A Machine Learning Approach for Automated Sunquake Detection

Vanessa Mercea

Technical University of Cluj-Napoca

7 September 2022

Work with: A.R. Paraschiv, D.A. Lăcătuș, A. Mărginean, D. Beșliu-Ionescu

Outline

Sunquakes

Definition

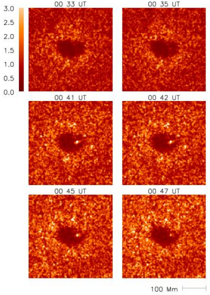
Sunquakes are progressive circular waves
observed on the fotosphere, produced by solar flares.

Interest

Manual detection is often laborious.
Manifestation circumstances are not entirely known.
Several detection methods are available but none are automatic.

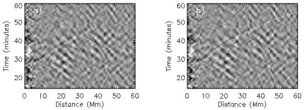
Sunquakes

Helioseismic Holography



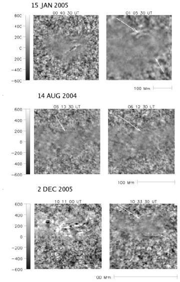
©Moradi et al., 2007

Time Distance



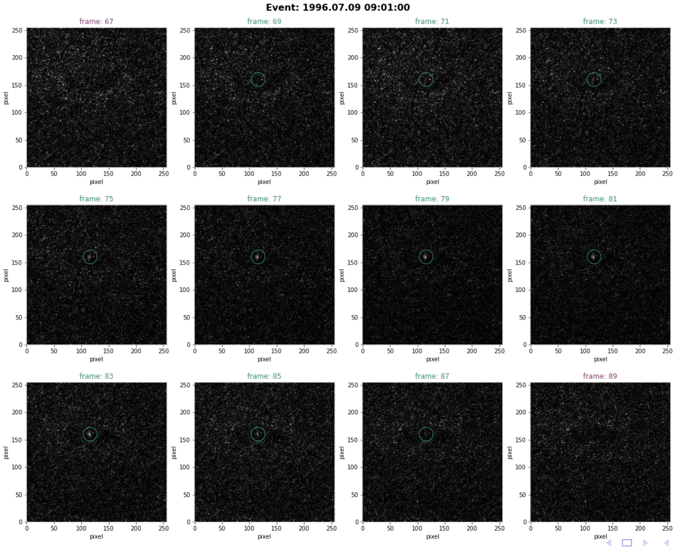
©Kosovichev and Zharkova, 1998

Movies/Wave Detection

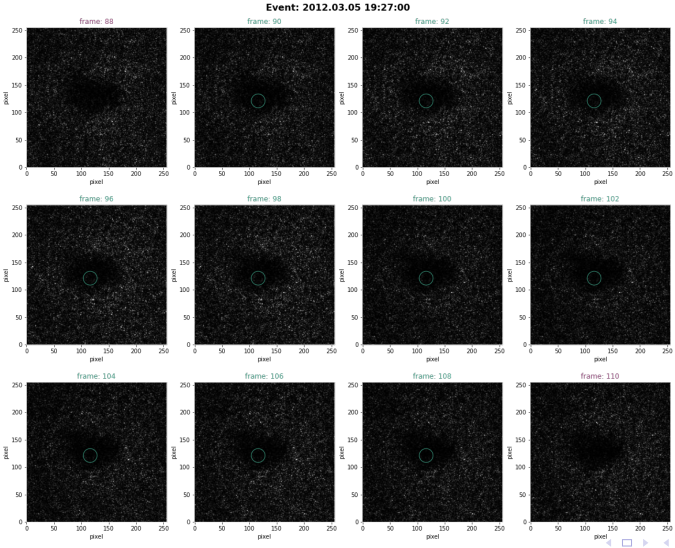


©Moradi et al., 2007

Data Morphology



Data Morphology



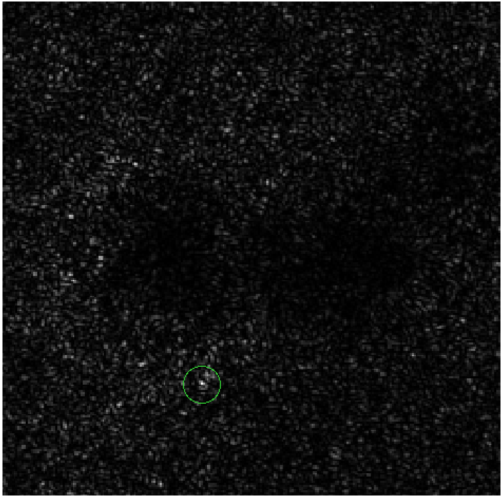
Challenges

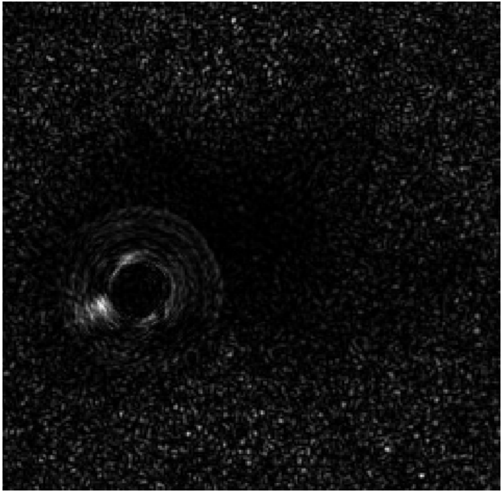
imbalance: on average, 4% of all individual data samples are SQ

complexity: intensity and variety of SQ signature patterns

artefacts: "eye", AR shadow signatures, intense solar storms

Artefacts





Dataset Construction

- sources: SC23 (1), SC24 (2)
- MDI (3) and HMI (4) Dopplergram download (batch script)
- coordinate conversion
- Holography method application (5)
