Neuron Model for Filtering Noisy Speech

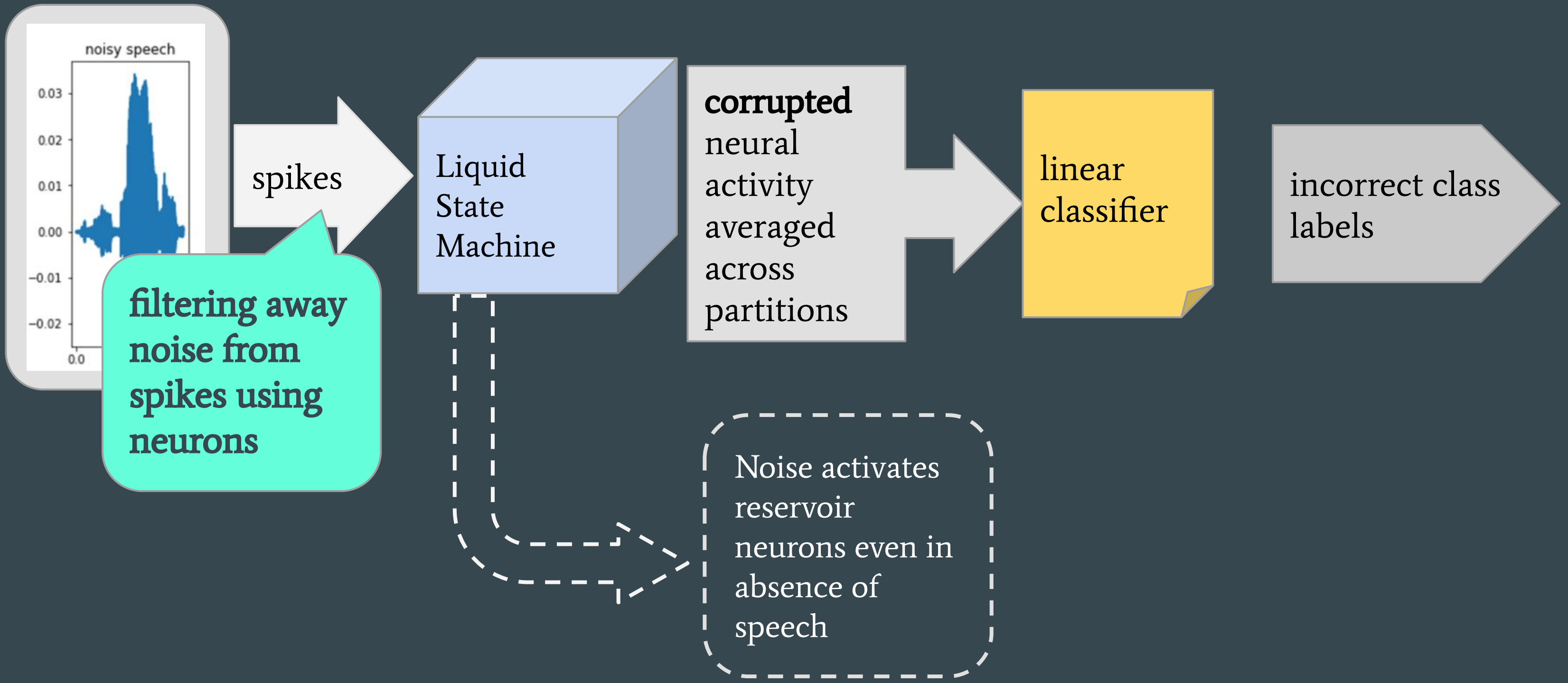
BTP-2, Spring 2023

Nivedya S Nambiar, 190070039
Guided by Prof. Udayan Ganguly and Anmol Biswas

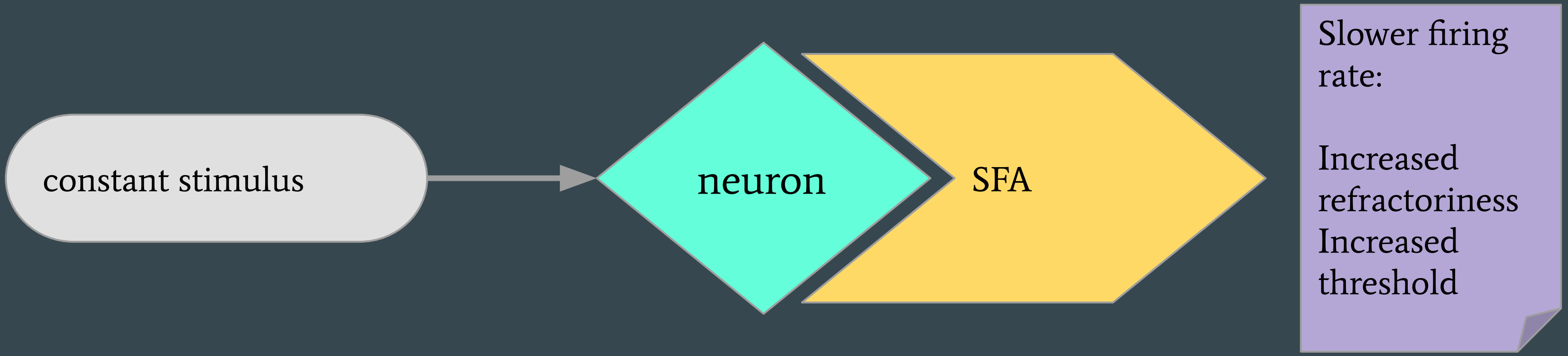
Outline

1. Introduction to problem statement
2. Overview of Spike Frequency Adaptation
3. Methodology
   a. LSM with SFA
   b. Modified dataset
   c. Gating as a filtering technique
   d. Developing energy detector, adding SFA
   e. Adding a voice bar detector
4. Testing with a reservoir
5. Conclusions, main learnings
6. Limitations and future work

Filtering away noise in speech using neurons



A tool - Spike Frequency Adaptation (SFA)



$$\frac{dV_m}{dt} = -g * a_K * V_m - g_K * a_K * (V_m - V_K) + u$$

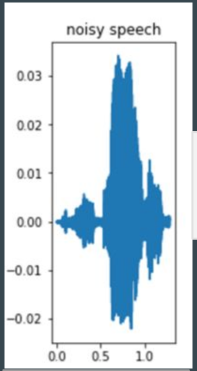
$$\frac{dg_K}{dt} = -\frac{g_K}{\tau_K} + dg * \Sigma_l \delta(t - t_l)$$

V_m: neuron’s potential
g_K: responsible for SFA
u: input current stimulus
