.	23:30	mercedes	957.39	18.32
25.9.	00:00	mercedes	746.68	19.4
23.9.	23:30	elleoncito	2231.54	14.65
24.9.	23:30	elleoncito	1142.2	17.55
25.9.	00:00	elleoncito	1043.11	17.95
23.9.	00:00	jicamarca	4694.22	11.41
24.9.	23:30	jicamarca	13008.54	6.99
23.9.	00:00	riogrande	2038.37	15.04
24.9.	23:30	riogrande	3796.96	12.34
25.9.	00:00	riogrande	2030.56	15.05
4.12.	23:00	arecibo	4341.75	11.75
5.12.	23:30	arecibo	4739.99	11.37
6.12.	23:30	arecibo	1554.91	16.21

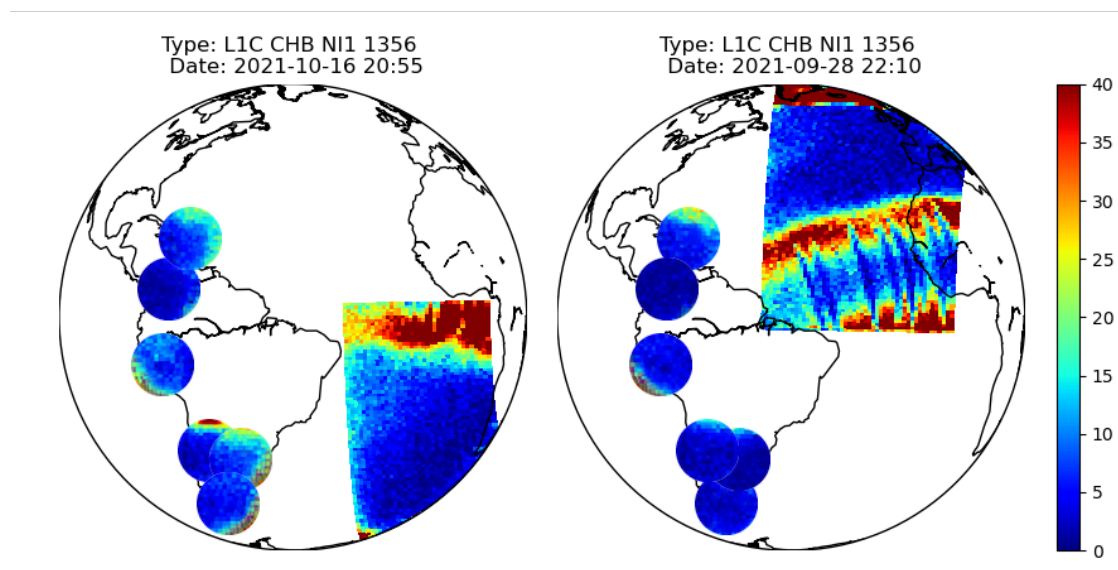
**Tabuľka 3–8** Výsledky presnosti finálneho modelu so zvolenými základnými nastaveniami. Snímky pre predikciu sú zvolené na základe ich postupnosti v dňoch

na Obrázku 3–8, na ktorom je jasne vidieť, že tieto oblasti sa aj napriek horším výsledkom nášmu modelu podarilo aspoň čiastočne predikovať.

Cieľom tejto práce bolo využiť metódy hlbokého učenia na dátach javu Airglow. K dosiahnutiu výsledku bolo potrebné oboznámiť sa s javom samotným ako aj dátami vytvorenými pre skúmanie javu. Podarilo sa nám osvojiť si túto problematiku a vytvoriť množinu údajov, ktorú sme mohli použiť pre splnenie celkového cieľa. Ďalšou úlohou bolo vybrať a oboznámiť sa s modelom hlbokého učenia. Ako model sme vybrali Pix2Pix, pri ktorom sme zistili jeho optimálne nastavenia a vytvorili model pre predikciu hodnôt Airglow. Výsledky modelu sme taktiež porovnali s predošlými výsledkami. Cieľ práce sa nám podarilo splniť aj napriek horším metrikám. Podarilo sa nám doplniť existujúce dáta o oblasti, ktoré by nebolo predtým možné získať. Ide

