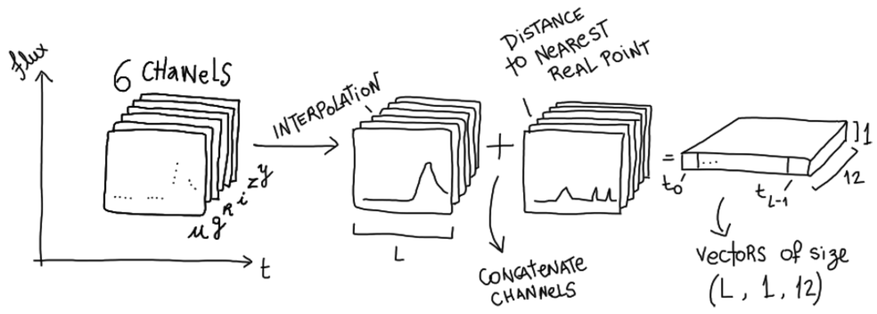
Pay attention to the Flare

Overview

Using simulated photometry, we attempt to perform light curve classification of transients as early as possible, without additional features (such as host galaxy photometric redshift) that might not be available when an event is first observed.We make use of a technique used in Natural Language Processing, referred to as **Self-Attention**, which makes a classifier learn to focus on important parts of a sequence, hoping that in our case, the relevant segment is the flare (increase in brightness) or the bit just before it. We use 3 different models known to perform well on the time series classification as baselines and find that by adding a *self-attentive layer* to one of them, classification results improve.

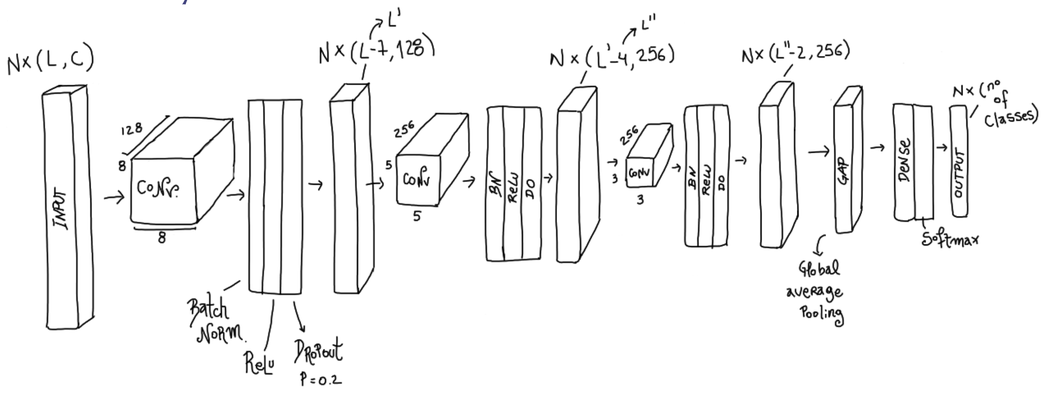
Dataset



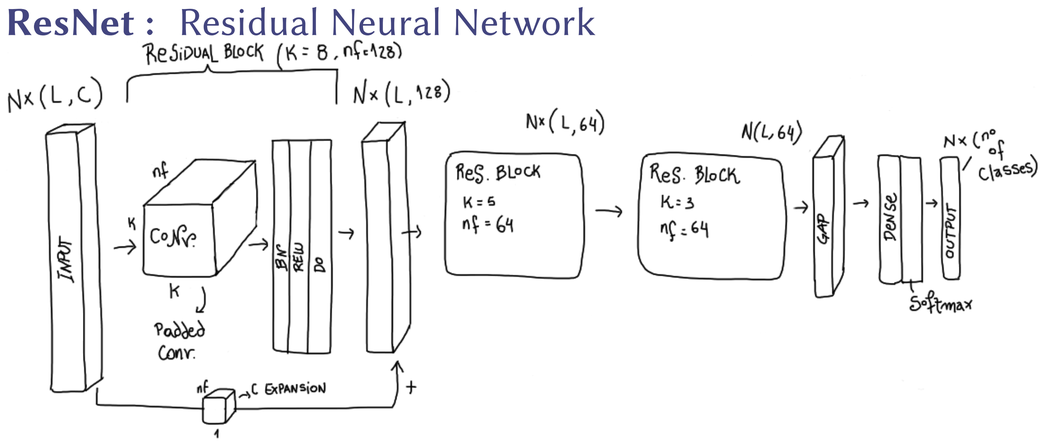
We use 300,000 samples of extragalactic objects in the PLAsTiCC dataset to train/validate/test our classifier. All light curves were linearly interpolated. Only the training set was balanced.