- two obtained datsets (SunquakeNet;DOI10.34740):
    1. acousic emission maps in FITS format: 53 (15 + 38)
    2. grayscale 2D images in JPEG format:
        - positive class: 845 (SC23: 205 + SC24: 640)
        - negative class: 13.055 (SC23: 3891 + SC24: 9164)

,

Experiment Areas

Reconstruction-based learning

unsupervised AutoEncoder feature extraction
feature classification

Contrastive learning

unsupervised and/or supervised CL feature extraction
feature classification

Object detection

region proposal
candidate region classification

Reconstruction-based learning

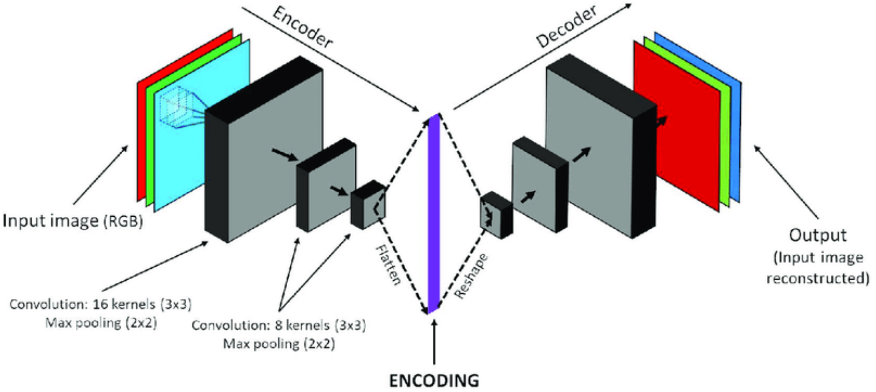


Figure: AutoEncoder architecture

Reconstruction-based experiments (AE & VAE)

- AutoEncoder
  - fully black reconstructions (ls $\leq$ 512)
  - poor classification results
- Variational AutoEncoder (VAE) (6)
  - attempt to capture distinct characteristics
  - almost fully black reconstructions (ls $\leq$ 512)
  - little improvement in classification

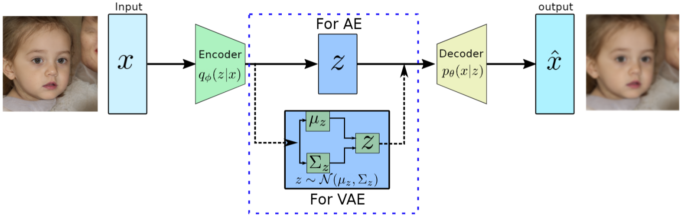
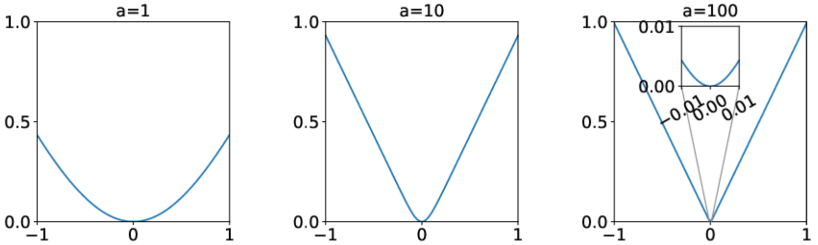


Figure: AE vs. VAE architecture

- Log-Cosh VAE (7)
  - $\uparrow$ values: L_1 loss + $\downarrow$ values: L_2 loss $\Rightarrow$ Log-cosh loss
  - attempt to mitigate the impact of noise
  - poor reconstructions (ls $\leq$ 512)
  - consistent improvement in classification