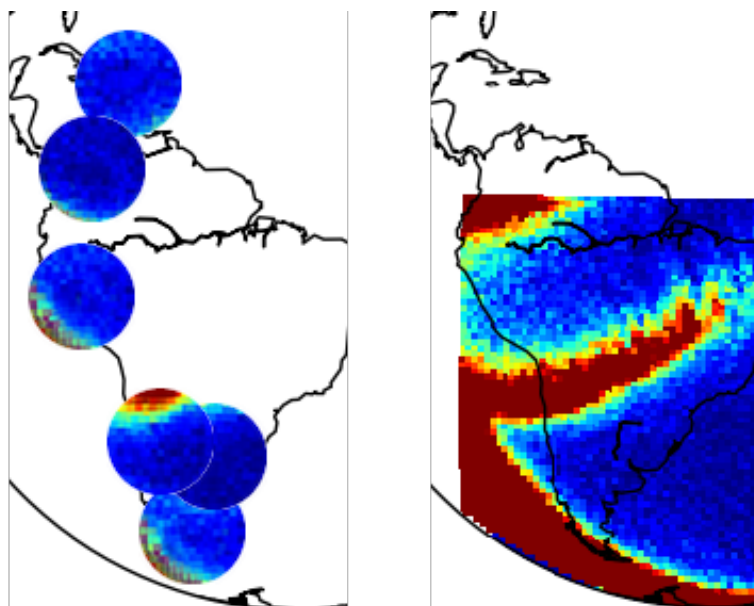
Severná pologuľa			Južná pologuľa		
Čas	MSE	PSNR	Čas	MSE	PSNR
05.01.2020			05.01.2020		
22:10	129.66	27.01	22:25	117.19	27.44
22:40	172.58	25.76	22:55	182.18	25.53
23:10	247.51	24.19	23:10	138.88	26.7
23:25	350.16	22.69	23:25	137.61	26.74
23:40	355.01	22.63	23:40	217.11	24.76
23:55	366.92	22.49	23:55	210.99	24.89

**Tabuľka 3–9** Výsledky metrík 6 po sebe idúcich snímok pre Severnú a Južnú pologuľu vytvorených ako výsledok bakalárskej práce.



**Obrázok 3–7** Ukážka doplnenia 2 snímok z NASA misie GOLD

o pokrok v predpovedaní tejto oblasti. Existuje niekoľko faktorov v tejto práci, ktoré by vhodné v budúcnosti preskúmať a vylepšiť. Jedným z nich je aplikovanie väčšej množiny údajov, ktoré by bolo možné získať pomocou klasifikačného modelu založenom na dátach inej vlnovej dĺžky celooblohových snímačov. Takáto klasifikácia



**Obrázok 3 – 8** Porovnanie predikovaných dát Airglow pomocou našej metódy kombinácie zdrojov dát

by mohla priniesť zlepšenie metrík. Ďalším možným vylepšením by bolo otestovanie iných modelov a prístupov strojového učenia, a to aj pomocou zakomponovania číselných hodnôt javu Airglow namiesto grafických.

## 4 Záver

V práci sme sa zaoberali oblasťou výskumu Airglow. Pri tomto jave sme preskúmali teoretické poznatky o jeho výskyte, rozdelení, histórii skúmania a venovali sme sa aj spôsobom získavania dát. Prvým spôsobom, ktorým sme sa zaoberali boli celoblokové dáta snímačov od Bostonskej Univerzity. Tieto snímače poskytujú pohľad na jav Airglow zo zemského povrchu zachytávaním jeho štruktúr na nočnej oblohe. Ako druhý spôsob sme popísali spôsob snímania pomocou špeciálnych snímačov orbitujúcich našu planétu na satelitoch. Takýmto princípom sa zaoberá NASA misia GOLD, ktorú sme bližšie predstavili.

Taktiež sme sa venovali strojovému učeniu, pri ktorom sme vytvorili teoretický prehľad o jeho fungovaní, rozdelení a využití. Obzvlášť sme sa venovali generatívnym adverzným sieťam, ktorých využitie pri generovaní obrázkov sme aplikovali na predikovanie javu Airglow. Ako model pre predikciu sme vybrali model Pix2Pix, ktorý sme využili na predikovanie dát z misie NASA GOLD. Dopĺňanie dát bolo možné vďaka použitiu dát z celoblokových snímačov. Dáta sme pred použitím v neurónovej sieti klasifikovali podľa výskytu oblačnosti a upravovali do vhodnej podoby.

Pri práci s vybraným modelom sme preskúmali jeho možnosti použitia ako aj jeho vhodné nastavenia pre nami používané dáta. Takto nastavený model sme aplikovali na dáta javu Airglow s cieľom získať nové poznatky a prispieť tak k výskumu oblasti výskytu javu Airglow.

