Glitch classification feature selection of Q-transform parameters

Guillem Fernández-Rodríguez
Master's Student in Mathematics at University of Valencia

VNIVERSITAT DE VALÈNCIA

VIRGO

GWƎ

University of Minho
School of Sciences

Utrecht University

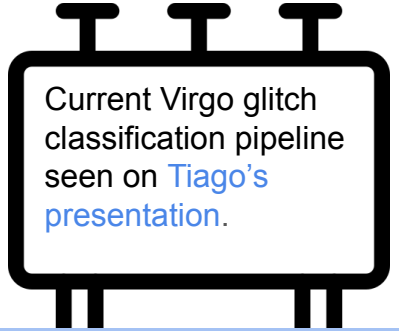
Collaborators: Isabel Cordero-Carrión, Tiago Fernandes, Melissa López

Introduction

- Glitches are transient noise signals in GW detectors.
- The Gravity Spy dataset provides glitch classifications with CNNs that use Q-transform spectrograms. arXiv:1611.04596

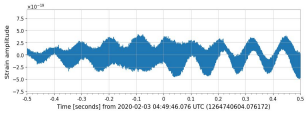
Motivation

- Study how Q-transform parameters and visualization choices influence glitch classification with CNNs.

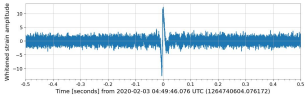


Introduction

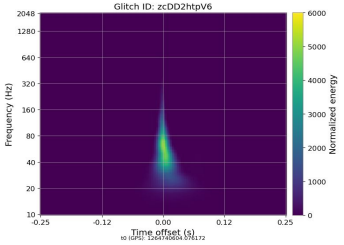
How exactly are glitch spectrograms generated?



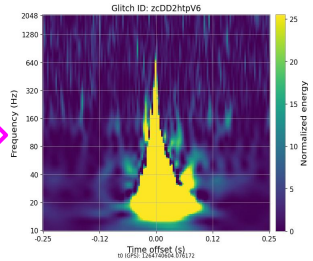
Signal



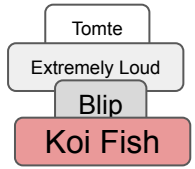
Whitened Signal



Q-transform
(interpolated)



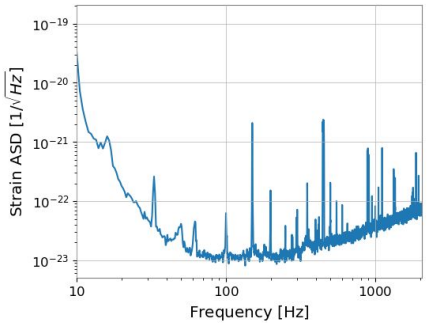
Visualize better

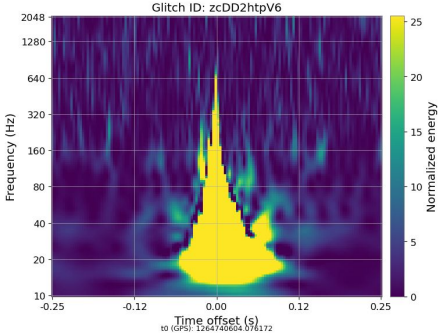


Classification

Signal Whitening

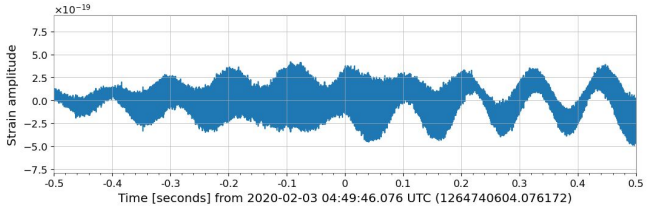
Background ASD



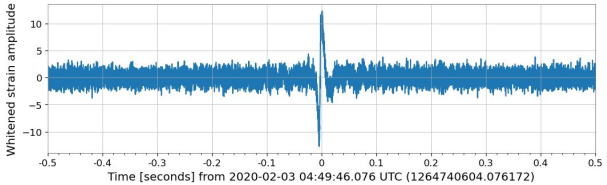


Welch method
or median

Q-transform



Inverse PSD
Whitening

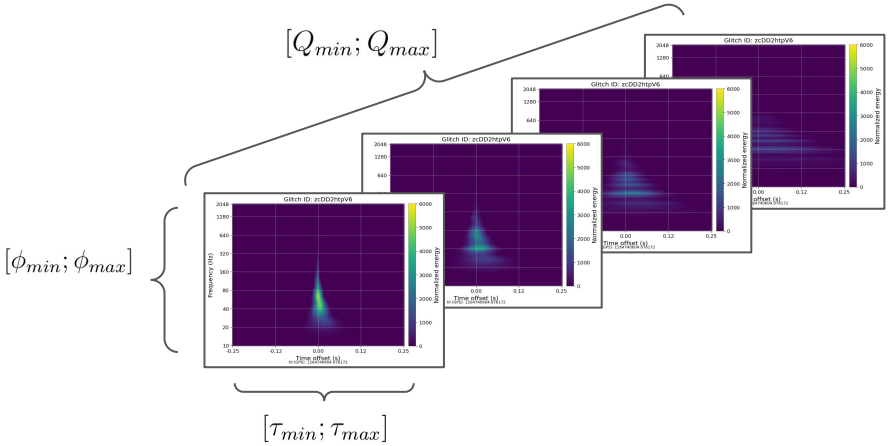


Q-transform

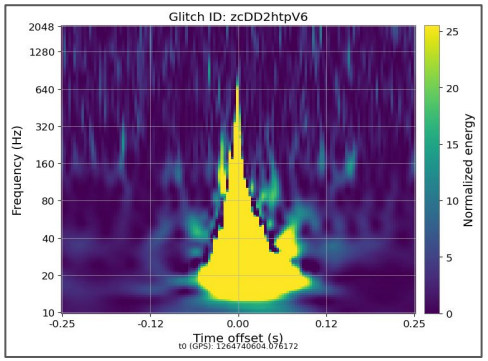
$$X(\tau, \phi, Q) = \int_{-\infty}^{+\infty} x(t) w(t - \tau, \phi, Q) e^{-2i\pi\phi t} dt$$

τ = central time $\quad$ ϕ = central frequency $\quad$ Q = quality factor

Time-frequency plane is tiled



Maximum energy Q-tile (in a relevant range)

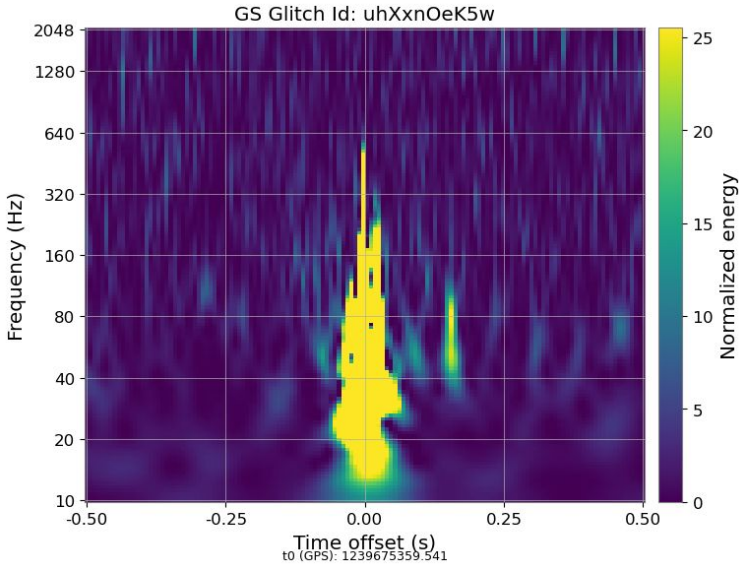


What choices influence glitch visualization in significant ways?