1

Figure: Log-cosh loss plots

Reconstruction-based learning

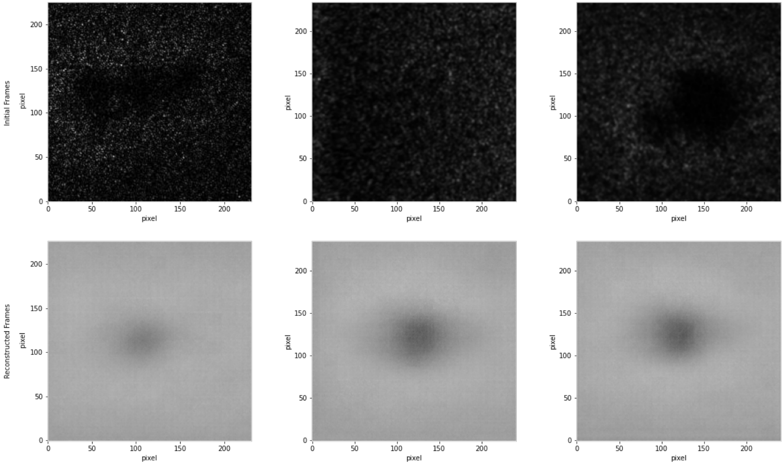
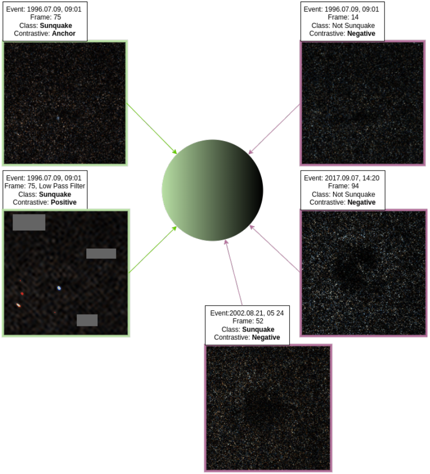


Figure: Log-Cosh VAE reconstructions

Contrastive Learning



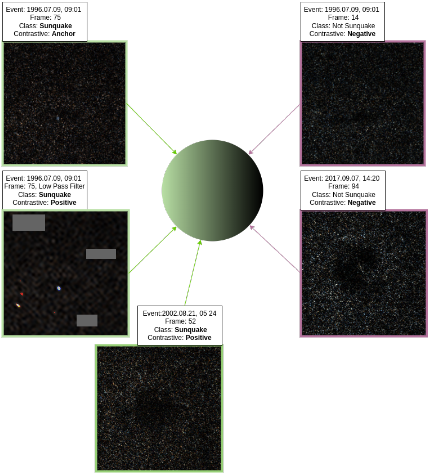


Figure: self-supervised CL (left), supervised CL (right) applied to SunquakeNet

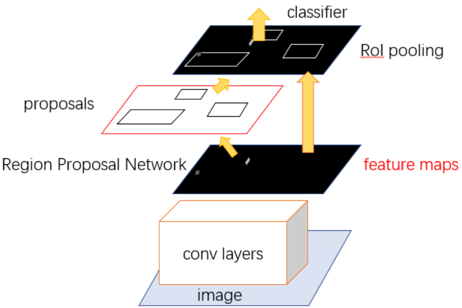
CL Experiments

- self-supervised CL (8)
  - results similar to Log-Cosh VAE
  - sunquake signatures are not consistently captured
- supervised CL (9)
  - attempt to capture distinct characteristics
  - tedious and unstable training due to imbalance
- self-supervised CL using upsampling $\Rightarrow$ supervised CL with a weighted loss (10):
  - attempt to mitigate the impact of imbalance
  - significantly improved results

Object Detection

Steps

identify multiple candidate object regions
classify whether or not each region is a Sunquake
return regions classified as Sunquake



Augmentation Methods

- Custom domain-specific Augmentations:
  - Altered Random Erase:
    - forces the network to be attentive to all areas inside the image
    - replaces the typically used Random Crop
    - decreases the probability of occluding a Sunquake
  - Solarized Low Pass Filter:
    - enhances high frequency signals and fades the others out
    - amplifies details for some Sunquake signatures
  - Time Based Mixing:
    - combines successive grayscale frames into a single 3D image
    - maintains the sequence property of data
- General Augmentations:
  - Geometric transforms:
    - flips (horizontal and vertical)
    - rotations (90, 180, 270)

