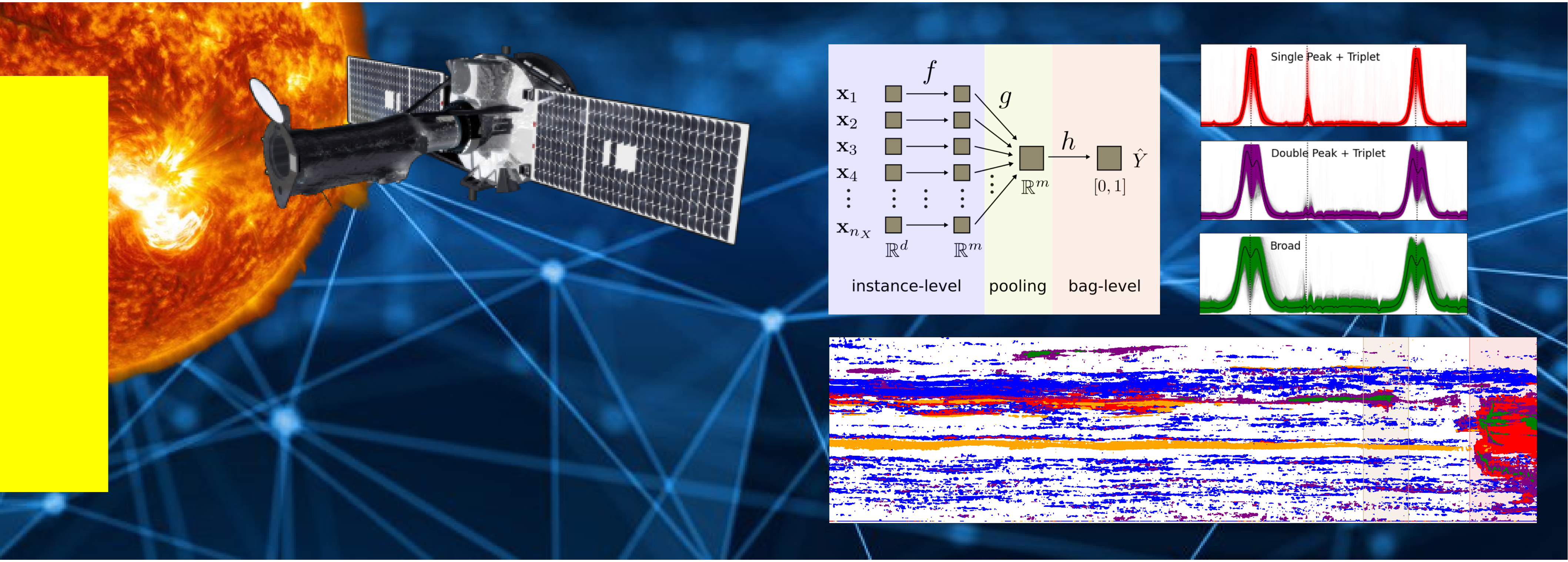
University of Applied Sciences and Arts
Northwestern Switzerland

