Weekly Report

Wangwon Lee, 2019/04/27

This week

- **Visualization**
  - Apply AM method to 1D-model
  - Process detail
  - In output layer
  - Analyze like SampleCNN paper

Next week

- **Visualization**
  - Compare AM result with filter
  - Find another method for listening
- **Study**
  - Brain data
  - AFNI

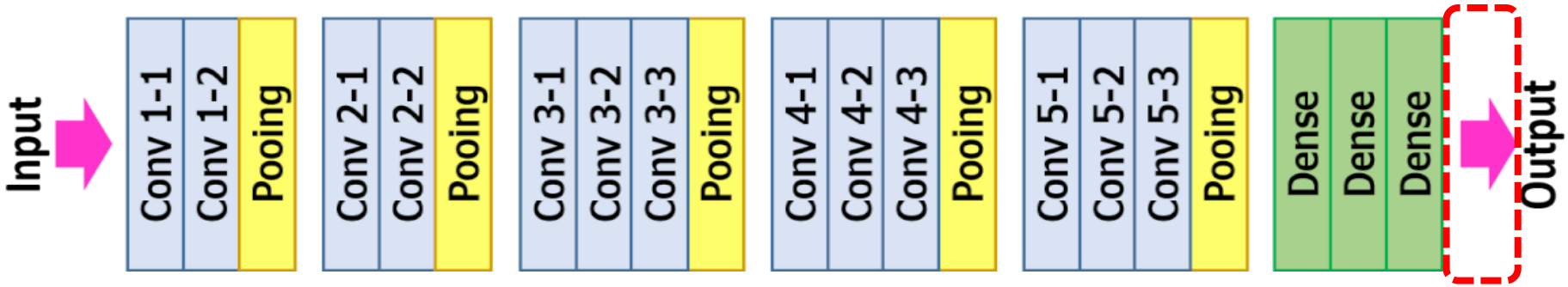
Interesting and new finding

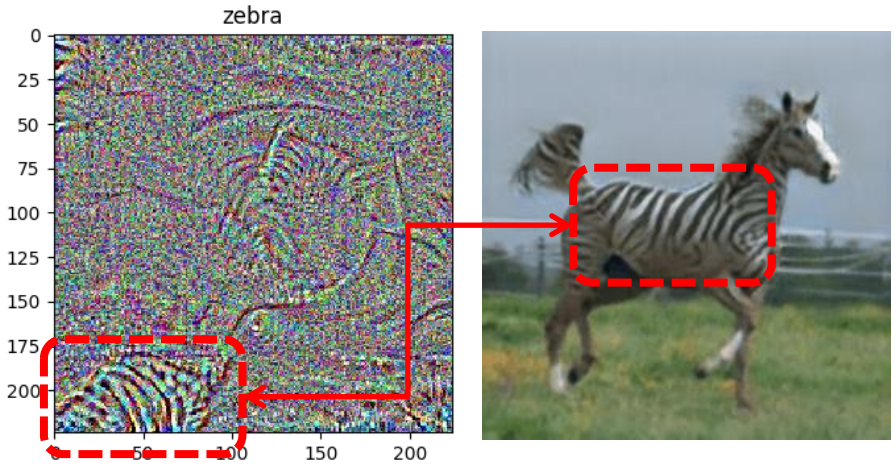
- Visualization
- Activation Maximization

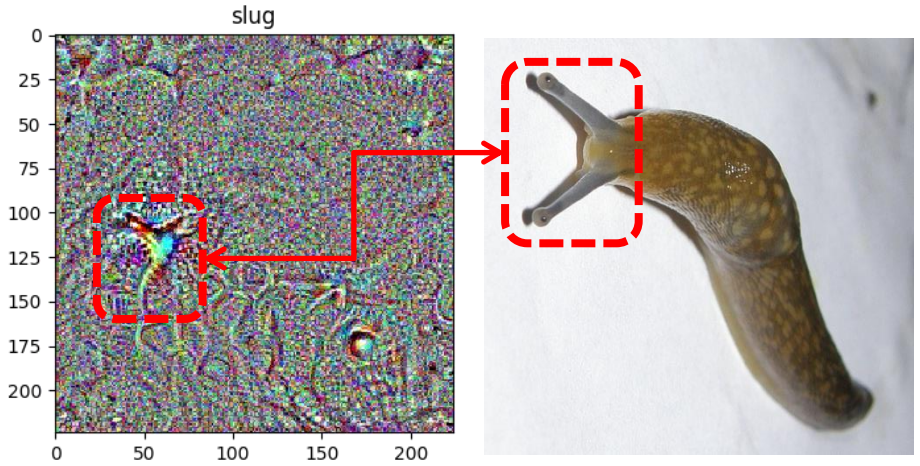
The aim of this month / Discussion

- **The aim of this month:** To investigate about CNN and visualization.