t_l: spike times
g, a_K, V_K, dg, τ_K: constants

Methodology - Overview

1. LSM with SFA
2. Developing an energy detector
   a. Adding frequency adaptation to improve results
3. Including voice bar detector
   a. Addition of frequency adaptation
4. Testing without and with a reservoir

LSM with SFA



spikes

10

Liquid State Machine

(with SFA in neurons)

10

10

SFA expected to make neurons resistant to noise

linear classifier

correct class labels?

max test accuracy : 23%

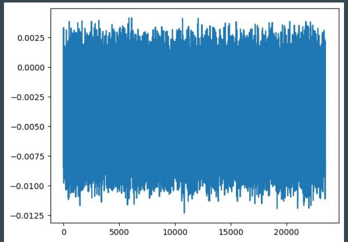
Marginally better than without SFA

DATA: spike arrays after mixing TI-46 samples with autorickshaw noise

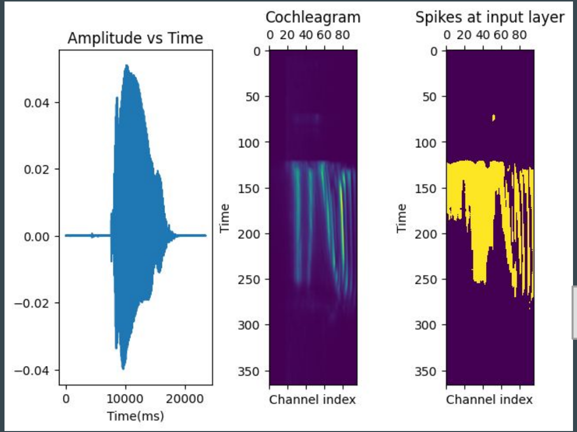
LSM with SFA - learnings for next attempt

- Further inspection revealed that data had very low SNR
- Need to work with simpler dataset
  - Higher SNR, simpler noise characteristics
- Model improvements
  - Build from bottom up
  - Focus on filtering