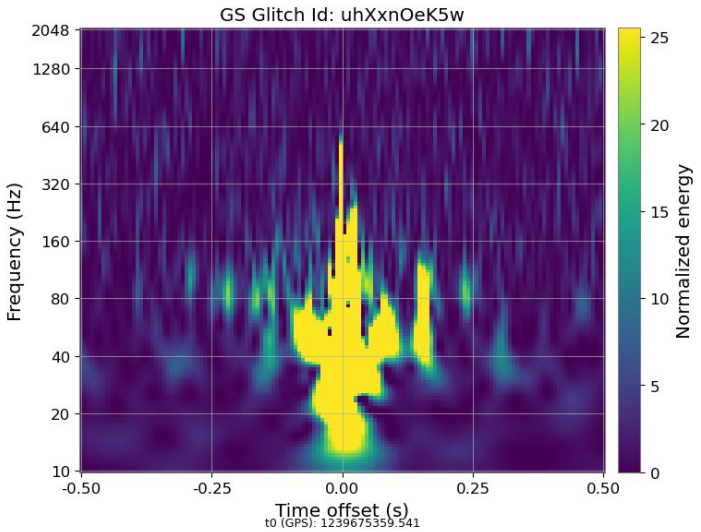
What choices influence glitch visualization in significant ways?

1. Signal **whitening method**:
   - Welch (mean) or median.

Examples: Whitening method



Welch method (mean)

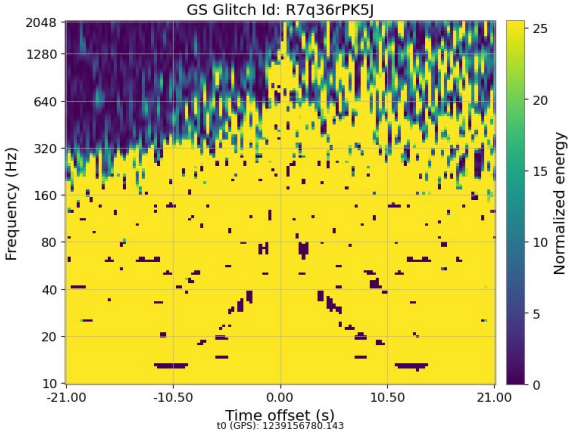


Median

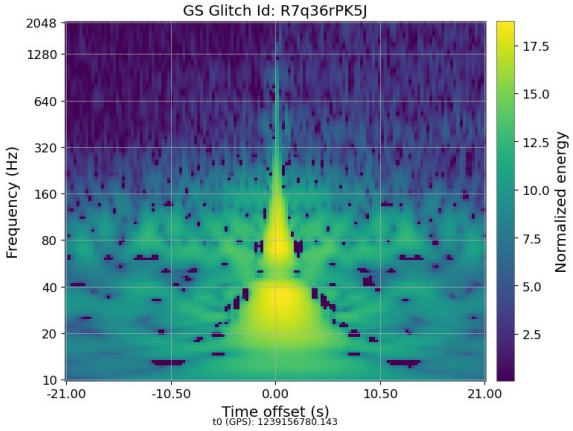
What choices influence glitch visualization in significant ways?

1. Signal **whitening method**:
   - Welch (mean) or median.
2. **Normalized energy scale**:
   - Restricted to range (0, 25.5).
   - Logarithmic scale.
   - Normalized images using a quantile scale.

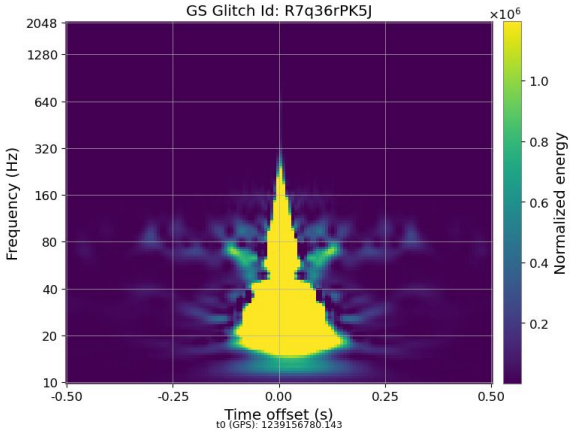
Examples: Normalized energy scale



Restricted scale



Log scale



Quantile normalized scale

A certain visualization might be better for human recognition but not necessarily for a DL model!

What choices influence glitch visualization in significant ways?

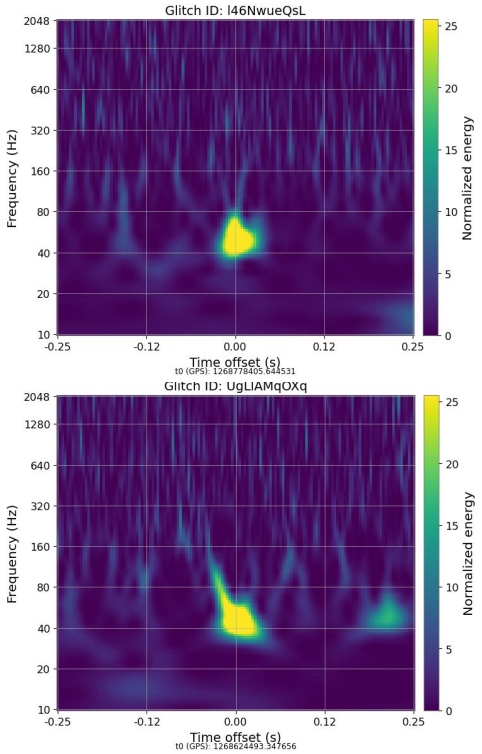
1. **Whitening method**:
   - Welch (mean) or median.
2. **Normalized energy scale**:
   - Restricted to range (0, 25.5).
   - Normalize images using a quantile scale.
   - Logarithmic scale.
3. **Quality factor**: Q; **time window:** T
   - Use a single fixed Q and T.
   - Use multiple fixed Qs and Ts.
   - Use a single random Q and fixed T.

