Sveučilište u Rijeci

Fakultet informatike i digitalnih tehnologija

Sveučilišni diplomski studij informatike

**DETEKCIJA SANTI LEDA**

**izvještaj**

IZRADIO: Marko Ribarić

prosinac, 2024. godine

**Sadržaj**

[1. PODACI 3](#_Toc185520765)

[2. MODELI 5](#_Toc185520766)

[3. PREDVIĐANJE NAJSIGURNIJEG PUTA NA KARTI 8](#_Toc185520767)

[Popis slika 9](#_Toc185520768)

# PODACI

Ideju sam dobio iz „Statoil/C-CORE Iceberg Classifier Challenge“ izazova. Podaci u train.json i test.json sadrže slike s nekoliko polja. Polje id označava ID slike, dok band\_1 i band\_2 sadrže podatke o radarskom signalu slike u obliku lista s 75x75 piksela. Inc\_angle označava kut pod kojim je slika snimljena, a može biti nedostajući ("na"). Polje is\_iceberg označava je li slika ledenjak (1) ili brod (0). Polje is\_iceberg postoji samo u train.json. U testnim podacima nalaze se i računalno generirane slike koje nisu ocjenjivane. Zbog tog razloga sam za učenje i testiranje koristio podatke iz train.json.

A screenshot of a computer screen

Description automatically generated

Slika Prikaz skupa podataka u formatu tablice

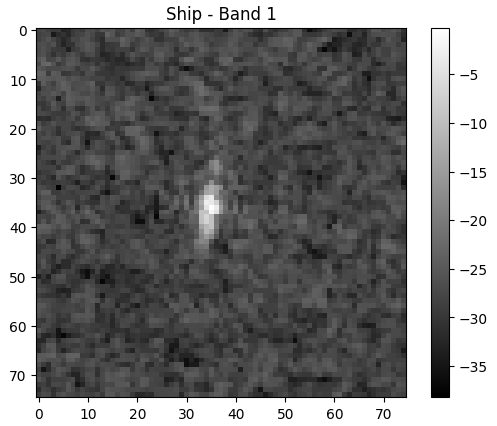
Skup podataka ima 1604 unosa i 5 stupaca. Na slici 2. se nalazi prikaz band 1 slike ledenjaka. Ledenjak se nalazi na sredini slike.

A close-up of a polarized image

Description automatically generated

Slika Prikaz band 1 ledenjaka

Na slici 2 nalazi se prikaz band 1 slike broda. Možemo primijetiti da je oblik broda dulji, dok je oblik ledenjaka sa slike 1 više kružan.



Slika Prikaz band 1 broda

# MODELI

U kodu sam prvo obradio podatke tako da sam zamijenio nedostajuće vrijednosti kuta snimanja (inc\_angle) s prosječnom vrijednošću kako bih izbjegao probleme pri treniranju modela. Zatim sam kombinirao dvije radarske slike (band\_1 i band\_2) u dvoslojnu sliku, pripremajući podatke za daljnju obradu. Nakon toga sam podijelio podatke na trening i validacijski skup kako bih mogao procijeniti performanse modela.

Za modeliranje sam prvo koristio Random Forest Classifier, gdje sam spljoštio slike u 1D vektore i trenirao model na tim podacima. Rezultati su prikazani na slici 4. Možemo vidjeti kako je prosjek svih metrika 0.74

A screenshot of a computer screen

Description automatically generated

Slika Rezultati Random Forest Classifier modela

Nisam bio zadovoljan ovom brojkom te sam sam implementirao Convolutional Neural Network (CNN), koristeći dvoslojne slike za bolje prepoznavanje uzoraka. U CNN modelu sam primijenio konvolucijske slojeve, max pooling, i potpuno povezane slojeve, a kako bih spriječio prekomjerno učenje, dodao sam „early stopping“.

Model je napravljen na 20 epocha te se rezultati nalaze na slici 5.

A screenshot of a computer

Description automatically generated

Slika Rezultati CNN modela

Napravio sam vizualizaciju točnosti na trening skupu, točnosti na validacijskom skupu, gubitak na trening skupu i gubitak na validacijskom skupu kroz povijest treniranja modela. Vizualizacija je prikazana na slici 6.

A graph of a graph

Description automatically generated with medium confidence

Slika Vizualizacija točnosti i gubitka kroz povijest treniranja CNN modela

Postotak točnosti predviđanja mi je bio konzistentno 88-91% s manje promjene ovisno o pokretanju modela.

Za bolji prikaz rezultata sam nasumično odabrao 10 slika s validacijskog skupa i za svaku predvidio oznaku pomoću treniranog modela. Zatim sam prikazao slike u dvije vrste: na gornjoj su predviđene oznake, a na donjoj stvarne oznake. Kroz grafički prikaz, za svaku sliku sam prikazao vjerojatnost predikcije zajedno sa stvarnom oznakom, omogućujući vizualnu usporedbu između predviđanja i stvarnog stanja.

A collage of images of light

Description automatically generated

Slika 10 slika sa validacijskog skupa

# PREDVIĐANJE NAJSIGURNIJEG PUTA NA KARTI

U aplikaciji „SEASENSE“ je glavni cilj bio navigiranje kroz vode u slučaju da se radar i primarne tehnologije pokvare. Rješenje je bila aplikacija za analiziranje najboljeg puta ovisno o različitim tehnologijama poput bivših ledenjaka, predviđanje ledenjaka ovisno o strujama i vjetrovima i dubina. Za pronalazak puta sa najmanje rizika se koristi dijkstra s nekim promjenama. Na slici 8. je prikazano kako izgleda pronalazak puta nakon 5 pokreta.

A screenshot of a game

Description automatically generated

Slika Prikaz pronalaska puta nakon 5 pokreta

Na slici 9. je prikazano kako izgleda pronalazak puta nakon što je došao do cilja.

A screenshot of a game

Description automatically generated

Slika Prikaz pronalaska puta na cilju

Uz pomoć korištenja slika oceana sa satelita možemo također izbjegavati polja na kojima smo skoro sigurni da postoje ledenjaci.

# Popis slika

[Slika 1 Prikaz skupa podataka u formatu tablice 3](#_Toc185520749)

[Slika 2 Prikaz band 1 ledenjaka 3](#_Toc185520750)

[Slika 3 Prikaz band 1 broda 4](#_Toc185520751)

[Slika 4 Rezultati Random Forest Classifier modela 5](#_Toc185520752)

[Slika 5 Rezultati CNN modela 6](#_Toc185520753)

[Slika 6 Vizualizacija točnosti i gubitka kroz povijest treniranja CNN modela 6](#_Toc185520754)

[Slika 7 10 slika sa validacijskog skupa 7](#_Toc185520755)

[Slika 8 Prikaz pronalaska puta nakon 5 pokreta 8](#_Toc185520756)

[Slika 9 Prikaz pronalaska puta na cilju 8](#_Toc185520757)