Modified Dataset

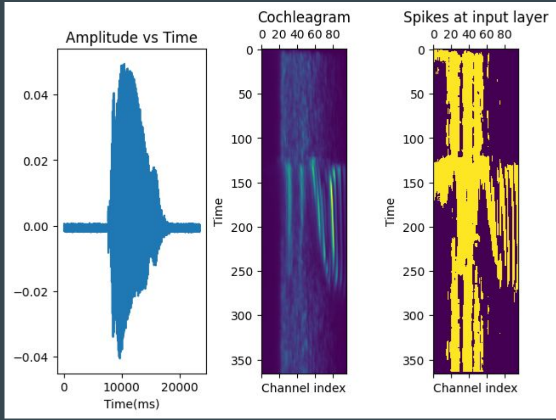


NOISE
Uniform and predictable segment of indoor home noise



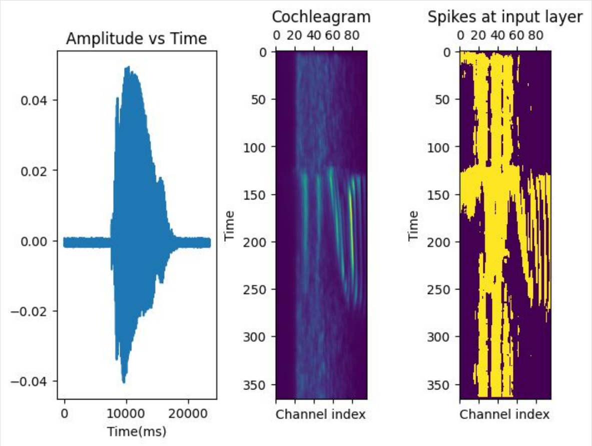
SNR=20dB

SPEECH
TI-46 spoken commands (10 classes)



Gating as a filtering technique

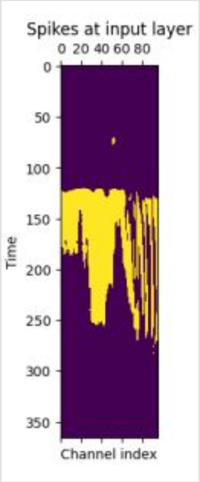
Input spike array (with noise)



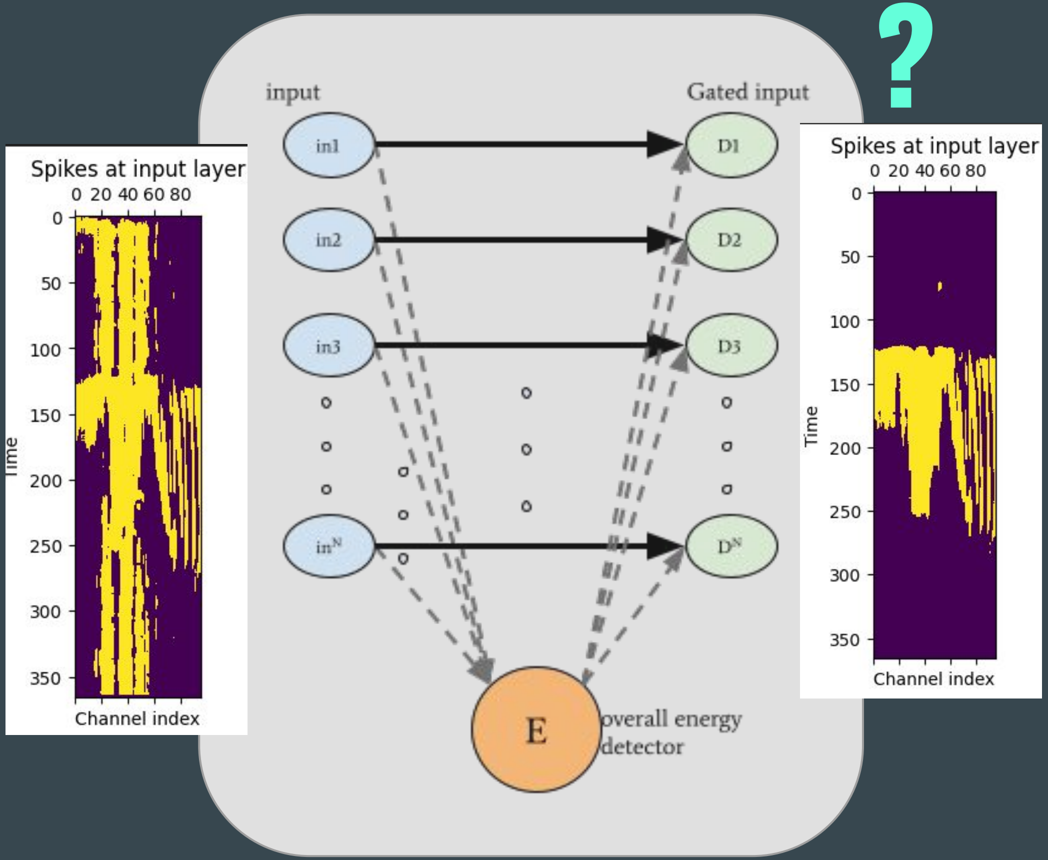
Gating in time - detect the golden window(s) for speech