Examples: Q-values and time windows

Air compressor

Blip

Q = 5.65, T= 0.5s



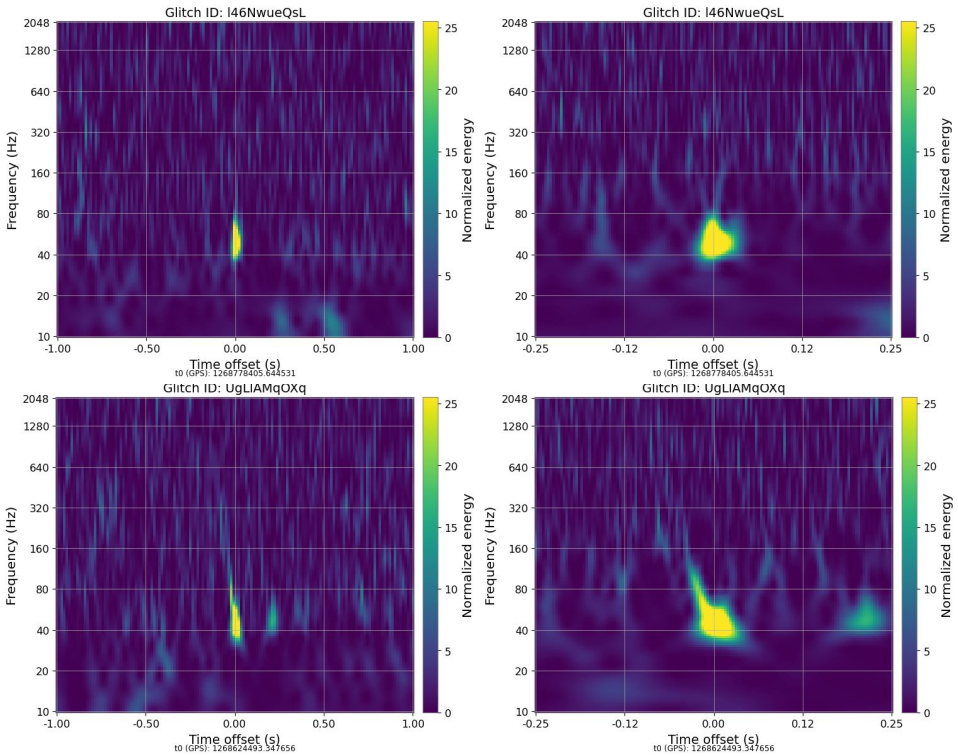
Examples: Q-values and time windows

Q = 5.65, T= 2s

Q = 5.65, T= 0.5s

Air compressor

Blip



Examples: Q-values and time windows

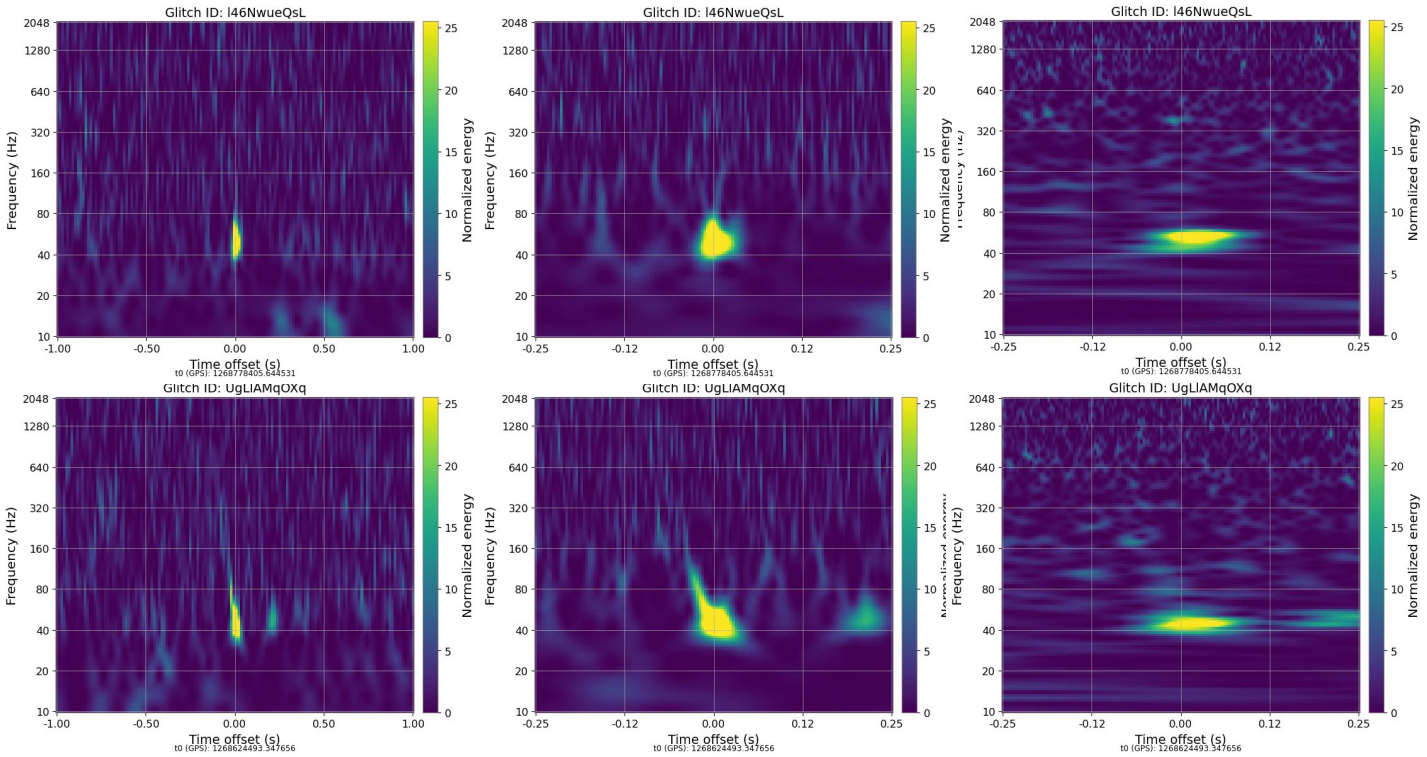
Q = 5.65, T= 2s

Q = 5.65, T= 0.5s

Q = 22.62, T= 0.5s

Air compressor

Blip



Dataset

Gravity Spy Database

↓ Filtered

Run **O3b**
Confidence >= 0.95
Max: 2000
Min: 200
Remove:
- None of the above
- Whistle

↓ Split in 70/10/20%

Train: 9529, Val: 1219, Test: 2438
11 classes

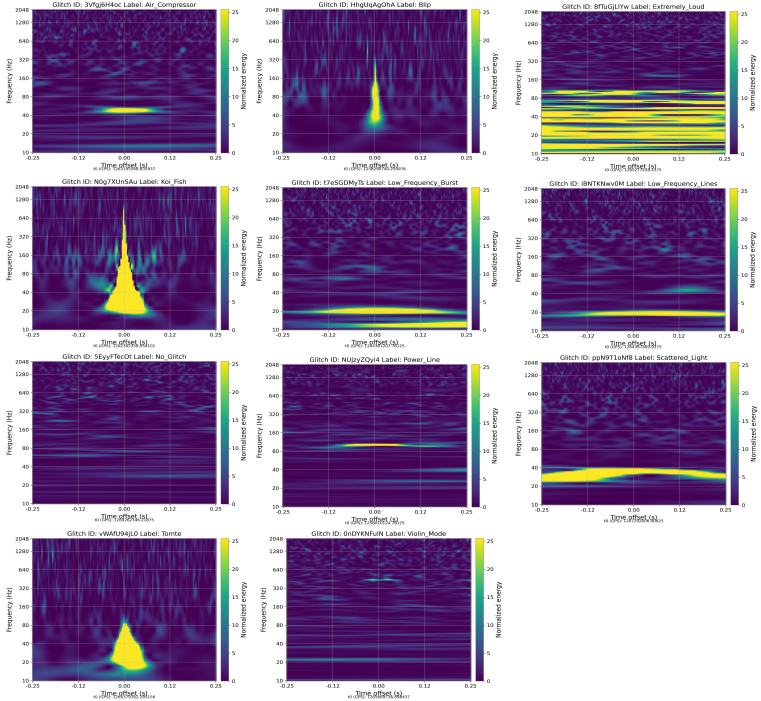
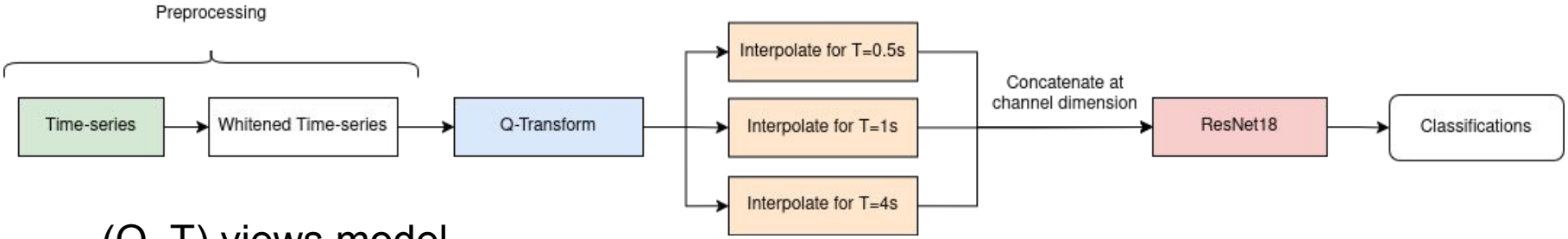


Fig. Example of each class in the dataset.

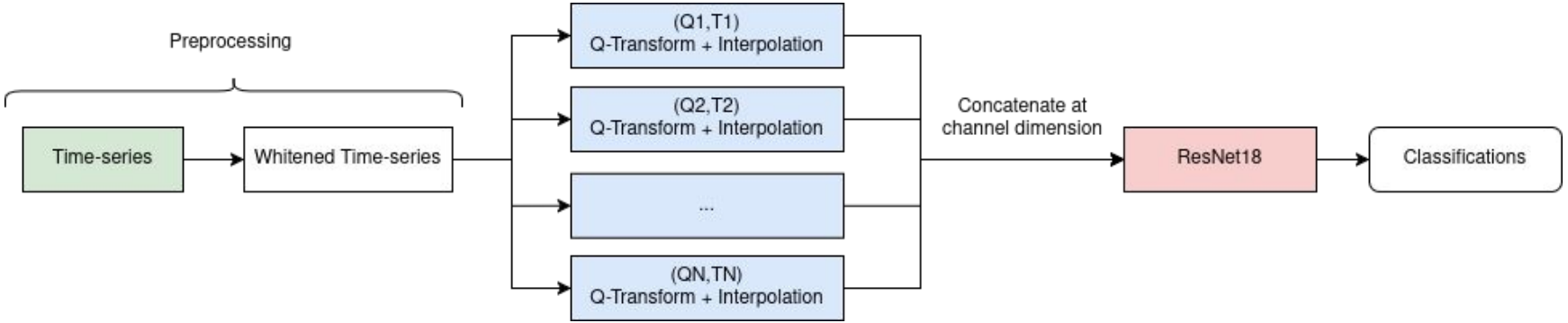
Methodology

Setup a Q-transform in an inference time pipeline to test various settings.