Visualization - Previous Work

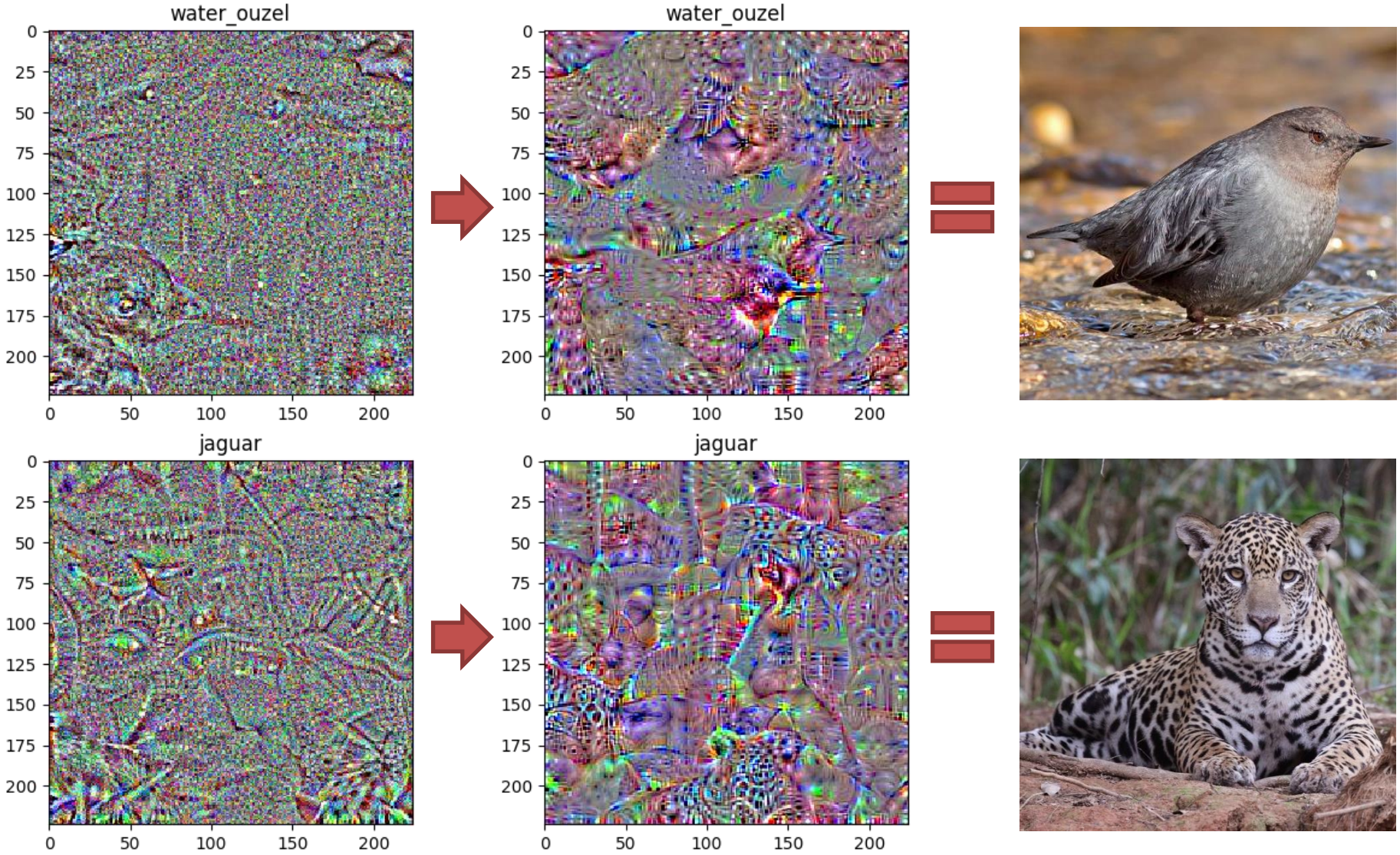




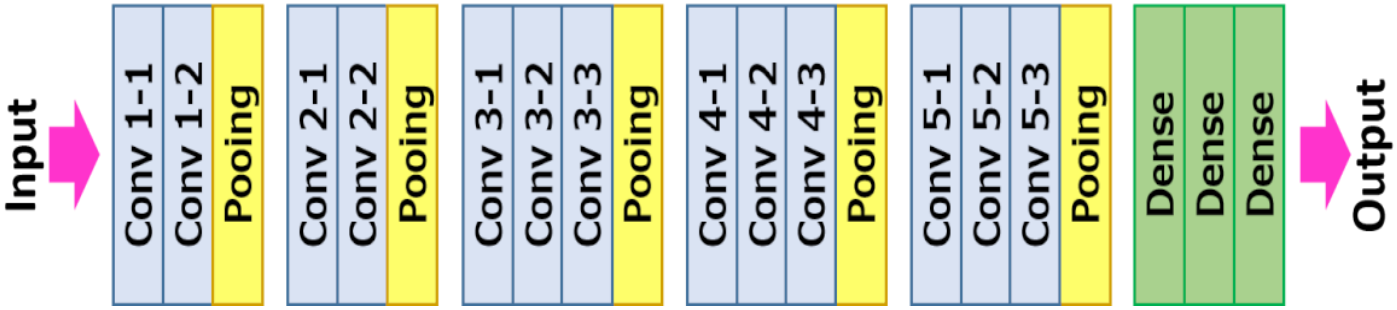


- In Softmax layer, if target node's activation is higher, that is representation of target label

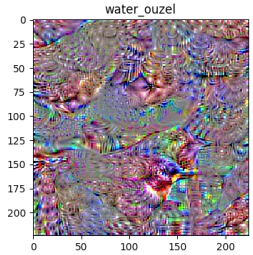
Visualization - Previous Work

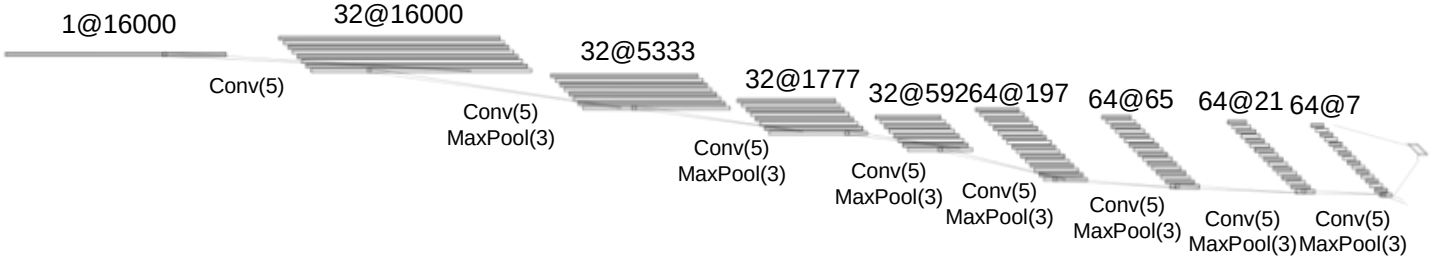


AM in 1D Model



EX)

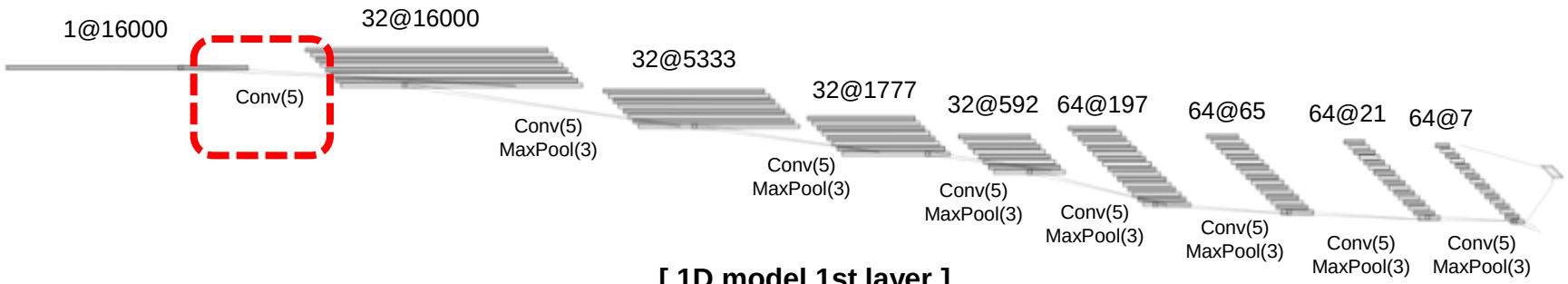




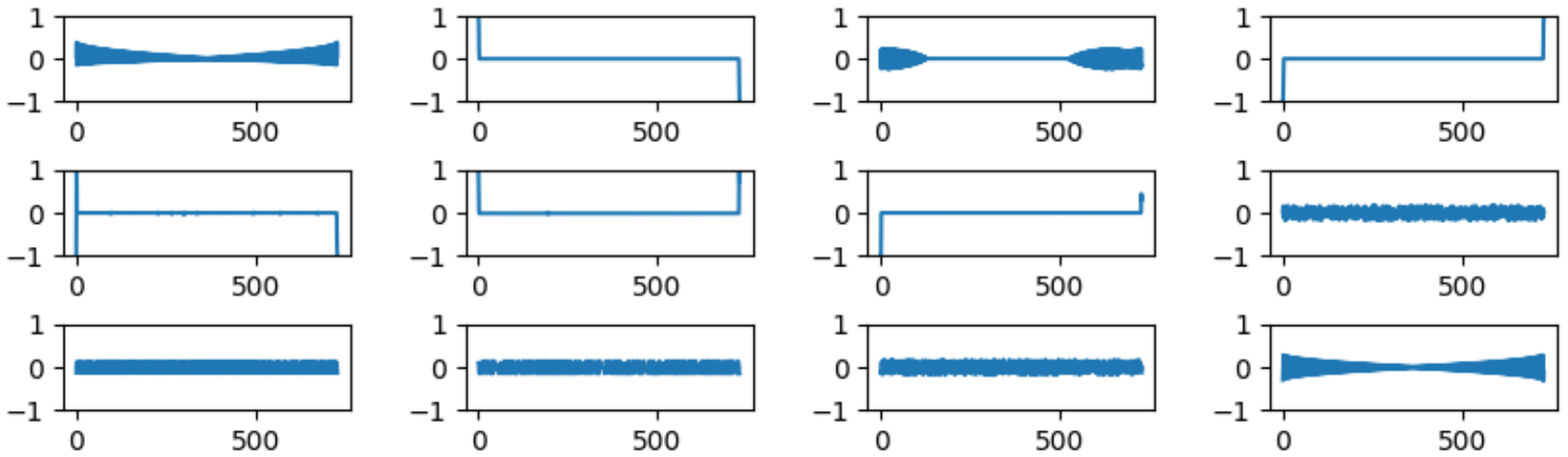
EX)

?

AM in 1D Model – 1st layer

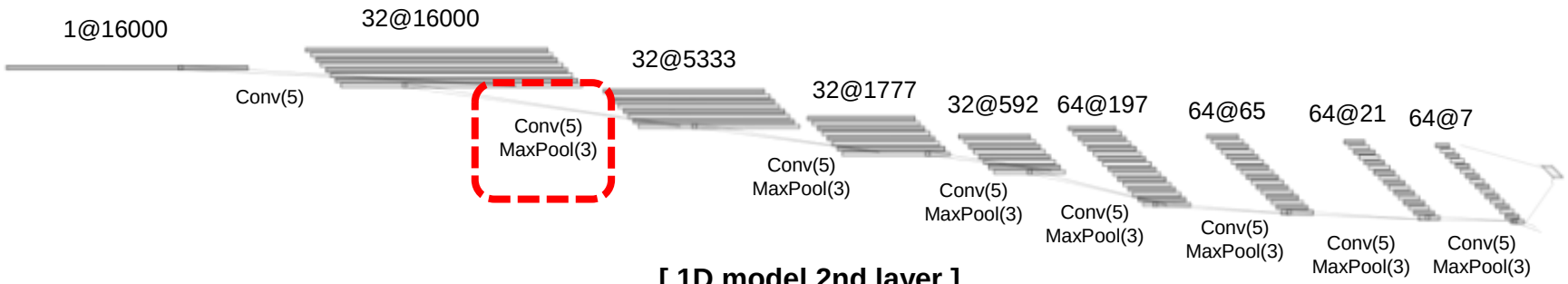


[1D model 1st layer]