member of
swissuniversities



Using Multiple Instance Learning for Explainable Flare Prediction

Cédric Huwyler

Institute for Data Science, FHNW Windisch, Switzerland

ML-Helio, 21 March 2022

Outline

Using Multiple Instance Learning for Explainable Solar Flare Prediction

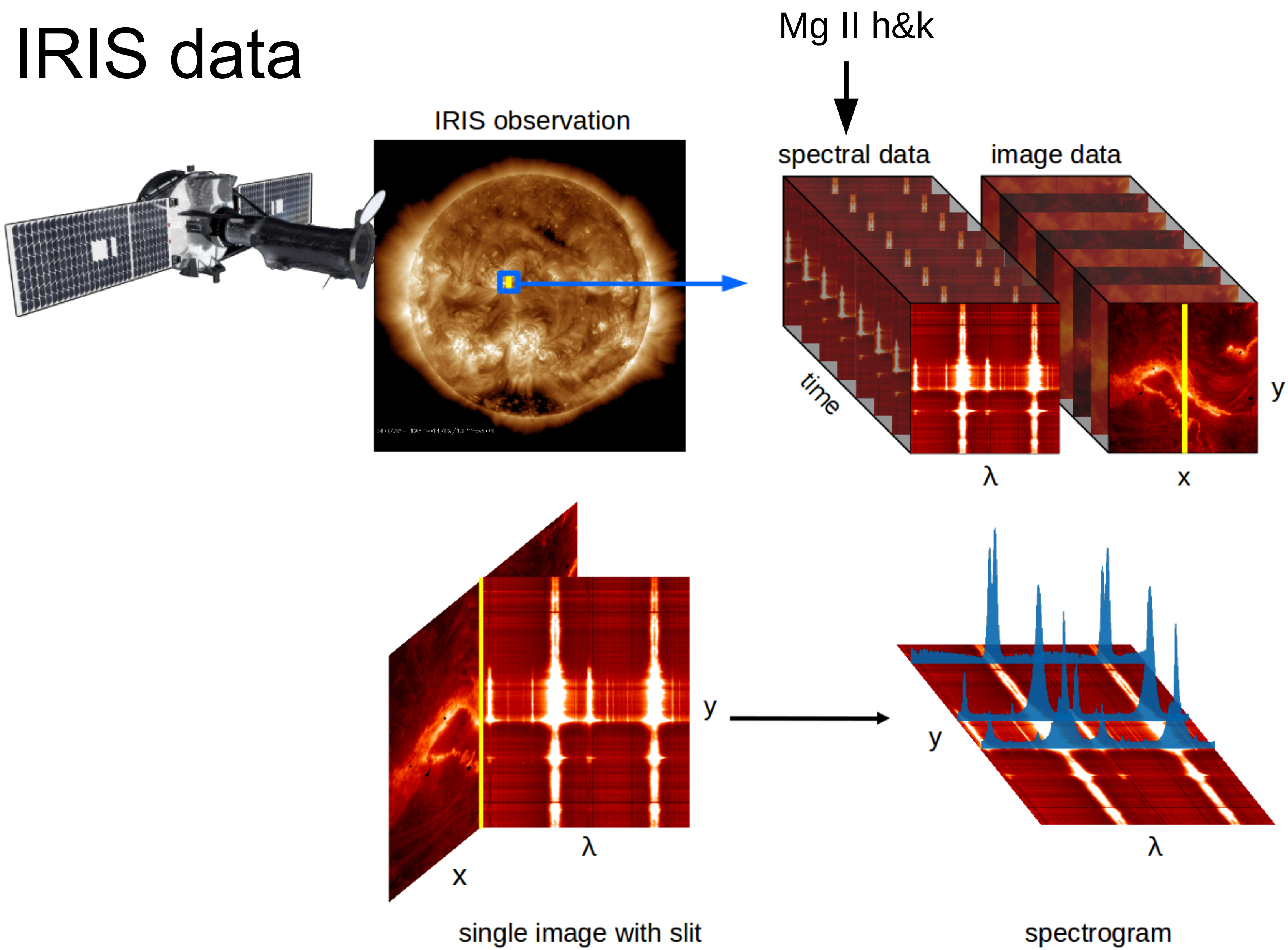
C. Huwyler[1], M. Melchior

Institute for Data Science, Fachhochschule Nordwestschweiz (FHNW)
Bahnhofstrasse 6, 5210 Windisch
Switzerland

Main message:
When labels are only available on a higher hierarchical level,
use Multiple Instance Learning! (also works for images)

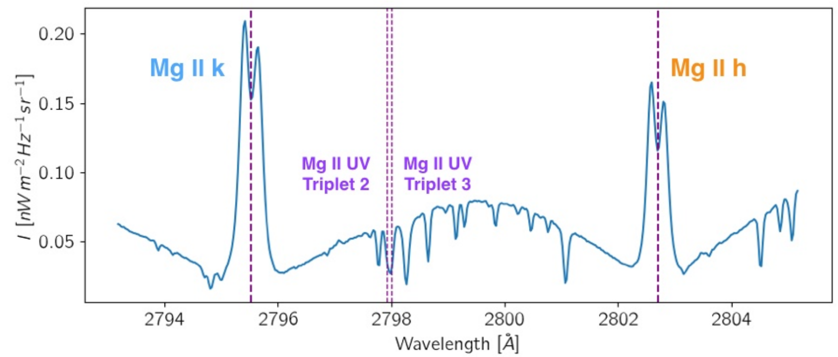
Use case:
Explainable flare prediction with IRIS spectral data

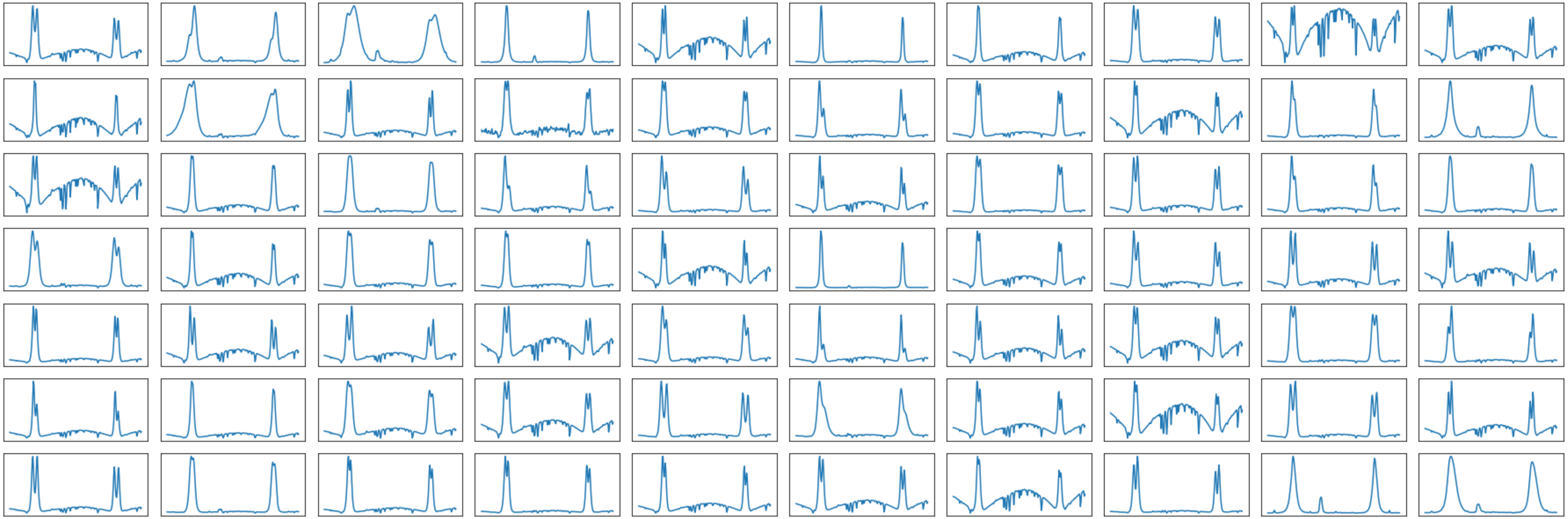
IRIS data


Mg II h&k
IRIS observation
spectral data
image data
time
λ
x
y
y
x
λ
y
λ
single image with slit
spectrogram