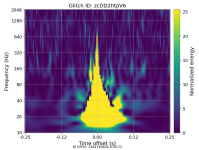
- Benchmarking model (welch whitening, (0,25,5) restricted range)



- (Q, T) views model



CNN feature embeddings



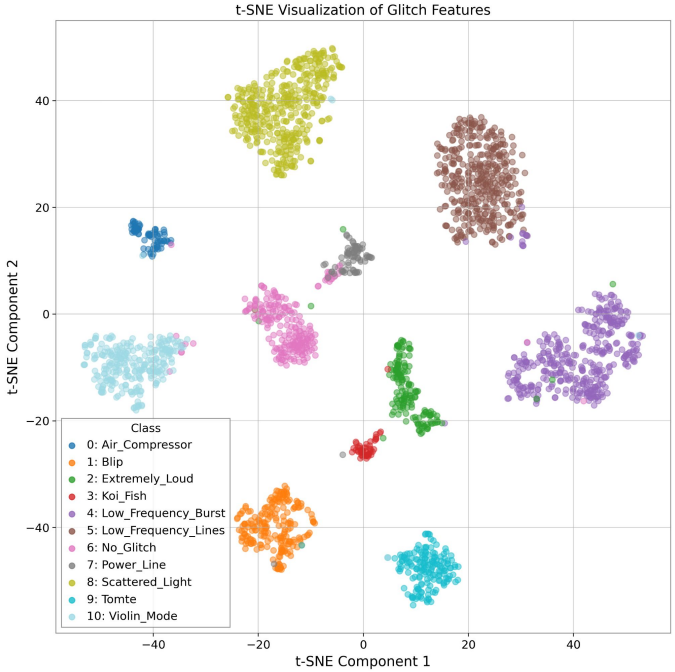
ResNet18
(- FC layer)

Extract features

t-SNE

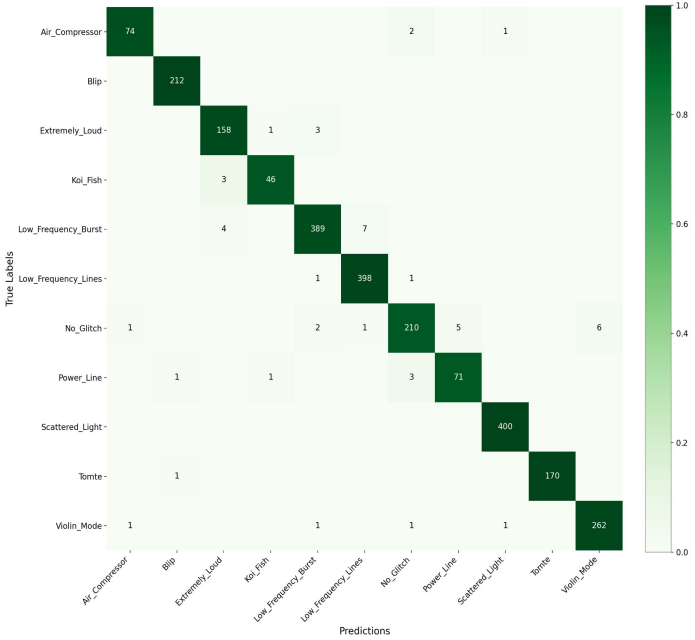
[Maaten, L.V., & Hinton, G.E. (2008)]

Embedding of features

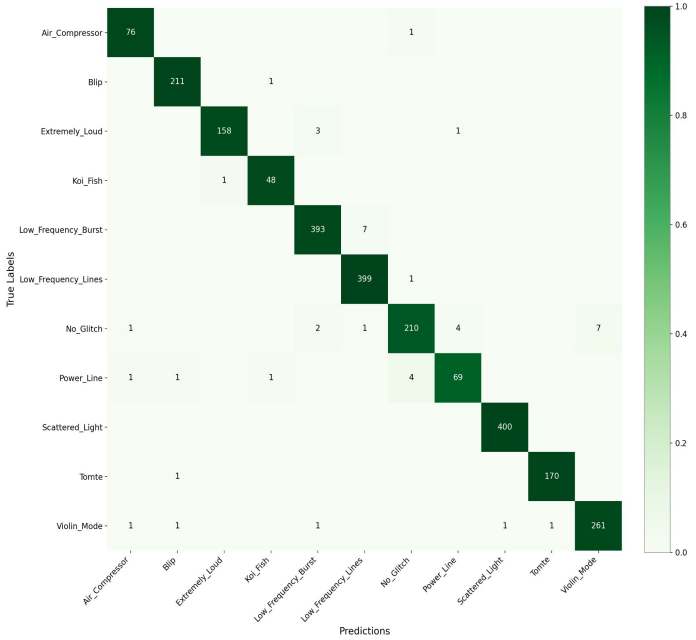


Results: whitening

Very similar!



