

Abstract

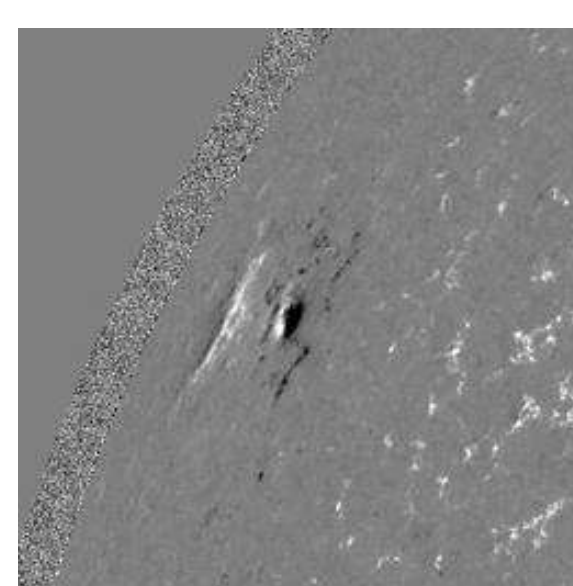
Acest proiect abordează problema predicției erupțiilor solare, un fenomen cu implicații majore asupra sănătății umane și a integrității tehnologiei satelitare și planetare.

Obiectivul principal al proiectului este dezvoltarea unui sistem capabil să calculeze predicția erupțiilor solare în timp.

Dintre metodele testate: custom CNN și VGG16, modele cu rezultate pozitive, VGG16 a oferit o acuratețe superioară.



Imagine NASA dataset 1102



Imagine SDO dataset

Set de date

Metodologia propusă este bazată pe un set de date NASA, care este format din foldere ce reprezintă mai multe zone ale soarelui.

Fiecare imagine conține în denumirea sa codificarea datei și a orei la care poza a fost făcută, de unde reiese timpul până la erupție.

Eruption code	Photo time	Eruption time	Time difference
1066	2010-05-03 00:00:00	2010-05-03 21:53:00	21:53:00
1066	2010-05-03 00:12:00	2010-05-03 21:53:00	21:41:00
1066	2010-05-03 01:24:00	2010-05-03 21:53:00	20:29:00
1102	2010-08-30 00:24:00	2010-08-30 15:22:00	14:58:00
1102	2010-08-30 00:12:00	2010-08-30 15:22:00	15:10:00
1102	2010-08-30 00:00:00	2010-08-30 15:22:00	15:22:00

Metode

Cea mai optimă metodă identificată este rețeaua neuronală VGG16 cu o acuratețe de 98% respectiv 89% rezultate obținute prin folosirea a două zone(1102, 1066).