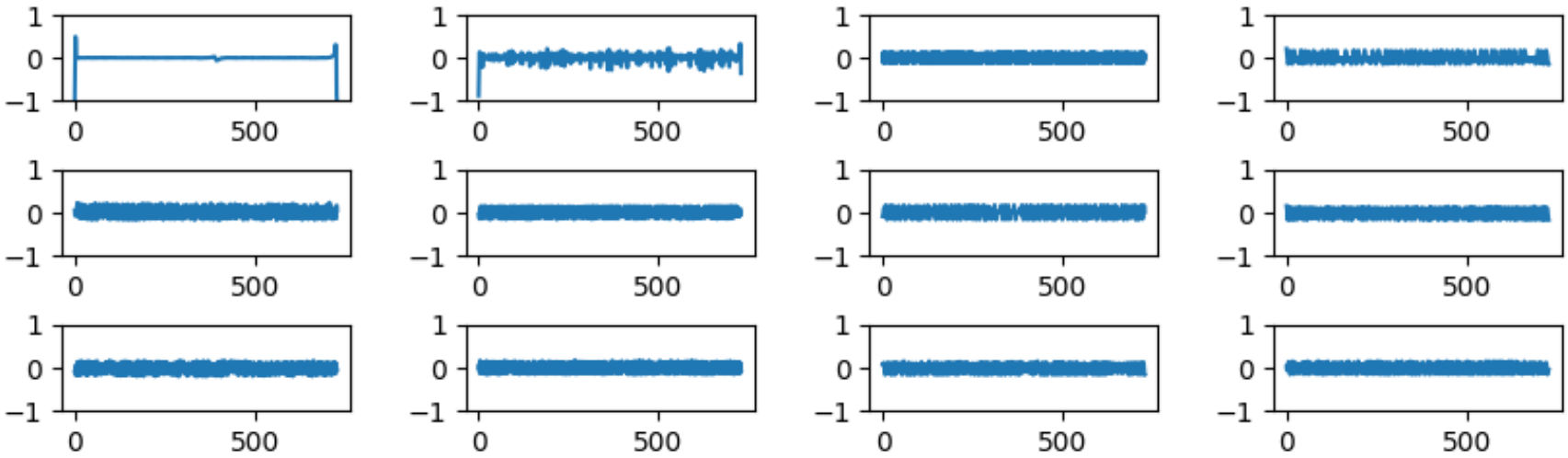


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 2nd layer

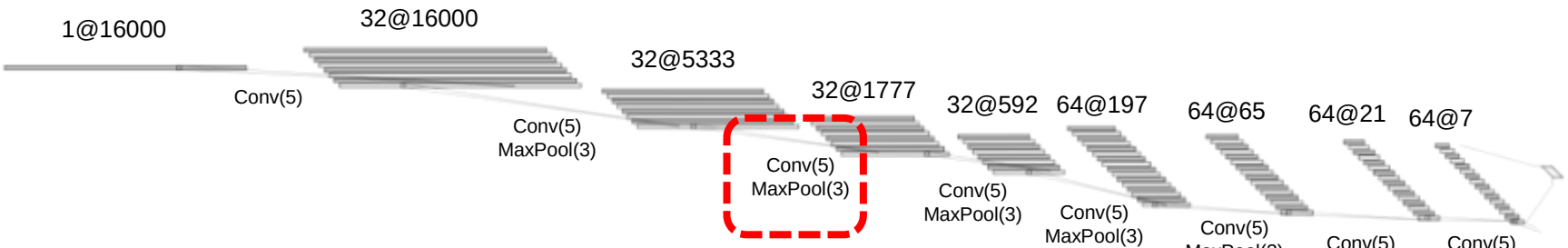


[1D model 2nd layer]

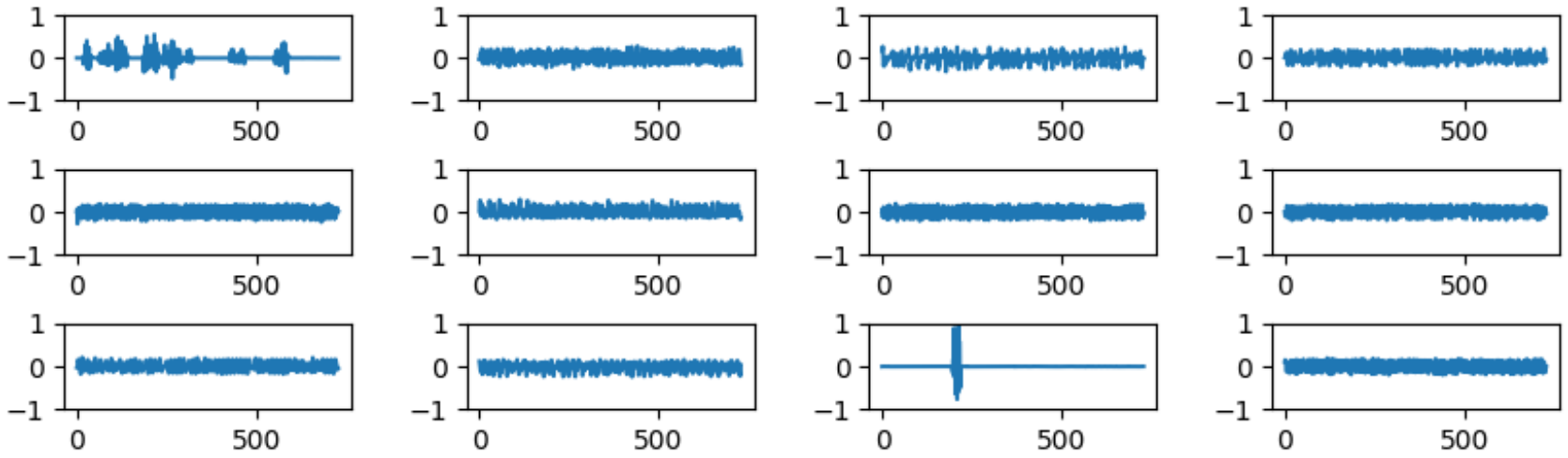


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 3rd layer

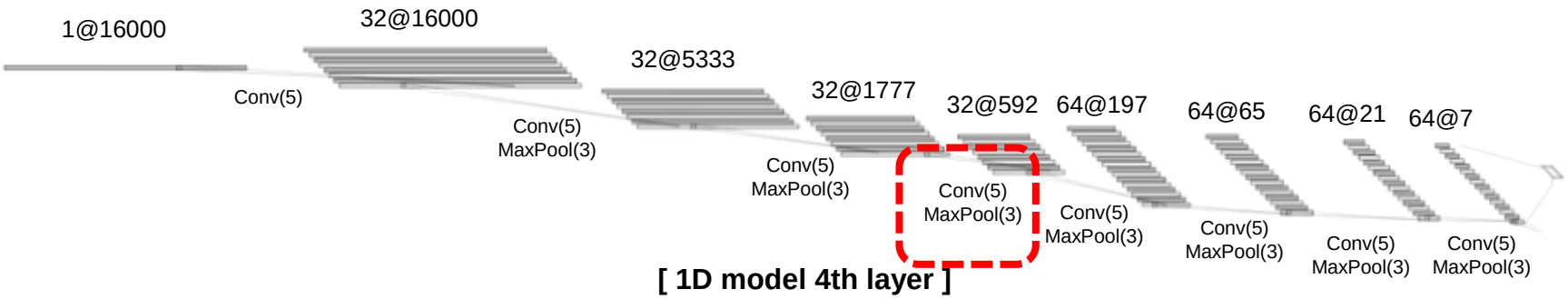


[1D model 3rd layer]

