

Development of Unsupervised-Learning Based Glitch Classification System

Gen Ueshima¹ (¹Nagaoka University of Technology)

Keiko Kokeyama², Chihiro Kozakai³, Shoichi Oshino², Yutaka Shikano⁴, Hirotaka Takahashi⁵, Takashi Uchiyama², Tatsuki Washimi³, Takahiro Yamamoto², Takaaki Yokozawa²
(²ICRR, ³NAOJ, ⁴Keio U, ⁵TCU)

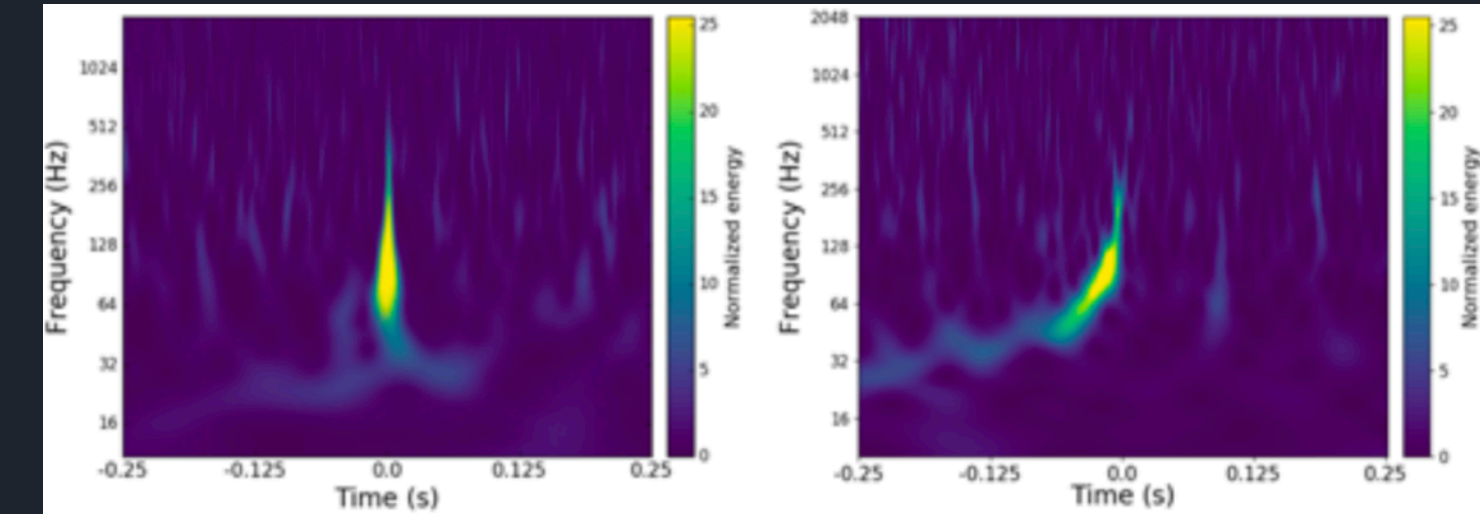
KAGRA International Workshop (KIW)

Background

- Gravitational wave telescopes have a large variety of glitch sources with corresponding time-frequency characteristics.

- It is important to **classify** the time-frequency characteristics around glitches.

- **We can get statistical information for each source of glitch to take some measures.**



Different types of glitch, Blip (left) and Chirp (left)

- Applying **Supervised Learning** based on deep learning [1] to glitch have some aspects of effective, or issues (#1, #2)

#1 Issues in general (glitch classification)

- We don't know the true number of glitch sources.
- Glitch has different frequencies (occur) depending on its source (Imbalanced data).

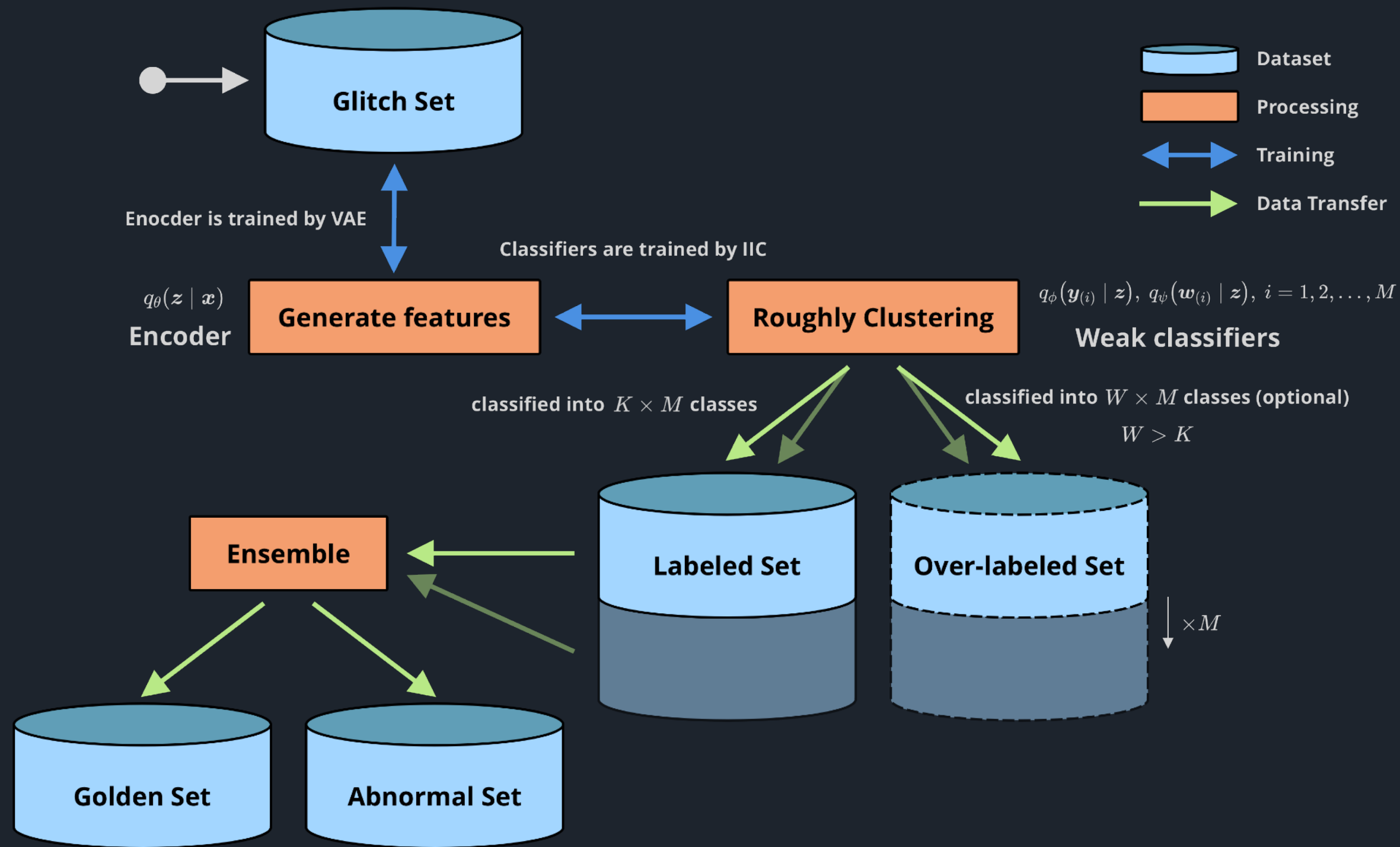
#2 Issues depending on Supervised Learning

- Highly dependent on skill of human to labeling. (requires a lot of work for labeling)
- Subclasses or abnormalities may be included.

We need some human-independent system for classification overcoming #1, #2

Overview of our system

- We develop a system to classify glitches based on **Unsupervised Learning**.
- We tested the system by applying to *Gravity Spy Dataset* [1].



Architecture of our system

Classification strategy

1. Generate features from glitch set using **Variational Autoencoder (VAE)** [2]
2. Roughly classify features into arbitrary number of classes (K, W) by multiple weak classifiers ($\times M$) using **Invariant Information Clustering (IIC)** [3]
 - **Labeled Set** : Using K classes, M patterns
 - **Over-labeled Set (optional)** : Using $W (> K)$ classes, M patterns
3. Ensemble (Consensus) multiple clustering results (preliminary)
 - **Golden Set** : Reliable glitches with reliable labels
 - **Abnormal Set** : Unreliable glitches

[1] M. Zevin, et al., "Gravity Spy: Integrating Advanced LIGO Detector Characterization, Machine Learning, and Citizen Science", Classical and Quantum Gravity, Vol.34, 064003 (2017).

