Solar Synthetic Imaging: Introducing Denoising Diffusion Probabilistic Models on SDO/AIA Data

European Space Weather Week ESWW

Francesco Pio Ramunno, André Csillaghy
Fachhochschule Nordwestschweiz FHNW

Toulouse, France

20-24 November, 2023

Goal of the project

Machine Learning Goal

The goal of this project is to utilize generative models, specifically diffusion models, to produce images of the Sun with a specific amount of activity present.

Why do we want to do this analysis?

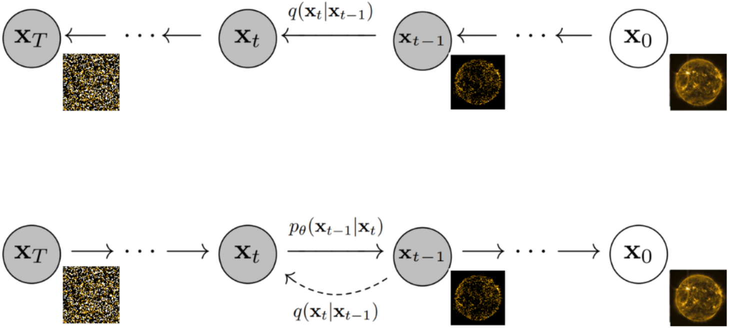
Generate the rarest events (e.g., M- or X-flares) to solve the problem of the unbalanced dataset, being able to investigate these phenomena more extensively with more data.

Data sources

We used three data sources:

- **SDOMLv2**: a subset of the SDO data already prepared for machine learning studies (Galvez, Richard et al 2019),
- **GOES X-Ray Sensor (XRS)**: soft X-ray measurements in the XRSB (1-8 Å) band,
- **Heliophysics Events Knowledgebase (HEK)**: peak time and GOES labels of flaring events.

What are the DDPMs?



Adapted from Ho et al. 2020

Experiments setup and labelling systems

We use the following setup:

- Image resolution: 64x64,
- Number of Epochs: 500 (for each model),
- Batch size = 12.

We perform the following distinct experiments:

- Discrete labels: A, B, C, M and X,
- Continous labels: X-ray values,
- Discrete labels + ceVAE embeddings

We compare the results with the following baseline model:

- ceVAE (Giger M., 2022)

Metrics

To evaluate the best model we used the following metrics:

- **Cluster Metrics** (Hackstein et al., 2023): determine if the generated distribution is similar to the true distribution. The cluster metrics can be divided into:
  1. Cluster Error (CE),
  2. Cluster Distance (CD),
  3. Cluster Standard Deviation (CS).
- **FID** (Heusel et al., 2017): determine the image quality level and the completeness of the generated distribution. The FID is computed using the following encoders:
  1. CLIP (Alec Radford et al., 2021)
- **F1 score**: check whether the generated image of a particular class (e.g. X) is similar to a true image of that class.

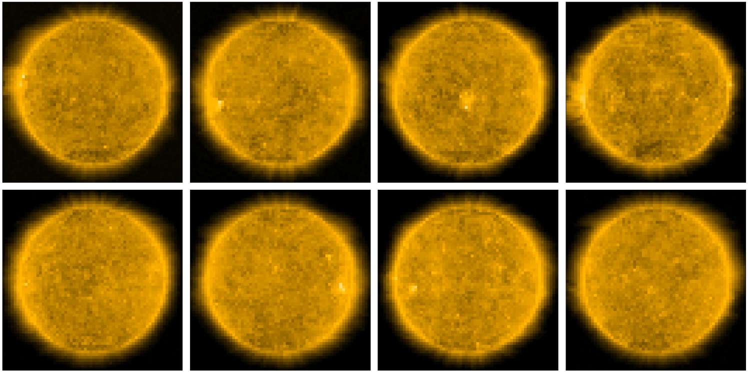
Metric Results

Metric	ceVAE	Discrete	Continous	cevae_emb
CE GEN $\downarrow$	7.948 ± 0.914	$\mathbf{0.130 \pm 0.036}$	1.503 ± 0.147	0.207 ± 0.036
CD GEN $\downarrow$	2.206 ± 0.009	0.921 ± 0.004	$\mathbf{0.934 \pm 0.002}$	0.838 ± 0.005
CS GEN $\downarrow$	3.239 ± 0.009	1.211 ± 0.004	$\mathbf{1.098 \pm 0.002}$	1.480 ± 0.005
FID CLIP $\downarrow$	5.05	0.122	**0.057**	0.39
F1 score $\uparrow$		**0.7**	0.34	0.6
Precision $\uparrow$		**0.73**	0.35	0.6
Recall $\uparrow$		**0.74**	0.37	0.7