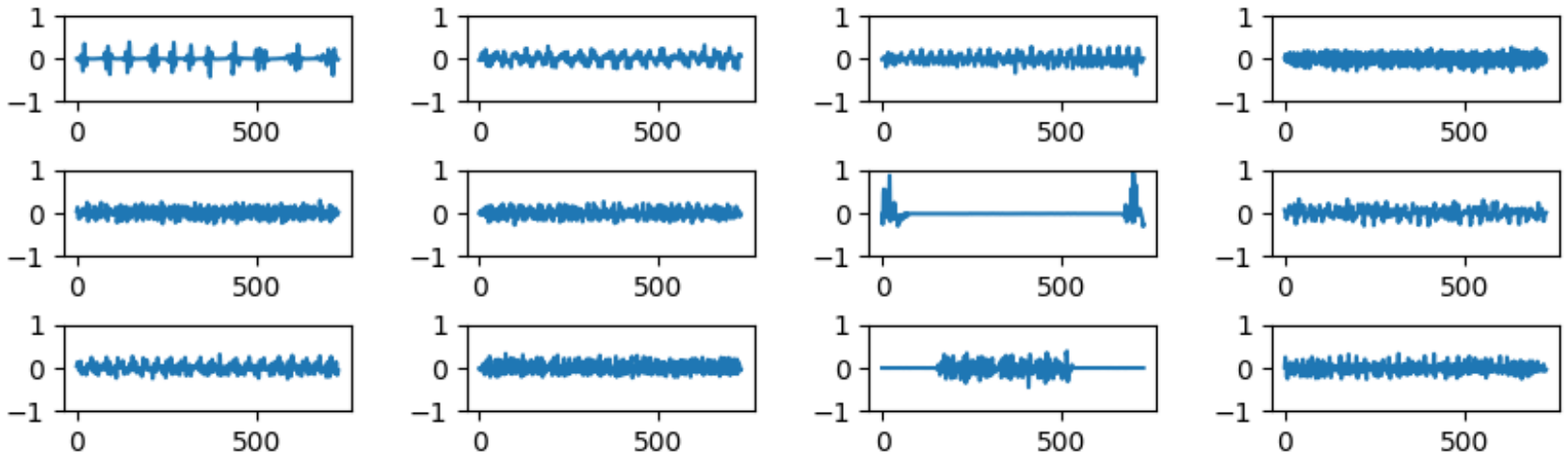


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 4th layer

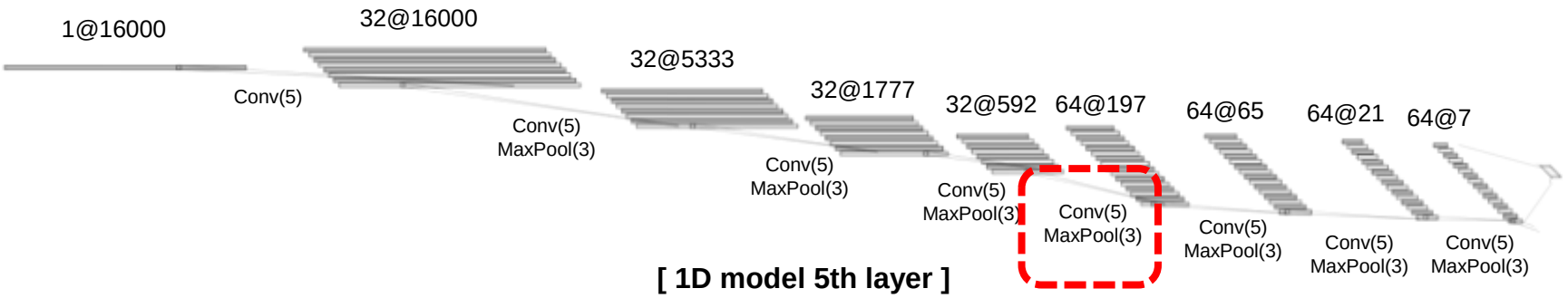


[1D model 4th layer]

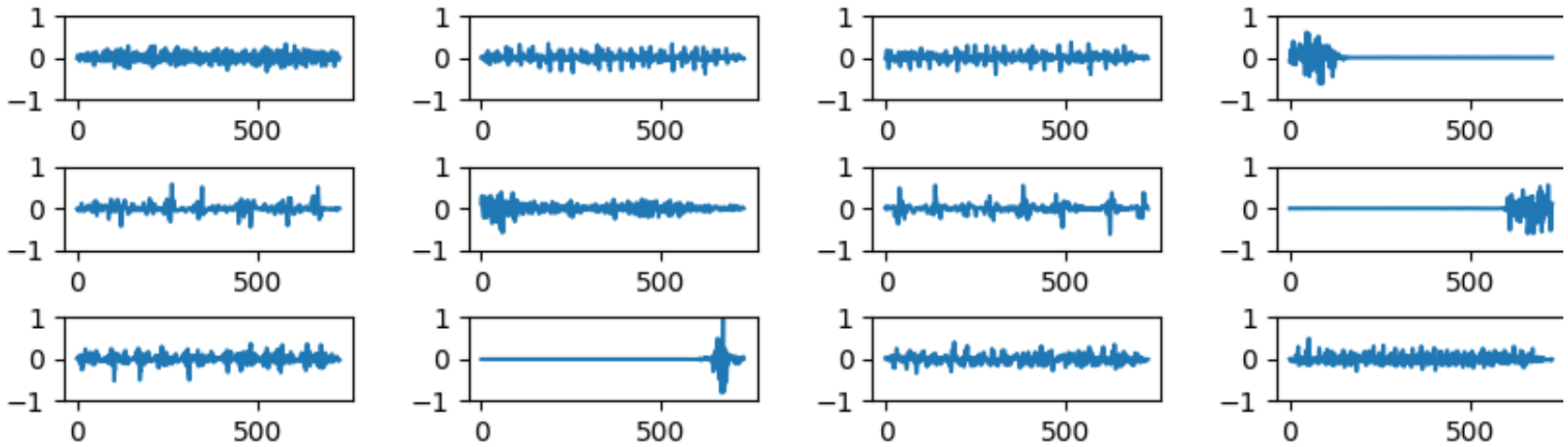


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 5th layer

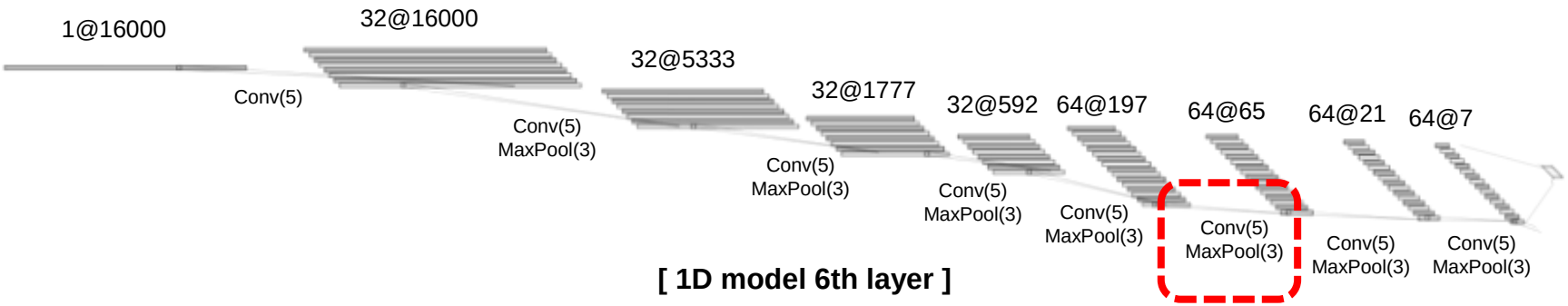


[1D model 5th layer]