[2] D. Kingma, et.al, "Auto-Encoding Variational Bayes", arXiv:1312.6114v10 (2014).

[3] Xu Ji, et.al, "Invariant Information Clustering for Unsupervised Image Classification and Segmentation", arXiv:1807.06653v4 (2019).

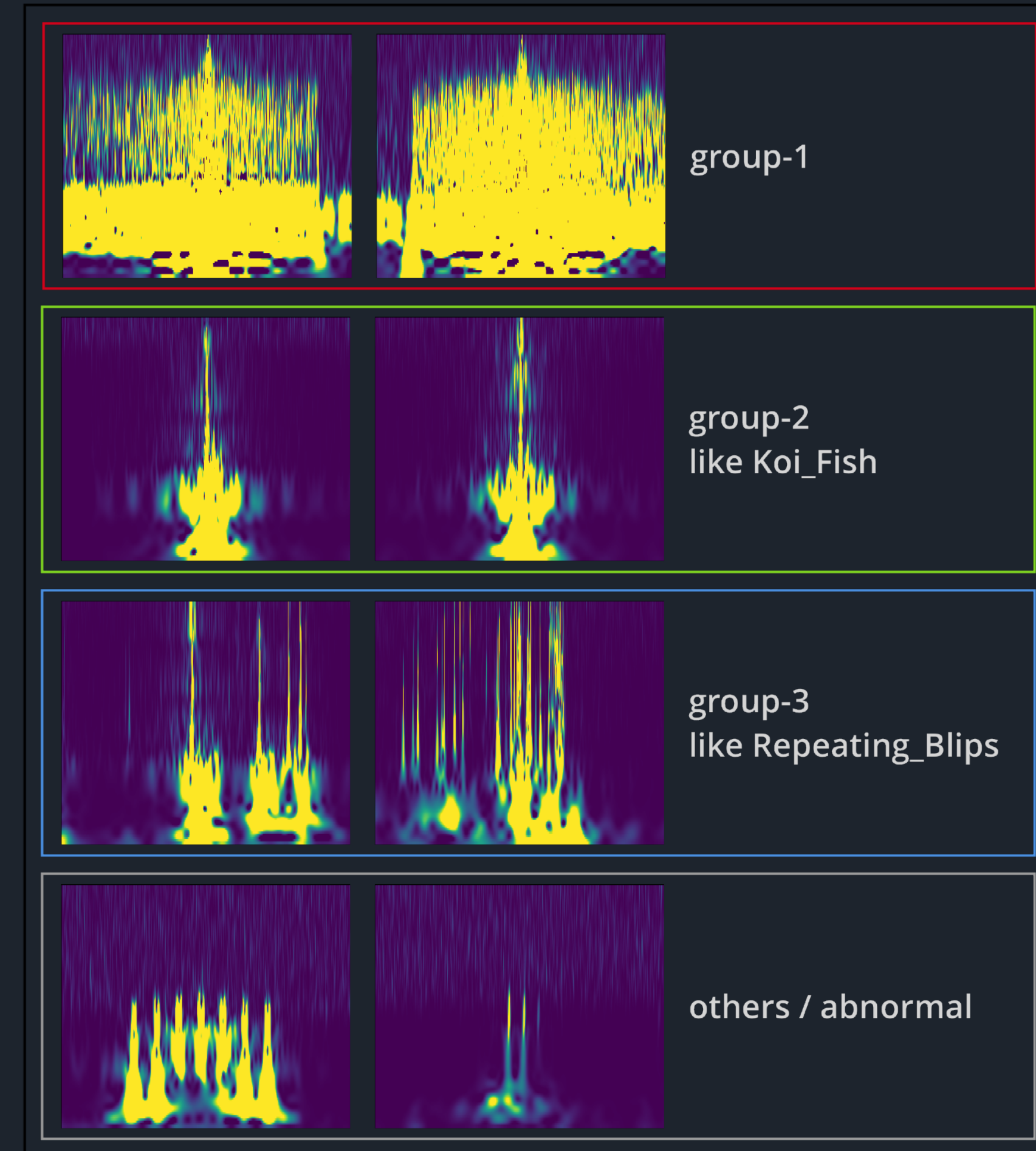
Consideration from *Gravity Spy Dataset*

- *Gravity Spy Dataset* have 22 types of glitches labeled by unique pipeline (human with CNN) [1].

Consideration

- Abnormals may be included
- Certain classes may be divided into multiple subclasses
- Some classes have too much diversity

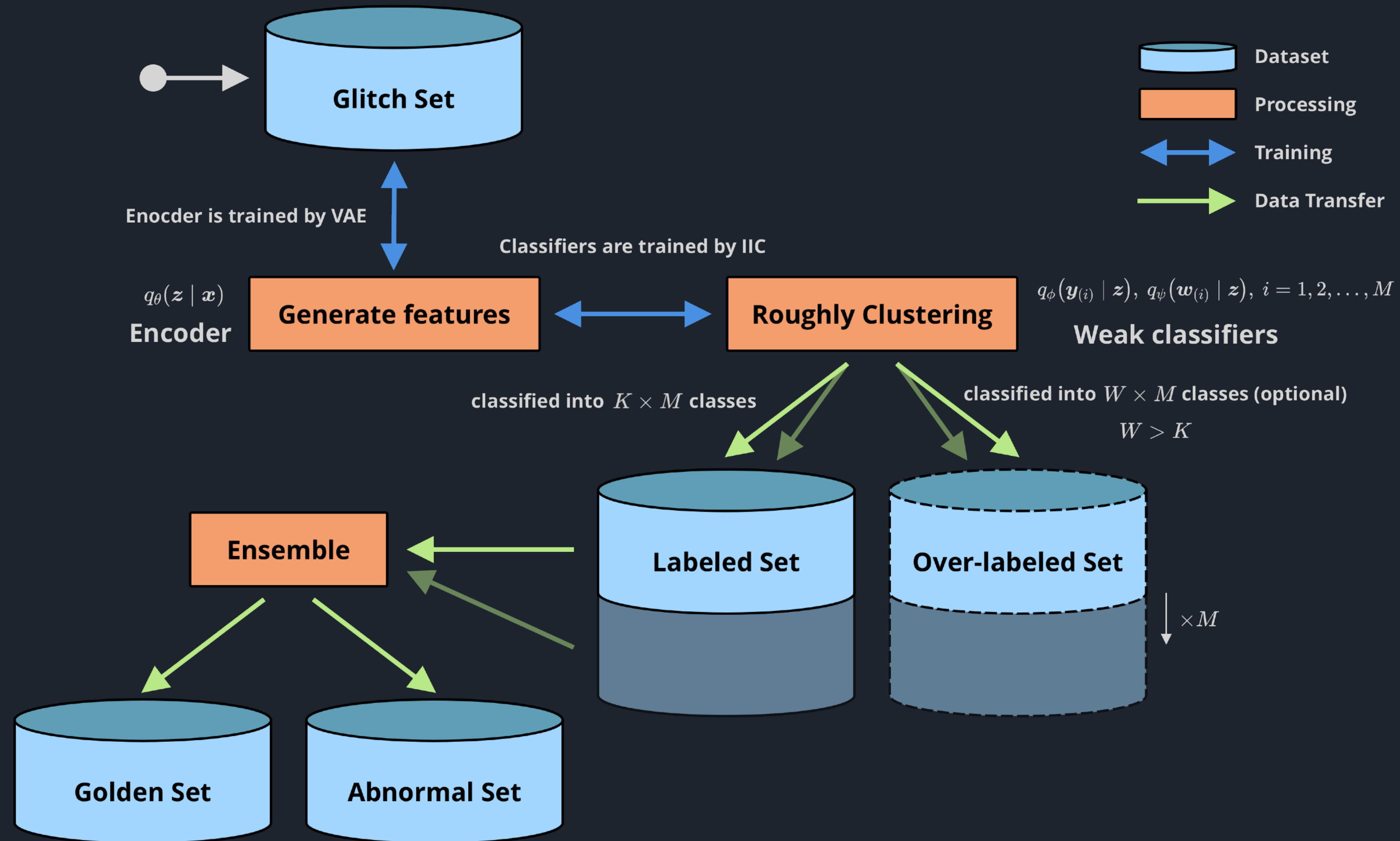
We designed the system to overcome these issues !