Altered Random Erase

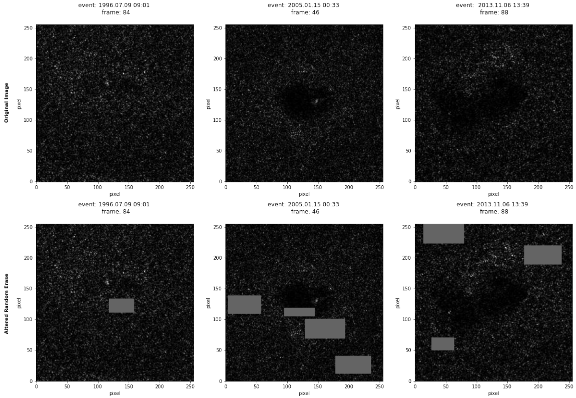


Figure: Visualization of the effect of the Altered Random Erase Augmentation

Solarized Low Pass Filter

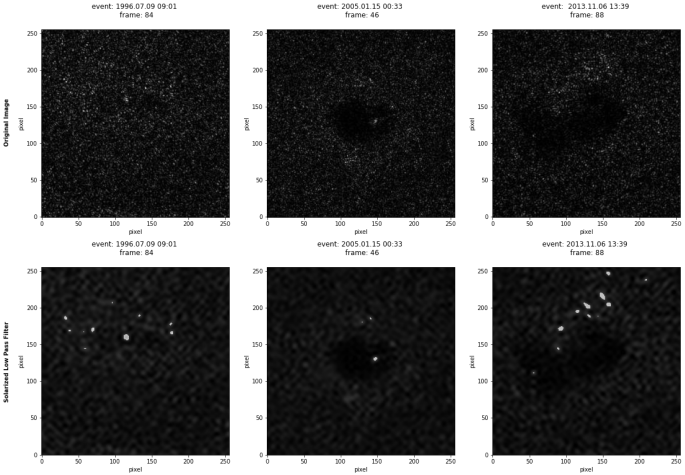


Figure: Visualization of the effect of Solarized Low Pass Filter Augmentation

Time Based Mixing

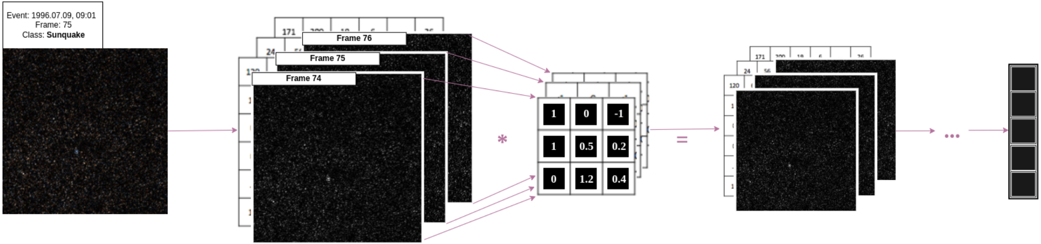


Figure: Visualization of the effect of Time Based Mixing Augmentation

Techniques to address imbalance

loss weighting

positive upsampling

SMOTE

Loss Weighting

Regular supervised CL loss

$$l_i = \sum_{i \in I} \frac{-1}{|P(i)|} \sum_{p \in P(i)} log \frac{e^{\frac{sim(z_i, z_p)}{\tau}}}{\sum_{k=1}^{2N} mask_{[k \neq i]} e^{\frac{sim(z_i, z_k)}{\tau}}} \tag{1}$$

Weighted supervised CL loss

$$l_i = \sum_{i \in I} \frac{-w_{y_i}}{|P(i)|} \sum_{p \in P(i)} log \frac{e^{\frac{sim(z_i, z_p)}{\tau}}}{\sum_{k=1}^{2N} mask_{[k \neq i]} e^{\frac{sim(z_i, z_k)}{\tau}}} \tag{2}$$

$$w_{y_i} = \frac{1 - \beta}{1 - \beta^{N_{y_i}}} \tag{3}$$

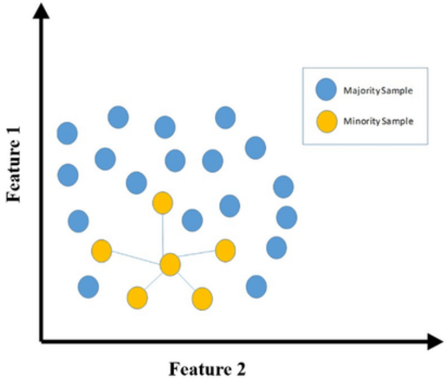
Positive upsampling

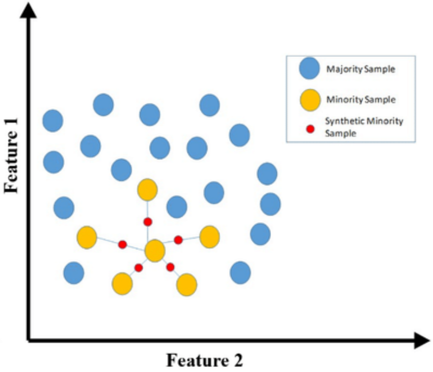
- the number of positive image samples is increased for unsupervised methods
- five extra copies are generated for each image (flips and rotations)
- may impose a transformation bias to the model if used in supervised methods