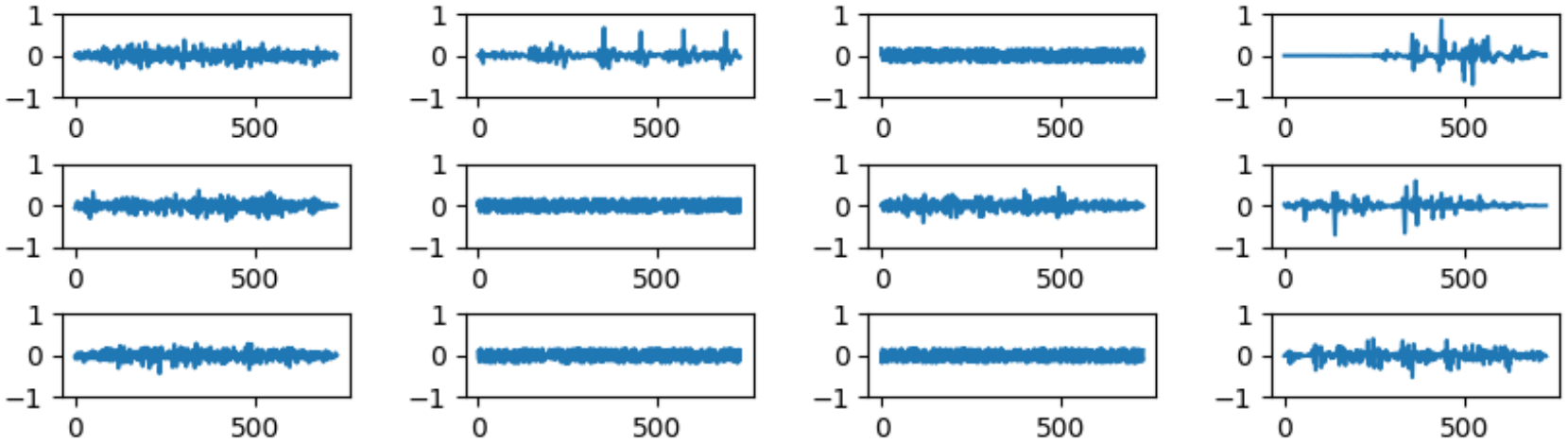


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 6th layer

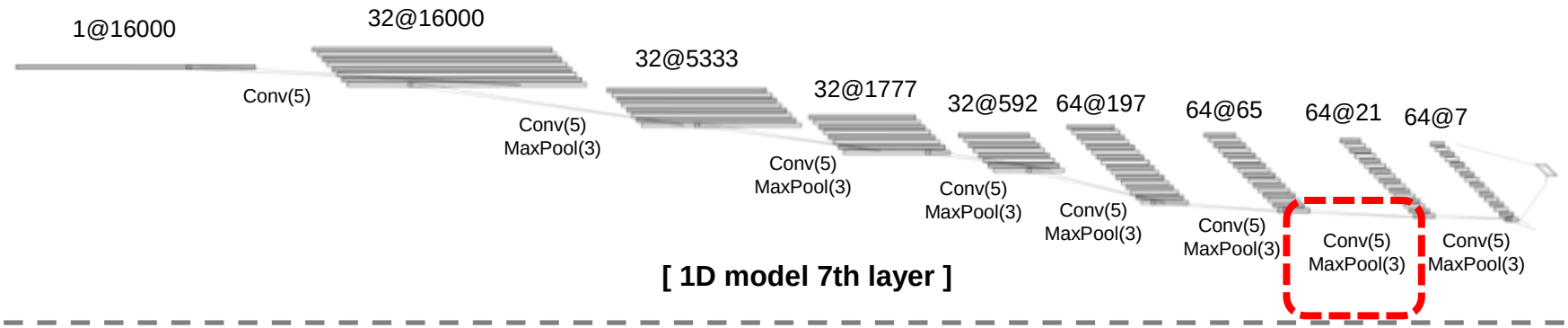


[1D model 6th layer]

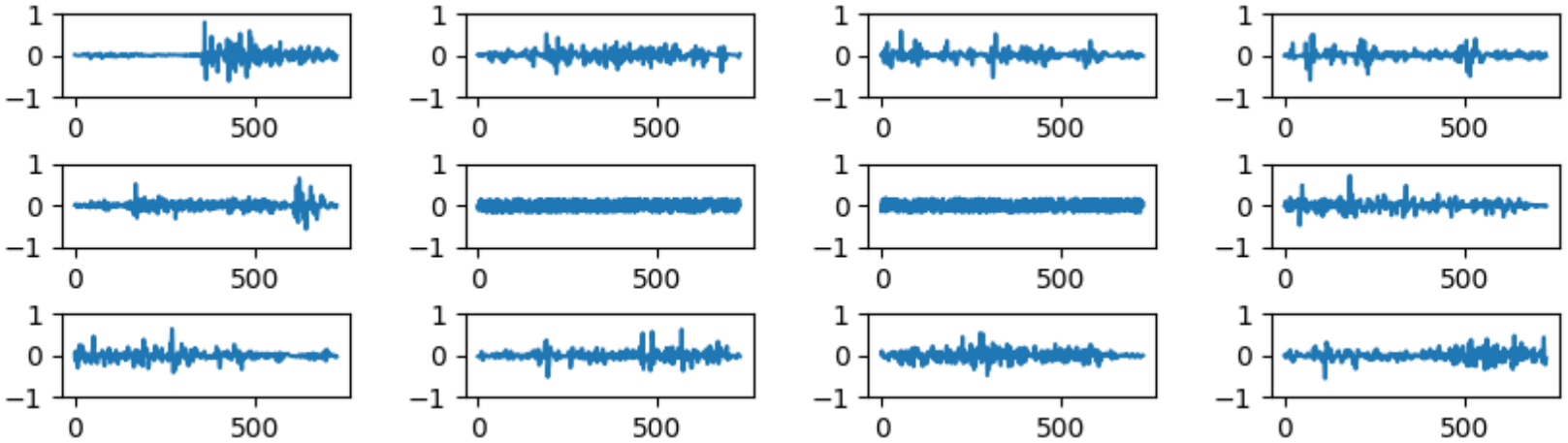


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 7th layer

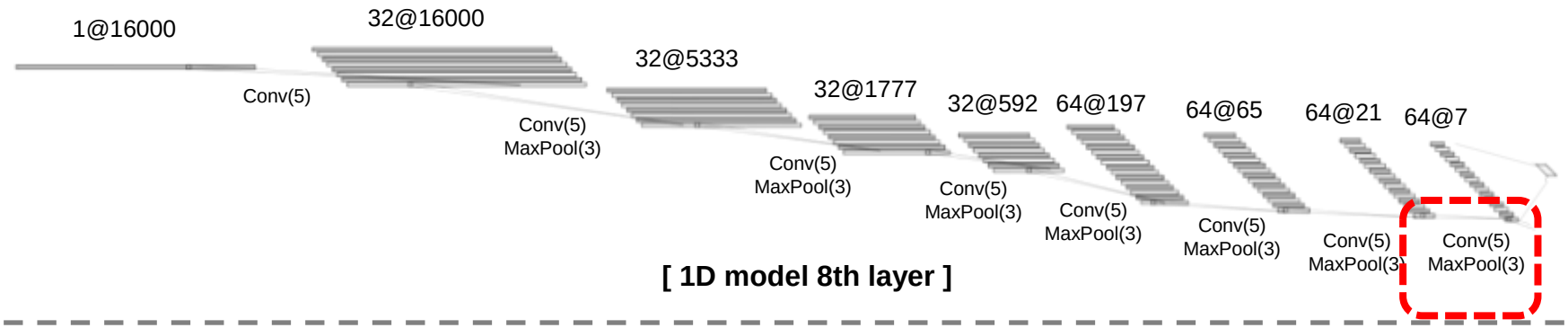


[1D model 7th layer]

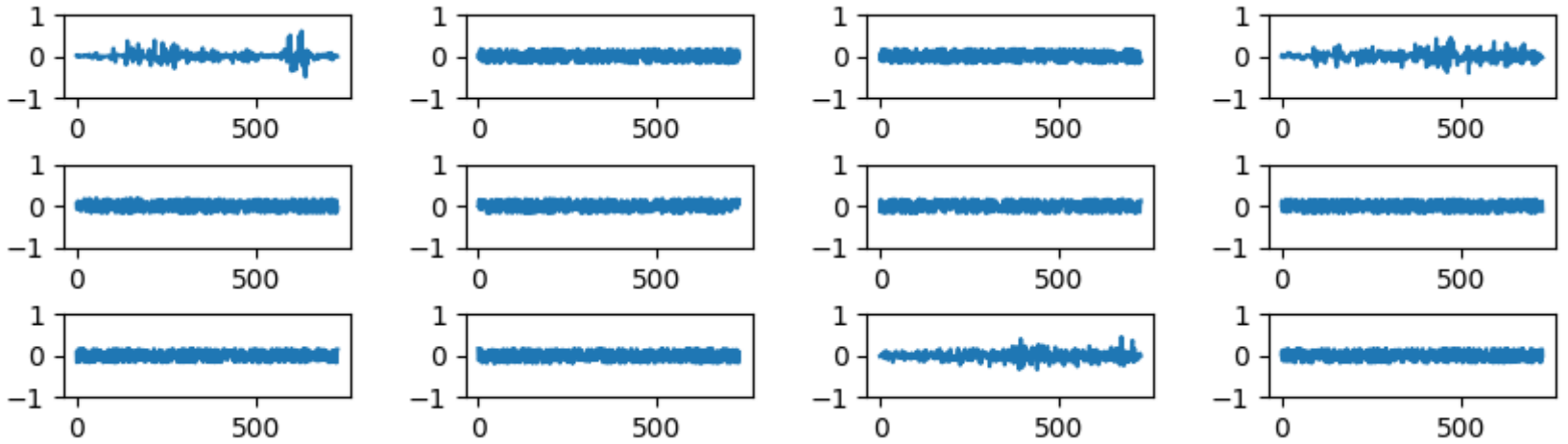


- As the layer becomes deeper, the shape of the waveform becomes more clear.