allow liquid to listen in only during golden window(s)

?



Developing an energy detector for gating



Motivation:

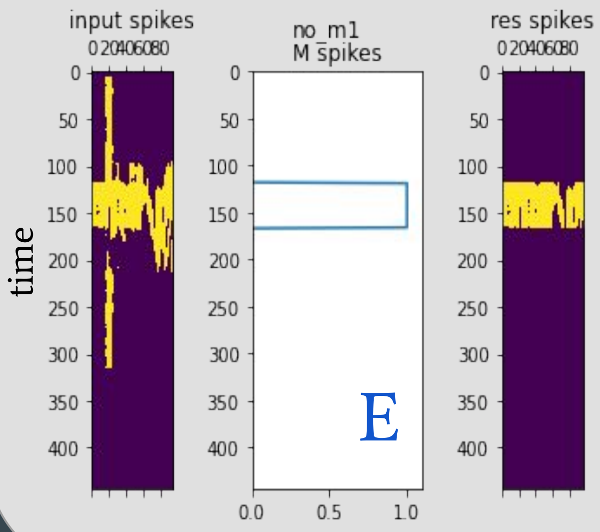
- **Speech is broadband relative to noise**
- Most channels activated in presence of speech

-“AND”xNoisySignal

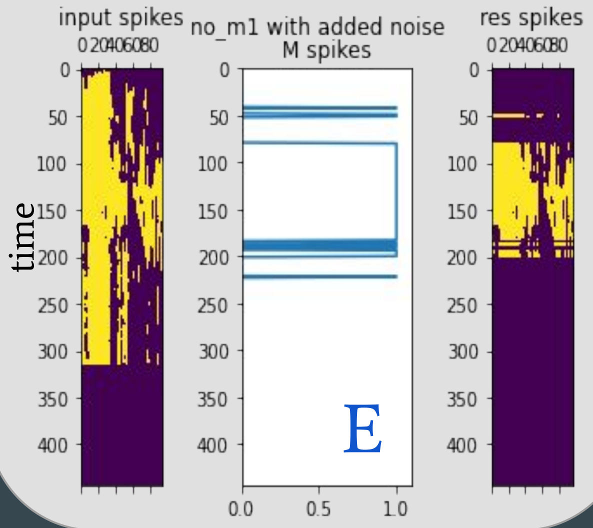
->GatedSignal