Baselines

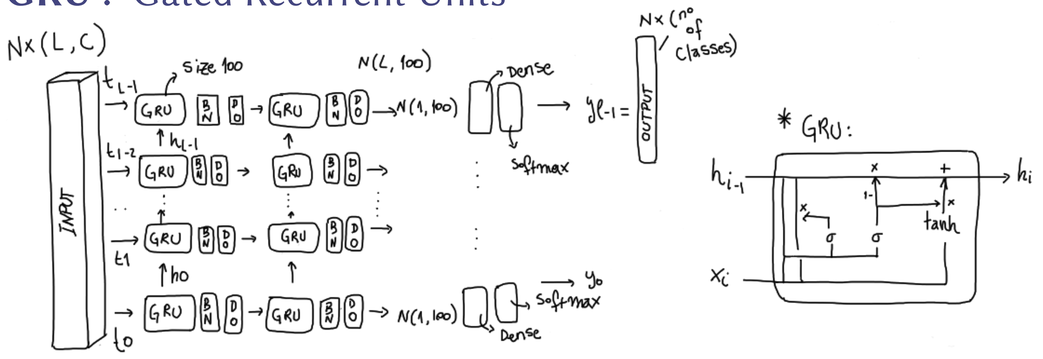
- **FCN :** Fully Convolutional Neural Network



- **ResNet :** Residual Neural Network



- **GRU :** Gated Recurrent Units

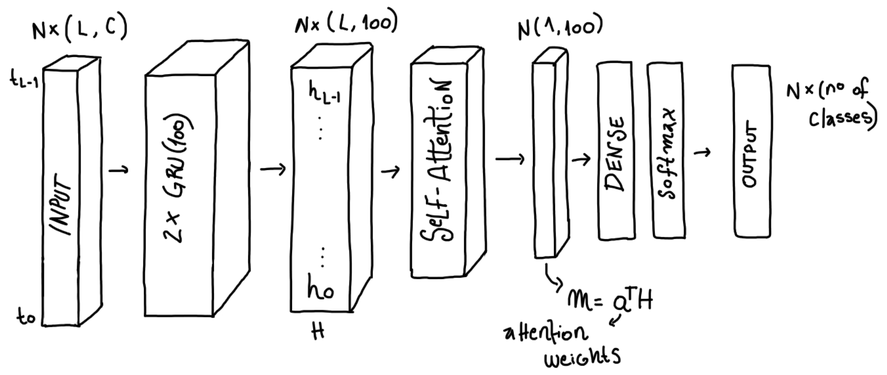


Future Work

- Construct a training set that resembles ZTF data, so we can test the classifier with real data.
- Adjust $W1$ and $w2$ dimensions to see which perform best.
- Construct visualization of what the attention layer is focusing on.
- Embed classifier on Lasair (https://lasair.roe.ac.uk)

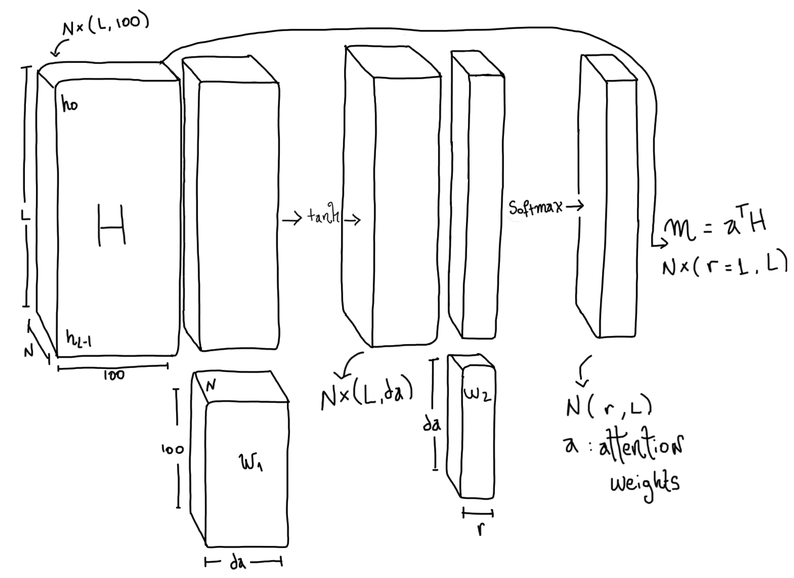
Self-Attentive Model

- **Architecture**: GRU baseline with an added self-attention layer.