AM in 1D Model – 8th layer



[1D model 8th layer]



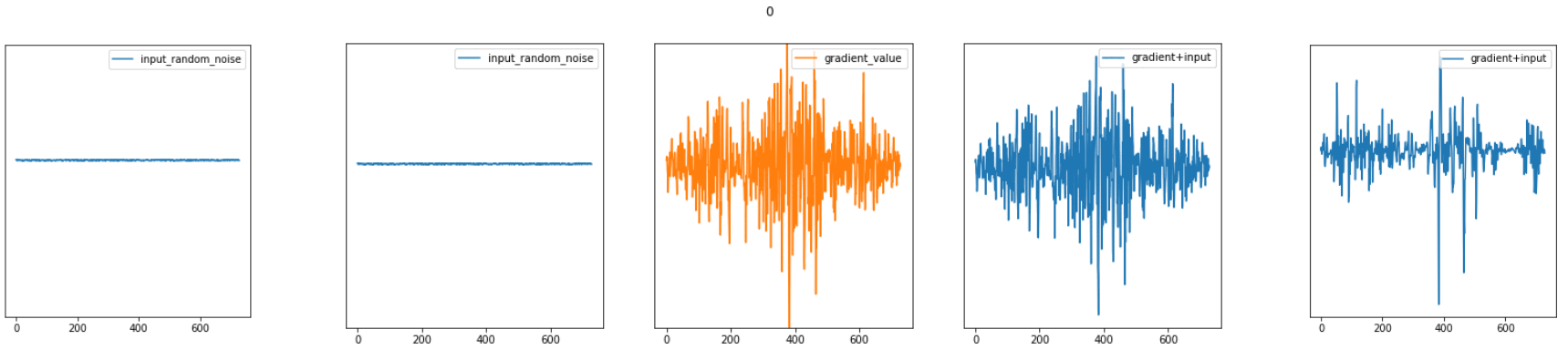
- A lot of waveform becomes flat again.
- Although each shape is similar, but the frequencies of each waveform are different.

AM in 1D Model - detail

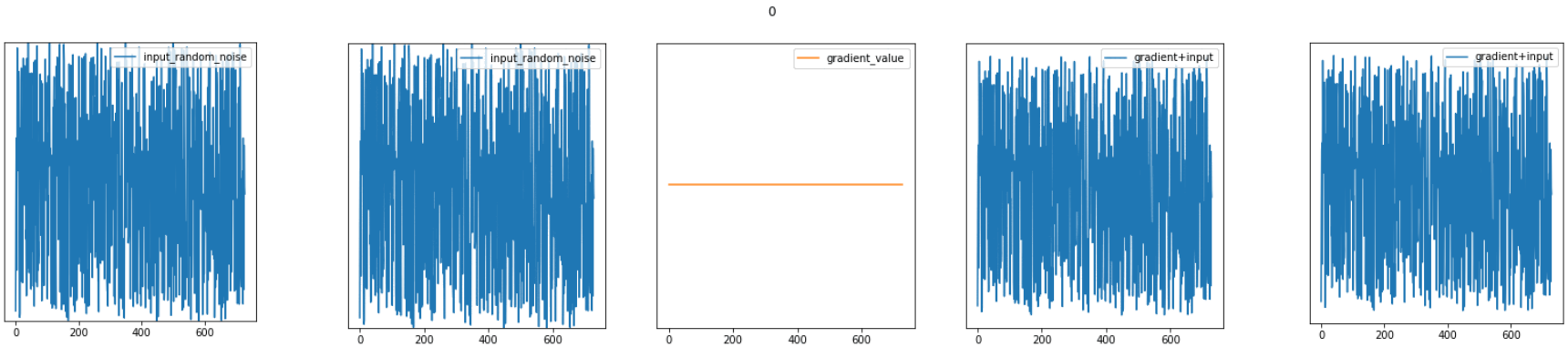
- Generate process of waveform
- Depending on the generated noise, the update may not be performed.

Start | Each iteration | End

1st Filter

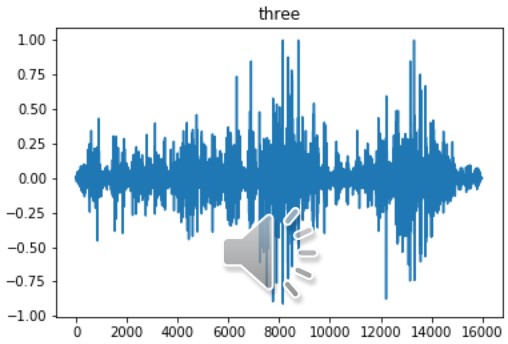


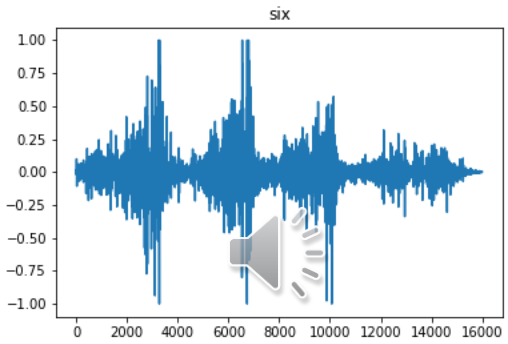
2nd Filter

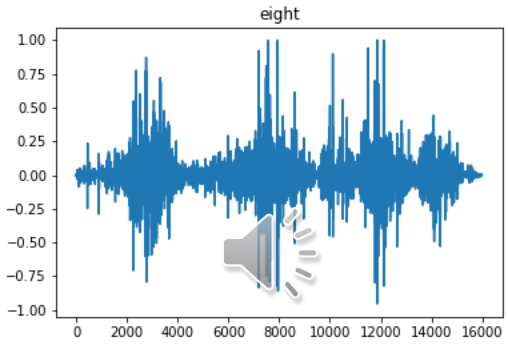


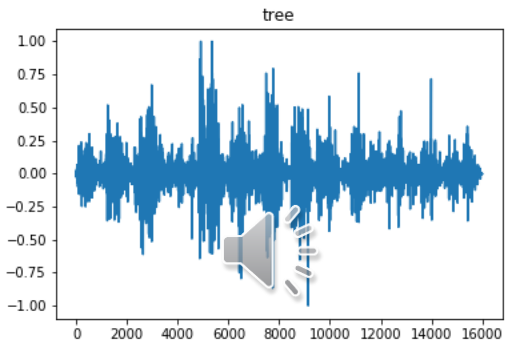
AM in 1D Model – output layer

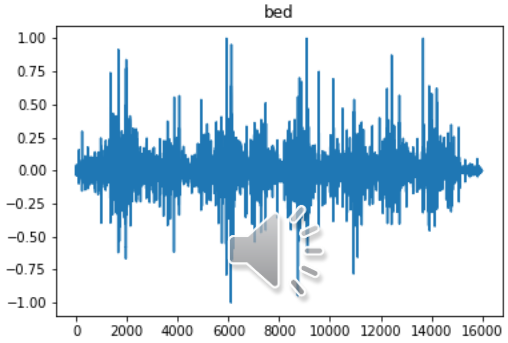
- It seems like real sound waveform
- What if we listen that?

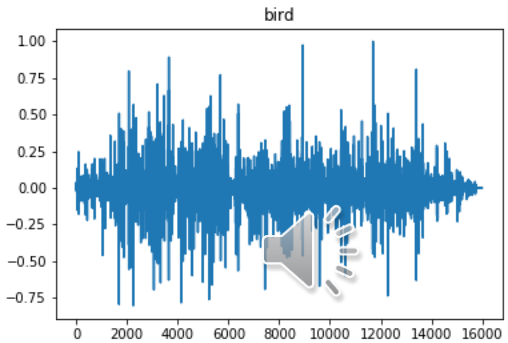






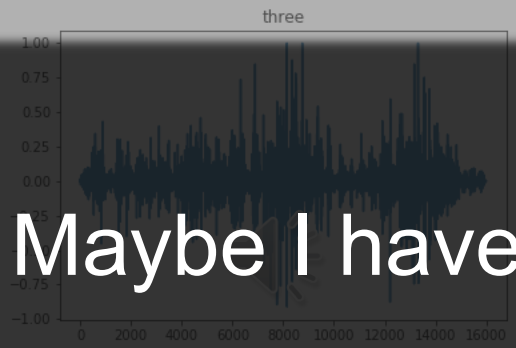


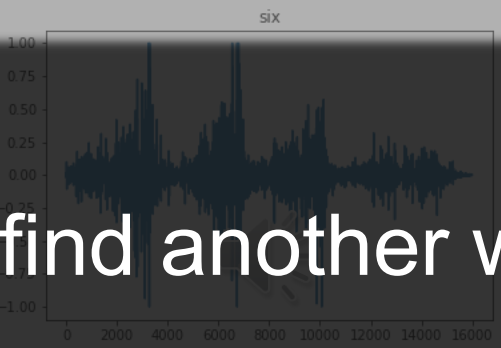


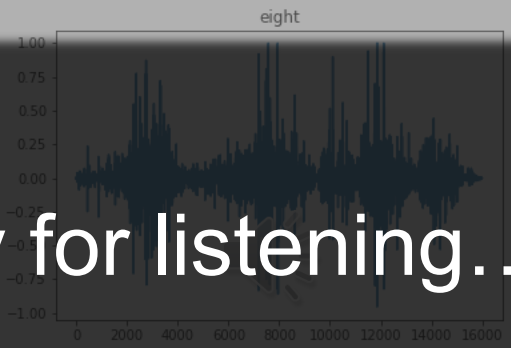


AM in 1D Model – output layer

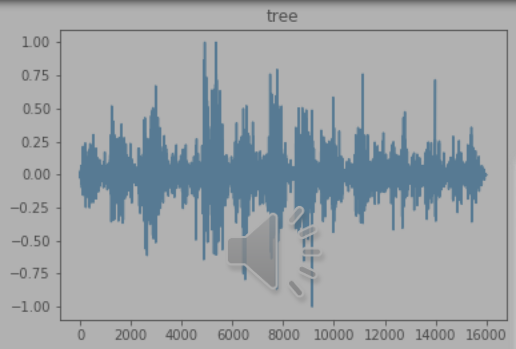
- It seems like real sound waveform
- What if we listen that?