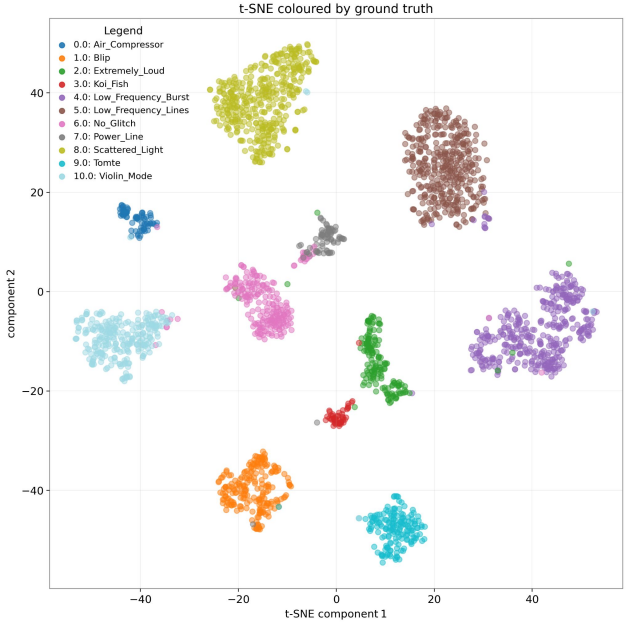
Welch - **97.30%** F1-score

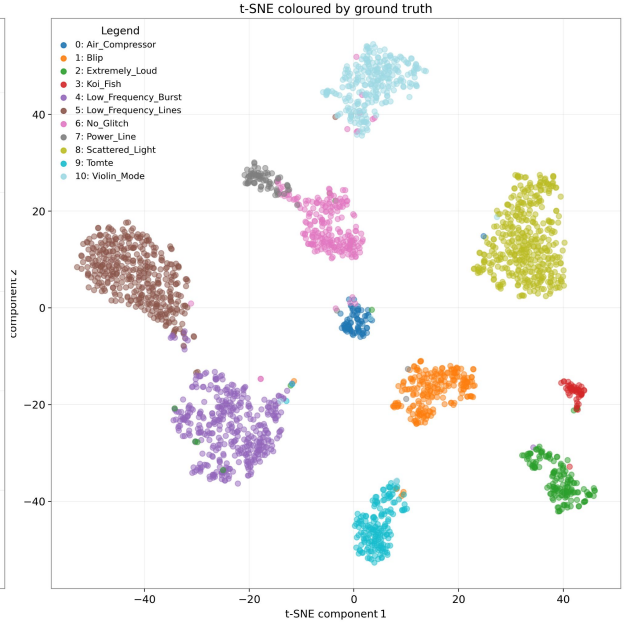


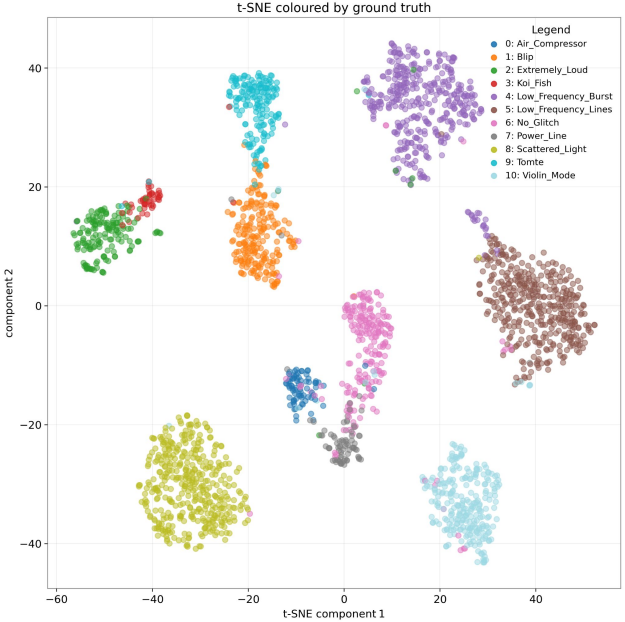
Median - **97.59**% F1-score

Test data

Results: normalized energy scale







Restricted range - **97.30** % F1

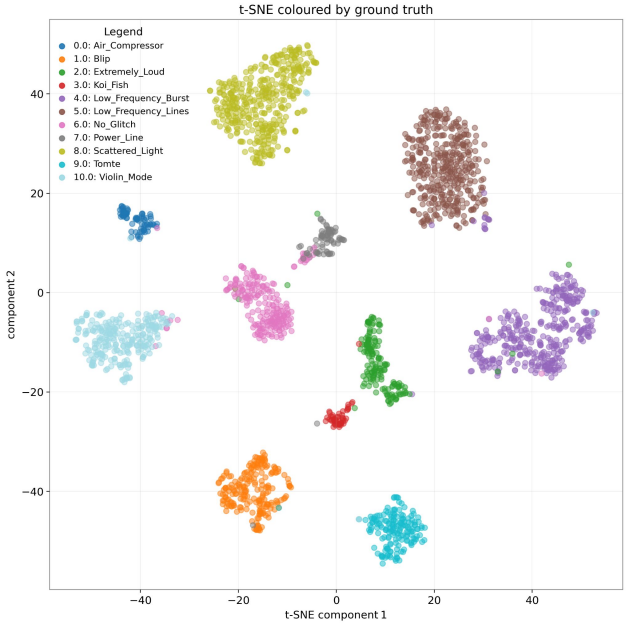
Log scale - **96.56** % F1

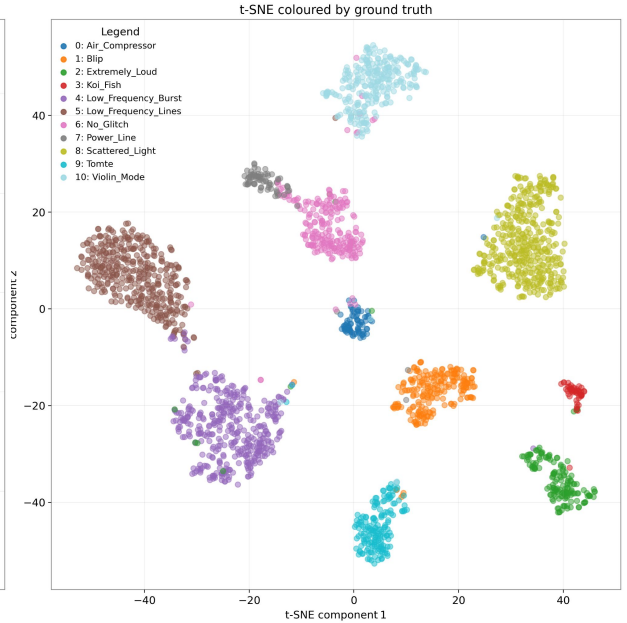
Quantile **normalized** scale - **93.06**% F1