Výsledky modelu sme porovnali pomocou metrík MSE a PSNR. Hodnoty, ktoré sme dostali, ale neboli uspokojivé. Naše výsledky sme následne porovnali aj s predošlým výskumom. Pri porovnaní sme zistili, že naše nové výsledky vykazovali horšie hodnoty metrík. Tieto hodnoty boli spôsobené aj tým, že pri predošlom spôsobe metriky porovnávali aj časti bez predikcie teda dosahovali lepšie hodnoty aj vďaka tomu. Naš princíp, ale priniesol predikciu v častiach GOLD dát, kde predtým nebolo možné získať hodnoty emisií Airglow. Ide o nový princíp predikcie tohto javu a

podarilo sa nám potvrdiť predpoklad o možnosti využitia takejto predikcie. Rozšírenie oblasti skúmania tak vďaka nášmu princípu prispelo ku lepšiemu pochopeniu správania javu na väčšej ploche, ako bolo doteraz možné. Keďže ide o nový princíp, nebolo naše výsledky možné porovnávať priamo s podobnými princípmi predikovania javu. Plán pre túto prácu v budúcnosti je pokračovať vo výskume a vylepšiť spôsob predpovedania hodnôt Airglow. Celkovým cieľom dopĺňania dát misie NASA GOLD je získať hodnoty emisií na celej ploche Zeme.

Pre vylepšenie výsledkov by bolo v budúcnosti možné vyskúšať iné princípy hlbokého učenia. Spojenie fyzikálnych modelov a metód strojového učenia by taktiež mohli priniesť nové výsledky. Možnosťou pre zlepšenie je aj využitie dát s filtrom zameraným pre zachytávanie hviezd. Pri klasifikácii jasných nocí by sme dokázali získať presnejšie výsledky a získali by sme väčšie množstvo použiteľných dát.

Výskum Airglow je naďalej dôležitým javom pre pochopenie atmosféry našej Zeme. Využitie metód hlbokého učenia má pre túto oblasť veľký potenciál a počas posledných rokov už začal prinášať prvé výsledky. Aplikácia týchto metód je prospešná nie len v predpovedaní hodnôt javu, ale aj v detekcii jeho útvarov. S ďalším vývojom budú tieto výsledky lepšie a prínos bude nenahraditeľný.



## Literatúra

*Aeronomy* (n.d.). <http://sirius.bu.edu/aeronomy/>.

Albawi, S., Mohammed, T. A. and Al-Zawi, S. (2017). Understanding of a convolutional neural network, *2017 International Conference on Engineering and Technology (ICET)*, IEEE. <http://dx.doi.org/10.1109/ICEngTechnol.2017.8308186>.

Alpaydin, E. (2021). *Machine Learning, revised and updated edition*, MIT Press, London, England. [https://books.google.sk/books?id=2nQJEAAAQBAJ&lpg=PR7&ots=fI\\_4TaXFjm&dq=machine%20learning&lr&hl=sk&pg=PP1#v=onepage&q=machine%20learning&f=false](https://books.google.sk/books?id=2nQJEAAAQBAJ&lpg=PR7&ots=fI_4TaXFjm&dq=machine%20learning&lr&hl=sk&pg=PP1#v=onepage&q=machine%20learning&f=false).

Chakrabarti, S., Paresce, F. and et al. (1983). The extreme ultraviolet day airglow, *Journal of Geophysical Research: Space Physics* **88**(A6): 4898–4904. <http://dx.doi.org/10.1029/JA088iA06p04898>.

Duann, Y., Chang, L. C. and et al. (2022). A methodology of retrieving volume emission rate from limb-viewed airglow emission intensity by combining the techniques of abel inversion and deep learning, *Atmosphere* **14**(1): 74. <http://dx.doi.org/10.3390/atmos14010074>.

Eastes, R. W., McClintock, W. E. and et al. (2017). The global-scale observations of the limb and disk (gold) mission, *Space Science Reviews* **212**(1–2): 383–408. <http://dx.doi.org/10.1007/s11214-017-0392-2>.

Eastes, R. W., McClintock, W. E. and et al. (2020). Initial observations by the gold mission, *Journal of Geophysical Research: Space Physics* **125**(7): e2020JA027823. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JA027823>.

Fradkov, A. L. (2020). Early history of machine learning, *IFAC-PapersOnLine* **53**(2): 1385–1390. 21st IFAC World Congress.

- 
- François-Lavet, V., Henderson, P. and et al. (2018). An introduction to deep reinforcement learning, *Foundations and Trends in Machine Learning* **11**(3–4): 219–354. <http://dx.doi.org/10.1561/22000000071>.
- Gattinger, R. L. (1968). Observation and interpretation of the airglow emissions, *Canadian Journal of Physics* **46**(14): 1613–1626. <https://doi.org/10.1139/p68-489>.
- Gold Instrument* (n.d.). <https://gold.cs.ucf.edu/science-mission/gold-instrument/>.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*, MIT Press. <http://www.deeplearningbook.org>.
- Goodfellow, I., Pouget-Abadie, J. and et al. (2014). Generative adversarial nets, in Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence and K. Weinberger (eds), *Advances in Neural Information Processing Systems*, Vol. 27, Curran Associates, Inc. [https://proceedings.neurips.cc/paper\\_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf).
- Goodfellow, I., Pouget-Abadie, J. and et al. (2020). Generative adversarial networks, *Communications of the ACM* **63**(11): 139–144. <http://dx.doi.org/10.1145/3422622>.
- Hore, A. and Ziou, D. (2010). Image quality metrics: Psnr vs. ssim, *2010 20th International Conference on Pattern Recognition*, IEEE. <http://dx.doi.org/10.1109/ICPR.2010.579>.
- Isola, P., Zhu, J.-Y., and et al. (2016). Image-to-image translation with conditional adversarial networks. <https://arxiv.org/abs/1611.07004>.
- Khomich, V. Y. and et al. (2008). Springer Berlin Heidelberg. <http://dx.doi.org/10.1007/978-3-540-75833-4>.
-