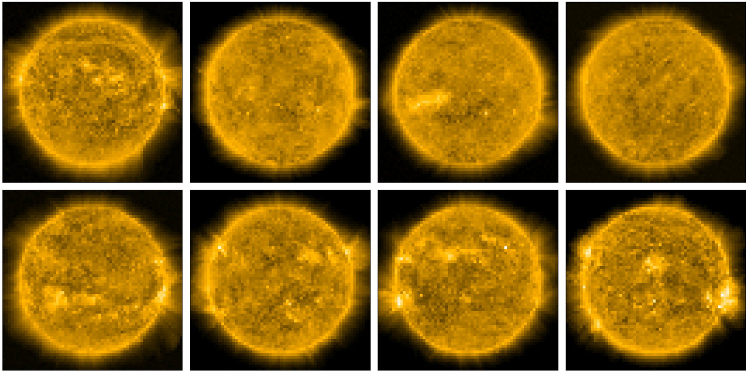
The benchmark values for the cluster metrics are:

- Cluster Error (CE): 0.002,
- Cluster Distance (CD): 1.001,
- Cluster Standard Deviation (CS): 0.998.

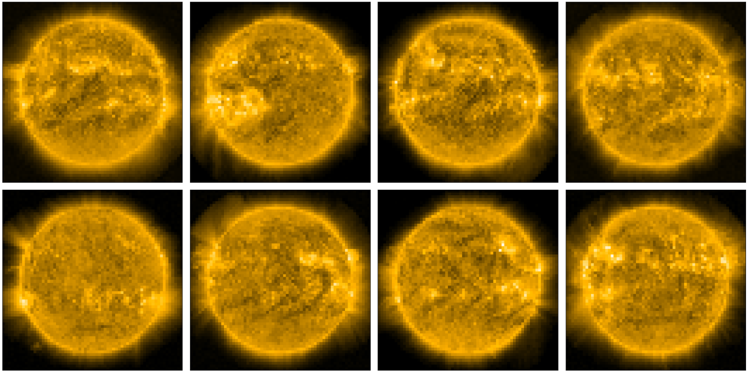
A



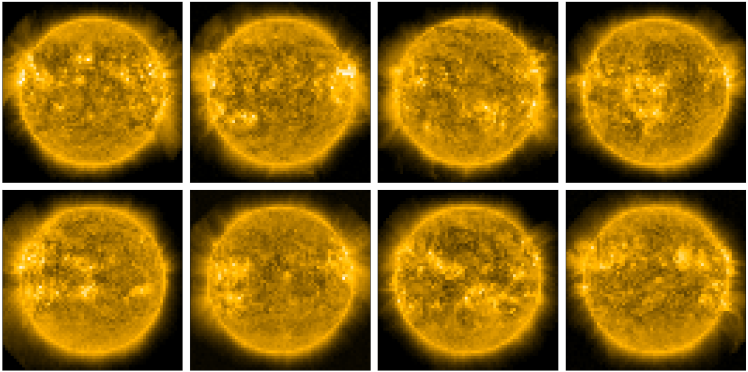
B



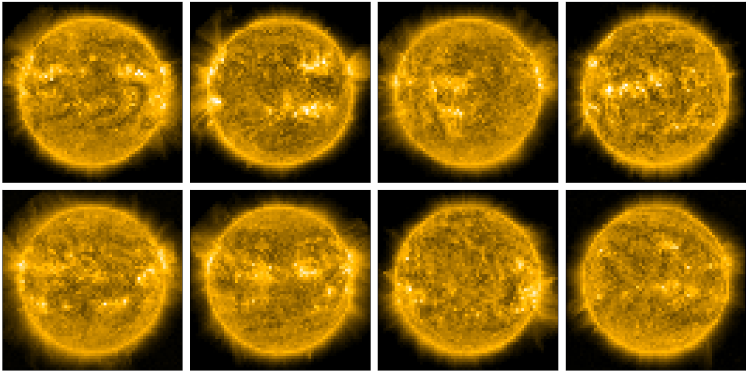
C



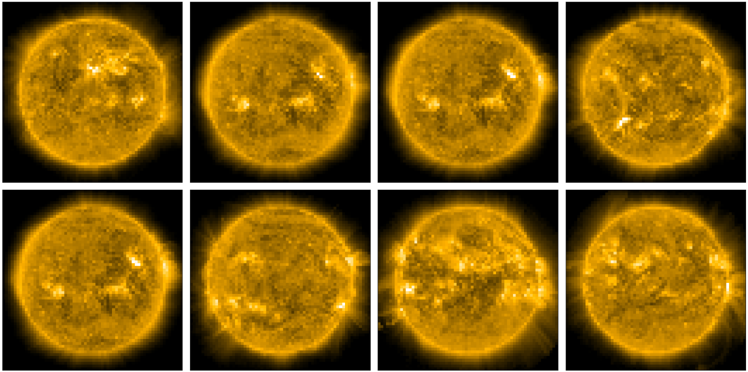
M



X



True X images



Application

Can we use these generated images to train a classifier and increase the accuracy?

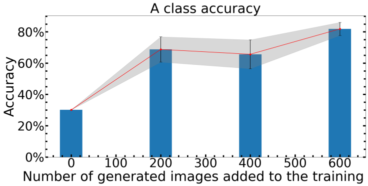
Application

What do we do?

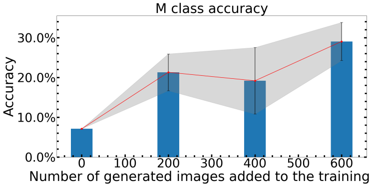
We train a DeIT (Touvron et al 2021) model for supervised classification with and without the addition of generated data to see if this helps the increase of the classification accuracy.

- We train a DeIT (Touvron et al 2021) model for supervised classification without any fine tuning because we want to test the impact of the added images only,
- This is not yet flare prediction.

Accuracy of lower represented classes