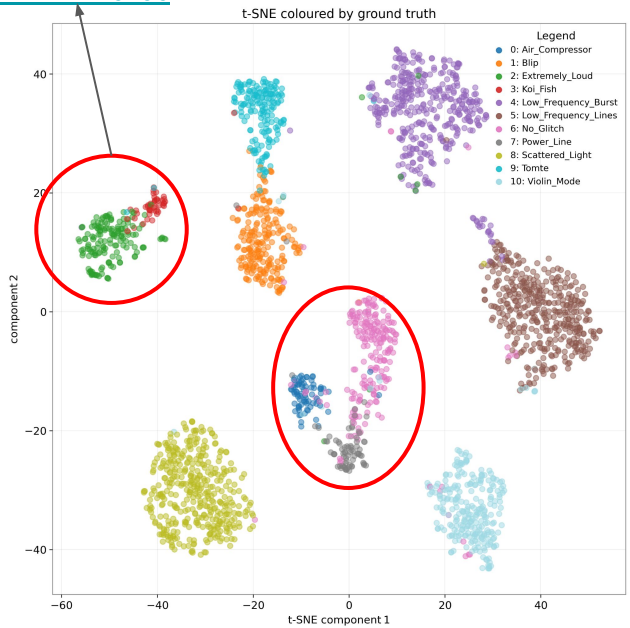
Test data

Results: normalized energy scale

arXiv:2412.16796







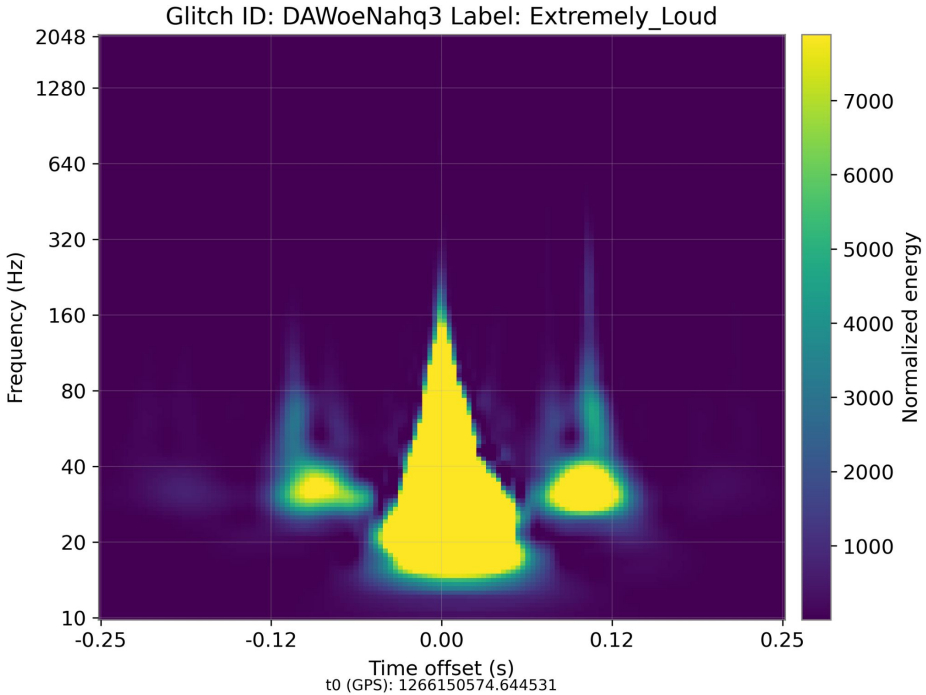
Restricted range - **97.30** % F1

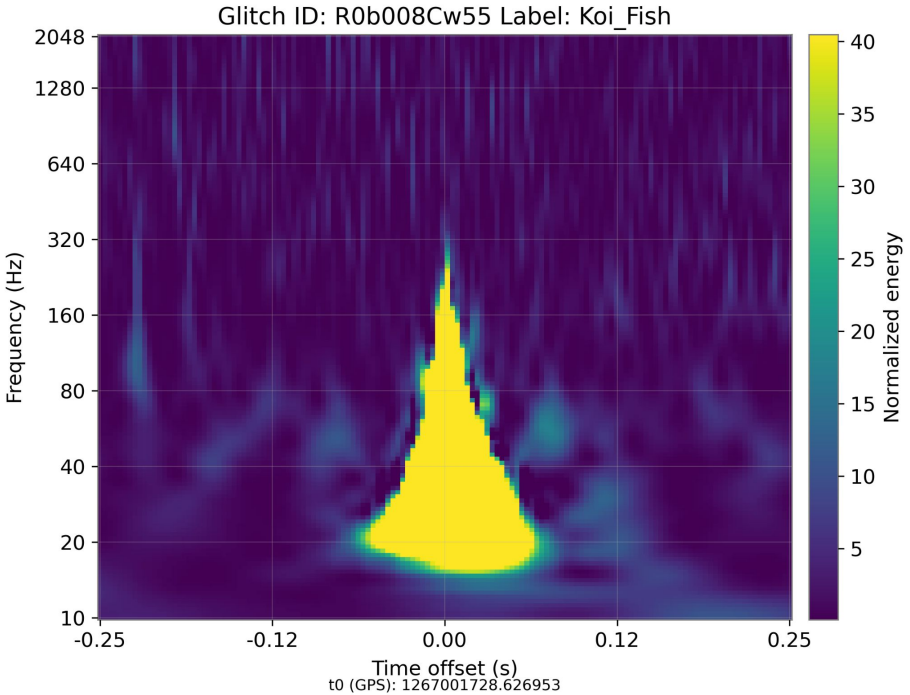
Log scale - **96.56** % F1

Quantile **normalized** scale - **93.06**% F1

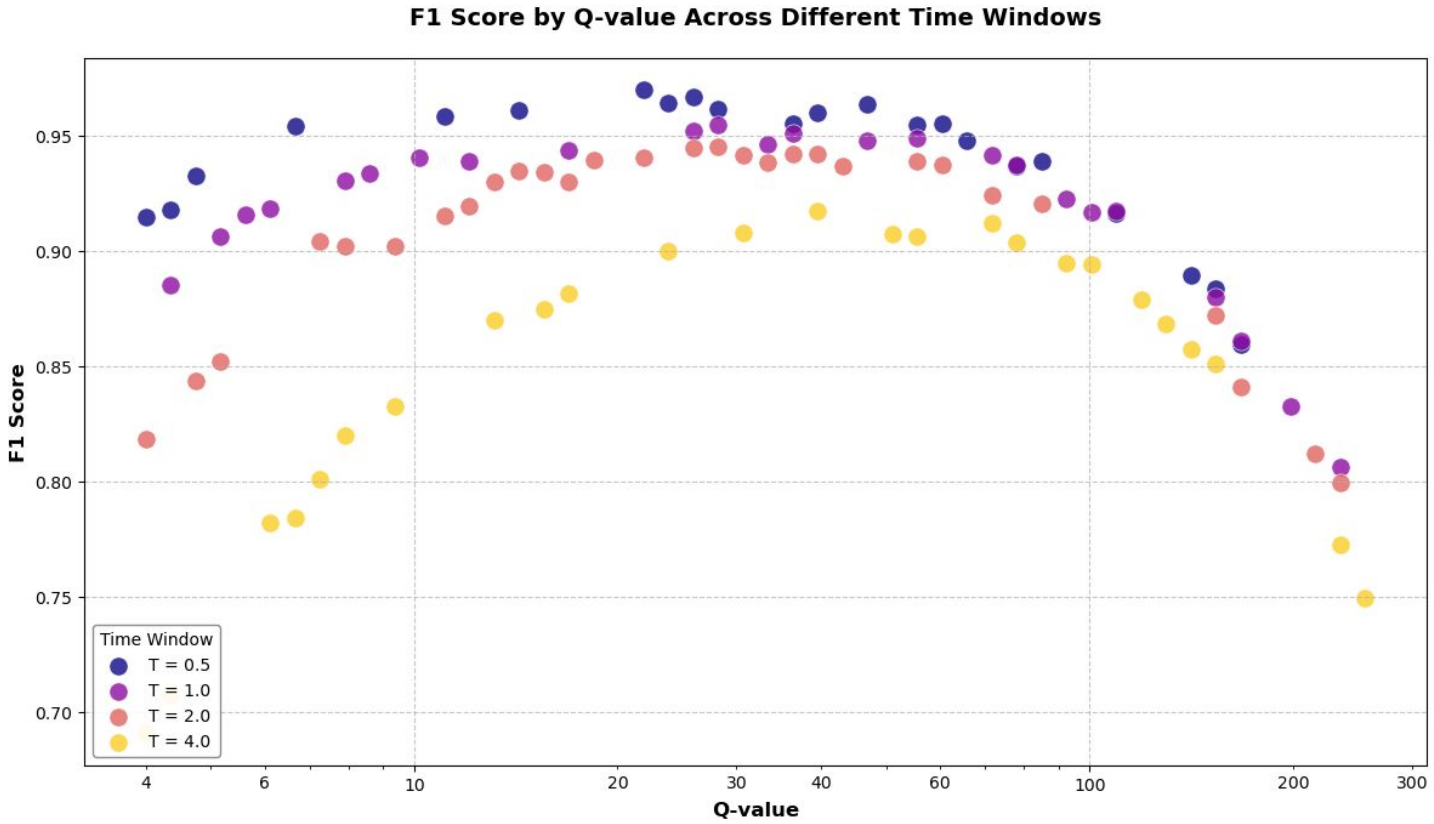
Test data

Results: normalized energy scale



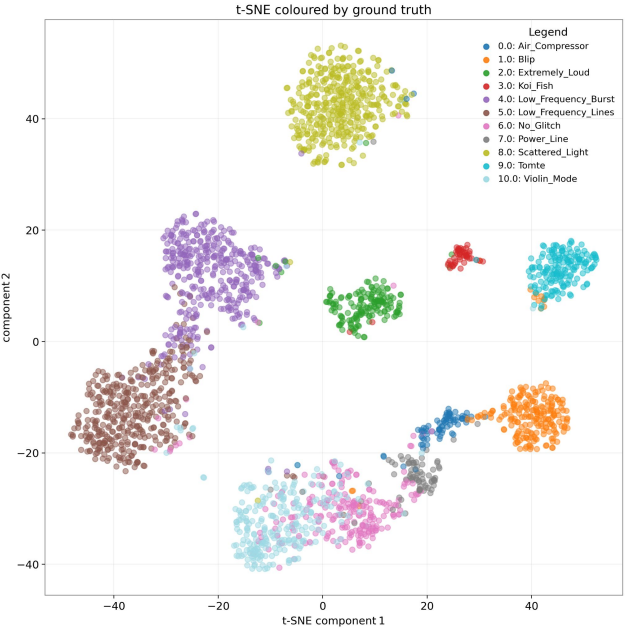


Results: one view (Q, T)

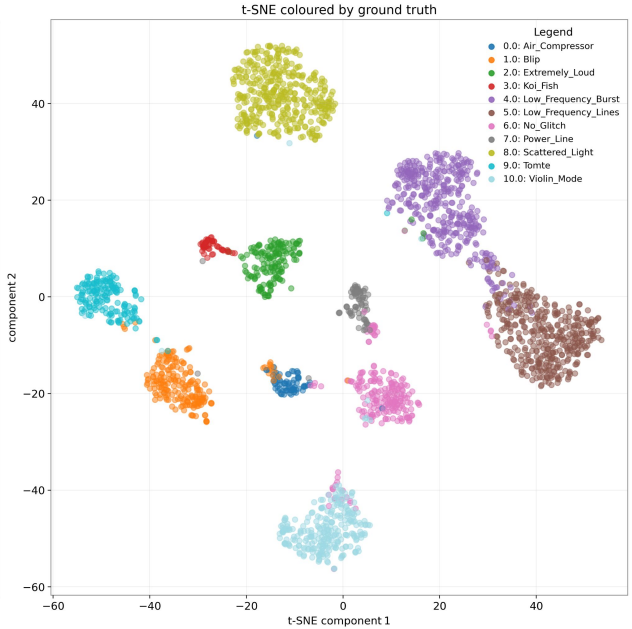


Val data

Results: one view (Q, 1s)