SMOTE

Definition

synthetic minority oversampling technique





Best models (so far)

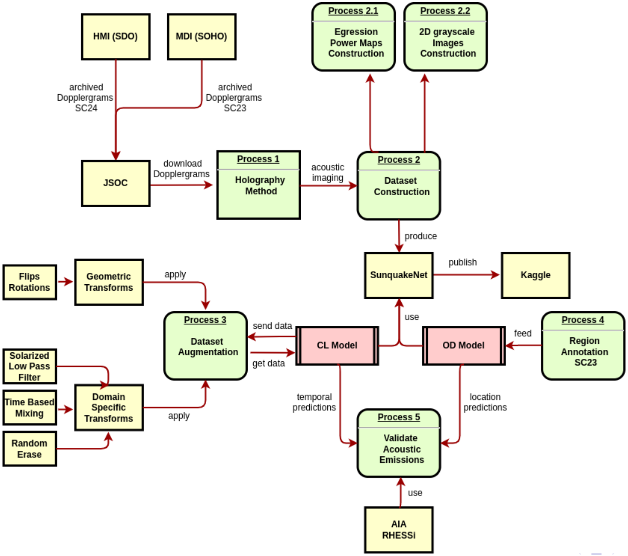
Sunquake time prediction

- CL Model with shuffle bias
  - frame level shuffling ⇒ shared Active Regions between the training and validation sets
  - training specifications (two-step):
    - supervised CL, 50 epochs: DenseNet-121, temperature 0.1, positive upsampling
    - embeddings classification: various classifiers trained using SMOTE
- CL Model with no bias
  - event level shuffling ⇒ unique Active Regions in the training and validation sets
  - training specifications (three-step):
    - self-supervised CL, 500 epochs: ResNet-18, positive upsampling
    - supervised CL 100 epochs: weighted loss, temperature 0.07
    - embeddings classification: various classifiers trained using SMOTE

Sunquake location prediction

- Object Detection Model
  - faster R-CNN, 50 epochs, trained only on positive SC23 image data and regions

Process Diagram



Numerical Results

Classifier	K-NN (Bagging)		SVC (Poly)		SVC (RBF)		Logistic Regression		SGD	
Augmentation	None	SMOTE	None	SMOTE	None	SMOTE	None	SMOTE	None	SMOTE
Precision	0.97	0.80	0.97	0.98	**0.99**	0.94	0.97	0.91	0.91	0.89
Recall	0.54	0.96	0.96	0.97	0.86	**0.99**	0.98	**0.99**	0.94	0.93
F1-Score	0.55	0.86	0.97	0.97	0.91	0.96	**0.98**	0.95	0.94	0.93
Accuracy	0.94	0.95	**0.99**	**0.99**	0.98	**0.99**	**0.99**	**0.99**	0.98	0.98
Metrics Avg	0.750	0.892	0.947	**0.977**	0.935	0.97	**0.98**	0.96	0.942	0.932

Classifier	K-NN (Bagging)		SVC (Poly)		SVC (RBF)		Logistic Regression		SGD	
Augmentation	None	SMOTE	None	SMOTE	None	SMOTE	None	SMOTE	None	SMOTE
Precision	0.63	0.65	0.66	**0.84**	0.49	0.59	0.64	0.64	0.54	0.62
Recall	0.54	0.54	0.54	0.54	0.50	0.54	0.54	0.54	0.54	0.54
F1-Score	0.55	0.55	0.55	0.55	0.49	0.55	0.55	0.55	0.54	0.55
Accuracy	0.93	0.93	0.93	**0.94**	0.93	0.92	0.93	0.93	0.89	0.93
Metric Avg	0.662	0.667	0.67	**0.715**	0.605	0.65	0.665	0.665	0.627	0.66

Table: Macro Average performance of different classifiers over embeddings produced by the CL model with **top: shuffle bias**, **bottom: no bias**, trained with and without SMOTE augmentation, for the test data in **SC23 & SC24** (2622 negative and 186 positive samples)

Prediction Distribution

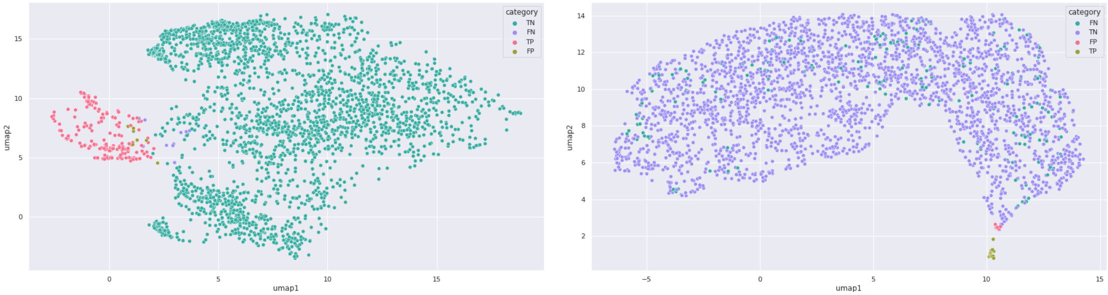


Figure: Predictions of the *left*: shuffle bias model; *right*: unbiased model, for the test data in SC23 & SC24, clustered by UMAP components