Energy detector for gating - Results

Clean signal



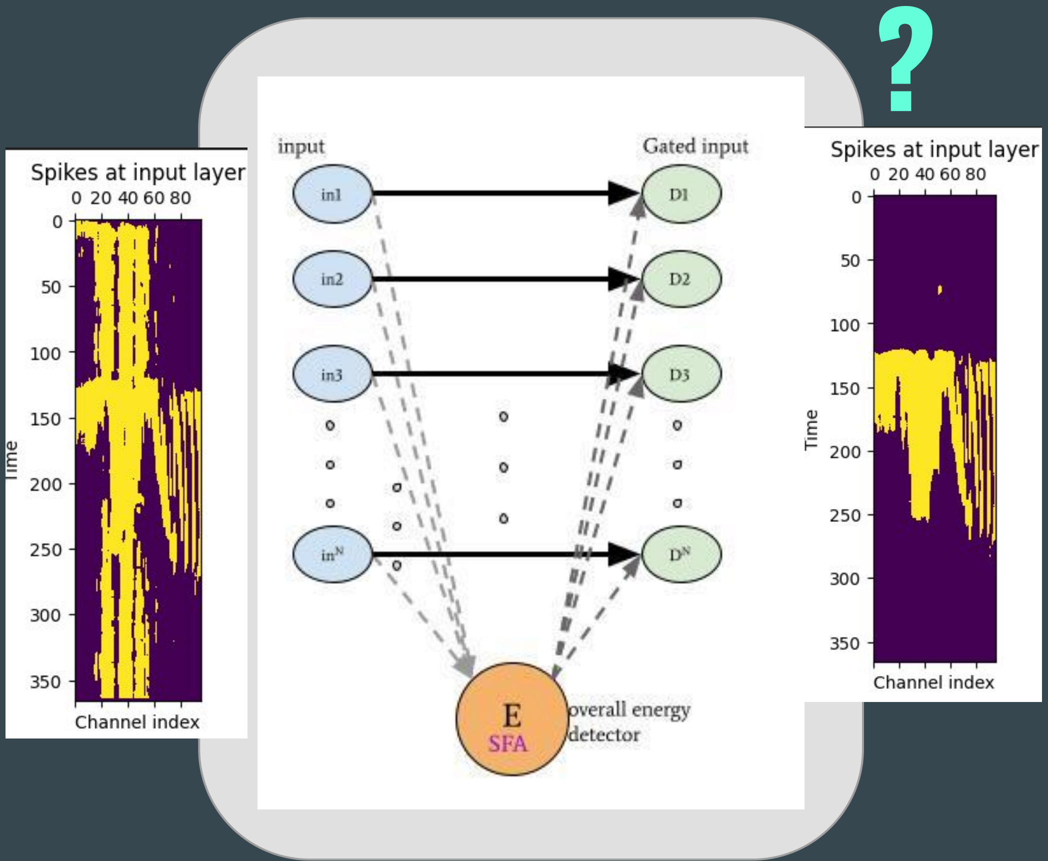
Noisy signal



Increasing threshold in E clips signal of interest

Decreasing threshold lets in noise

Energy detector with SFA

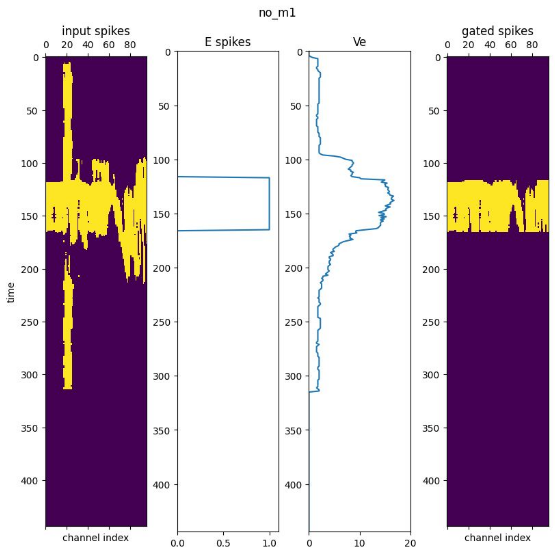


Motivation:
- With SFA, E becomes resistant to noise even when threshold is lowered