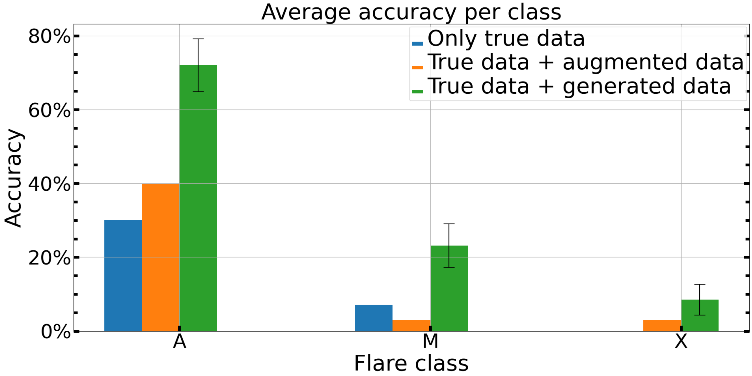
A class



M class

X class

Comparison

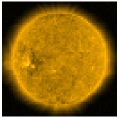

Average accuracy per class
Only true data
True data + augmented data
True data + generated data
Accuracy
80%
60%
40%
20%
0%
A
M
X
Flare class

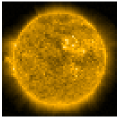
Conclusion

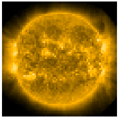
- The best model to generate SDO/AIA images is the model guided with the discrete GOES labels
- It is possible to control the level of activity on top of the sun thanks to the labelling system that we adopted,
- It is possible to apply the generated images to manage the unbalanced dataset in a classifier and increase the accuracy per class,
- As future work, we would like to test it on other deep learning tasks (e.g., obtain the magnetograms of the generated images and increase the image resolution and build a solar flare predictor on full disk images),
- The paper is in peer review.

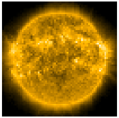
Time for Questions!