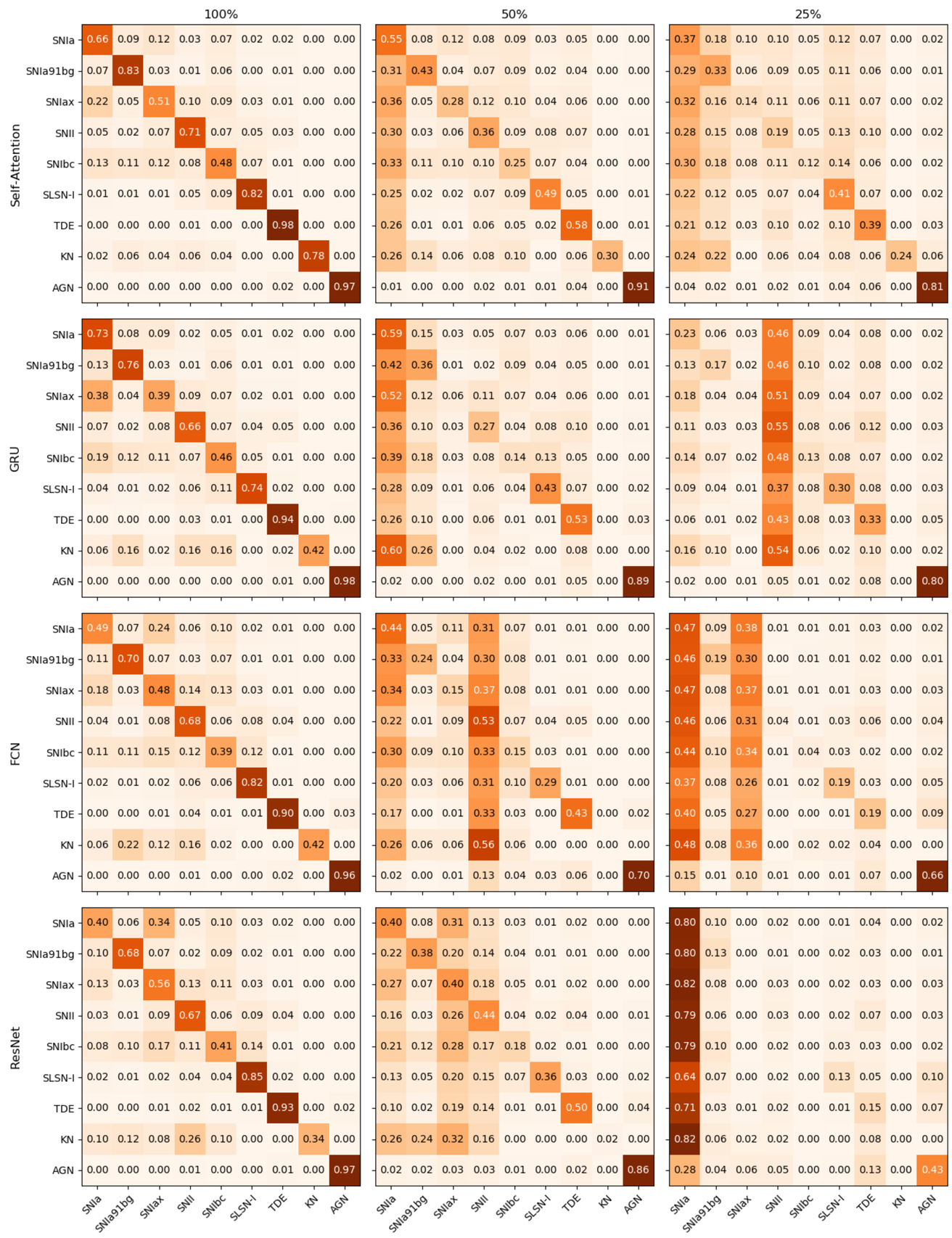


- **Self-attentive Layer**:

Given the outputs of the baseline GRU $H = (h_0, h_2, \ldots, h_{L-1})$, it computes *attention weights* $a = softmax(w_2 \tanh(W_1) H^T)$, where W_1 and w_2 are learnt and their dimensions are hyper-parameters than can be chosen arbitrarily.



Classification Results



Confusion matrices for all models tried with 100%, 50% and 25% of the light curve. Y axis is the true label while X axis is the predicted label. Numbers in the diagonal represent the portion of objets correctly classified for a given class.

Amanda Ibsen aibsen@roe.ac.uk https://github.com/aibsen/pay_attention_to_the_flare IfA - University of Edinburgh