Energy detector with SFA - Results

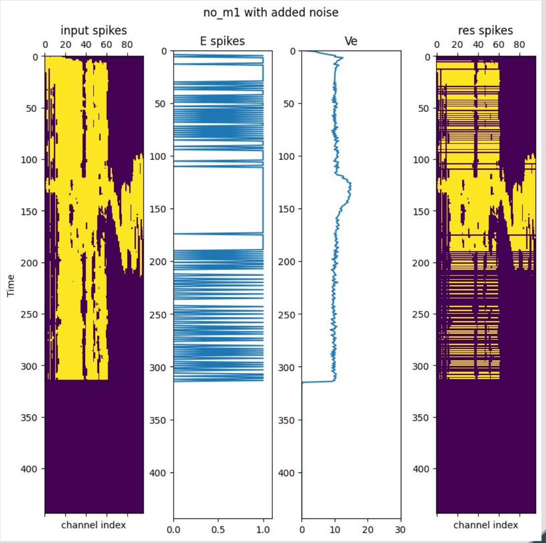
Clean signal

time



Noisy signal

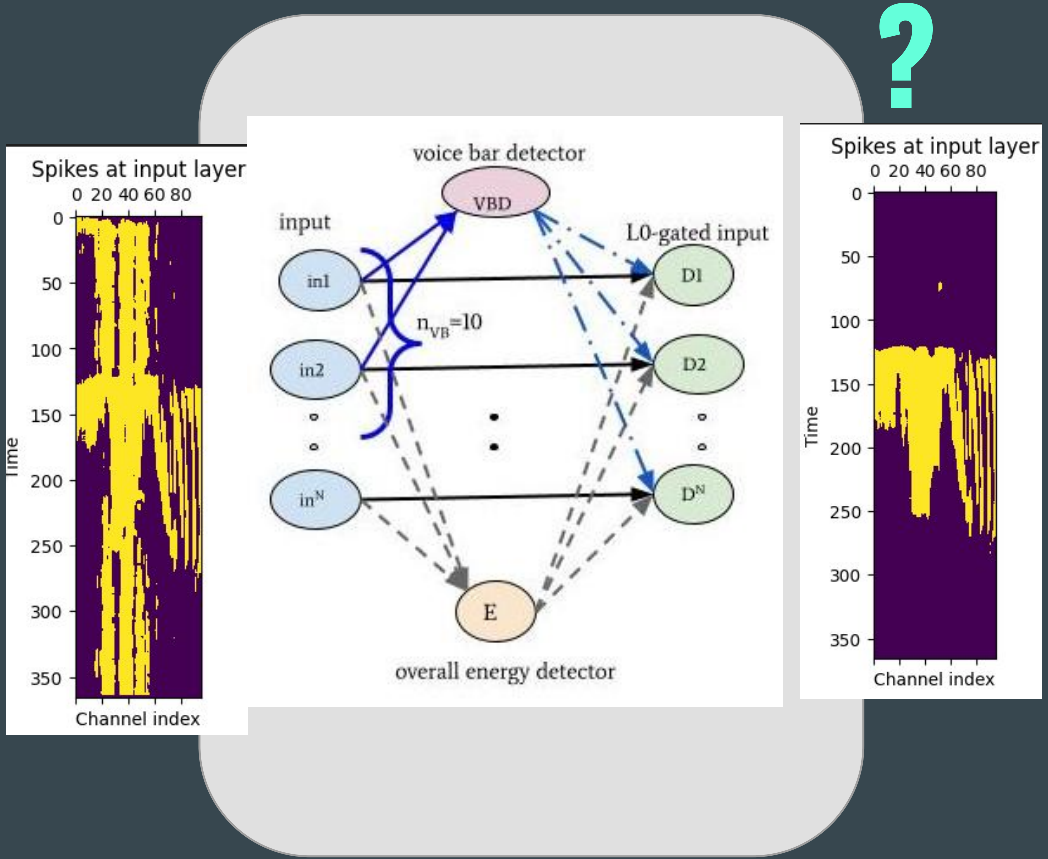
time



E not accurate for noisy signals
- clipping of useful signal at times

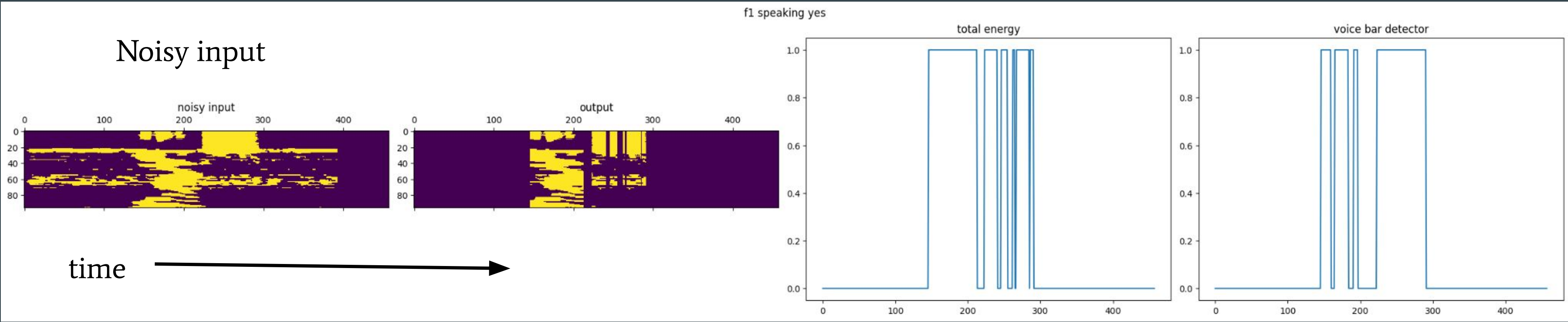
Finding the sweet spot of parameters requires manual tuning

