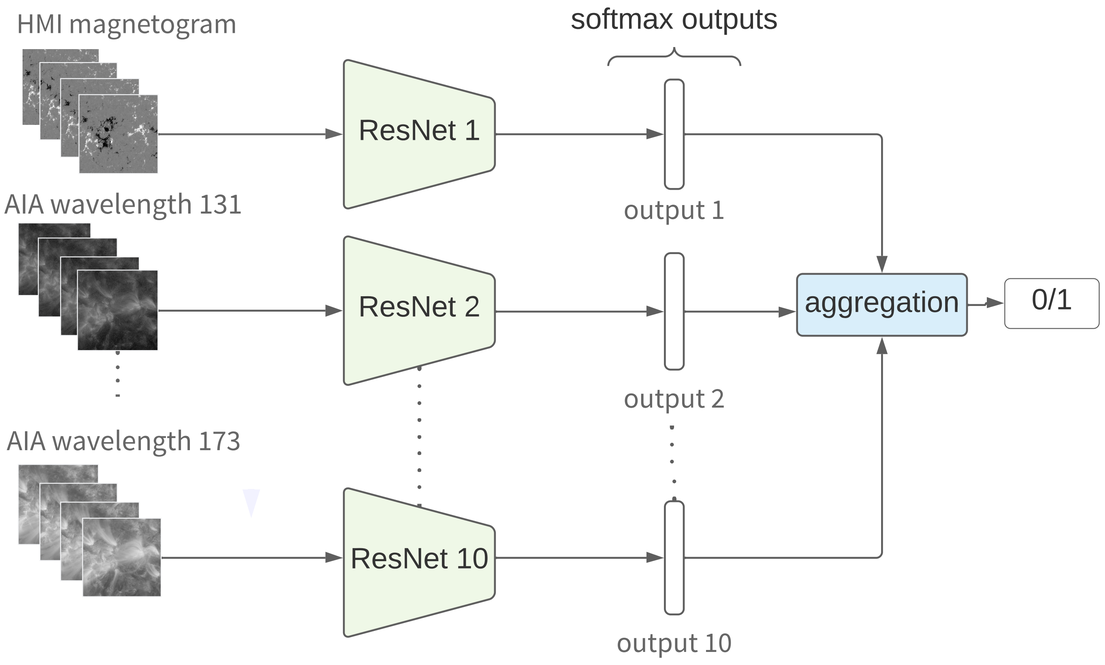
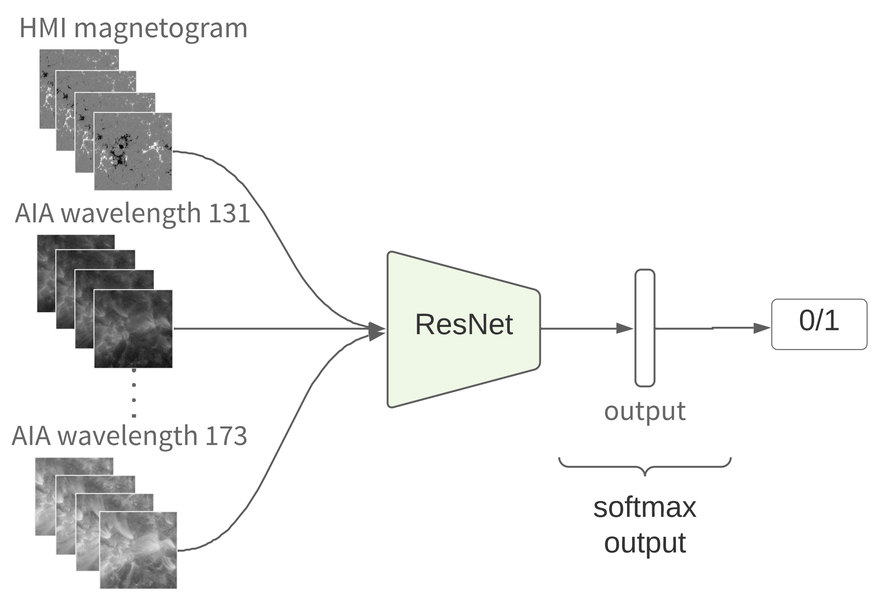
Solar flare prediction using multi-channel models

Yana Shtyk[1,2], Olga Taran[2], André Csillaghy[1] and Jonathan Donzallaz[3]

1 - FHNW Institute for Data Science, Switzerland, 2 - University of Geneva, Department of Informatics, Switzerland,
3 - iCoSys, HES-SO, Switzerland



(a)



(b)

Figure 1: Schematic representation of the proposed models: (a) single channel are first trained on individual ResNet models models and the resulting outputs are aggregated via majority voting or summation techniques; (b) the model with the deep aggregation uses all channels jointly during the training.

