project-9

June 11, 2021

Contents

Chapter 1

Introduction

In this project, we implement Convolutional and Pooling connections. In pooling layer, each neuron can only spike once. The input is a raw grayscale image encoded using `Time-to-First-Spike` encoder. In each iteration the filter of the convolution layer will be convolved over the input of that iteration, and the output will be added to the neurons' voltage in the that layer. Finally, the spikes of the conv-layer neurons (`simple cells`) will connect to pooling-layer neurons (`complex cells`).

In the first part we investigate the effect of neurons' thresholds in both convolution and pooling layers which represent `Simple cells` and `Complex cells` in the brain. Then in the second part, we investigate the effect of convolution and pooling parameters. (stride and kernel-size).

The filter we use is a gabor filter with two orientations (0° and 45°). Other parameters of the filter are as follows:

$FilterSize = 13 \times 13$ (default) , $\sigma = 3$, $\gamma = 1$, $\lambda = 2.5\pi$

Default neurons' parameters:

1. Convolution layer:
   - $u_{rest} = -60mv$
   - $threshold = -57mv$
   - $tau = inf$ (no decay)
   - $r = 1$
2. Pooling layer:
   - $u_{rest} = -60mv$
   - $threshold = -59mv$
   - $tau = inf$ (no decay)
   - $r = 1$

Default Convolution and Pooling params:

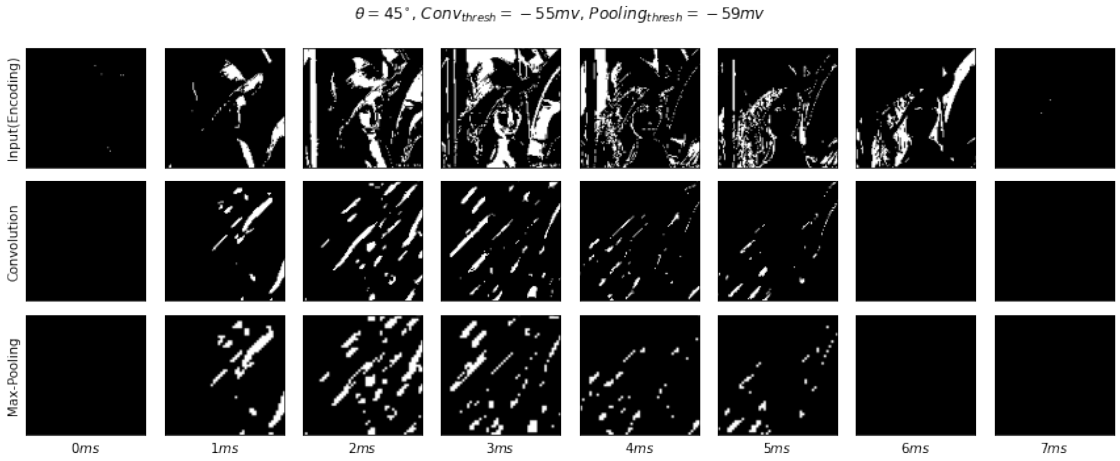
1. Convolution:
   - $Stride = 1 \times 1$
2. Pooling:
   - $KernelSize = 2 \times 2$
   - $Stride = 2 \times 2 (= KernelSize)$

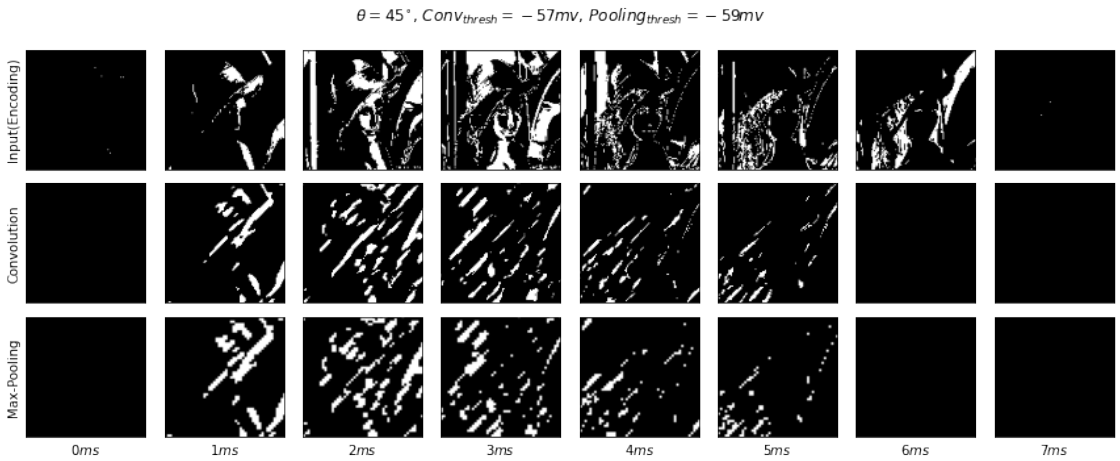
Note: In all plots we shift the convolution outputs by one step, and the pooling outputs by two steps to the left. The reason is that this way we can better see what each layer is doing.