Voice bar detector with Energy detector



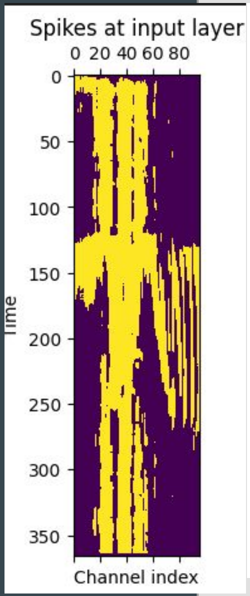
Motivation:
- Voiced sounds activate lower frequency range, but have low overall energy
- retain E for voiceless sounds (turbulent airflow)

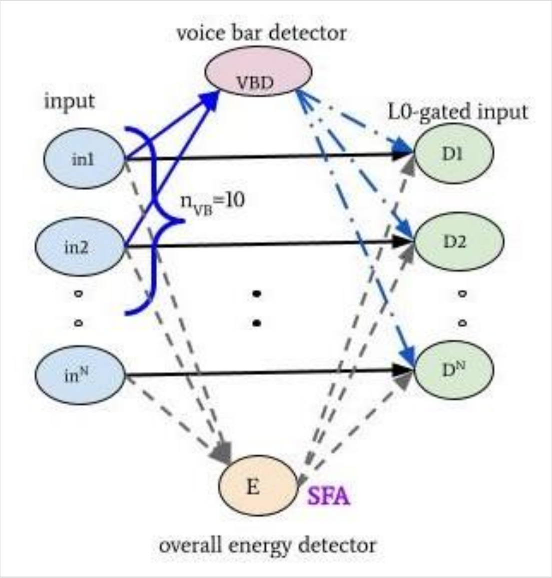
Voice bar detector with Energy detector - Results



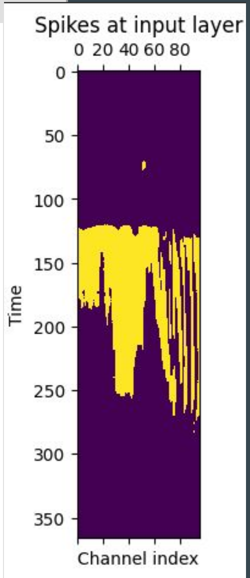
Weights to L0 from voice bar detector vs E had to be adjusted manually to achieve this
Some clipping still occurs as signal from E decreases