Glitches in Extremely_Loud class can divide into subclasses

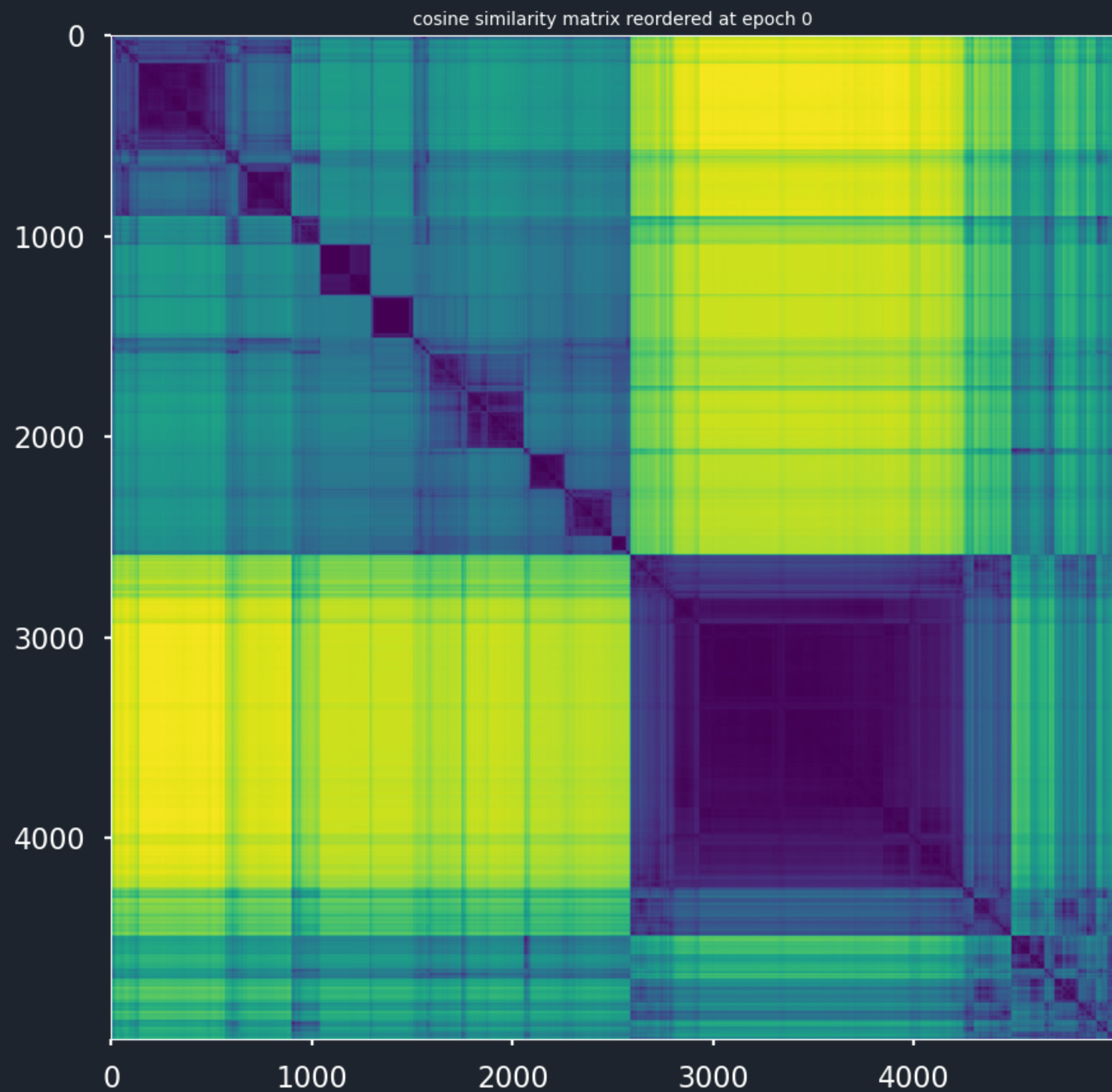
Overview of our system

- We develop a system to classify glitches based on **Unsupervised Learning**.