Mg II h&k spectral profiles

source: Lockheed Martin





Research Problem

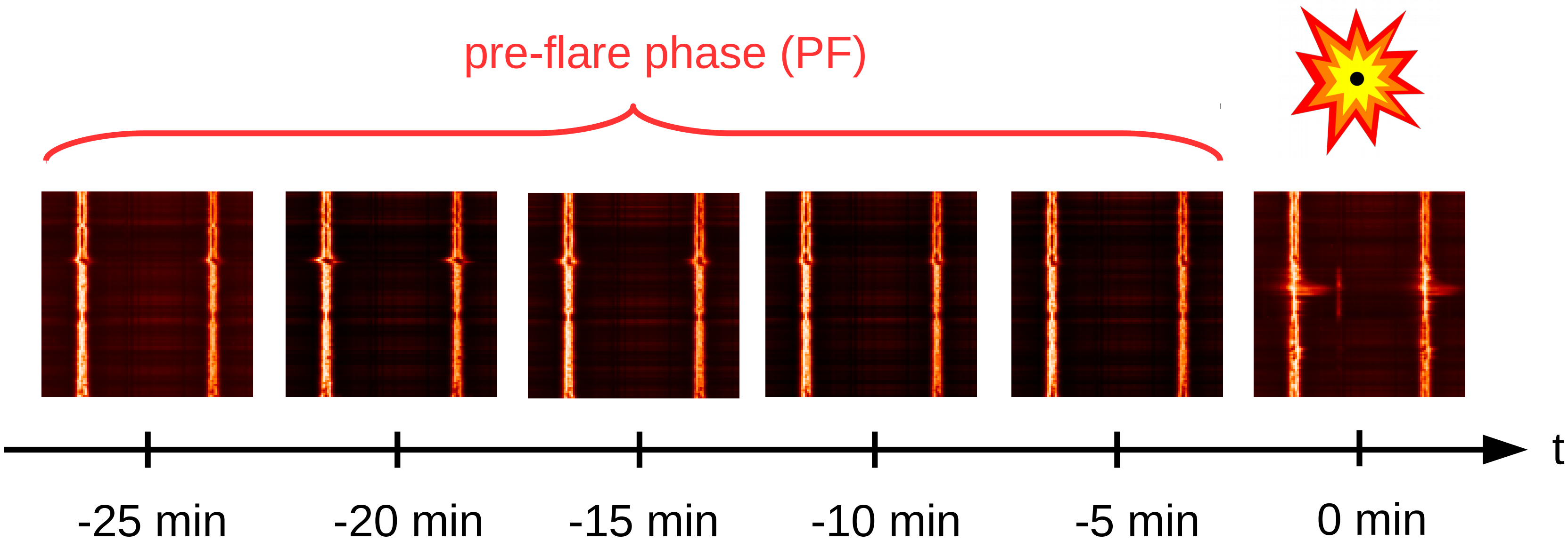
▶ **Which types of Mg II h&k spectra appear shortly before a flare?**

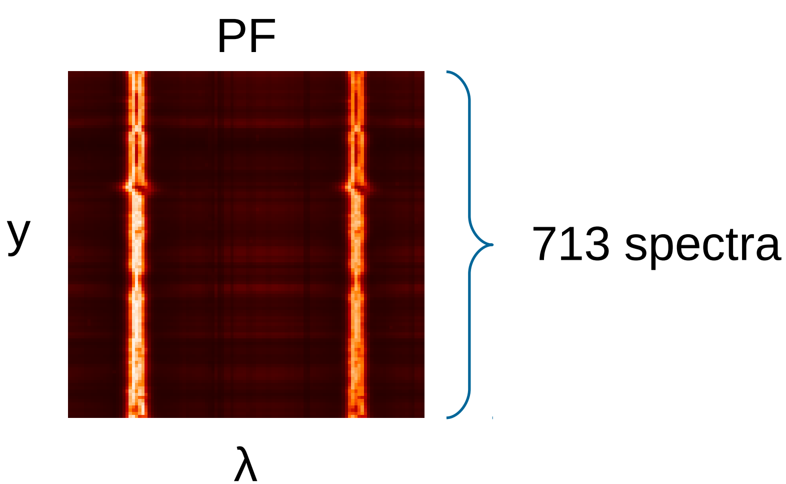
▶ **How well can they be used to predict flares?**

Dataset:

- 37 carefully selected IRIS observation windows
  - 18 non-flaring active regions (AR)
  - 13 M and 6 X-class pre-flare active regions (PF)
- Each of 25 minutes length
- Each with tens to hundreds of timesteps and a few hundred y-pixels
- Balanced classes (not an operational scenario)
- Spectra interpolated to 240 bins, stored as vectors
- Spectra normalized to range [0, 1]

Flare prediction with IRIS





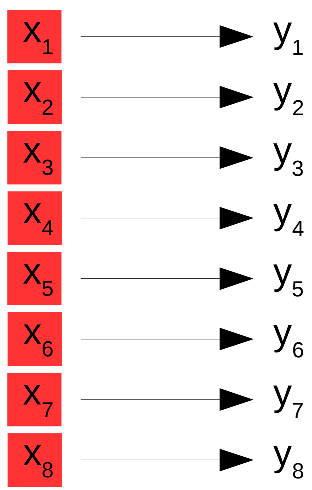
One timestep:

1 label, but ‘bag’ of 713 spectra

~10’000 timesteps in 37 observations

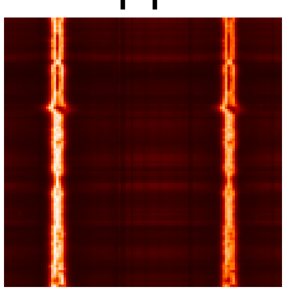
Multiple Instance Learning

Supervised Learning:



PF

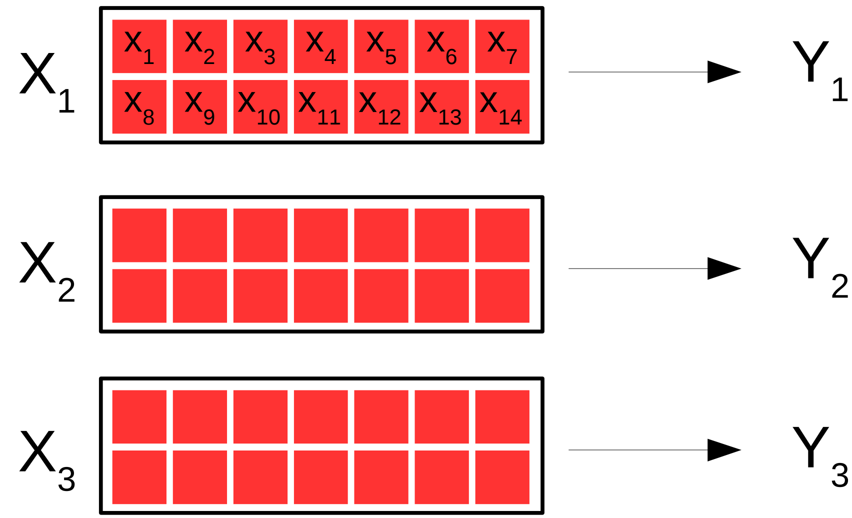
y



713 spectra

λ

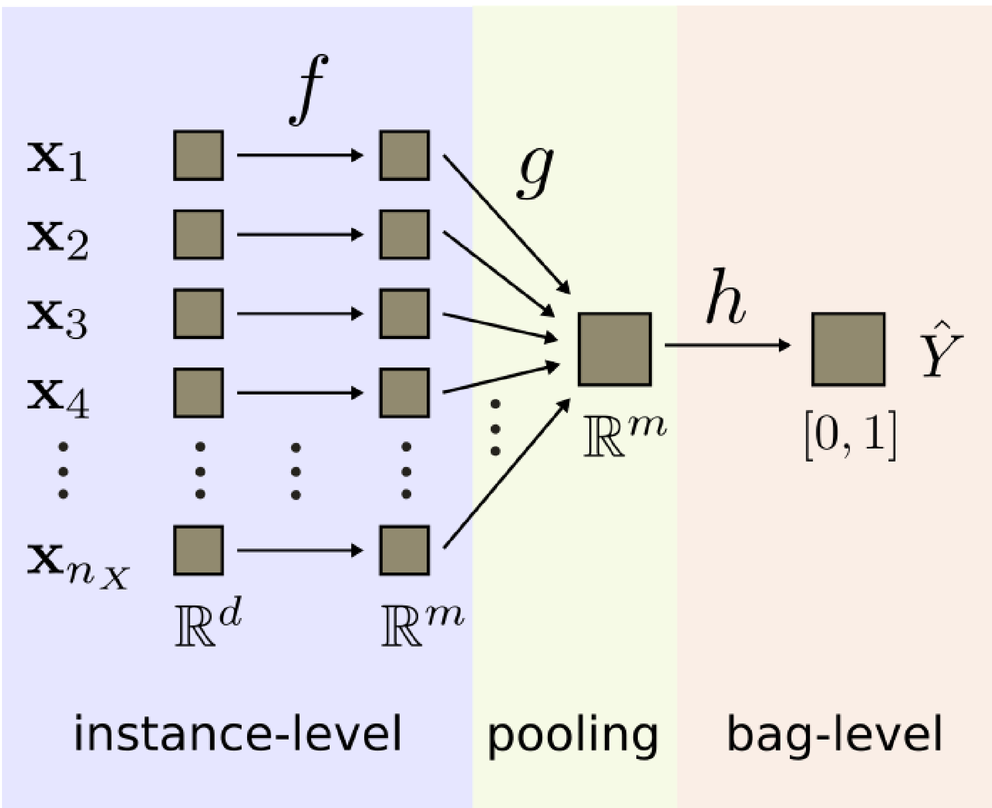
Multiple Instance Learning (MIL):



$$X = \{\underset{\mathbb{R}^d}{\overset{\Downarrow}{\mathbf{x}_k}}\}_{k=1,\dots,n_X}$$

$$Y \in \{0,1\}$$

Multiple Instance Learning



C. Huwyler, M. Melchior, 2021

$$X = \{\underset{\mathbb{R}^d}{\overset{}{\mathbf{x}_k}}\}_{k=1,\ldots,n_X}, \quad \mathbf{x}_k \in \mathbb{R}^d$$

$$Y \in \{0,1\}$$

$$\hat{Y} = S(X)$$

$$S = h \circ g \circ f$$

Neural networks:

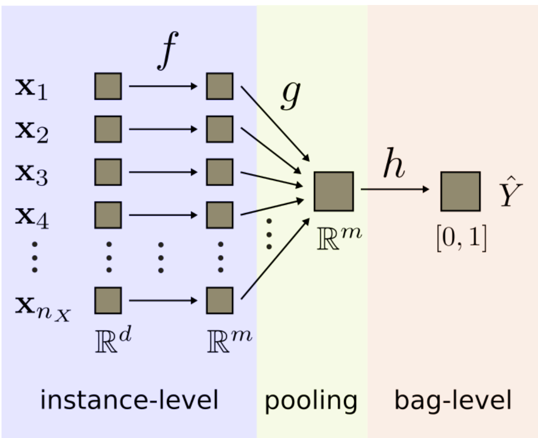
$f = f_\theta$ ('mapper')

$g = g_\varphi$ ('pooling')

$h = h_\psi$ ('classifier')

Instance-based MIL

$$S = h \circ g \circ f$$



Use f to assign a probability to every instance:

$$p_k = f_\theta(\mathbf{x}_k) \in [0, 1]$$

Log-Mean-Exp to interpolate between mininium and maximum pooling:

hyperparameter

$$g(p_k) = \frac{1}{r} \log \left(\frac{1}{n_X} \sum_k e^{r p_k} \right) \in [0, 1]$$

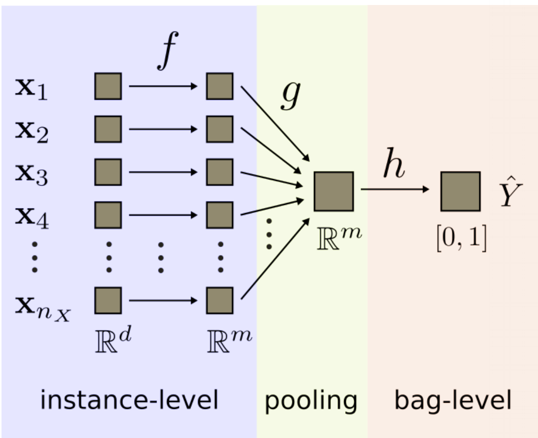
$$\lim_{r \to 0} g(p_k) = \frac{1}{n_X} \sum_k p_k$$