Q = 4



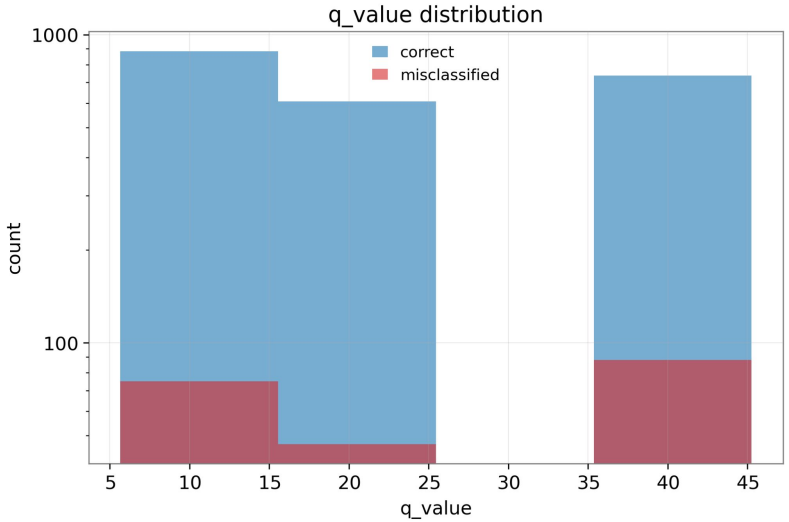
Q = 36.34

Q = 129.82

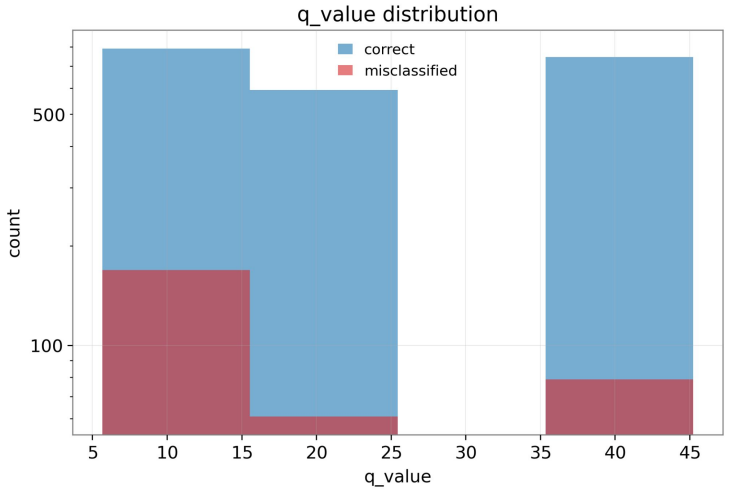
Test data

Results: one view (Q, 1s)

Distribution of classifications by Gravity Spy's Q value



Model (Q = **4**, 1s)



Model (Q = **129.82**, 1s)

Test data

Results: multiple views (Q, T)

N views	F1-score
1	95.9%
2	**96.81%**
3	96.11%
4	96.05%
5	96.37%

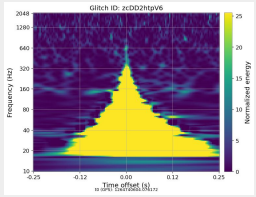
Greedy search

Randomly generate 100 candidate views with:

- T = [0.5, 1, 2, 4]
- 50 Qs chosen logarithmically between [4, 256]

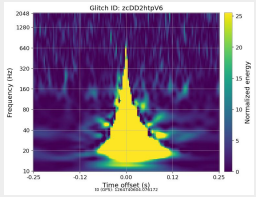
Select best model and search for next view.

N=1



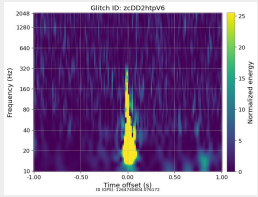
(21.84, 0.5)

N=2



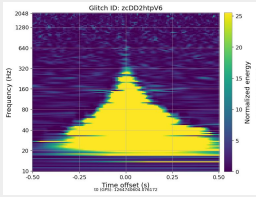
(7.89, 0.5)

N=3



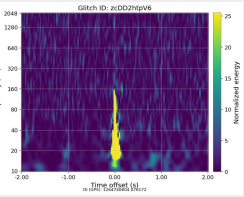
(5.62, 2.0)

N=4



(65.84, 1.0)

N=5

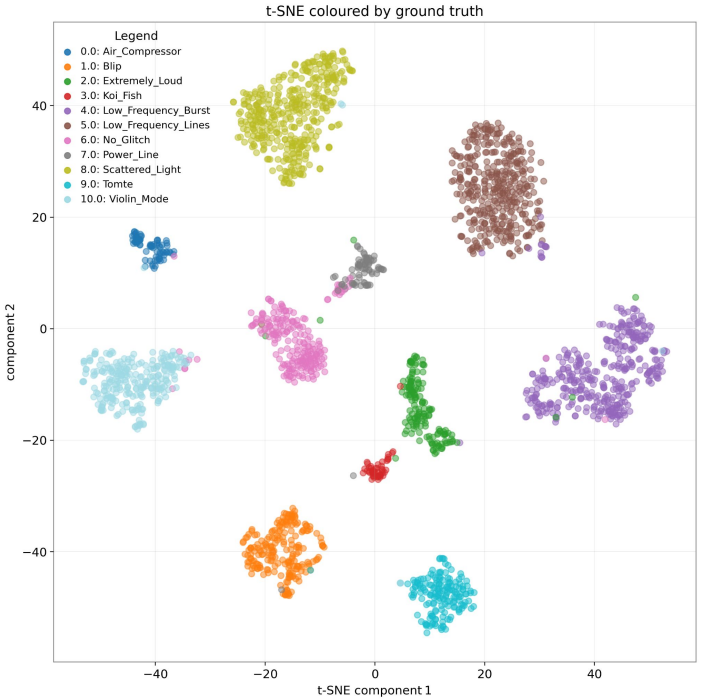


(8.59, 4.0)

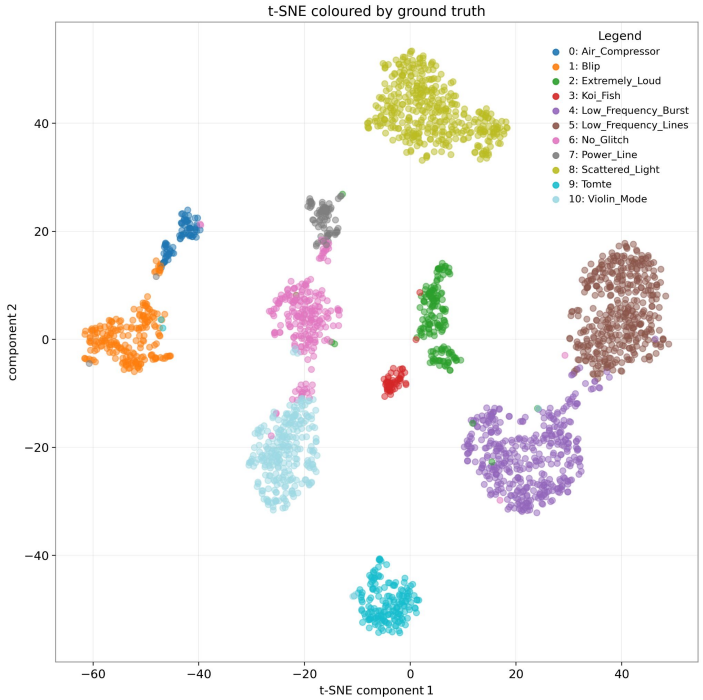
Test data

Results: random Q values (Q, 1s)