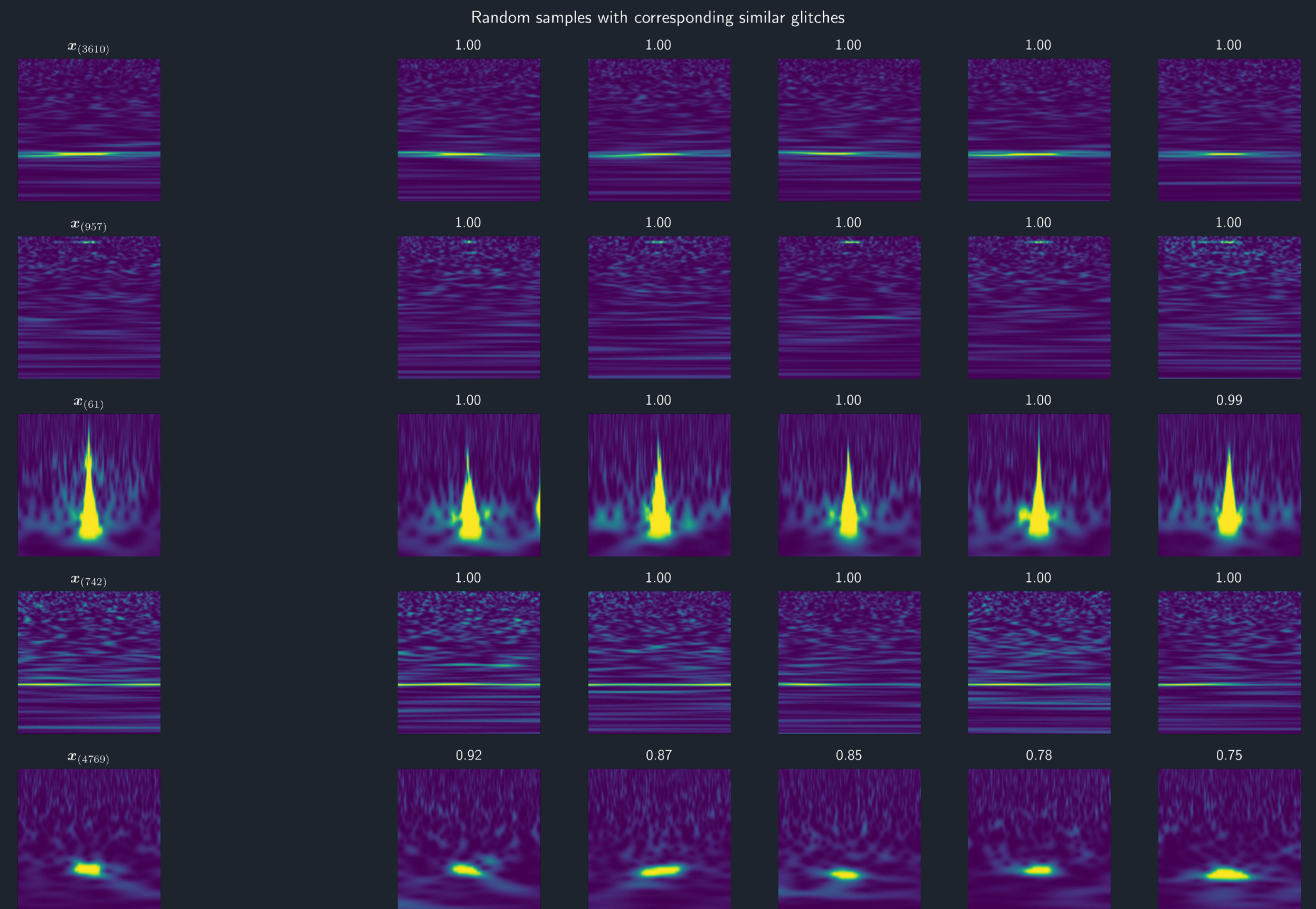


Architecture of our system

Overview of results



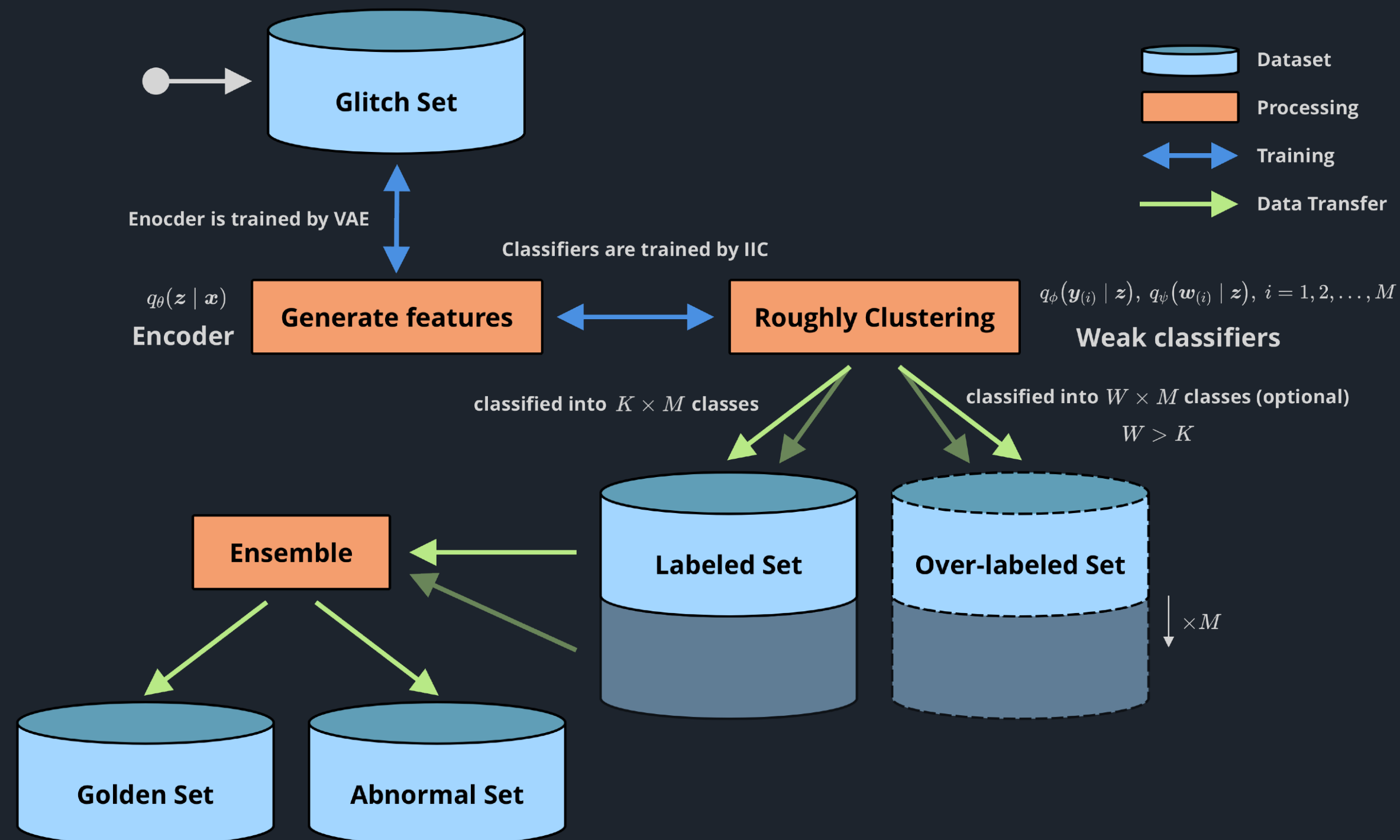
Sorted cosine distance matrix of hyper graph (test data)



Random samples (left) with the 5 most similar samples (right)

Overview of our system

- We develop a system (Fig. 1) to classify glitches based on **Unsupervised Learning**.



Architecture of our system

Classification strategy

- Generate features from glitch set using **Variational Autoencoder (VAE)** [2]
- Roughly classify features into arbitrary number of classes (K, W) by multiple weak classifiers ($\times M$) using **Invariant Information Clustering (IIC)** [3]
 - Labeled Set** : Using K classes, M patterns
 - Over-labeled Set (optional)** : Using $W (> K)$ classes, M patterns
- Ensemble (Consensus) multiple clustering results (preliminary)
 - Golden Set** : Reliable glitches and with hyper graph with reliable labels
 - Abnormal Set** : Unreliable glitches

Variational Autoencoder (VAE) [2]

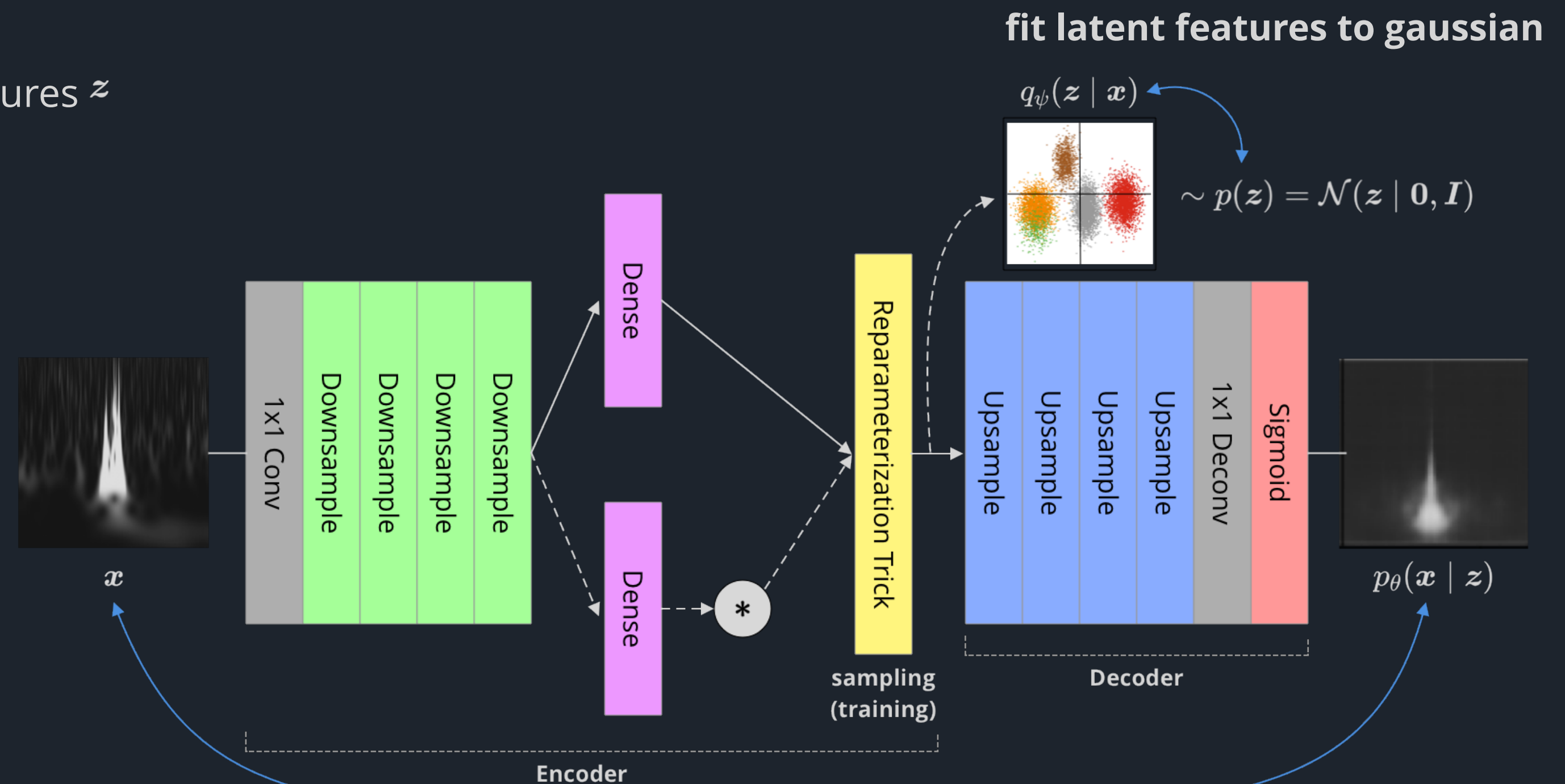
- Unsupervised or semi-supervised generative model by assuming, input data \mathbf{x} is generated by latent features \mathbf{z} (our system use unsupervised model)
- We can generate disentangled features $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ from input data \mathbf{x}