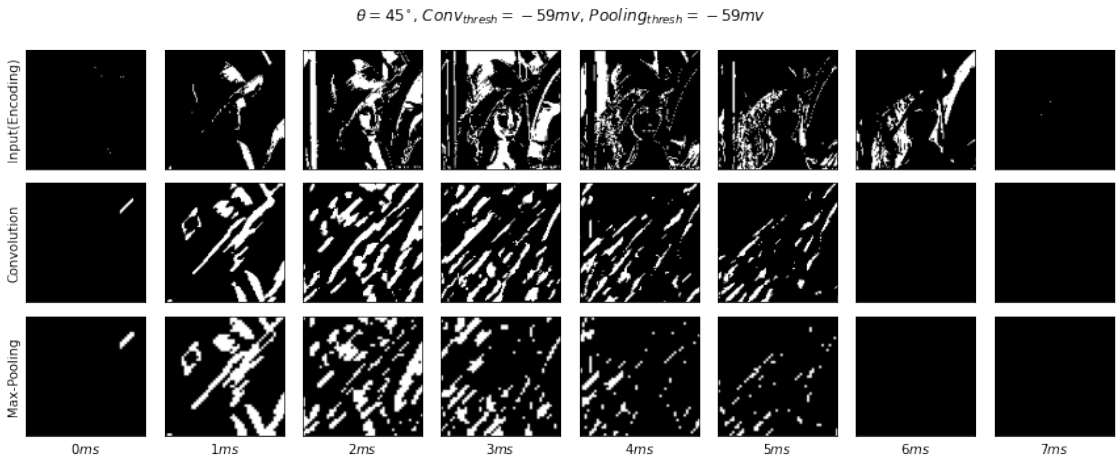
Chapter 2

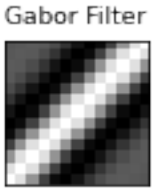
Part 1 (Neurons' thresholds)