Prediction Analysis (unbiased model)

Event Date	TP	FP	FN	GT	Event Date	TP	FP	FN	GT
1996-07-09 09:01	0	0	19	19	2012-07-04 09:47	2	0	17	19
2001-04-06 19:13	0	0	16	16	2012-07-06 13:26	2	0	15	17
2001-09-24 09:35	0	0	11	11	2013-11-08 04:20	2	0	18	20
2002-07-23 00:27	0	6	14	14	2015-03-11 16:11	2	0	14	16
2012-03-05 19:27	2	0	19	21	2015-09-28 14:53	2	0	18	20
2012-03-06 07:52	2	0	11	13					

Table: SVC (poly) predictions for embeddings produced by the unbiased CL model for the test data in SC23 & SC24 (2622 negative și 186 positive samples).

Cosine distance characteristic analysis (unbiased model)

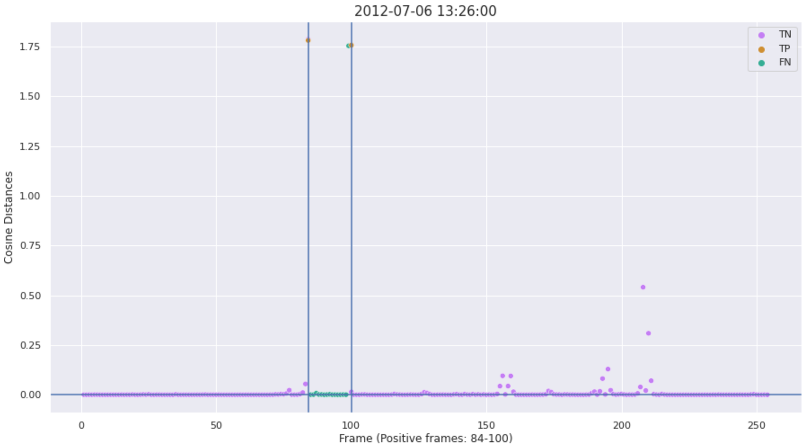


Figure: Cosine Distances computed between consecutive frames embeddings for the event at *2012-07-06 13:26* in the test set

Mean characteristic analysis (unbiased model)

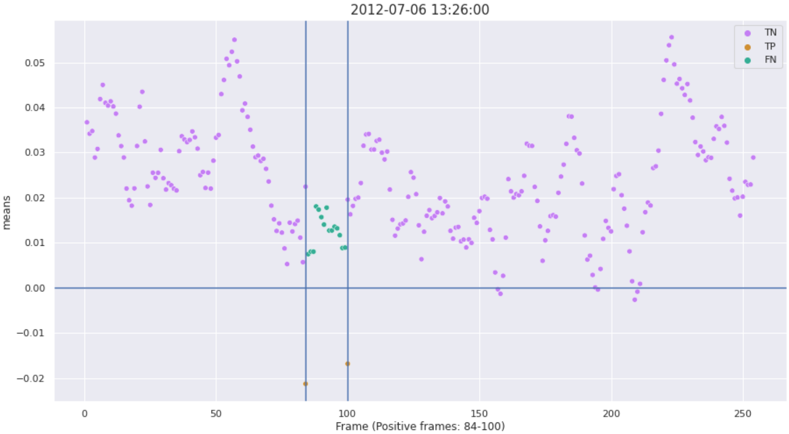


Figure: Means computed for frame embeddings for the event at *2012-07-06 13:26* in the test set

Cosine distance characteristic analysis (biased model)

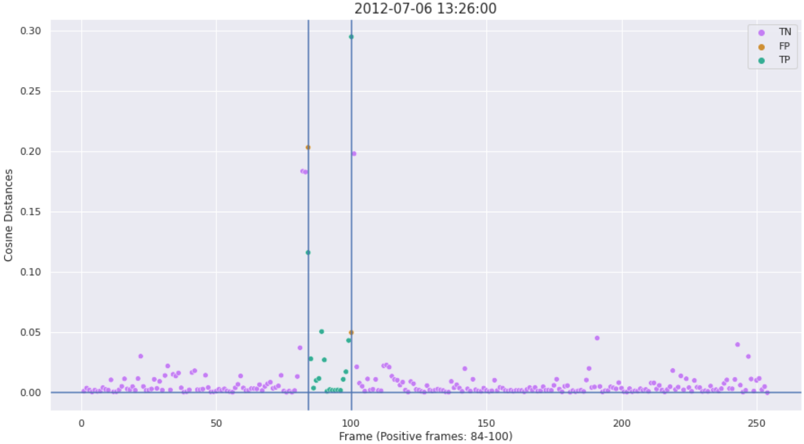


Figure: Cosine Distances computed between consecutive frames embeddings for the event at *2012-07-06 13:26* in the test set