Architecture

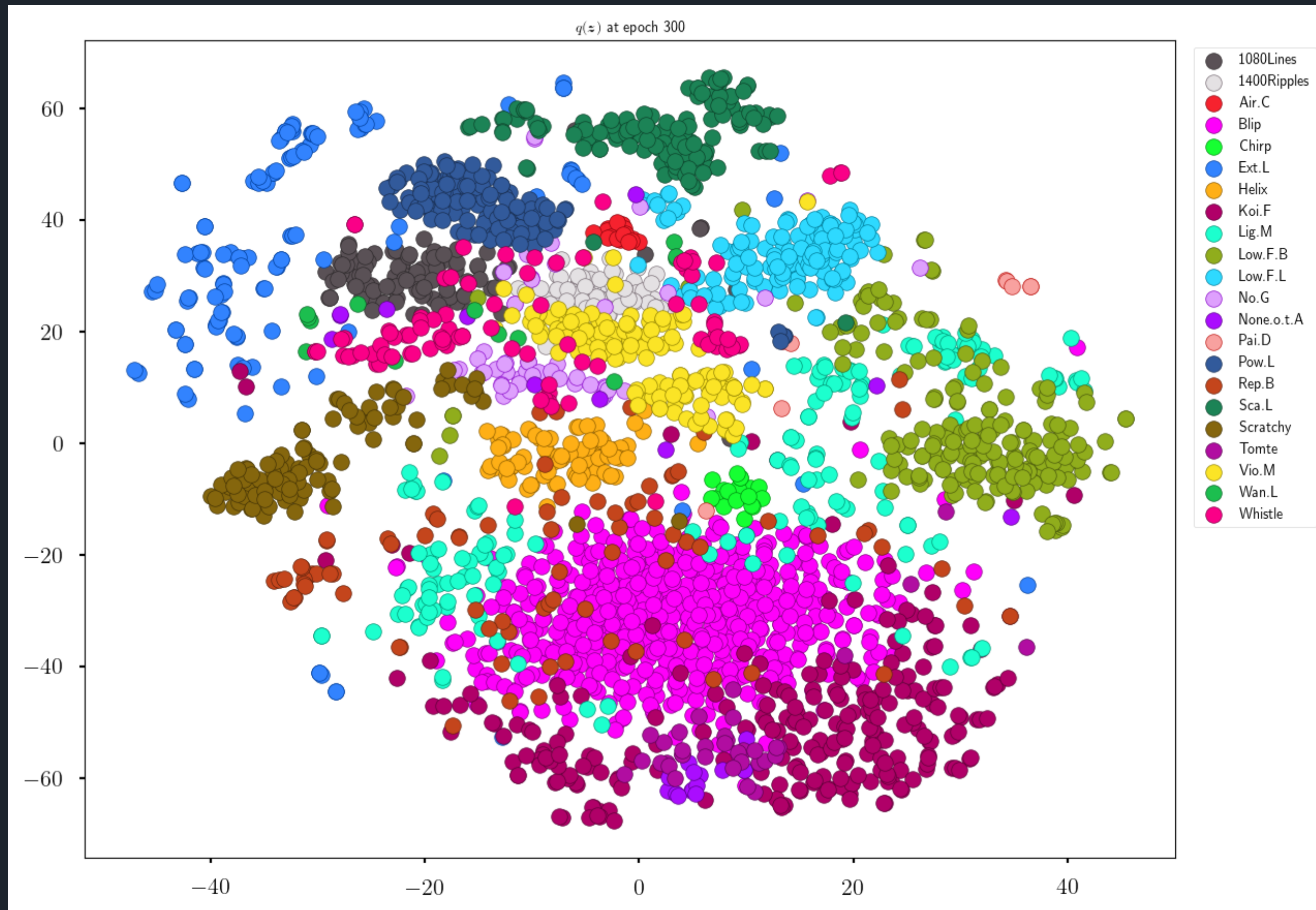
- **Encoder** $q_\psi(\mathbf{z} | \mathbf{x})$ infer latent features \mathbf{z} from input data \mathbf{x}
- **Decoder** $p_\theta(\mathbf{x} | \mathbf{z})$ generate reconstructed image \mathbf{x} from latent features \mathbf{z}

Objective

$$\begin{aligned}
 \text{maximize } \log p_\theta(\mathbf{x}) &= \log \mathbb{E}_{q_\psi(\mathbf{z}|\mathbf{x})} \left[\frac{p(\mathbf{z})p_\theta(\mathbf{x} | \mathbf{z})}{q_\psi(\mathbf{z} | \mathbf{x})} \right] \\
 &\geq \mathbb{E}_{q_\psi(\mathbf{z}|\mathbf{x})} \left[\log \frac{p(\mathbf{z})p_\theta(\mathbf{x} | \mathbf{z})}{q_\psi(\mathbf{z} | \mathbf{x})} \right] \\
 &= \mathbb{E}_{q_\psi(\mathbf{z}|\mathbf{x})} [\log p_\theta(\mathbf{x} | \mathbf{z})] - \mathbb{E}_{q_\psi(\mathbf{z}|\mathbf{x})} \left[\log \frac{q_\psi(\mathbf{z} | \mathbf{x})}{p(\mathbf{z})} \right] \\
 &= \frac{1}{L} \sum_{i=1}^L \log p_\theta(\mathbf{x} | \mathbf{z}_{(i)}) - \mathcal{D}[q_\psi(\mathbf{z} | \mathbf{x}) \| p(\mathbf{z})] = \mathcal{L}
 \end{aligned}$$



Result-1: Latent features generated by VAE



2D latent features by t-SNE with true labels (Gravity Spy)

Consideration

- We identified the presence of classes containing multiple subclasses such as Ext.L, Vio.M.
- The data distribution for each class partly overlaps with the others, but it is generally separated.
- We found that differences in the statistical variance between the classes by visualizing the distribution of the data,

Invariant Information Clustering (IIC) [3]

- Unsupervised classifier by maximize mutual information of classification results from two different types of features z, z'
- We can roughly classify features to arbitrary number of classes K or L and $W > K$, and can get different kinds M of results.

Architecture

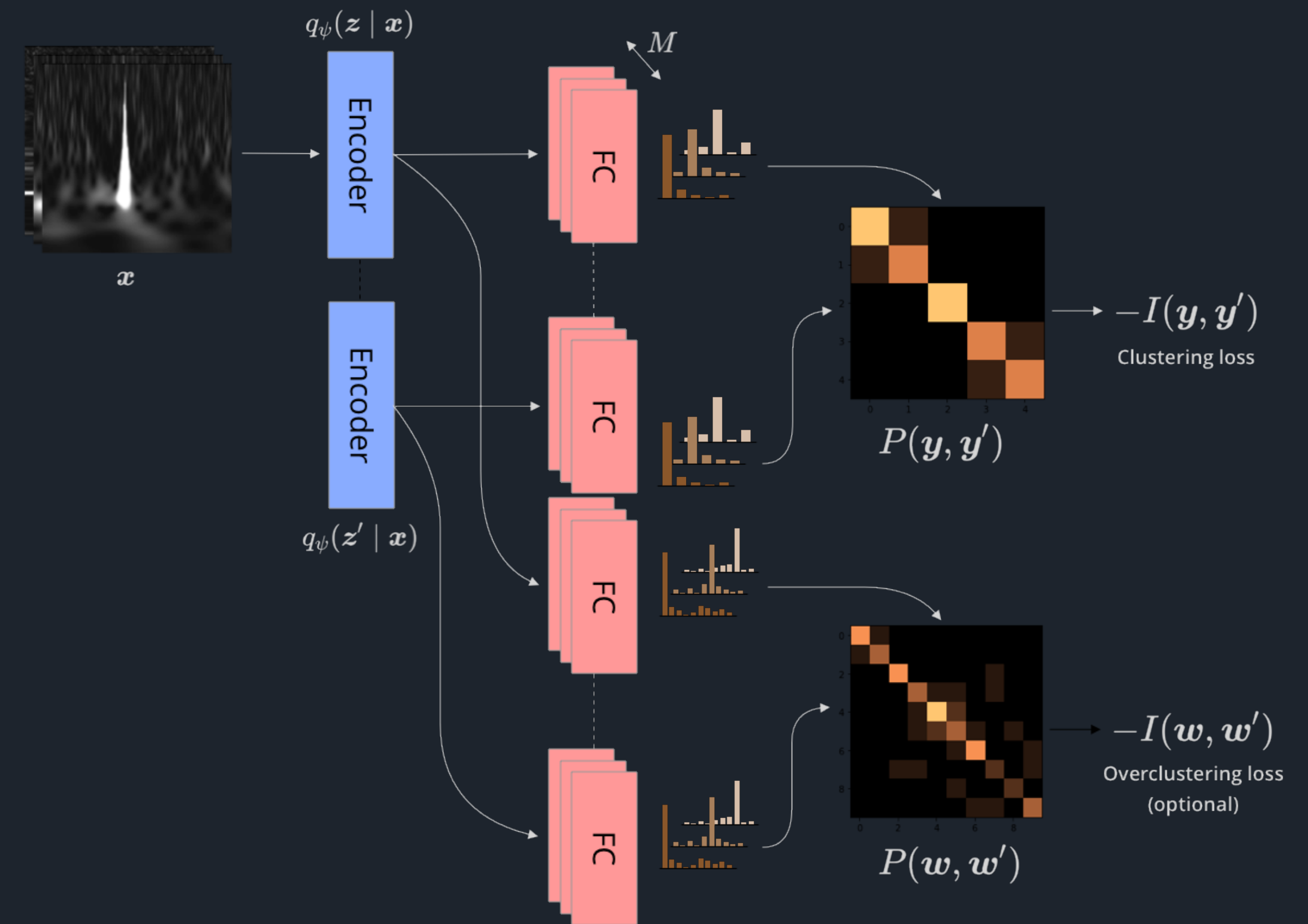
- **Encoder** $q_\psi(z | \mathbf{x})$ infer latent features z from input data \mathbf{x}
- **Multiple weak classifiers** classify z to arbitrary number of classes with different kinds of results.