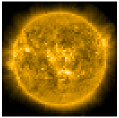
Generated Images

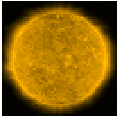

(a) A class

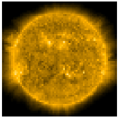

(b) B class

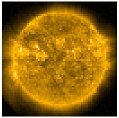

(c) C class

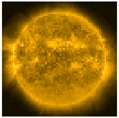

(d) M class

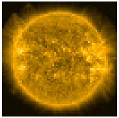

(e) X class


(f) A class


(g) B class


(h) C class


(i) M class


(j) X class

What are the DDPMs?

Diffusion Probabilistic models are very popular nowadays and we can summarize their usage into the following bullet points:

- Forward process or noising process (Ho et al., 2020):

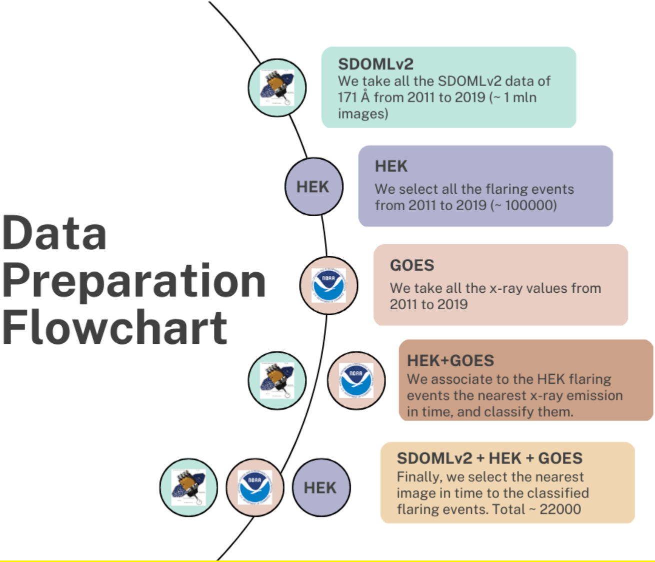
$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}),\ q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t\mathcal{I}) \tag{1}$$

- Reverse process or denoising (Ho et al., 2020):

$$p_\theta(x_{t-1}|x_t) := \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \tag{2}$$

- Classifier Free Guidance (Ho Salimans, 2022)

$$\tilde{\epsilon_\theta}(z, c) = \epsilon_\theta(z, c) + w \cdot (\epsilon_\theta(z, c) - \epsilon_\theta(z)) \tag{3}$$

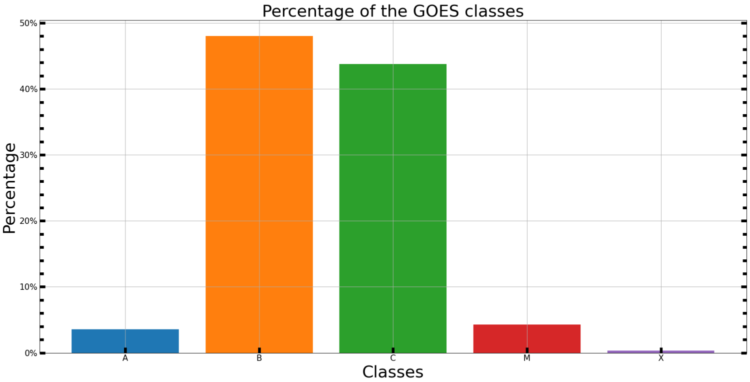

Data Preparation Flowchart
SDOMLv2
We take all the SDOMLv2 data of 171 Å from 2011 to 2019 (~ 1 mln images)
HEK
HEK
We select all the flaring events from 2011 to 2019 (~ 100000)
GOES
We take all the x-ray values from 2011 to 2019
HEK+GOES
We associate to the HEK flaring events the nearest x-ray emission in time, and classify them.
HEK
SDOMLv2 + HEK + GOES
Finally, we select the nearest image in time to the classified flaring events. Total ~ 22000