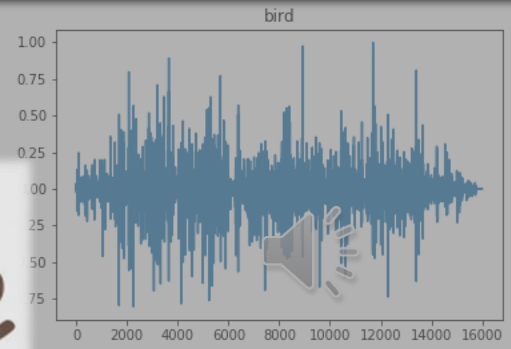






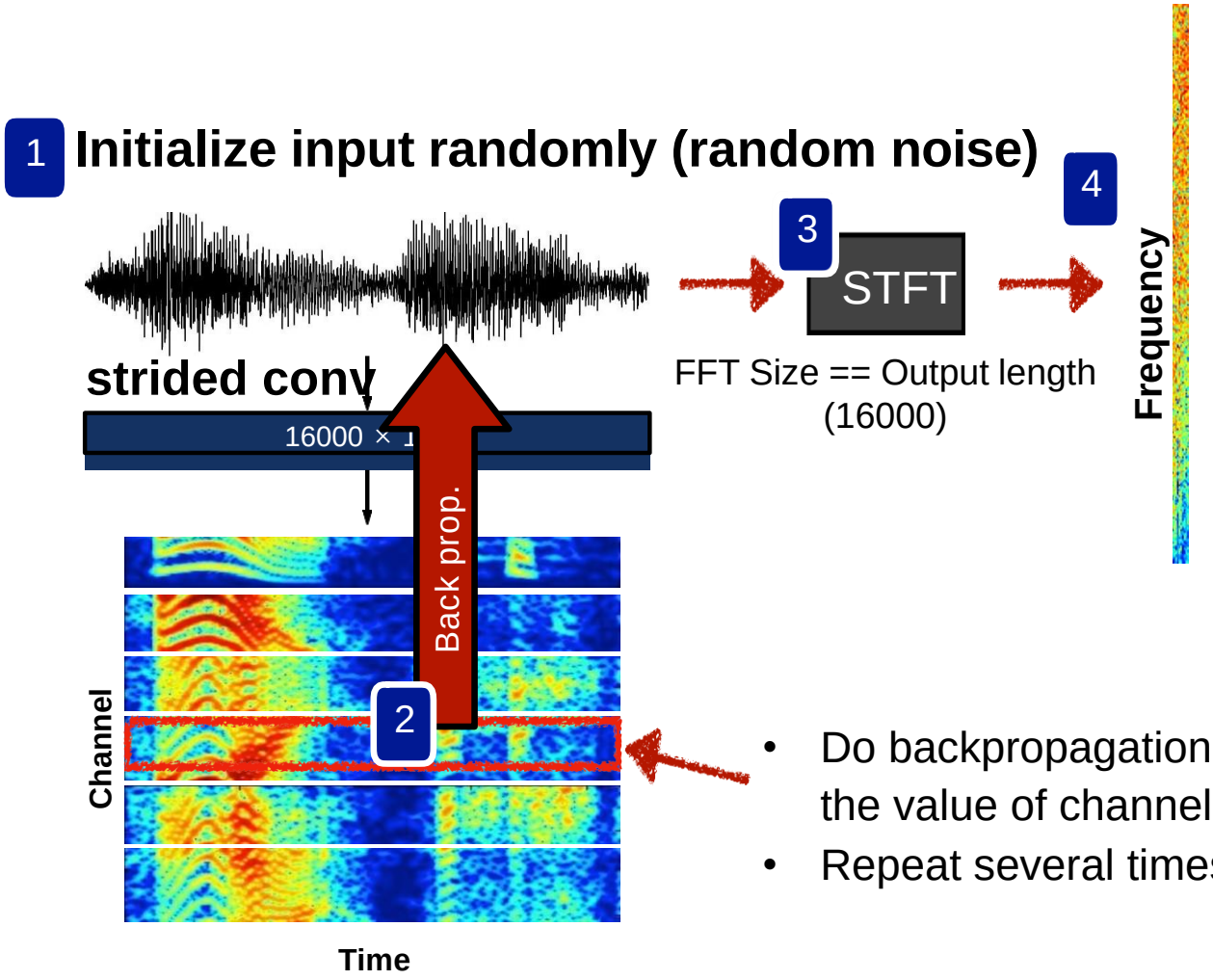
Maybe I have to find another way for listening…





AM in 1D Model – Visualize like SampleCNN

- Visualization process detail



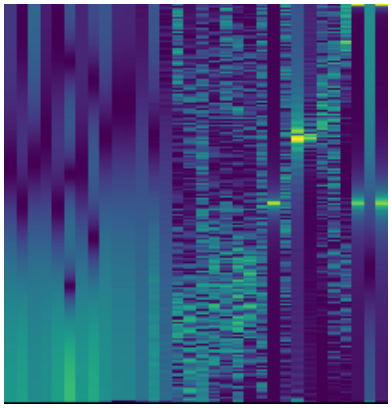
Do this for all the filters on all layers and combine them

- Do backpropagation such that the mean of the value of channel is maximum.
- Repeat several times to converge to input.

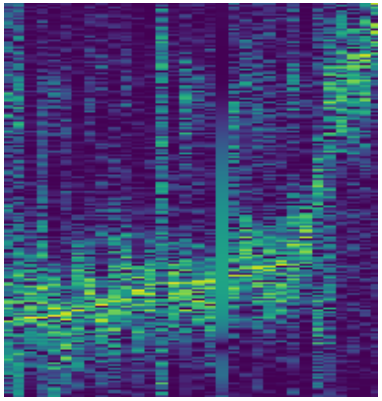
AM in 1D Model – Visualize like SampleCNN

- X axis: each filter (sorted by most highest frequency of each filer)
- Y axis: frequency