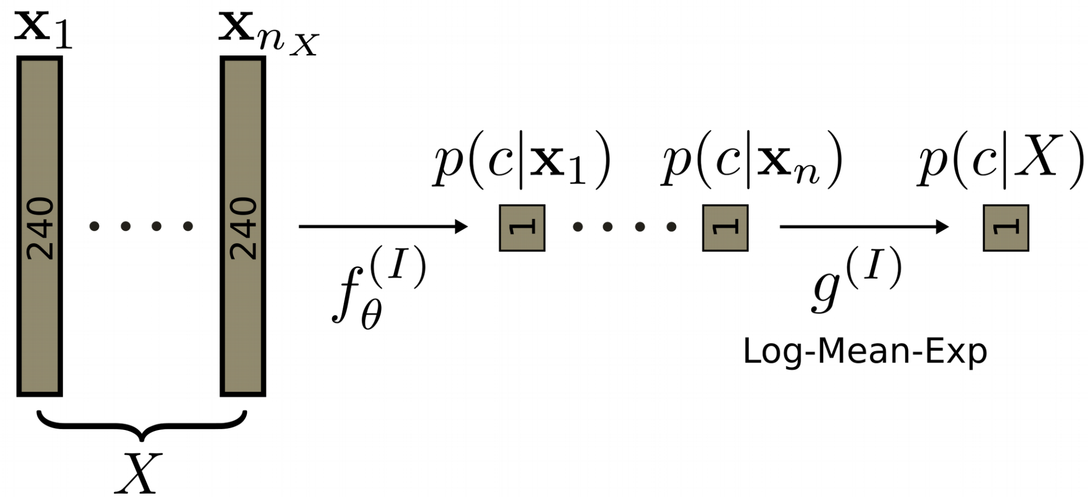
$$\lim_{r \to \infty} g(p_k) = \max_k p_k$$

The pooling g already returns a probability, the function **h can be set to the identity.**

Instance-based MIL

$$S = h \circ g \circ f$$

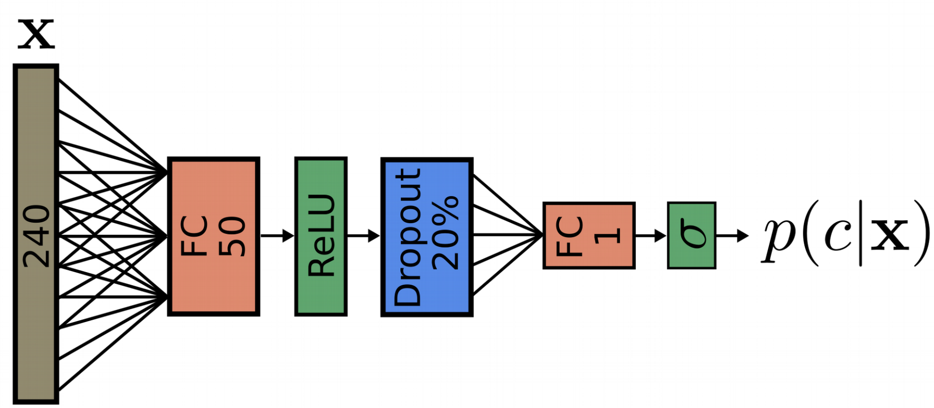




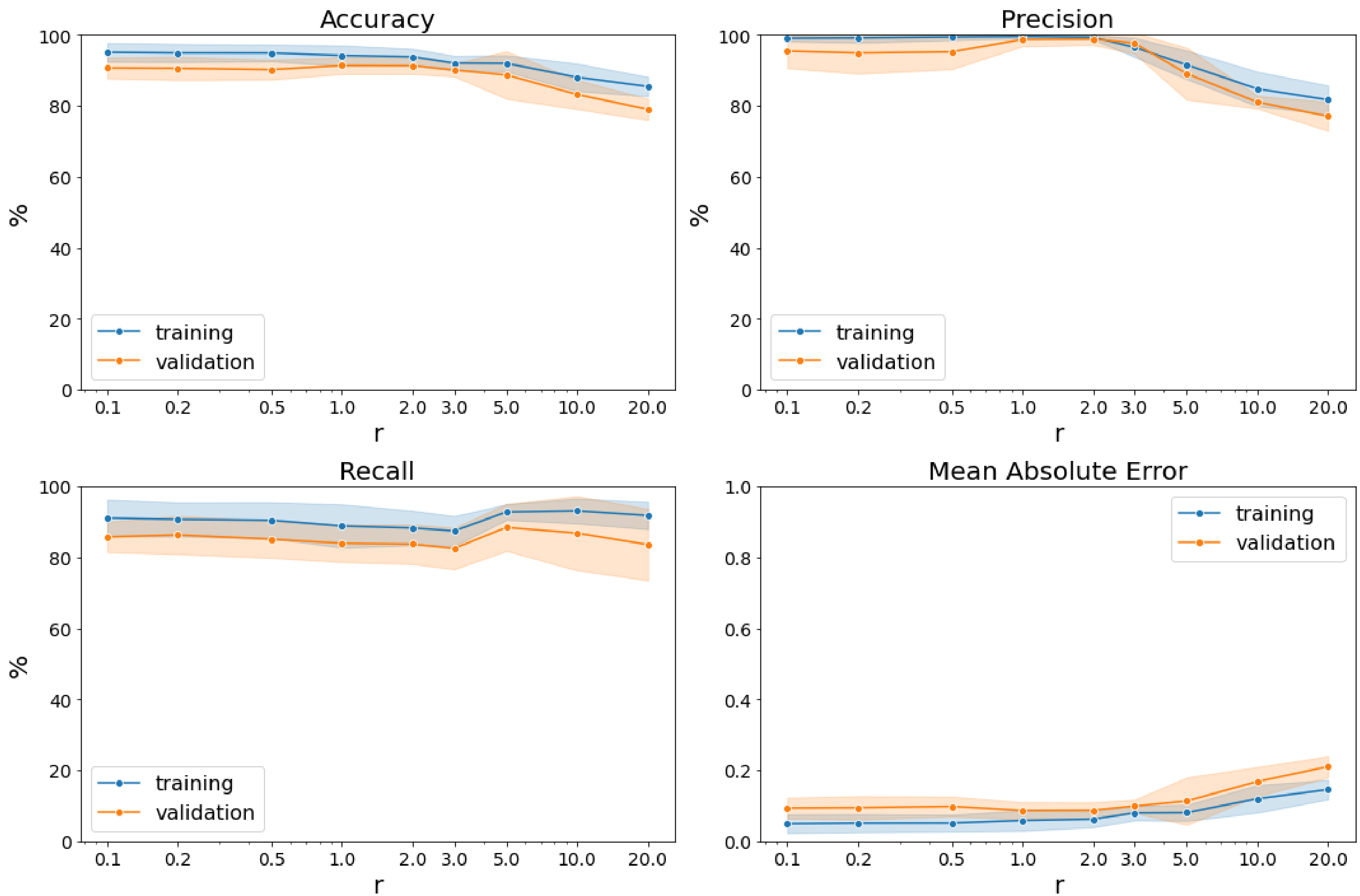
Training:

- Dataset with 10’000 labeled bags of spectra
- MAE loss
- Early stopping

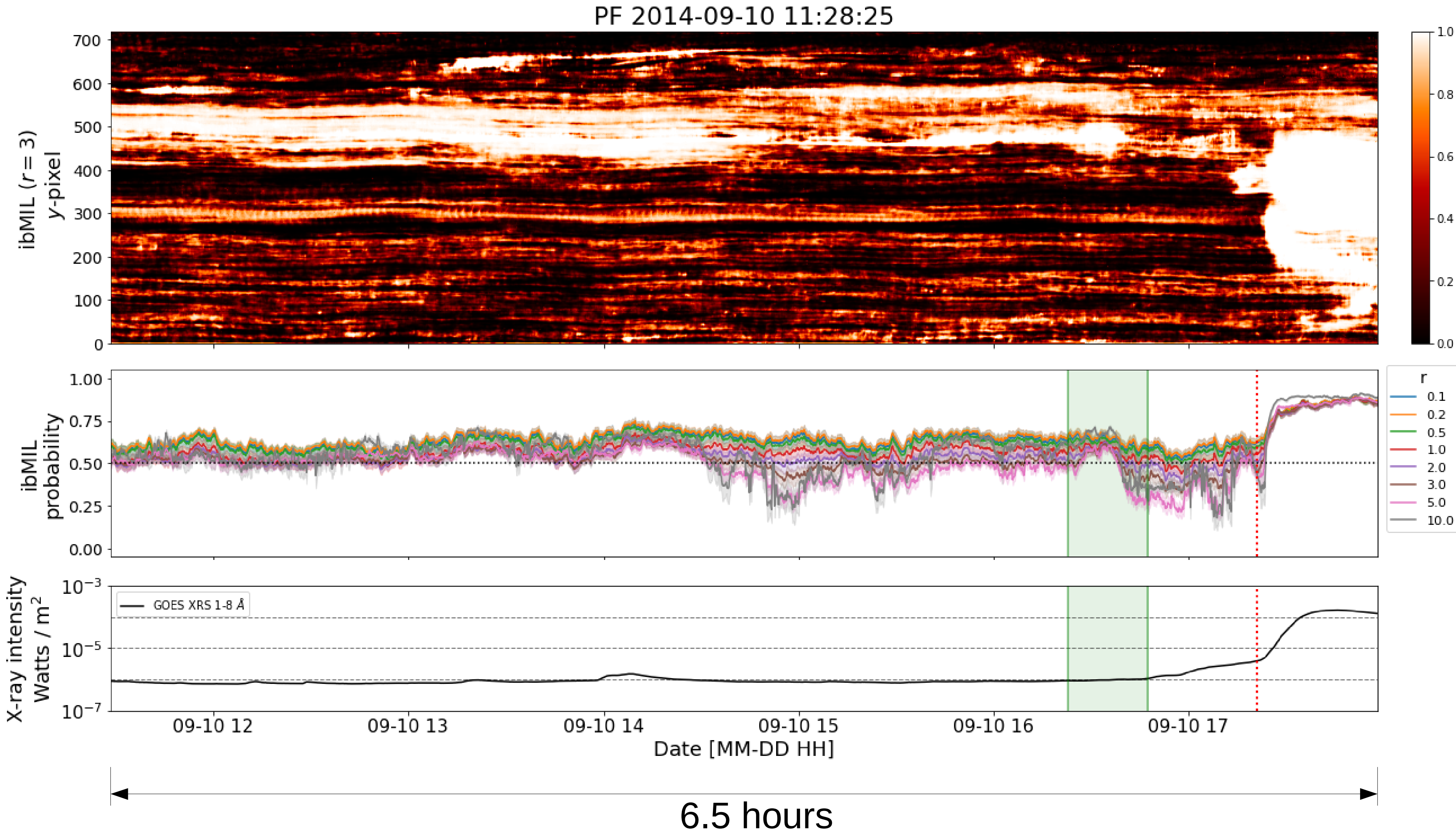
$f_\theta^{(I)}$:



Bag-Level: Prediction Results

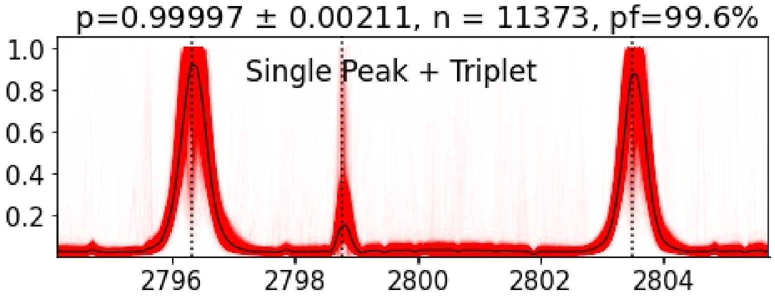


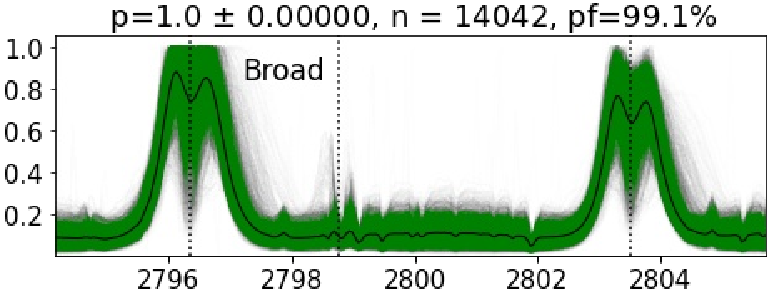
Instance-Level: Saliency Maps

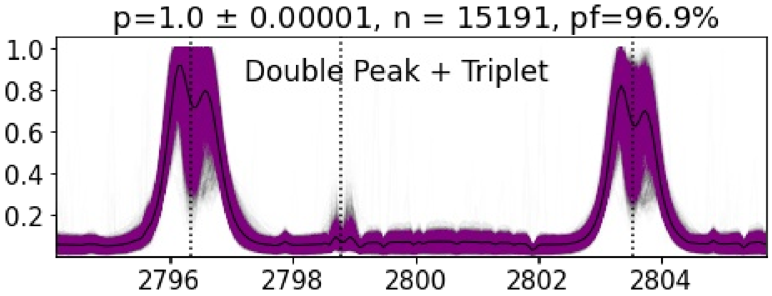


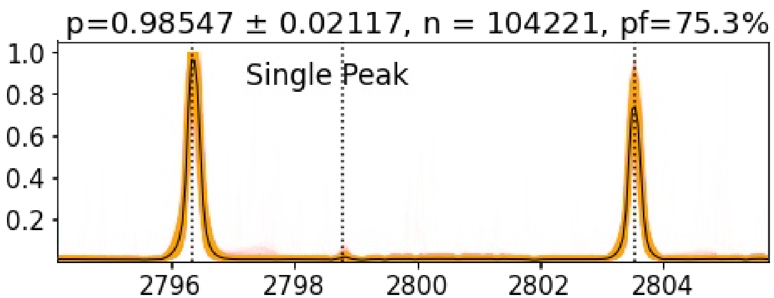
Instance-Level: Characteristic Pre-Flare Spectra

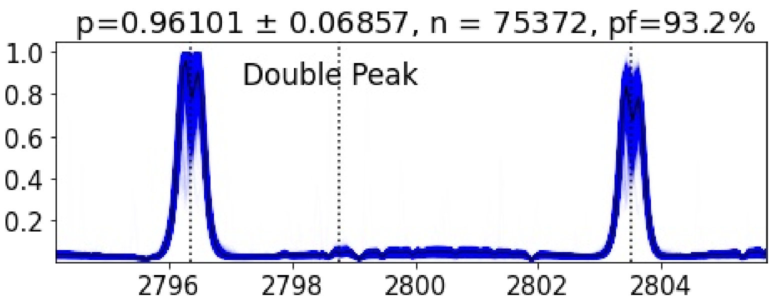
k-means clustering with k=128 – 5 particular types of spectra for high average probabilities:



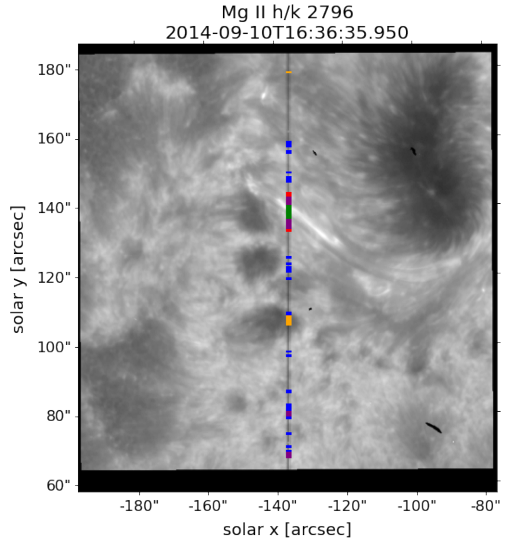


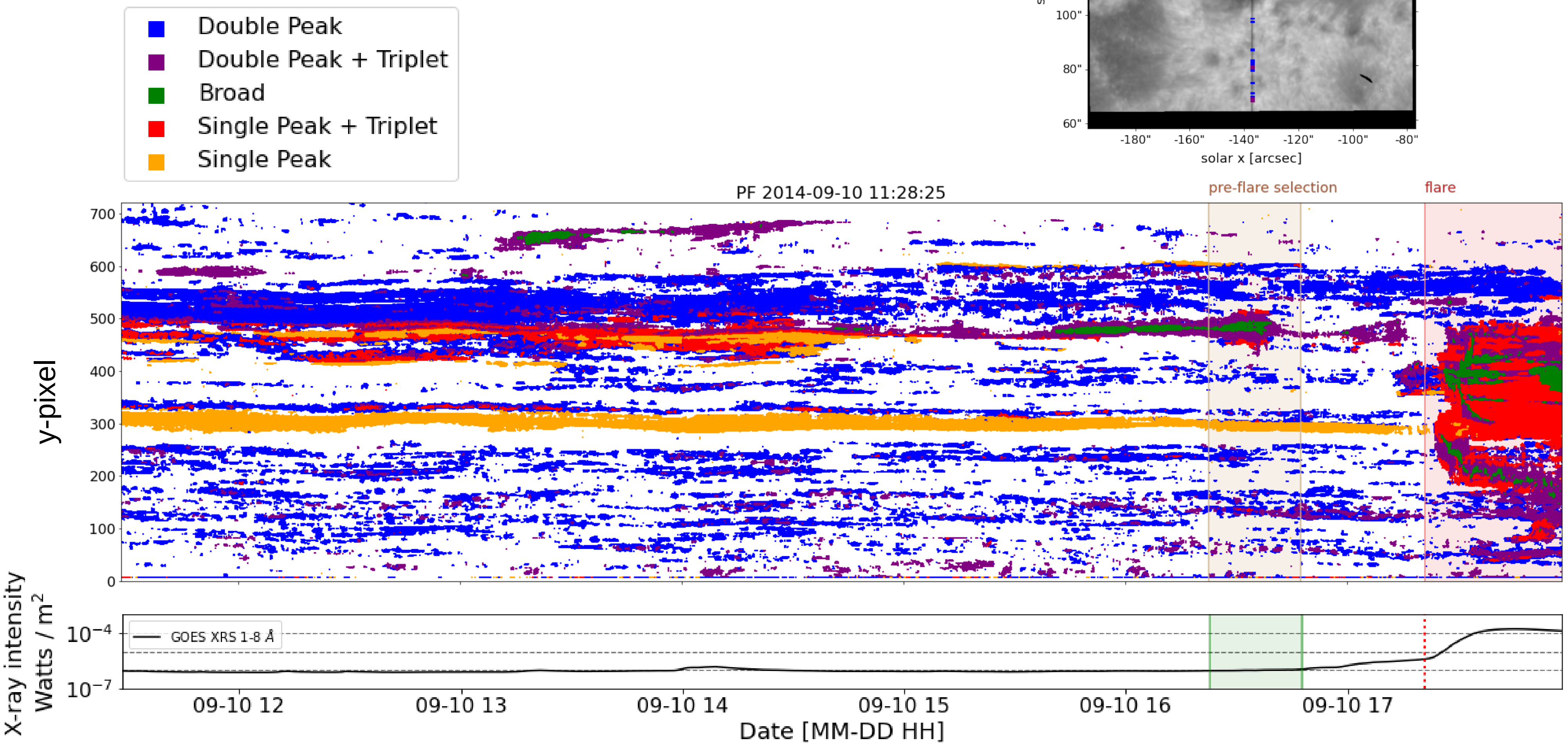




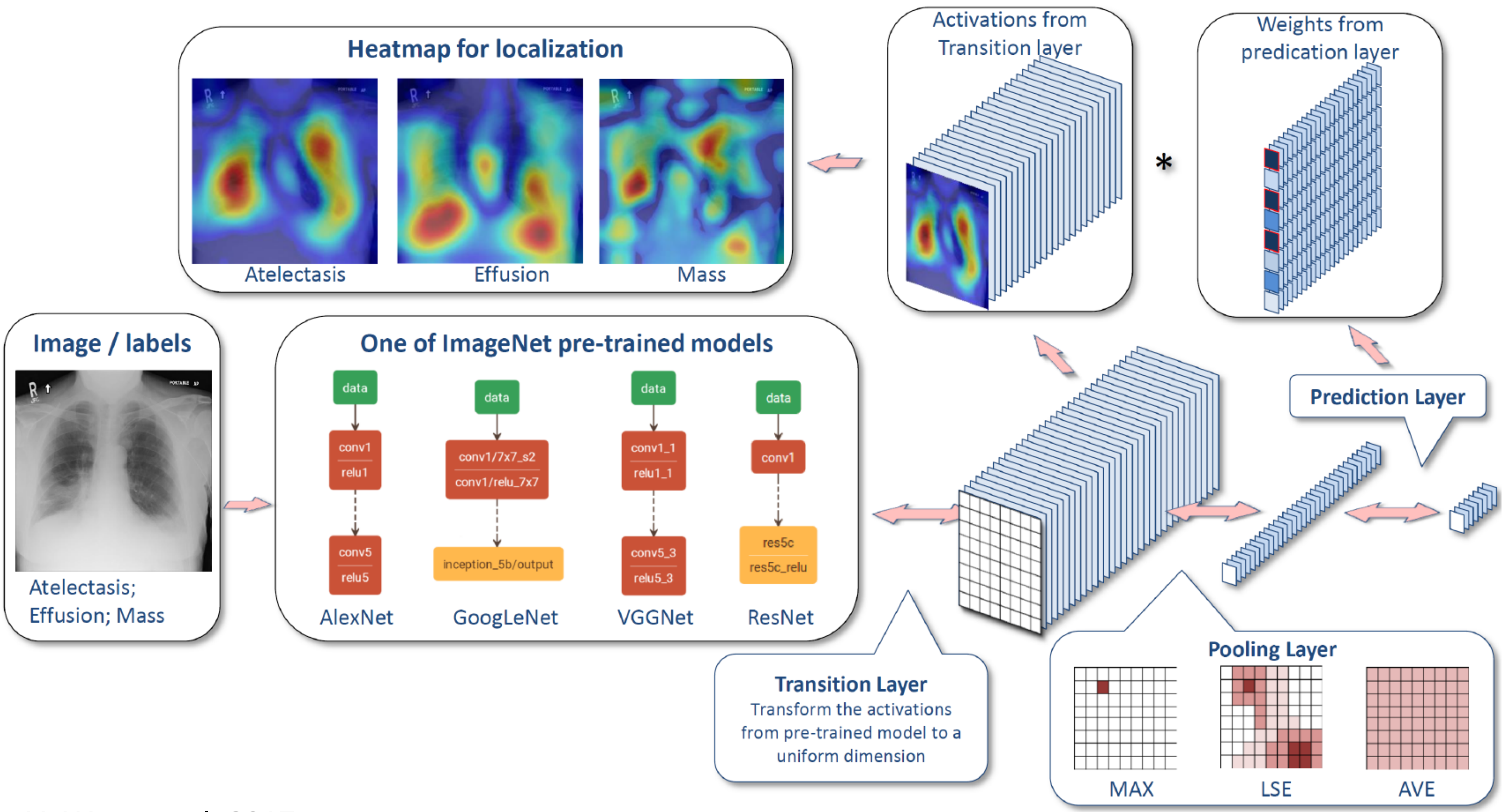


Visualizing Groups





Multiple Instance Learning on Images



X. Wang et al, 2017
ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases

Conclusion

- Use MIL when labels are not available at level of individual instances, but higher up.
- Use MIL when your problem can be divided into instances and you desire explainability of bag-level predictions.
- MIL also works for image segmentation tasks – to be used on image data, such as AIA / HMI!