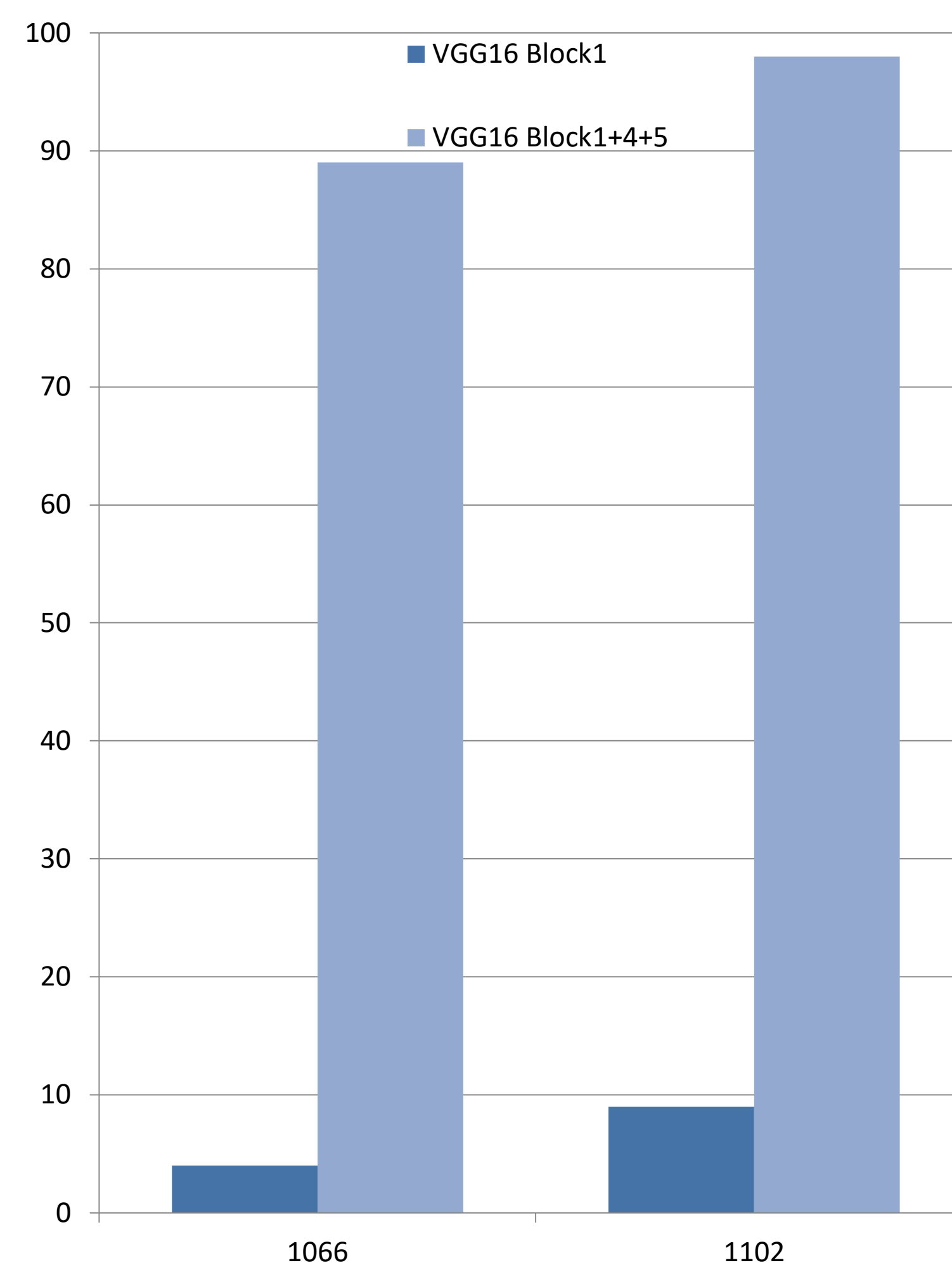
Proiectul abordează o problema de regresie liniară, întrucât variabila țintă este o valoare numerică continuă, reprezentată de timpul până la eveniment, spre deosebire de o problemă de clasificare abordată anterior pe dataset-ul SDOBenchmark.

Ca și metodologie de evaluare a rezultatelor avem:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$R2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

Rezultate antrenare



Arhitectura

Ca și arhitectură am folosit VGG16 preantrenat pe ImageNet cu care am făcut transfer learning astfel:

- am înghețat toate blocurile
- antrenare bloc 1 (învățare trăsături generale)
- antrenare bloc 4,5 (învățare trăsături complexe)
- adăugare straturi specifice regresiei(linear activation)

Rezultate

Rezultate obținute anterior cu dataset-ul SDO pentru clasificare:

✓ Acuratete: 55%

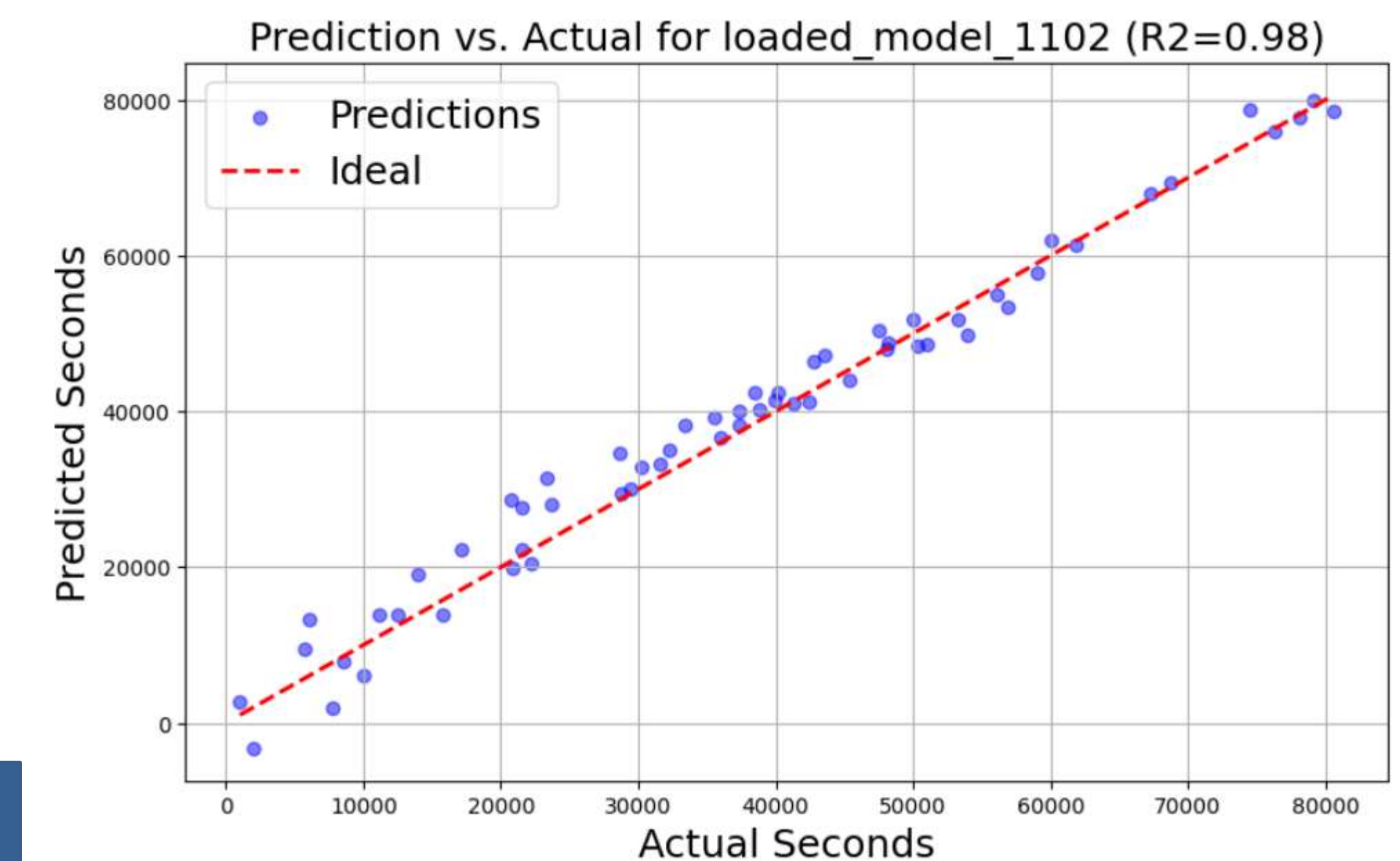
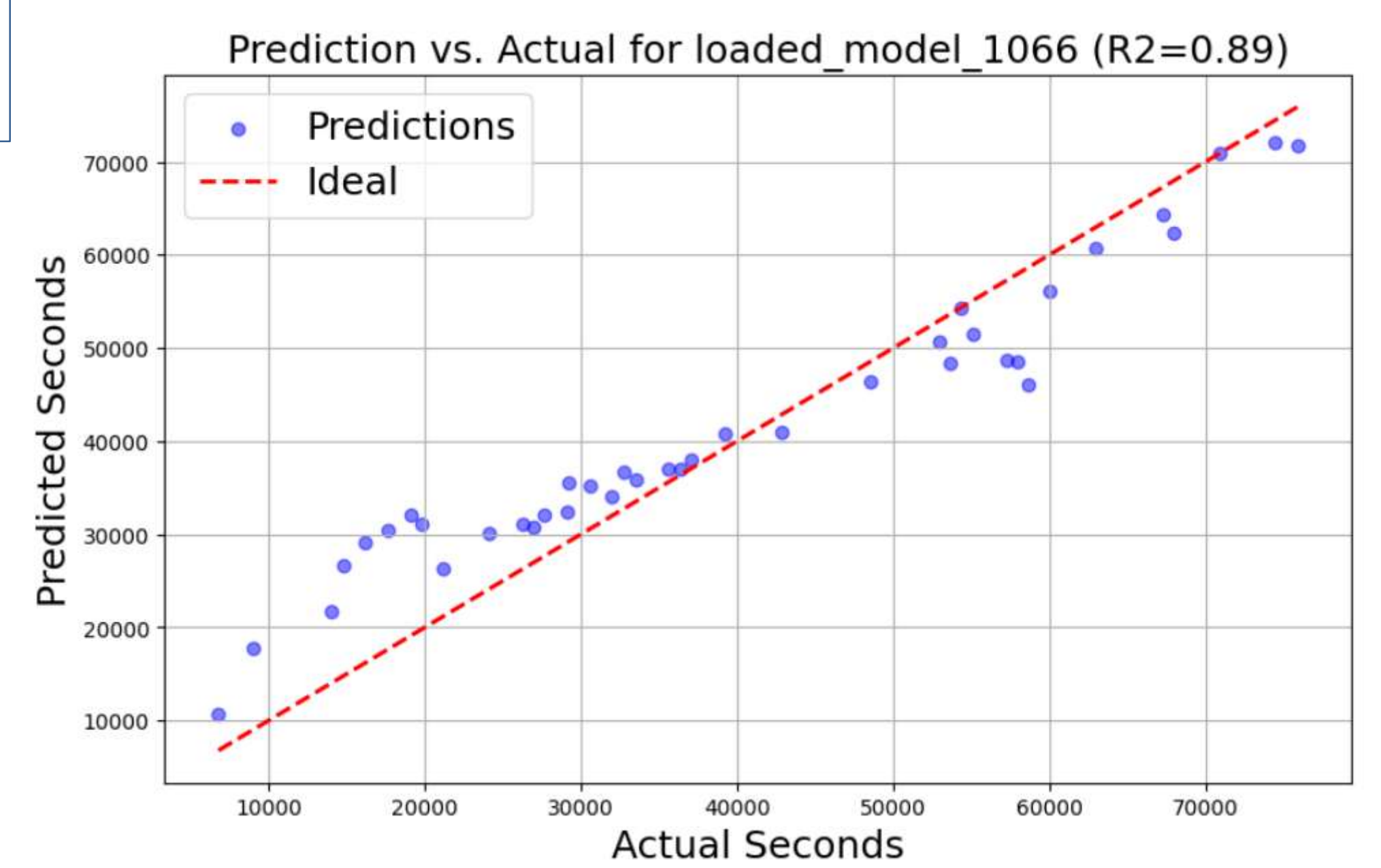
Rezultate finale obținute cu dataset-ul NASA folosind modelul VGG16 pentru regresie liniară în funcție de cele două zone:

✓ 1066 :

- R2 = 0.89
- RMSE = 6418s ~ 106 min

✓ 1102:

- R2 = 0.98
- RMSE = 3290s ~ 55 min

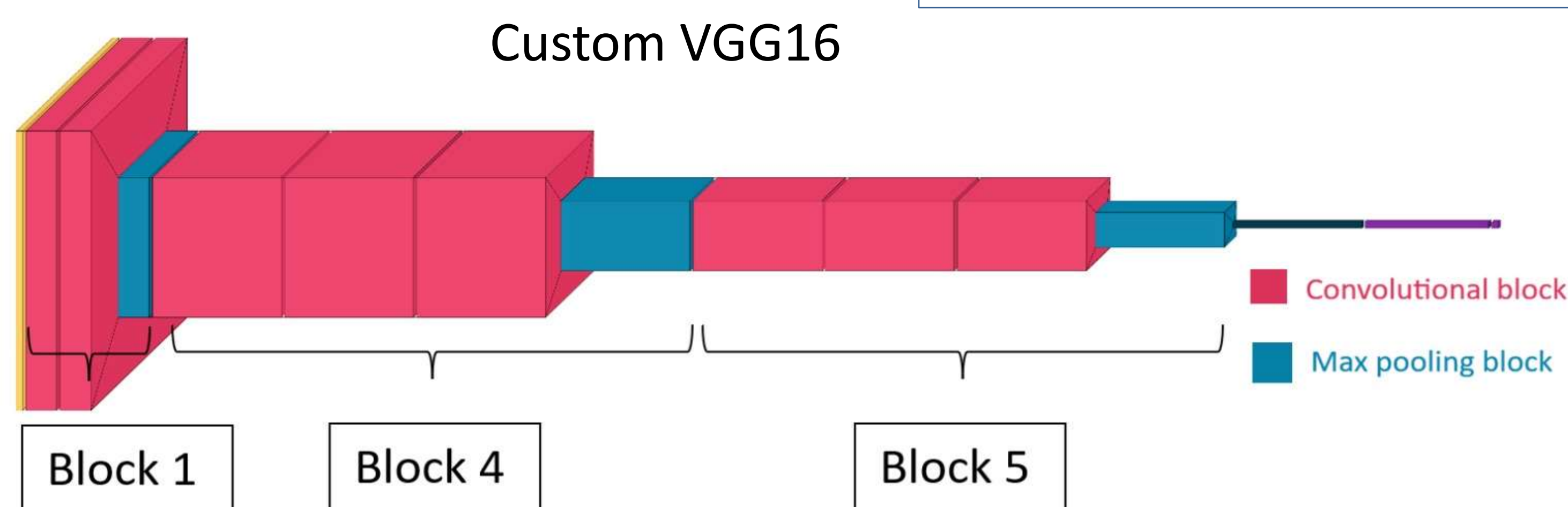


Concluzii

Limitări:

- imaginile provin din multe zone ale soarelui, de aici și incapacitatea modelului de a învăța eficient
- imaginile au puține caracteristici și detaliile sunt greu de distins

În concluzie, am reușit să avem o acuratețe crescută în predicția erupțiilor folosind zone distincte ale soarelui.



Contact

Mărguță Dan-Alexandru
Aparaschivei Teodor
Facultatea de "Automatică și Calculatoare" Iași
dan-alexandru.marguta@student.tuiasi.ro
teodor.aparaschivei@student.tuiasi.ro

References

1. Julia Bringewald, Solar Flare Forecast: A Comparative Analysis of Machine Learning Algorithms for Solar Flare Class Prediction
2. Vysakh P. A. , Prateek Mayank, Solar Flare Prediction and Feature Selection using Light Gradient Boosting Machine Algorithm
3. Xuebao Li, Yanfang Zheng, Xinshuo Wang, and Lulu Wang, Predicting Solar Flares Using a Novel Deep Convolutional Neural Network
4. Interactive multi-instrument database of solar flares
5. Eric Jonas, Monica G. Bobra, Vaishaal Shankar, J. Todd Hoeksema, Benjamin Recht Flare Prediction Using Photospheric and Coronal Image Data