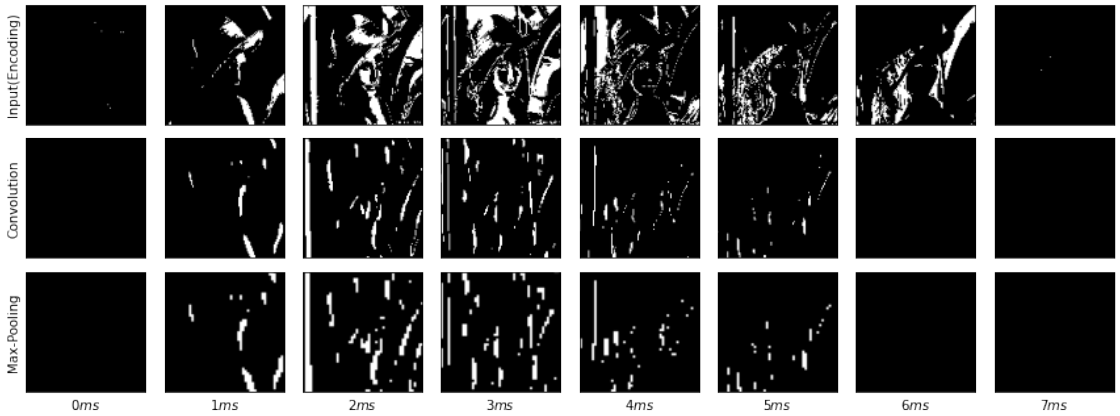
2.1 $Convolution_{threshold}$

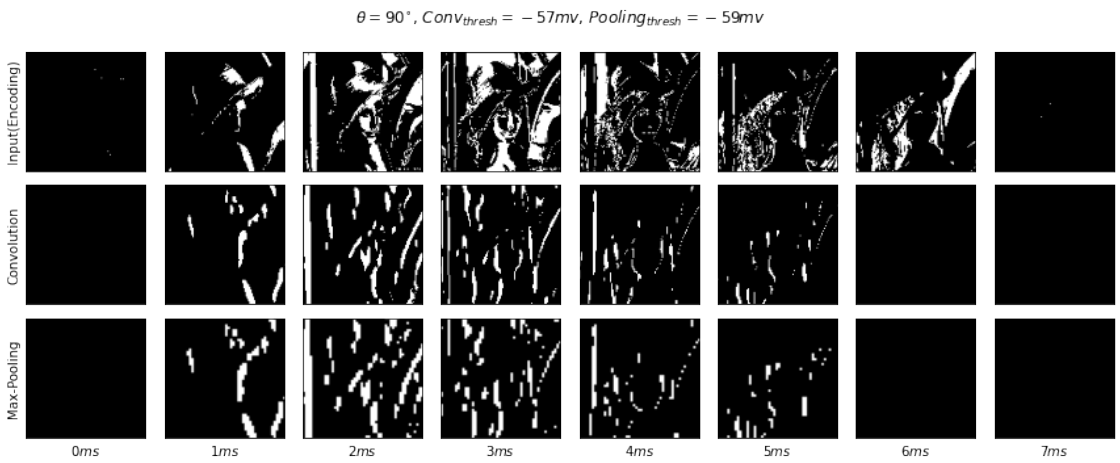


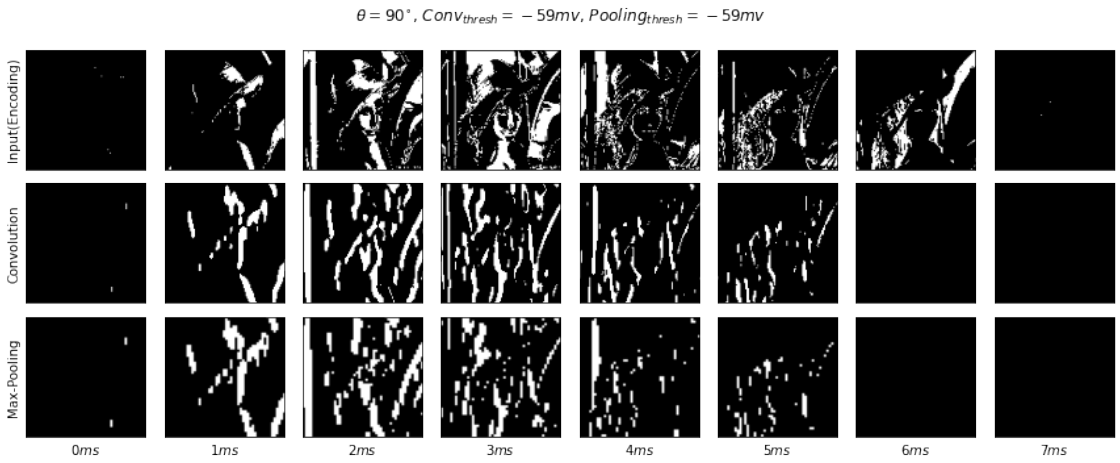



$\theta = 45^{\circ}, Conv_{thresh} = -59mv, Pooling_{thresh} = -59mv$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter

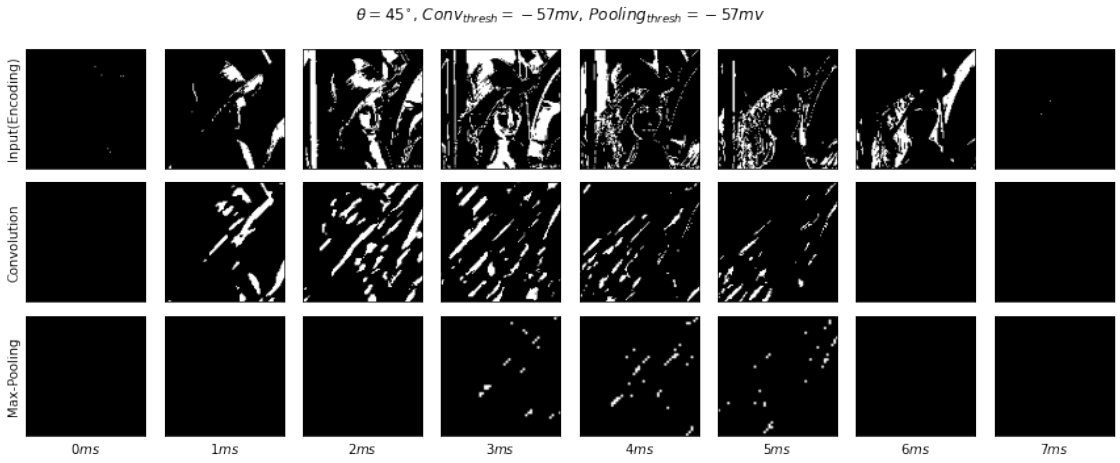

$\theta = 90^{\circ}, Conv_{thresh} = -55mv, Pooling_{thresh} = -59mv$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms

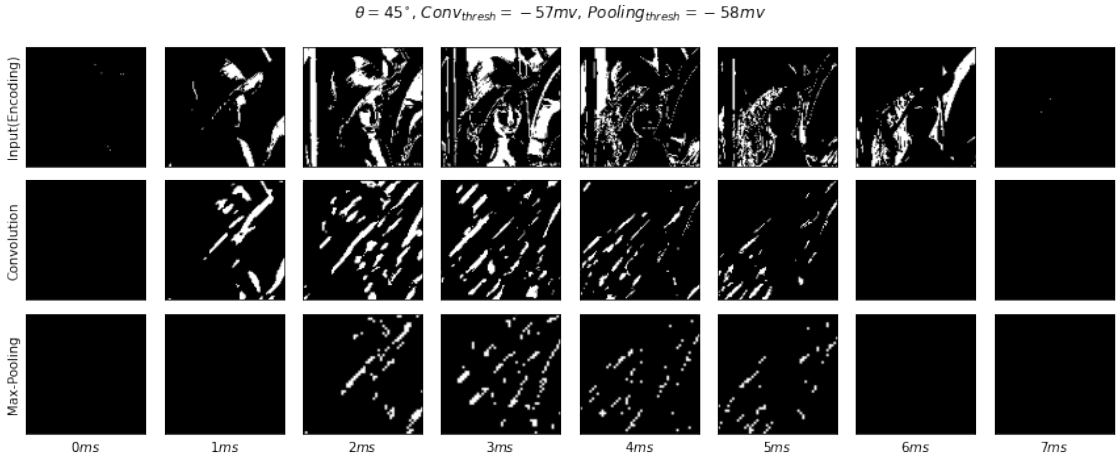


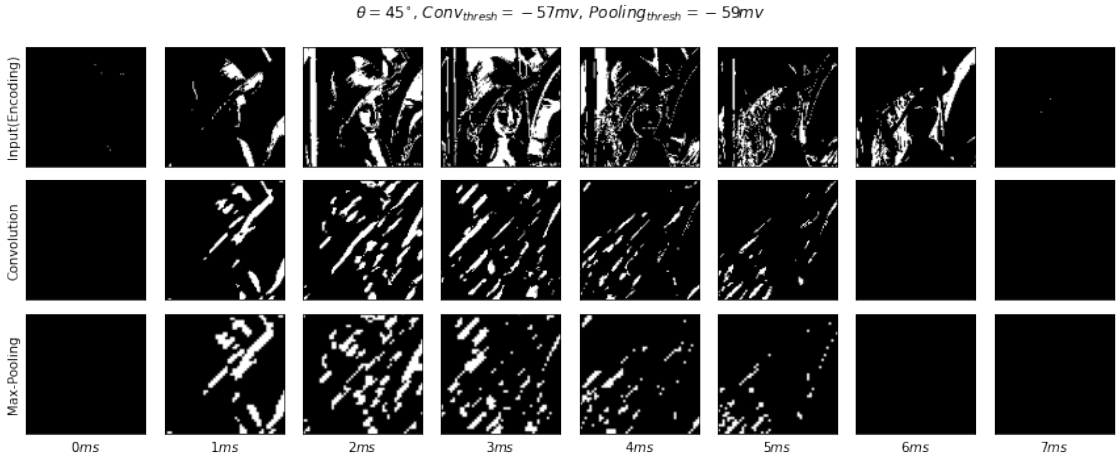


By lowering the threshold of convolution layer, some parts of the input that are not very similar to the filter also make the neurons spike. This is not desirable as it makes the output of the convolution layer prone to noise. However, if we set the threshold too high, only strong stimulus will activate the neurons, and we might lose some information in the early time steps.