Objective

$$\max[I(\mathbf{y}, \mathbf{y}') + I(\mathbf{w}, \mathbf{w}')]$$

$$I(\mathbf{y}, \mathbf{y}') = \sum_{k=1}^K \sum_{k'=1}^K P(\mathbf{y} = k, \mathbf{y}' = k') \log \frac{P(\mathbf{y}=k, \mathbf{y}'=k')}{P(\mathbf{y}=k|\mathbf{x})P(\mathbf{y}'=k'|\mathbf{x})}$$

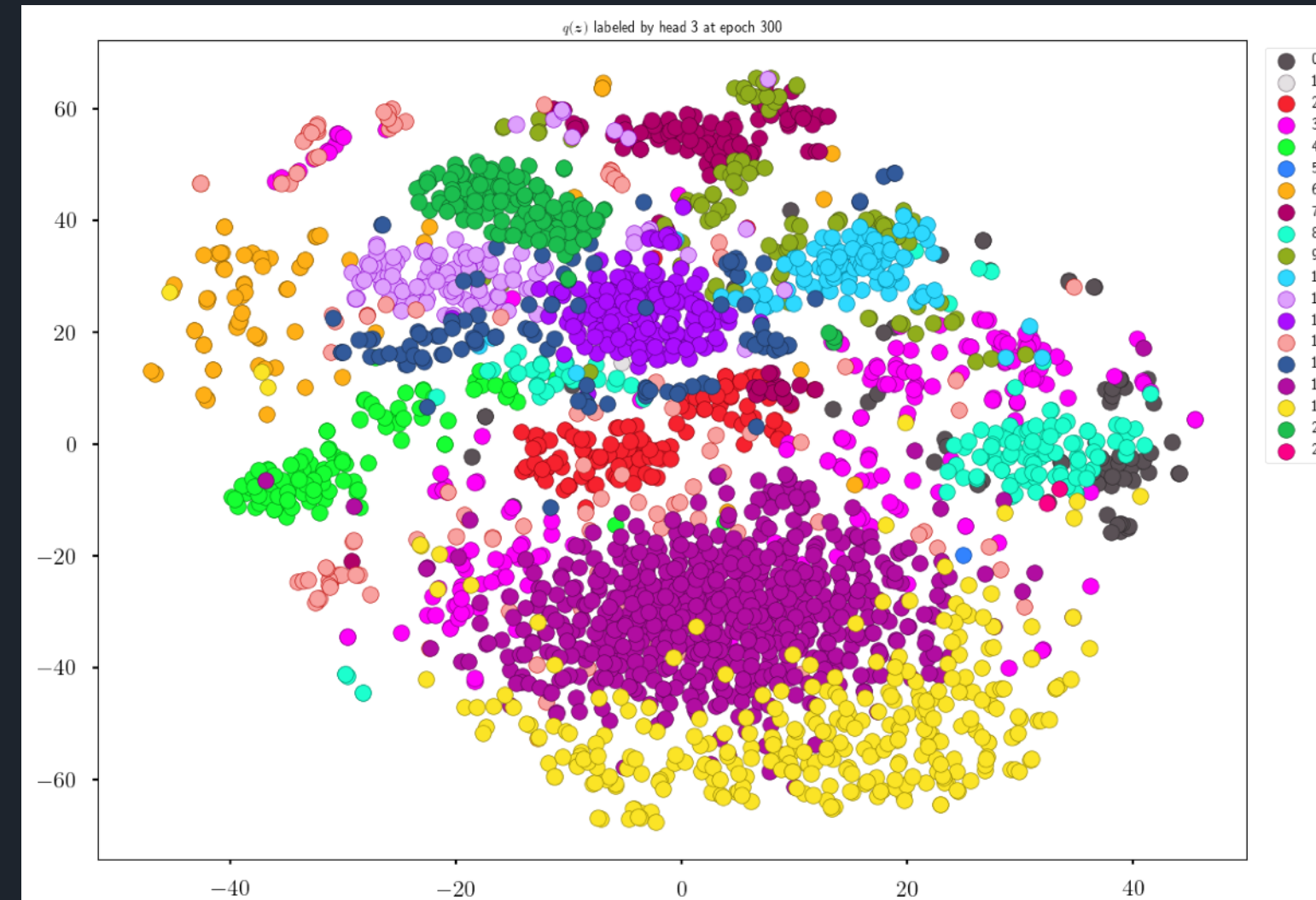
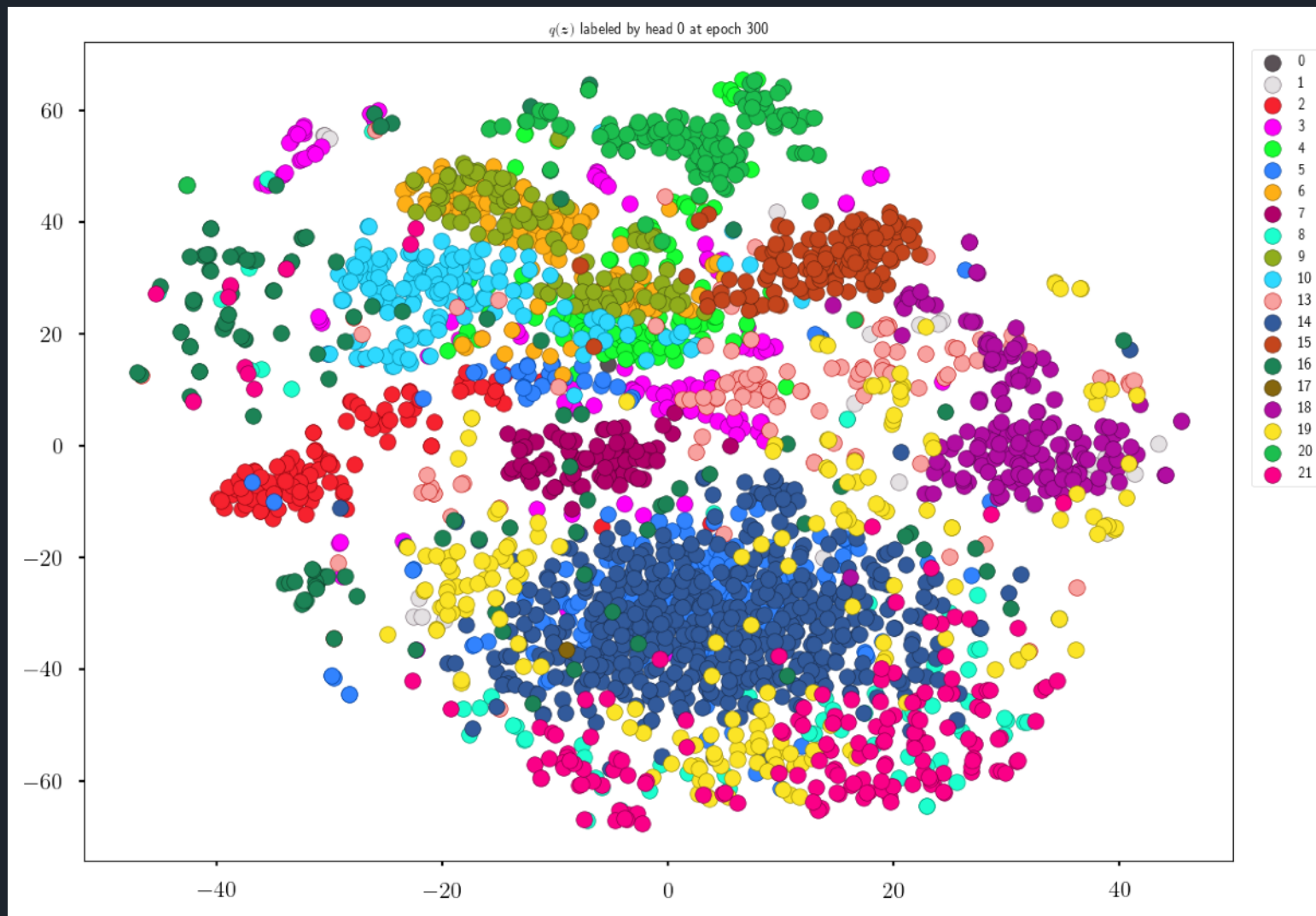
$$I(\mathbf{w}, \mathbf{w}') = \sum_{c=1}^C \sum_{c'=1}^C P(\mathbf{w} = c, \mathbf{w}' = c') \log \frac{P(\mathbf{w}=c, \mathbf{w}'=c')}{P(\mathbf{w}=c|\mathbf{x})P(\mathbf{w}'=c'|\mathbf{x})}$$



Architecture of IIC

[3] Xu Ji, et.al, "Invariant Information Clustering for Unsupervised Image Classification and Segmentation", arXiv:1807.06653v4 (2019).