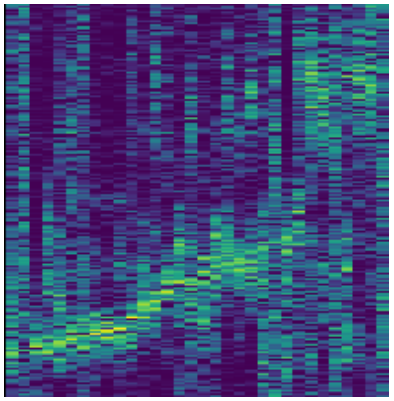
< 1st layer >



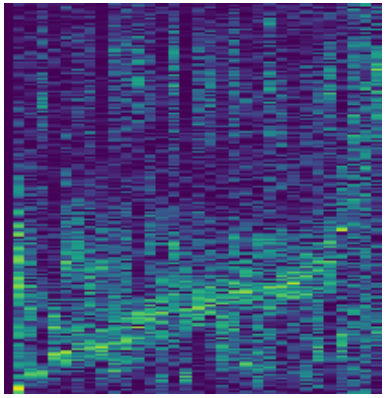
< 2nd layer >



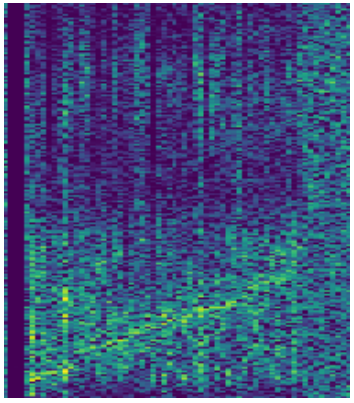
< 3rd layer >



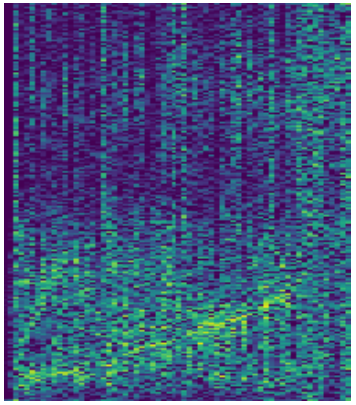
< 4th layer >



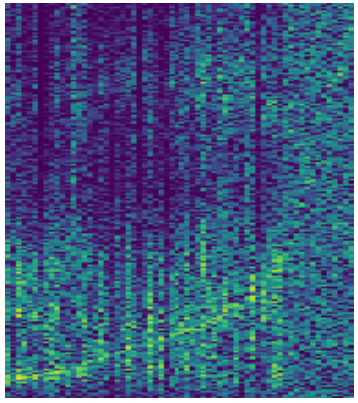
< 5th layer >



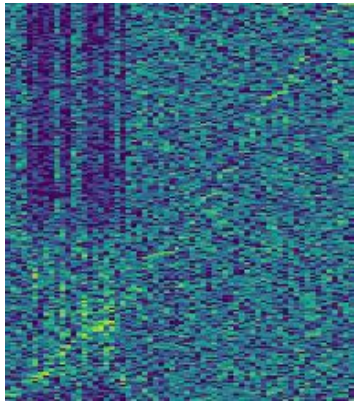
< 6th layer >



< 7th layer >



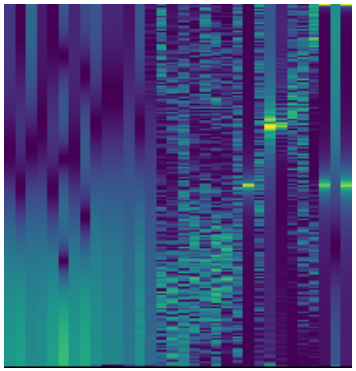
< 8th layer >



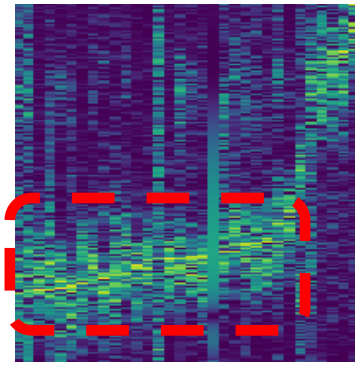
AM in 1D Model – Visualize like SampleCNN

- As the result, most of result focus on around 200Hz ~ 3400Hz
- In general, the frequency band between 300Hz and 3000Hz is called the voice frequency.
  - https://en.wikipedia.org/wiki/Voice_frequency
- Therefore, It means that the model focus on the voice frequency

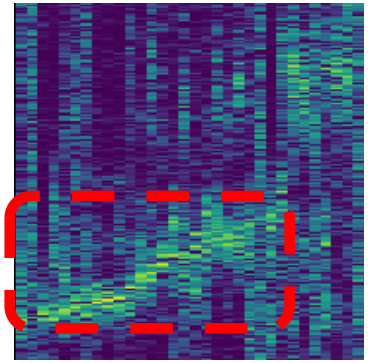
< 1st layer >



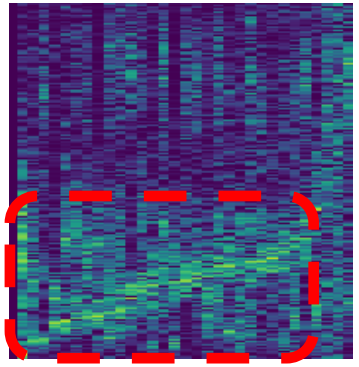
< 2nd layer >



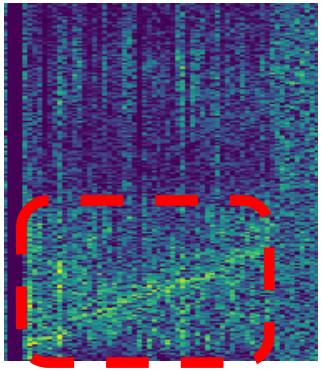
< 3rd layer >



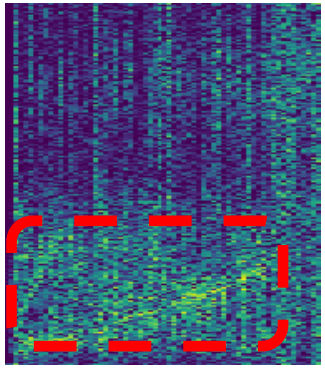
< 4th layer >



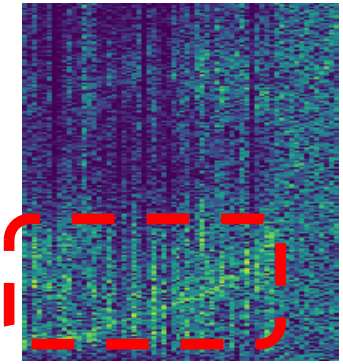
< 5th layer >



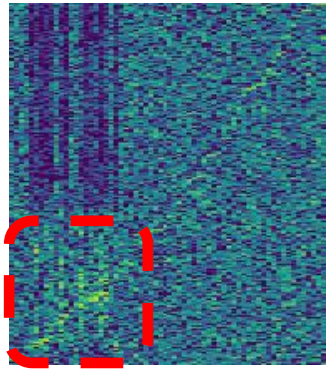
< 6th layer >



< 7th layer >



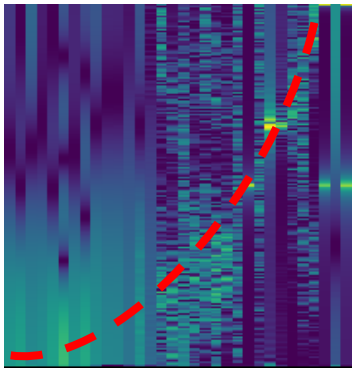
< 8th layer >



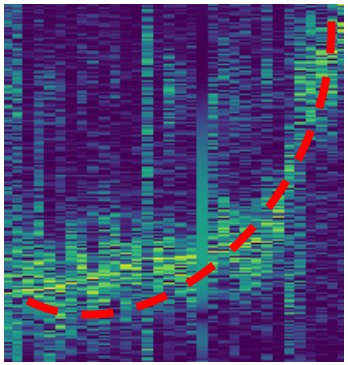
AM in 1D Model – Visualize like SampleCNN

- As we can see, each filter focus on each frequency.
- Therefore, It means that the channel is similar to the unsorted frequency

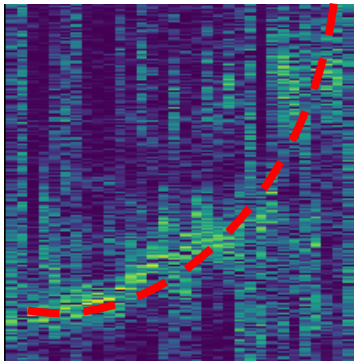
< 1st layer >



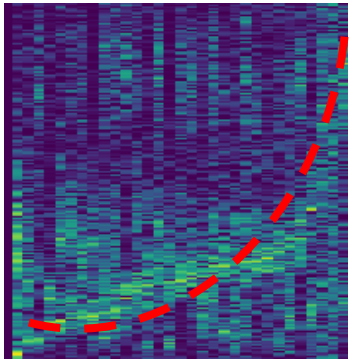
< 2nd layer >



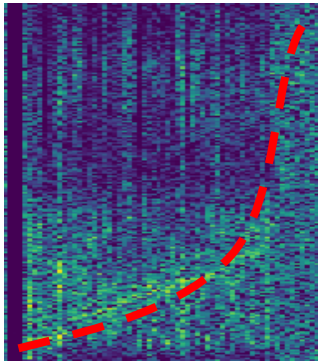
< 3rd layer >



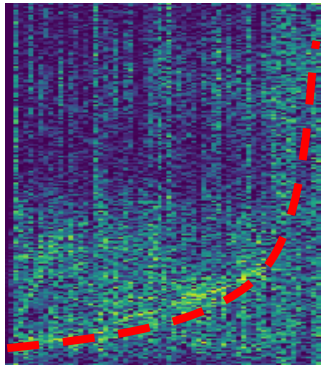
< 4th layer >



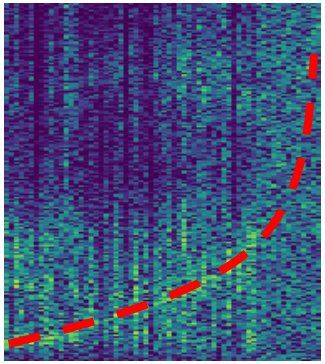
< 5th layer >



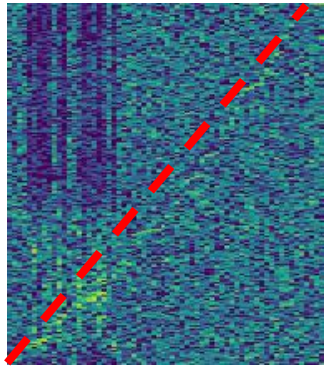
< 6th layer >



< 7th layer >



< 8th layer >



Any Question?

Thank you