Unlabeled (problematic) dataset analysis

problematic data cubes

AIA/RHESSI prediction analysis on additional datasets

Predictions for the problematic 2012.05.08 13:02 event

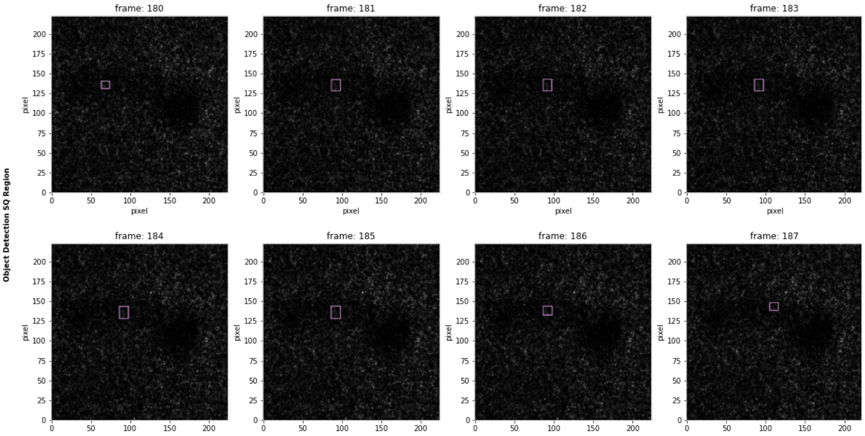


Figure: Position of candidate SQ signatures, at frames $[180, 188)$, predicted positive for the *2012.05.08 13:02* event in the unlabeled dataset.

Prediction validation for the problematic 2012.05.08 13:02 event

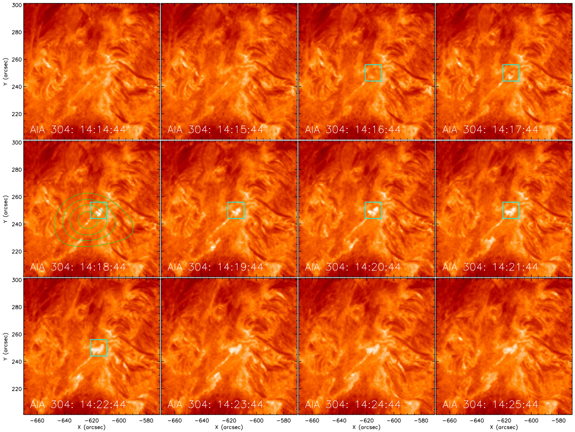


Figure: Flaring Activity observed in the AIA 304 Å channel and the RHESSI location of seismic signatures for the *2012.05.08 13:02* event in the unlabeled dataset.

Predictions for the problematic 2011.12.30 03:03 event

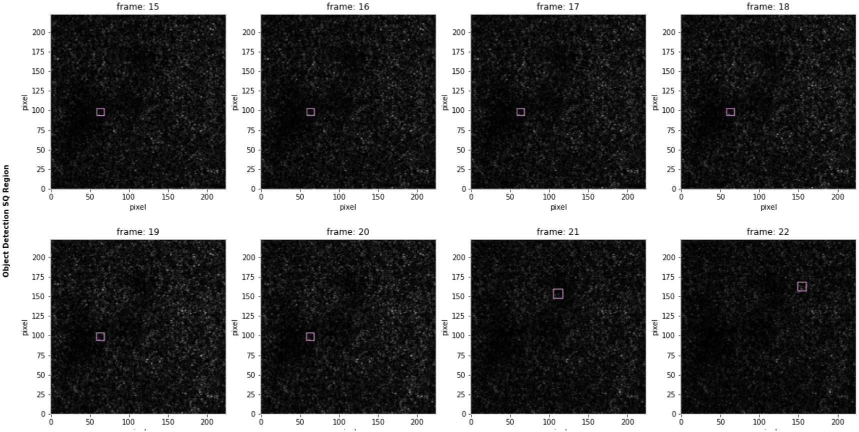


Figure: Position of candidate SQ signatures, at frames $[15, 22)$, predicted positive for the *2011.12.30 03:03* event in the unlabeled dataset.