Benchmark

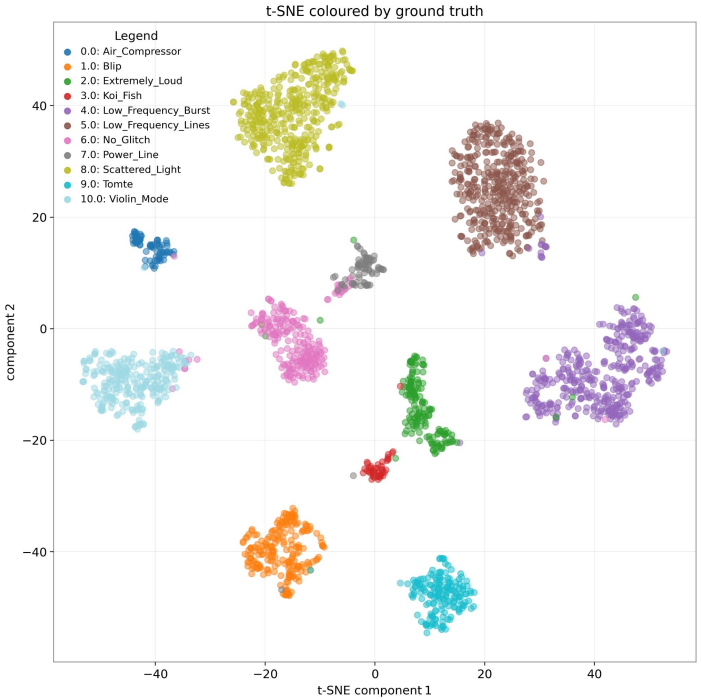


Random Q values

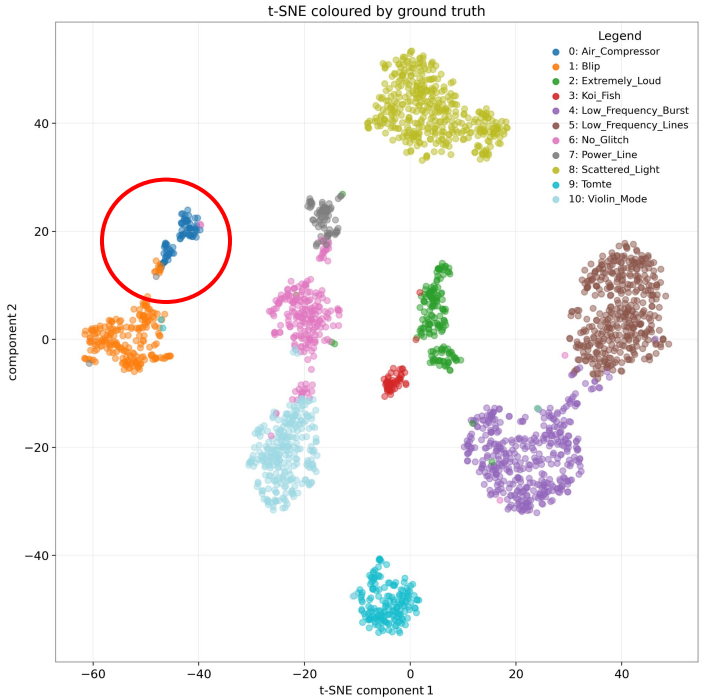


Results: random Q values (Q, 1s)

Benchmark



Random Q values



Summary and conclusions

- 🦷 Whitening method doesn't seem relevant for confidently classified glitches.
- 🎚️ Normalized energy scale is crucial for imposing SNR differences and uncover certain class similarities.
- 🎯 Q value positively biases classes and classifications that use a similar GS Q value.
- 🪄 A medium Q value range works best to cluster classes and a longer time window necessitates slightly higher Q values.
- 🔀 Different Q values also allow a good classification as an alternative to only different time windows.

THANK YOU FOR LISTENING!

…and to these awesome collaborators:

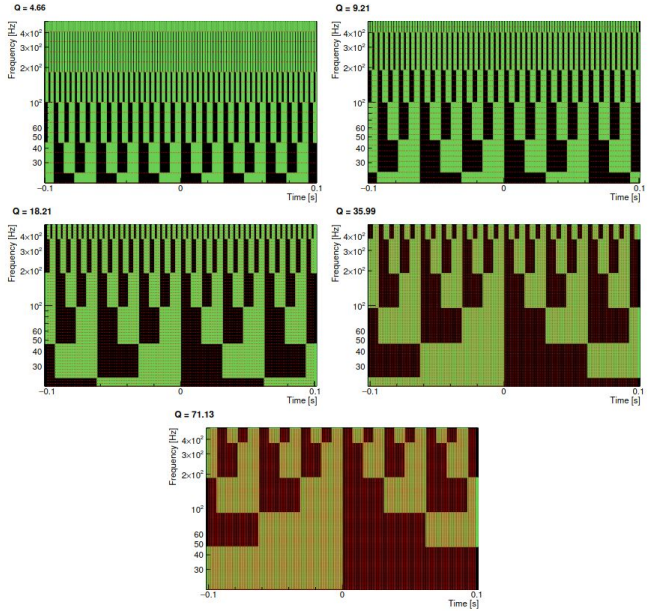
Extra slides: Q transform

$$X(\tau, \phi, Q) = \int_{-\infty}^{+\infty} x(t) w(t - \tau, \phi, Q) e^{-2i\pi\phi t} dt$$

τ, ϕ central frequency and time

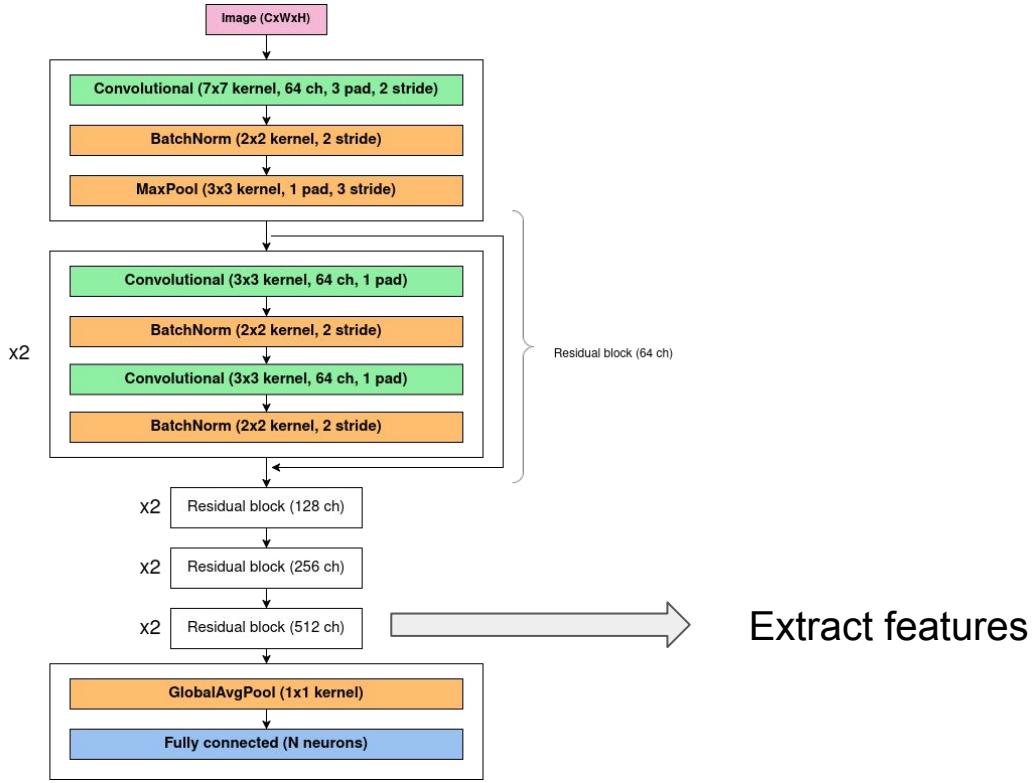
$\Delta t, \Delta f$ characteristic bandwidth and duration

$$\Delta t = \frac{Q}{\sqrt{2}\pi\phi} \quad \Delta f = 2\sqrt{2}\frac{\phi}{Q}$$



arXiv:2007.11374, https://doi.org/10.1016/j.softx.2020.100620

Extra slides: ResNet18 architecture



Extra slides: Dataset distribution

Glitch Family	Train	Val	Test
Low Frequency Burst	1400	200	400
Low Frequency Lines	1400	200	400
Scattered Light	1400	200	400
Violin Mode	931	133	266
No Glitch	787	113	225
Blip	741	106	212
Tomte	600	86	171
Extremely Loud	566	81	162
Air Compressor	267	38	77
Power Line	265	38	76
Koi Fish	172	24	49

Extra slides: unsupervised clustering of “Whistle”