Proposed method

- We apply the residual neural network ResNet18 to predict flares
- Single channel models are trained to investigate the predictive capabilities of each individual channel
- Aggregation techniques are applied to investigate the performance of multi-channel models

Aggregation types

- **Majority voting** predicts the same output as the majority of the individual channels as shown in Figure 1.a
- **Summation technique** makes a decision based on the sum of the softmax outputs of the individual channels as shown in Figure 1.a
- In the **deep aggregation** model, the different channels are jointly used for training as shown in Figure 1.b

SDOBenchmark Dataset

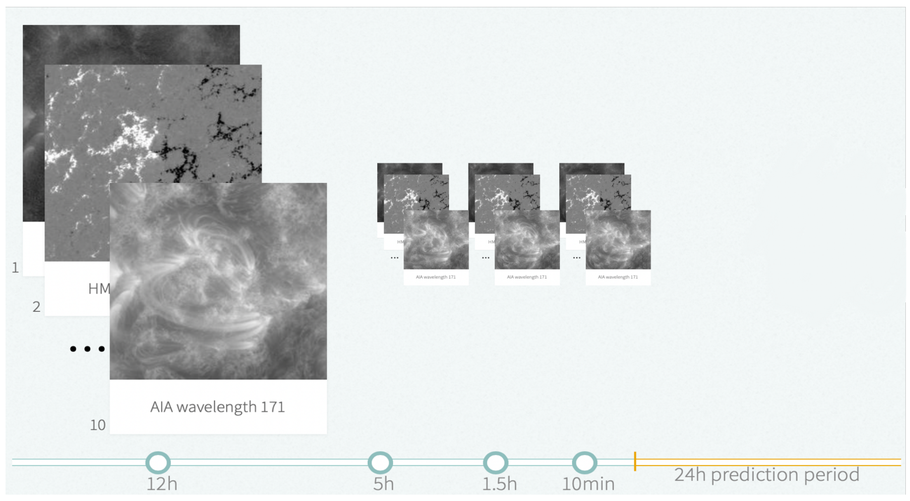


Figure 2: Dataset visualization (https://i4ds.github.io/SDOBenchmark/).

- Consists of active regions patches cropped from the SDO dataset (https://sdo.gsfc.nasa.gov/data/)
- Includes data from 10 channels
- Channels are formed by 8 AIA and 2 HMI detectors from SDO
- Each channel includes 8336 training samples and 886 test samples
- Each sample consists of 4 time steps images
- Each image has size 256x256
- Labels correspond to the 24-hour prediction horizon

Type	Non-flare	C	M	X
Train	4941	2885	475	35
Test	360	350	151	25

Table 1: Dataset statistics, number of samples

Results

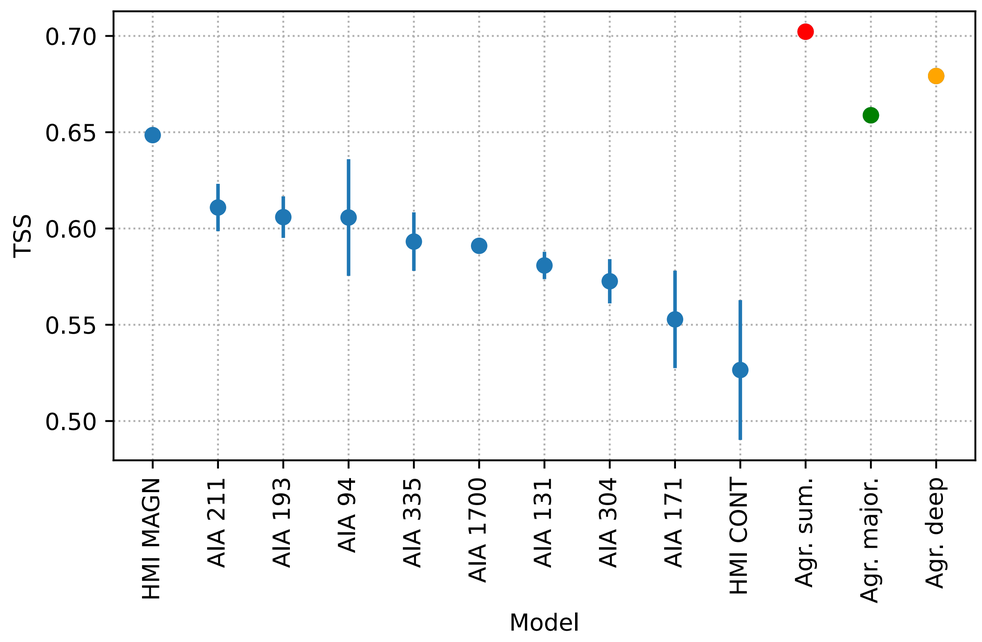


Figure 3: True Skill Score (TSS) for individual and multi-channel models.

Training process

- A stochastic gradient descend optimizer was used for training
- Early stopping was applied to prevent overfitting

Conclusions & future work

- The channels such as HMI magnetogram and AIA wavelenghts 211, 193 and 94 show good predictive capabilities
- In the multi-channel model, the agreggation via summation of the softmax outputs of the individual channels improves substantially the predictive capabilities
- In future work, we plan to investigate deeply the multi channel training with more complex aggregation techniques and to apply the proposed multi-channel approach to bigger datasets

This work was done as a part of the RODEM project:
https://rodem.ch