Different types of Solar Flare

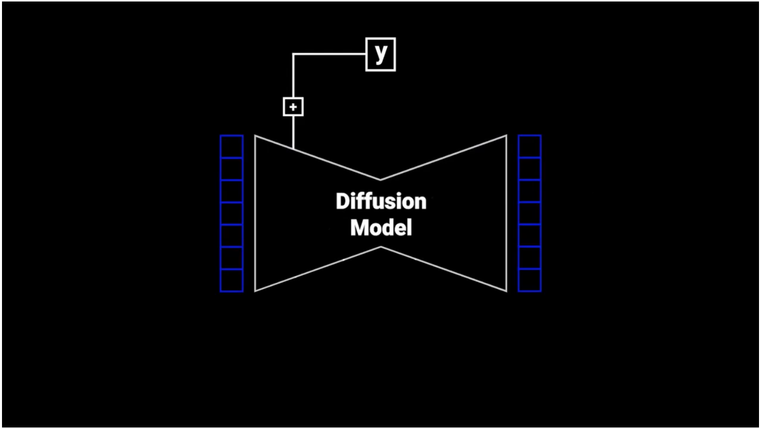
The modern classification system for solar flares uses the letters A, B, C, M, or X, according to the peak flux in watts per square metre (W/m2) of soft X-rays:

- A: $< 10^{-7}$
- B: 10^{-7} –10^{-6}
- C: 10^{-6} –10^{-5}
- M: 10^{-5} –10^{-4}
- X: $> 10^{-4}$

Distribution of the images per GOES class



Classifier Free Guidance (CFG)



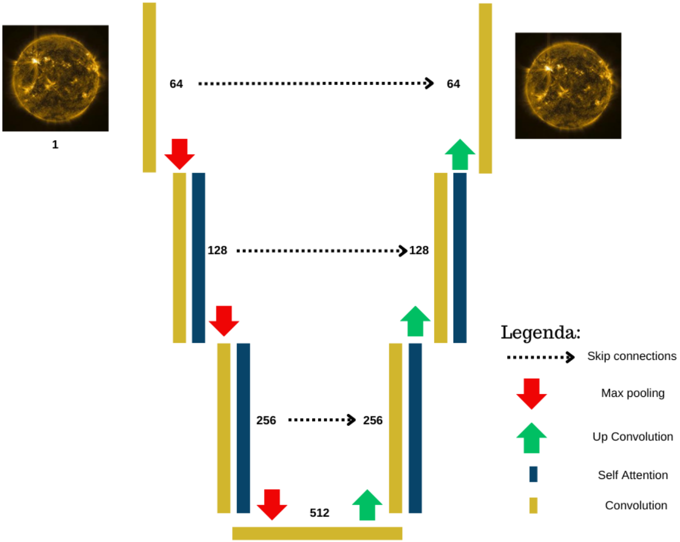
Credits: Outlier

Classifier Free Guidance (CFG)

- Guide the diffusion model with labels in order to be able to produce an image with a determined label,
- The idea is that you train it both unconditioned and conditioned and then interpolate between the two giving a weight to the conditioning in a way that you can direct the production towards that particular label space,
- The most important hyper-parameters are the:
  1. $\rho_c =$ probability of training with labels,
  2. $w =$ the CFG scale, the weight for the interpolation.

$$\tilde{(}\epsilon_\theta(z, c)) = \epsilon_\theta(z, c) + w \cdot (\epsilon_\theta(z, c) - \epsilon_\theta(z)) \tag{4}$$

Unet



1
64
64
128
128
256
256
512
Legenda:
Skip connections
Max pooling
Up Convolution
Self Attention
Convolution

Cluster Metrics

- Cluster Error:

1. $$\epsilon = \frac{1}{K}\sum_{c=1} K \frac{(\hat{n}_c - n_c)^2}{n_c^2} \tag{5}$$

   We count the number of samples $\hat{n}_c$ in each of K clusters and compute the difference to the target n_c. This metric measures whether the interesting regions in feature space, i. e. the clusters, are populated with the same number of samples as in the target distribution. A value of 0 indicates a perfect match. Larger values indicate deviation from the target, i. e. over- and underproduction of some type.