Voice bar detector with Energy detector - Adding SFA



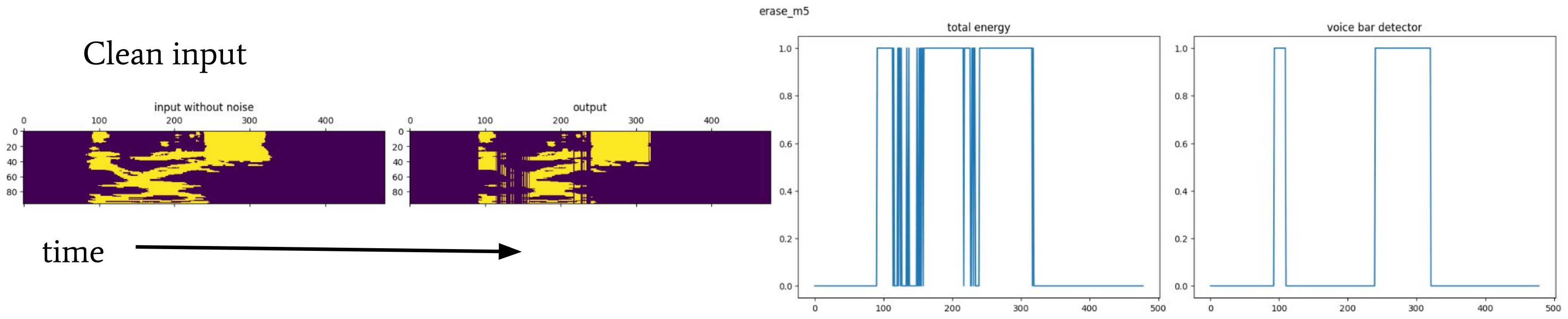


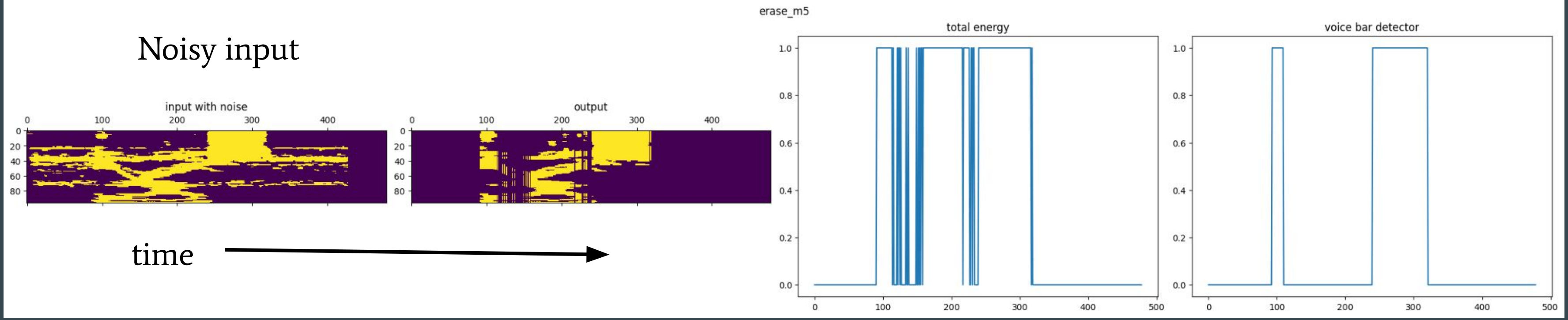
?



Motivation:
- Regulate activity of E after threshold is lowered

Voice bar detector with Energy detector (SFA) - Results





Test scores after adding linear classifier

Directly feeding gated spikes/noisy spikes to linear classifier. Train on clean data

Table 1: Test scores for different data types - noisy, with and without gating

Data Type	trained on same type	trained on clean data
Noisy data without gating	0.924	0.574
Noisy data after gating	0.880	0.701

Adding reservoir, and training on clean data

Table 2: Test scores after adding reservoir and training on clean data, tested on - noisy, with and without gating

Data Type	Train score	Test score
Noisy data without gating	0.969	0.425
Noisy data after gating	0.969	0.604

Conclusions, main learnings

- This project explores how artificial neuron models can work in gating to remove noise
- Introduced two neurons, energy detector E and voice bar detector VBD after studying the properties of speech vs noise
- SFA used to regulate spike frequency
- Gated version of input is passed onto liquid reservoir and tested with linear classifier (trained on clean data input)
- Test scores indicate gating scheme does take noisy input closer to clean input (of same class label)

Limitations and future work

1. Present scheme only includes filtering in time
   a. Individual spikes could be examined and eliminated
   b. Utilise correlation between channels activated at once during speech
2. Trade-off between inclusion of noisy spikes and clipping of useful signals
   a. Require scheme to adapt to both
   b. Parameter tuning needs to find more generic set of values across SNR, individual samples, different types of noise
3. Include learning
   a. E and VBD can be trained to take up their current roles through schemes like STDP
   b. Weights between channels in L0 for cross channel correlation in speech, utilising properties of formants
4. Use of log scale short-time Fourier transforms to potentially avoid any distortion by Lyon ear model

References

[1] Akshay Raj Gollahalli. Spike Encoders. https://github.com/ akshaybabloo/Spikes, 2020.

[2] Rob Hagiwara. How do I read a spectrogram? https://home.cc. umanitoba.ca/~robh/howto.html#intro, 2009.

[3] B. Schrauwen and J. Van Campenhout. Bsa, a fast and accurate spike train encoding scheme. In Proceedings of the International Joint Conference on Neural Networks, 2003., volume 4, pages 2825–2830 vol.4, 2003.

[4] Sciforce. Our Adaptation of Lyon's Auditory Model for Python. https://medium.com/sciforce/ our-adaptation-of-lyons-auditory-model-for-python-4f41adf55d4e, 2019.

[5] StackOverflow. Mix second audio clip at specific SNR to original audio file in Python. https://stackoverflow.com/questions/71915018/ mix-second-audio-clip-at-specific-snr-to-original-audio-file-in-python, 2022.

[6] Alessandro Treves. Mean-field analysis of neuronal spike dynamics. Network: Computation in Neural Systems, 4(3):259–284, 1993.

Thank you

Q&A