Prediction validation for the problematic 2011.12.30 03:03 event

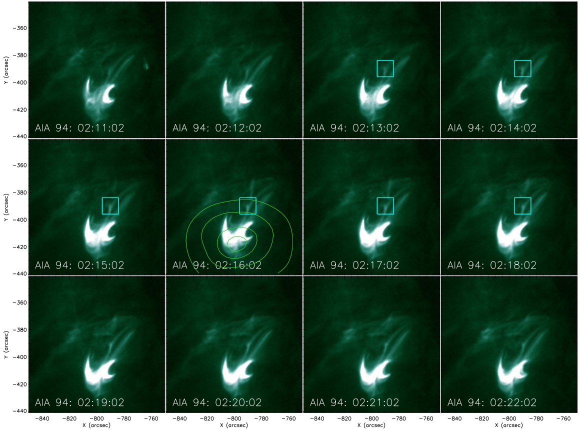


Figure: Flaring Activity observed in the AIA 94 Å channel and the RHESSI location of seismic signatures for the *2011.12.30 03:03* event in the unlabeled dataset.

Summary

- constructed and published two datasets of SC23 & SC24
  - strongly imbalanced
  - with artefacts
  - complex
- implemented multiple ML detection methods:
  - AutoEncoders (AE, VAE, Log-Cosh VAE)
  - CL (self-supervised, supervised, with weighted loss)
  - Object Detection (Faster R-CNN)
  - Various classifiers
- implemented 3 custom domain-specific augmentations:
  - Altered Random Erasing
  - Solarized Low Pass Filter
  - Time Based Mixing

Findings

AE results are poor, but a level of clustering was observed at event level

distinct CL models provide consistent results

distinct CL models capture a low similarity between
Sunquake transition frames and their neighbors

correlations are found between predicted lower acoustic emission sources
and Solar eruptions accompanied by high-energy X-Ray emissions

Prospect

Time and location prediction components are separated

improve CL models with explainability techniques

Models predict Sunquakes of a too short duration

mix more than 3 frames in the Time Based Mixing augmentation

Spurious correlations are not addressed

create a more fine-grained separation of signature pattern classes

Thank you!

Don't hesitate to reach out if you have any questions or ideas related to this topic!

Mercea.Fl.Vanessa@student.utcluj.ro / mercea.vanessa@yahoo.com

[1] D. Besliu-Ionescu, A. Donea, P. Cally, and C. Lindsey, “Web-based Comprehensive Data Archive of Seismically Active Solar Flares,” in *Advances in Solar and Solar-Terrestrial Physics*, 2012, p. 31.

[2] I. N. Sharykin and A. G. Kosovichev, “Sunquakes of Solar Cycle 24,” , vol. 895, no. 1, p. 76, May 2020.

[3] P. H. Scherrer, R. S. Bogart, R. I. Bush, J. T. Hoeksema, A. G. Kosovichev, J. Schou, W. Rosenberg, L. Springer, T. D. Tarbell, A. Title, C. J. Wolfson, I. Zayer, and MDI Engineering Team, “The Solar Oscillations Investigation - Michelson Doppler Imager,” , vol. 162, pp. 129–188, 1995.

[4] J. Schou, P. H. Scherrer, R. I. Bush, R. Wachter, S. Couvidat, M. C. Rabello-Soares, R. S. Bogart, J. T. Hoeksema, Y. Liu, T. L. Duvall, D. J. Akin, B. A. Allard, J. W. Miles, R. Rairden, R. A. Shine, T. D. Tarbell, A. M. Title, C. J. Wolfson, D. F. Elmore, A. A. Norton, and S. Tomczyk, “Design and Ground Calibration of the Helioseismic and Magnetic Imager (HMI) Instrument on the Solar Dynamics Observatory (SDO),” *"Solar Physics"*, vol. 275, pp. 229–259, Jan. 2012.

[5] C. Lindsey and D. C. Braun, “Basic Principles of Solar Acoustic Holography - (Invited Review),” , vol. 192, pp. 261–284, Mar. 2000.

[6] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[7] P. Chen, G. Chen, and S. Zhang, "Log hyperbolic cosine loss improves variational auto-encoder," 2019. [Online]. Available: https://openreview.net/forum?id=rkglvsC9Ym

[8] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proceedings of the 37th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, H. D. III and A. Singh, Eds., vol. 119. PMLR, 13–18 Jul 2020, pp. 1597–1607. [Online]. Available: https://proceedings.mlr.press/v119/chen20j.html

[9] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 18 661–18 673. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/d89a66c7c80a29b1bdbab0f2a1a94af8-Paper.pdf

[10] Z. Zhong, J. Cui, E. Lo, Z. Li, J. Sun, and J. Jia, "Rebalanced siamese contrastive mining for long-tailed recognition," *arXiv preprint arXiv:2203.11506*, 2022.