Cluster Metrics

- Cluster Distance:

  1.
$$D = \frac{1}{d}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{d}_i^2} \tag{6}$$

  We compute the distance $\hat{d}_i$ to the corresponding cluster center for each of N samples. Then, D is the normalized root-mean-square (RMS) of these distances. This metric measures whether the samples populate the correct regions in feature space with sufficient diversity. Values larger than 1 indicate that the sample contains images outside the target distribution.

- Cluster Std:

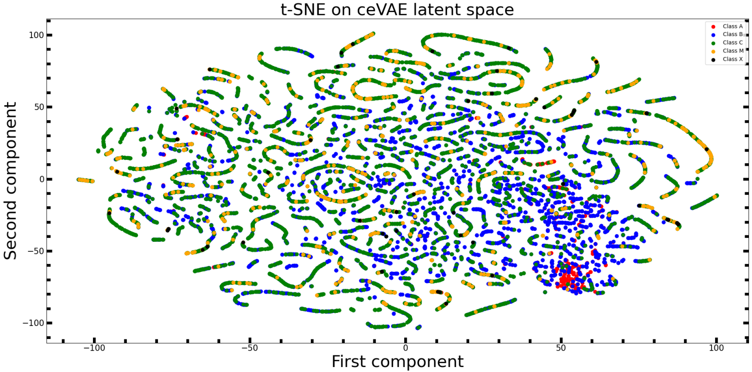
1.
$$S = \frac{1}{S_{target}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{d}_i - d \cdot D)^2} \quad (7)$$

For the distances to the cluster center $\hat{d}_i$, we further compute S as standard deviation from D.

FID

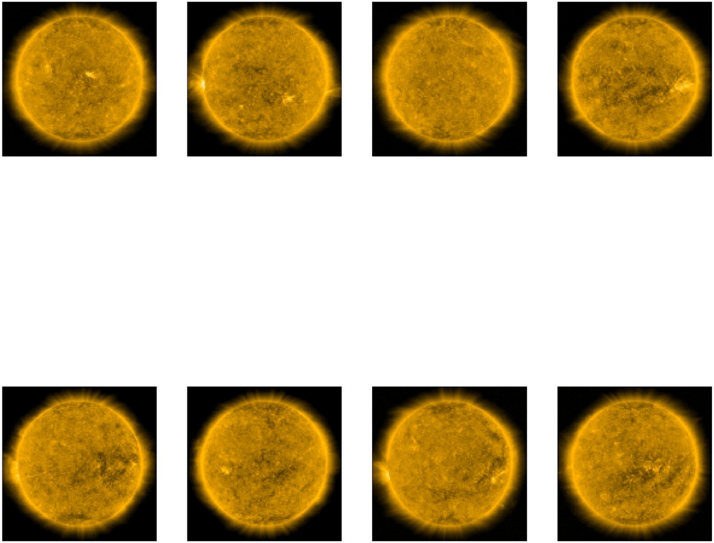
- FID:
  1. $FID(x, g) = ||\mu_x - \mu_g||^2 + \mathrm{Tr}(\Sigma_x + \Sigma_g - 2(\Sigma_x \Sigma_g)^{(1/2)})$,
- F1 score:
  1. $F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
- Precision and Recall:
  1. P $= \frac{TP}{TP+FP}$
  2. R $= \frac{TP}{TP+FN}$

t-SNE ceVAE

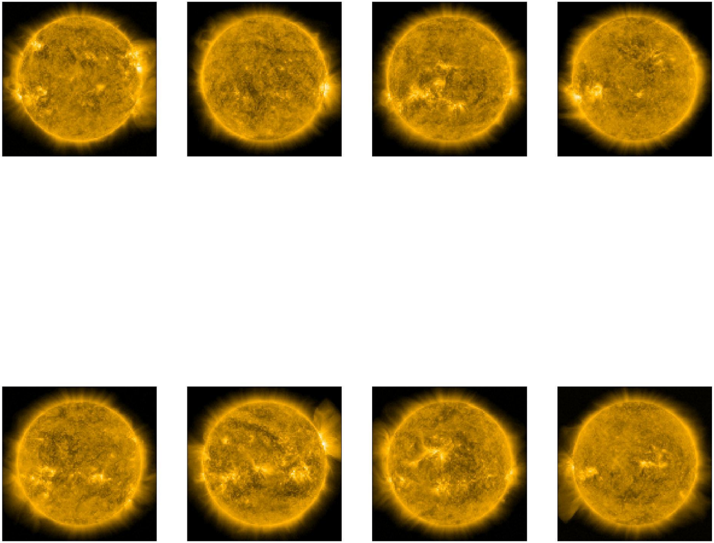


t-SNE on ceVAE latent space
Class A
Class B
Class C
Class M
Class X
100
50
0
−50
−100
−100
−50
0
50
100
Second component
First component