Result-2: Classify features by IIC



| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|-------------|---|-----|---|----|----|-----|----|----|---|----|----|----|----|-----|----|-----|----|-----|----|-----|----|----|
| 1080Lines | 0 | 0 | 0 | 2 | 7 | 0 | 1 | 0 | 0 | 0 | 86 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1400Ripples | 0 | 0 | 0 | 0 | 4 | 0 | 31 | 0 | 0 | 34 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Air_C | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 |
| Blip | 0 | 0 | 0 | 0 | 0 | 124 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 426 | 0 | 1 | 1 | 9 | 0 | 0 |
| Chirp | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ext.L | 0 | 5 | 0 | 35 | 0 | 5 | 0 | 11 | 0 | 0 | 0 | 0 | 7 | 0 | 59 | 0 | 0 | 0 | 0 | 6 | 8 | 0 |
| Helix | 0 | 0 | 0 | 0 | 0 | 0 | 84 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Koi.F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 52 | 0 | 0 | 0 | 4 | 20 | 0 | 1 | 0 | 0 | 19 | 0 | 153 | 0 | 0 |
| Lig.M | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 44 | 0 | 0 | 4 | 0 | 23 | 93 | 0 | 0 | 0 | 0 | 0 |
| Low.F.B | 0 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 146 | 18 | 0 | 0 | 0 | 0 | 0 |
| Low.F.L | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 121 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| No.G | 0 | 0 | 2 | 7 | 3 | 28 | 10 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| None.o.t.A | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 5 | 0 | 0 | 4 | 0 | 11 | 0 | 3 | 0 | 0 |
| Pai.D | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| Pow.L | 0 | 0 | 0 | 0 | 1 | 3 | 76 | 0 | 0 | 55 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Rep.B | 0 | 0 | 0 | 13 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 18 | 47 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Rep.B | 0 | 0 | 0 | 0 | 16 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 119 | 0 | 0 | 0 | 0 |
| Sca.L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Scratchy | 0 | 104 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Tomte | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 33 | 0 | 1 | 0 | 0 |
| Vio.M | 0 | 0 | 0 | 26 | 62 | 0 | 0 | 0 | 0 | 20 | 0 | 34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vio.M | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Wan.L | 1 | 0 | 1 | 30 | 11 | 0 | 11 | 0 | 0 | 34 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Whistle | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|-------------|----|----|---|-----|---|---|----|----|----|----|----|----|----|----|----|----|----|-----|-----|----|----|----|
| 1080Lines | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1400Ripples | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 70 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Air_C | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 1 | 3 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Blip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 555 | 6 | 0 | 0 |
| Chirp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 0 | 0 |
| Ext.L | 0 | 0 | 0 | 19 | 0 | 0 | 78 | 0 | 5 | 0 | 0 | 0 | 0 | 33 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Helix | 0 | 83 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Koi.F | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 7 | 237 | 0 | 0 | 0 | 0 |
| Lig.M | 8 | 0 | 0 | 140 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 8 | 9 | 0 | 0 | 0 |
| Low.F.B | 57 | 0 | 0 | 25 | 0 | 0 | 0 | 0 | 0 | 95 | 10 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| Low.F.L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 45 | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| No.G | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 29 | 3 | 2 | 2 | 1 | 0 | 13 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| None.o.t.A | 1 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | 0 | 0 | 1 | 15 | 0 | 0 | 0 | 0 |
| Pai.D | 5 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Pow.L | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 134 | 0 | 0 | 0 |
| Rep.B | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 62 | 0 | 0 | 0 | 19 | 3 | 0 | 0 |
| Rep.B | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 95 | 0 | 33 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| Sca.L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Scratchy | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 34 | 0 | 0 | 0 |
| Tomte | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vio.M | 0 | 35 | 0 | 0 | 0 | 0 | 16 | 0 | 0 | 0 | 1 | 82 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vio.M | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 8 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Wan.L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 7 | 0 | 78 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Whistle | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Consideration

- Each classifier has its own domain for good classification.
- Clustering works well even to a certain extent with imbalanced data.
- Classification performance of classes with extremely small numbers of data is poor (e.g. Chirp).