- Lai, C., Xu, J., and et al. (2023). Statistical characteristics of nighttime medium-scale traveling ionospheric disturbances from 10-years of airglow observation by the machine learning method, *Space Weather* **21**(5): e2023SW003430. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2023SW003430>.
- Lin, Z., Wang, Q. and Lai, C. (2023). Analysis of airglow image classification based on feature map visualization, *Applied Sciences* **13**(6): 3671. <http://dx.doi.org/10.3390/app13063671>.
- Mogren, O. (2016). C-rnn-gan: Continuous recurrent neural networks with adversarial training. <https://arxiv.org/abs/1611.09904>.
- Nasteski, V. (2017). An overview of the supervised machine learning methods, *HORIZONS.B* **4**: 51–62. <http://dx.doi.org/10.20544/HORIZONS.B.04.1.17.P05>.
- Otsuka, Y., Shiokawa, K. and et al. (2004). Geomagnetic conjugate observations of medium-scale traveling ionospheric disturbances at midlatitude using all-sky airglow imagers, *Geophysical Research Letters* **31**(15). <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GL020262>.
- Pfaff, R. F. (2012). The near-earth plasma environment, *Space Science Reviews* **168**(1–4): 23–112. <http://dx.doi.org/10.1007/s11214-012-9872-6>.
- Savigny, C. v. (2017). Airglow in the earth atmosphere: basic characteristics and excitation mechanisms, *ChemTexts* **3**(4). <http://dx.doi.org/10.1007/s40828-017-0051-y>.
- Sedlak, R., Welscher, A. and et al. (2023). Analysis of 2d airglow imager data with respect to dynamics using machine learning, *Atmospheric Measurement Techniques* **16**(12): 3141–3153. <http://dx.doi.org/10.5194/amt-16-3141-2023>.

- Serag, A., Ion-Margineanu, A. and et al. (2019). Translational ai and deep learning in diagnostic pathology, *Frontiers in Medicine* **6**. <http://dx.doi.org/10.3389/fmed.2019.00185>.
- Sherstinsky, A. (2020). Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network, *Physica D: Nonlinear Phenomena* **404**: 132306. <https://www.sciencedirect.com/science/article/pii/S0167278919305974>.
- Silverman, S. M. (1970). Night airglow phenomenology, *Space Science Reviews* **11**(2): 341–379. <https://doi.org/10.1007/BF00241526>.
- Slovinský, S. and Mackovjak, S. (2020). Airglow. <https://slovinsky.art/en/2020/09/03/airglow/>.
- Smith, S. M., Scheer, J. and et al. (2006). Characterization of exceptionally strong mesospheric wave events using all-sky and zenith airglow observations, *Journal of Geophysical Research: Space Physics* **111**(A9). <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JA011197>.
- Thanakulketsarat, T., Supnithi, P. and et al. (2023). Classification of the equatorial plasma bubbles using convolutional neural network and support vector machine techniques, *Earth, Planets and Space* **75**(1). <http://dx.doi.org/10.1186/s40623-023-01903-7>.
- Toshinori Maihara, F. I. and et al. (1993). Observations of the oh airglow emission, *Publications of the Astronomical Society of the Pacific* **105**(691): 940. <https://dx.doi.org/10.1086/133259>.
- Usama, M., Qadir, J. and et al. (2019). Unsupervised machine learning for networking: Techniques, applications and research challenges, *IEEE Access* **7**: 65579–65615. <http://dx.doi.org/10.1109/ACCESS.2019.2916648>.

Willmott, C. and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance, *Climate Research* **30**: 79–82. <http://dx.doi.org/10.3354/cr030079>.

Yntema, L. (1909). On the brightness of the sky and total amount of starlight, *Publications of the Kapteyn Astronomical Laboratory Groningen, vol. 22, pp. 1-55* **22**: 1–55. <https://adsabs.harvard.edu/full/1909PGro...22....1Y/0000001,000.html>.

## Zoznam príloh

**Príloha A** Kód vytvorený pri vypracovaní diplomovej práce je dostupný na GitHubu: <https://github.com/kkuichi/mb371om>

**Príloha B** CD médium obsahujúce všetky kódy a elektronickú verziu diplomovej práce