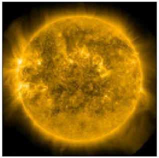
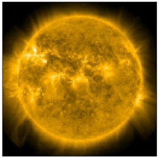
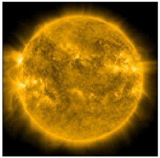
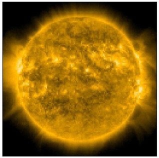
Low level activity 128x128

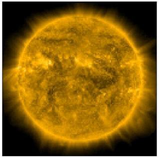
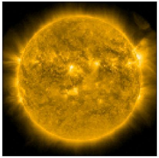
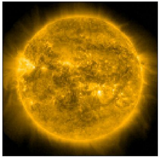
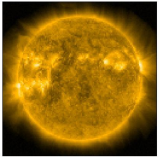


Medium level activity 128x128

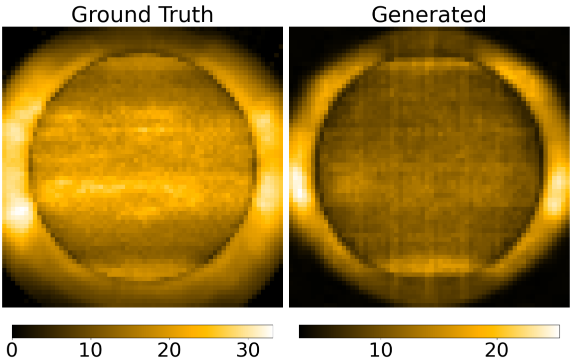


High level activity 128x128

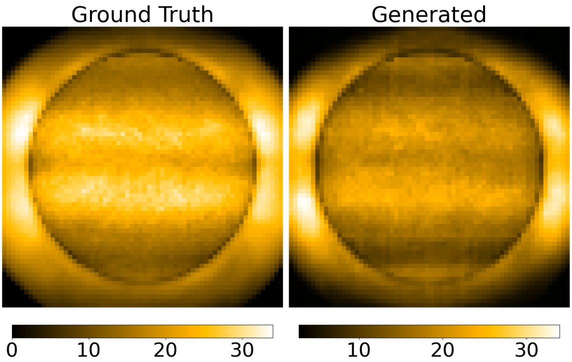





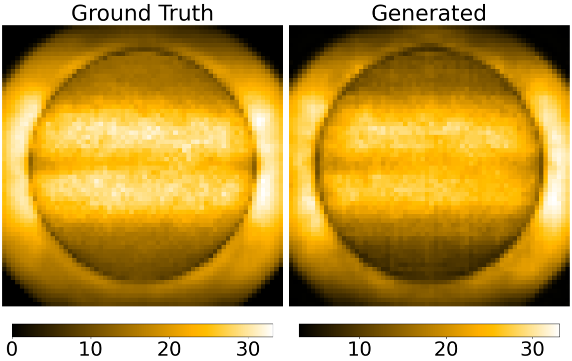
Std A



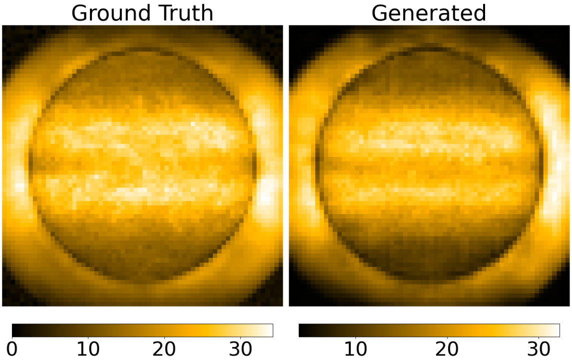
Std B



Std C



Std M



Std X