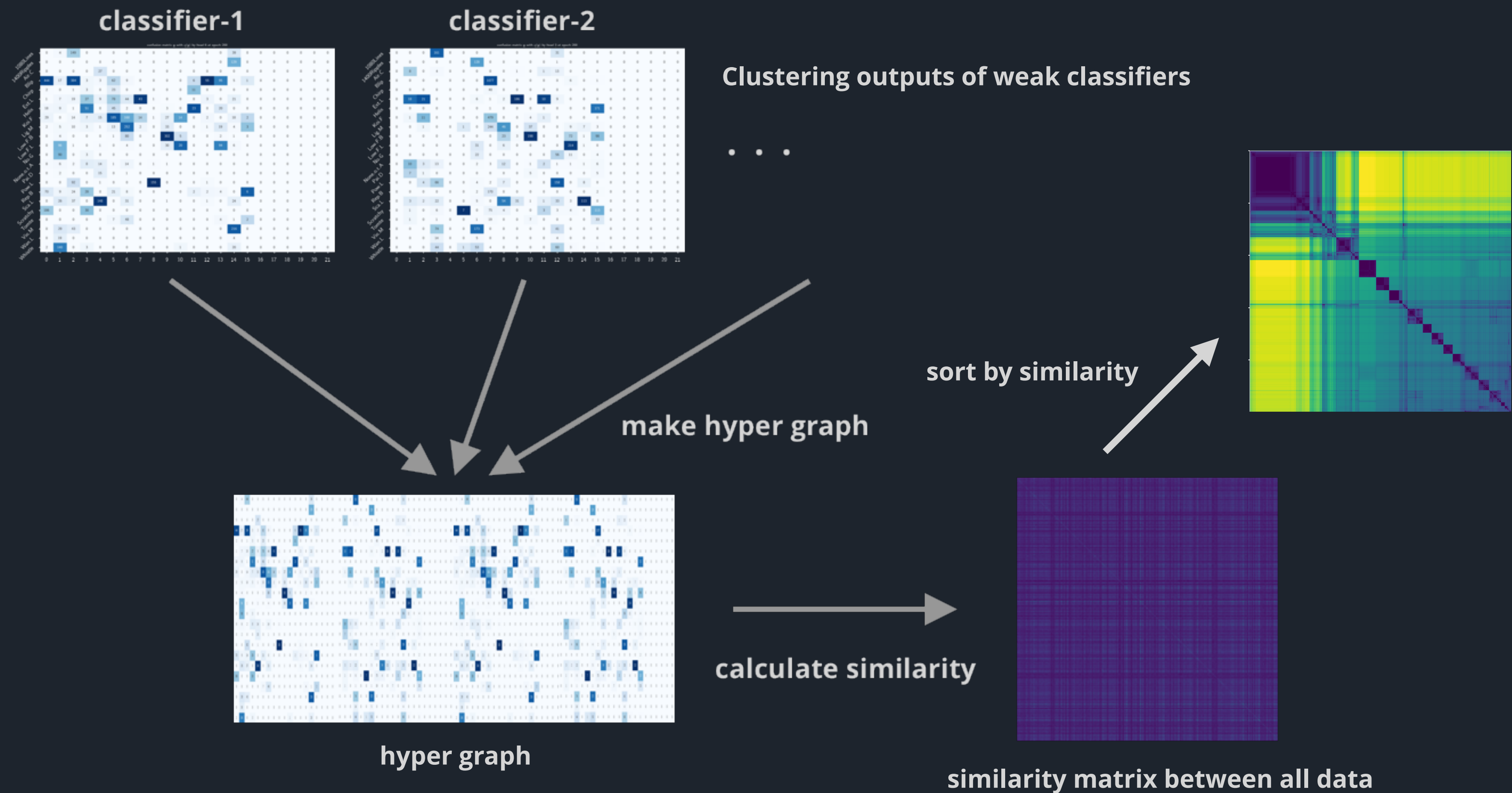
(a) classifier-0

(b) classifier-1

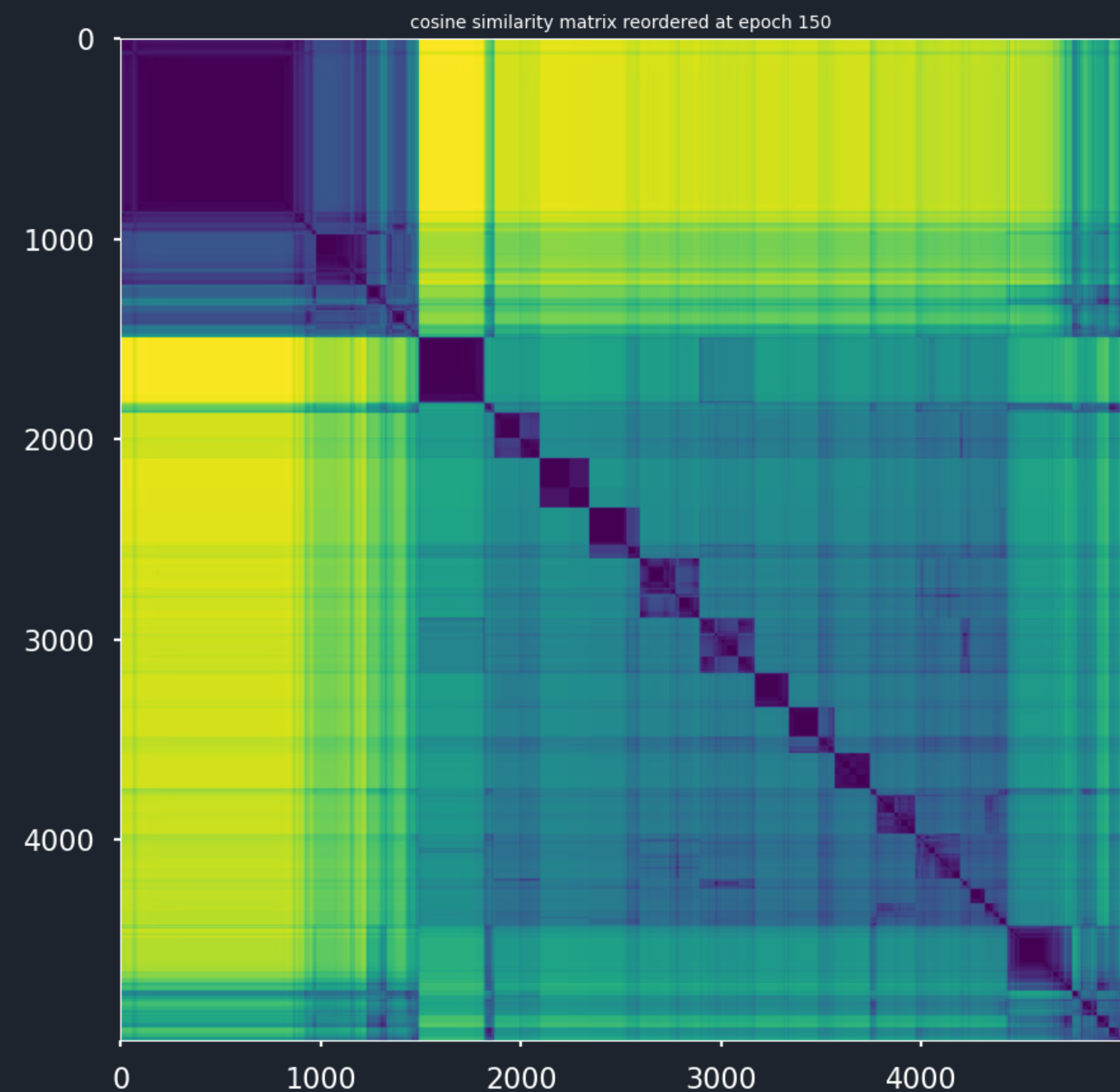
2D latent features with classifiers (upper), confusion matrix (lower)

Ensemble (preliminary)

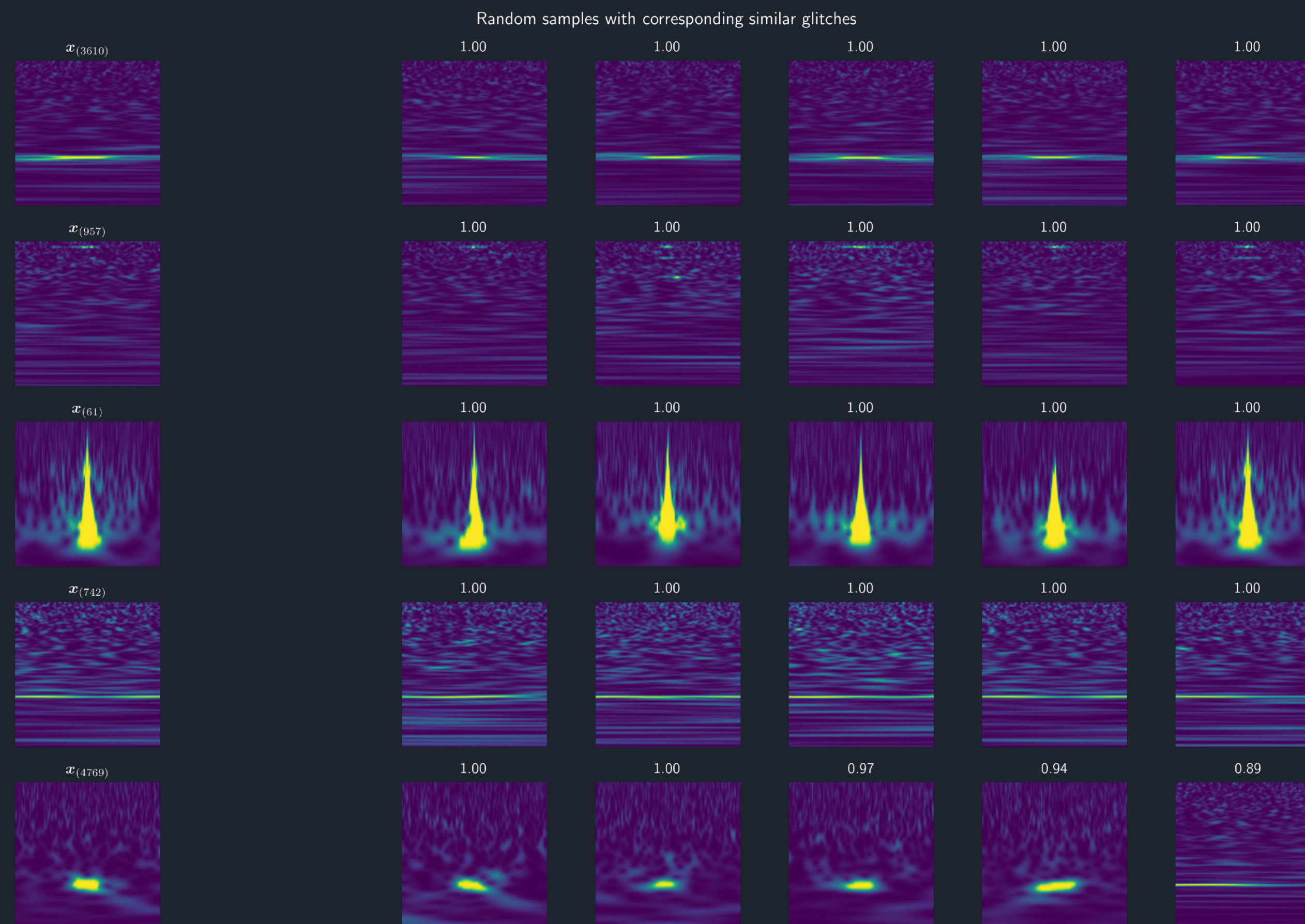
- Concatenate outputs from all weak classifiers to make **hyper graph**
- Calculate similarity matrix between all test data



Result-3: Ensemble outputs of all weak classifiers (preliminary)



Sorted cosine distance matrix of hyper graph (test data)



Random samples (left) with the 5 most similar samples (right)

Consideration & Future works

Consideration of results

- We could get class-dependent features from Gravity Spy Dataset by using VAE.
- We could see the different kinds of results from IIC weak classifiers.
- We could get similar glitches of any sample by ensemble all outputs from IIC weak classifiers.

Future works

- Consider methods to improve ensemble clustering.
- Apply our system to KAGRA data and evaluate its performance.