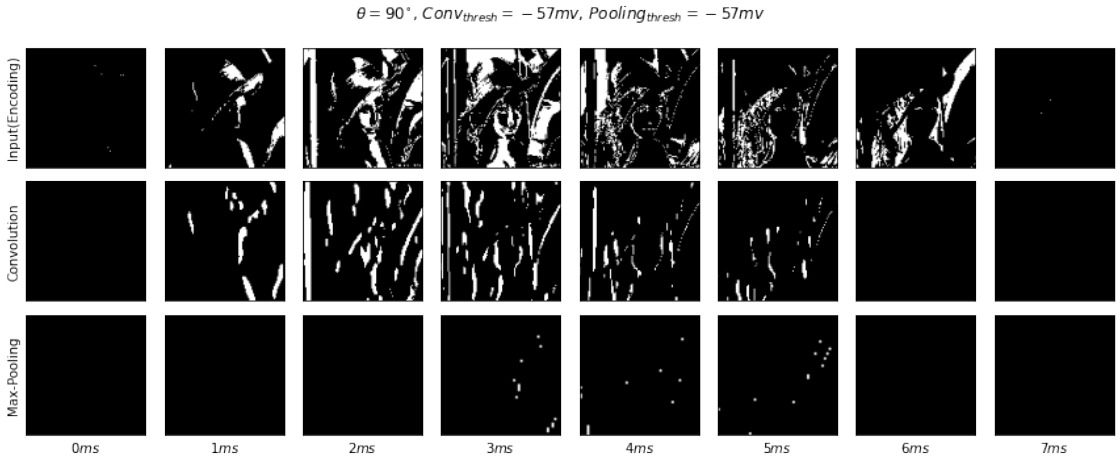
2.2 $Pooling_{threshold}$

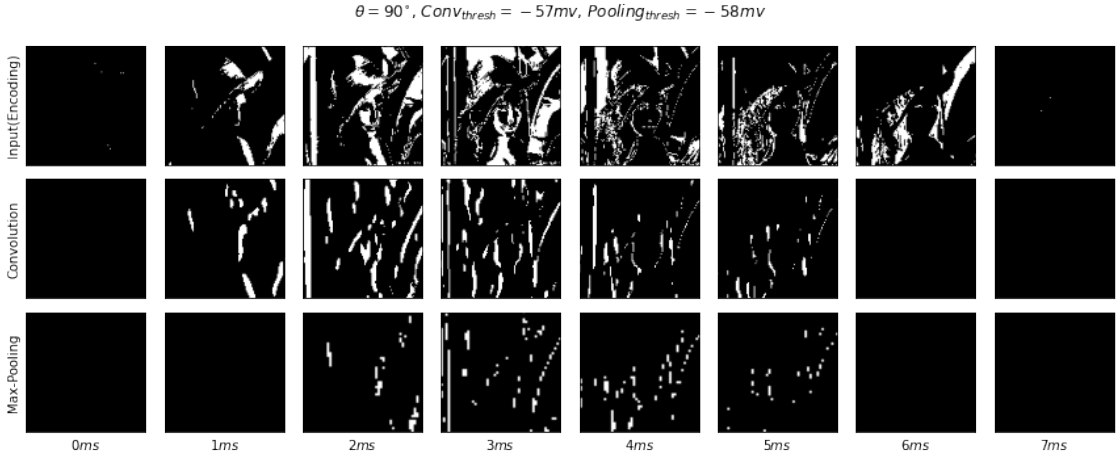


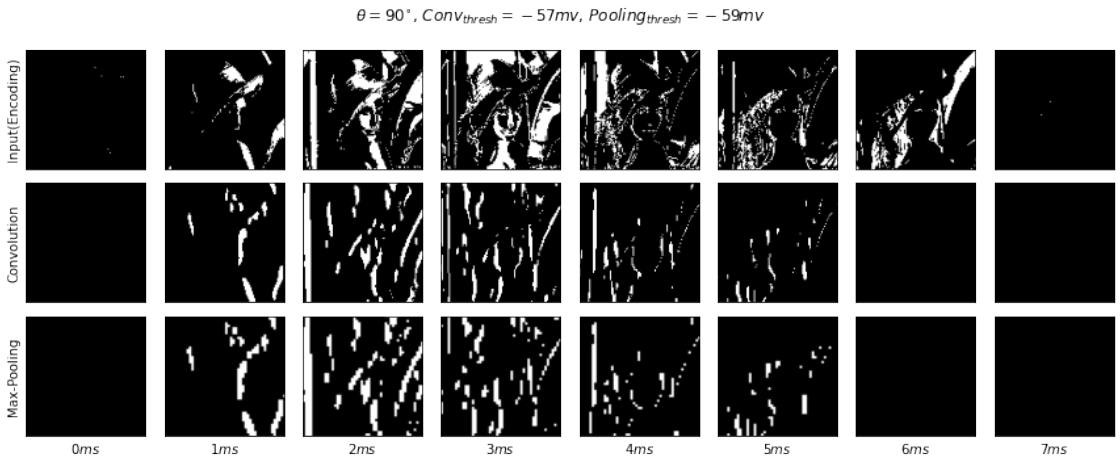





Gabor Filter

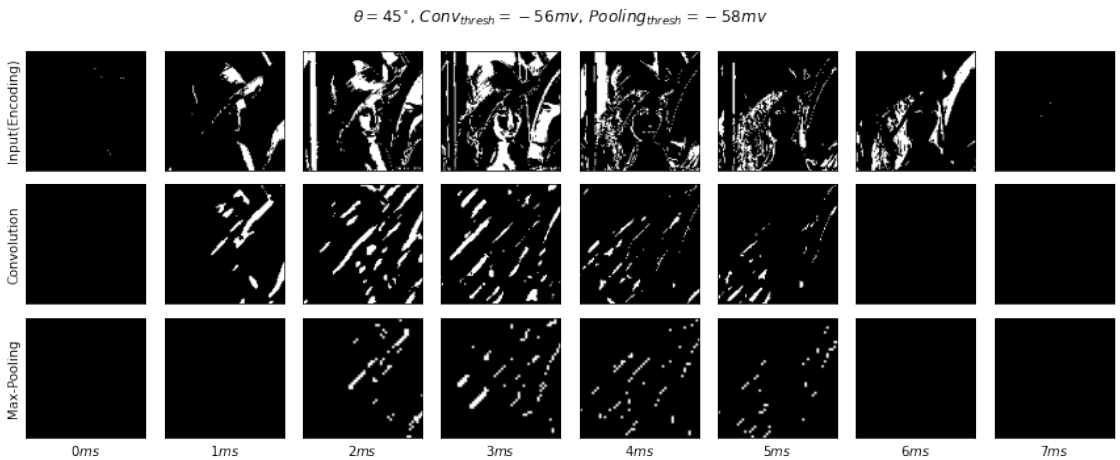

$\theta = 90^{\circ}, Conv_{thresh} = -57mv, Pooling_{thresh} = -57mv$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


$\theta = 90^{\circ}, Conv_{thresh} = -57mv, Pooling_{thresh} = -58mv$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms

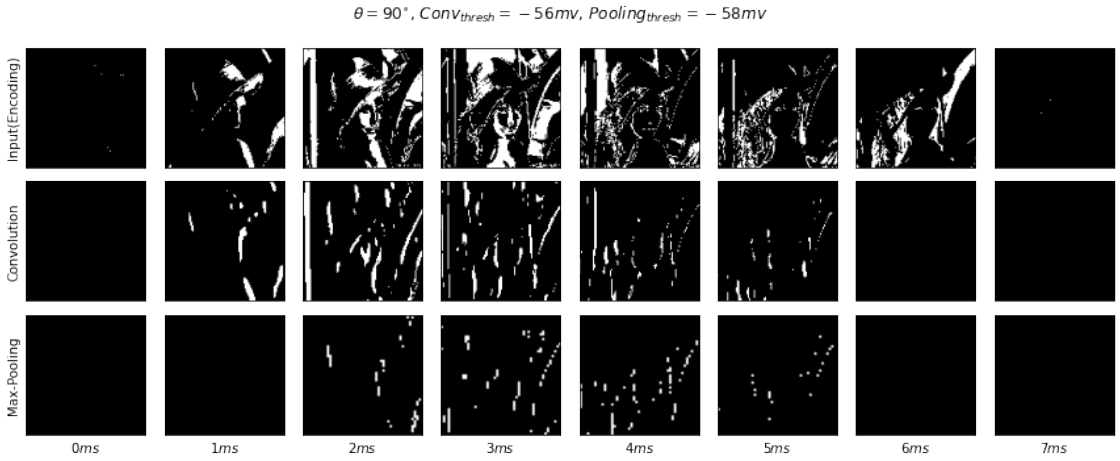


The effect of increasing threshold for Pooling neurons is not very different from the previous case. However, the input to the pooling neurons are only spikes (binary values). So, if we set the threshold to 2 units above u_{rest}, it means that 2 pre-synaptic neurons have to spike in order for the pooling neuron to spike. This way, much useful information will be lost. So, it's better to keep the pooling neurons' threshold to 1 unit above u_{rest} and change the convolution layer neurons' threshold. Another important thing to consider is that in the brain, complex cells that here we simulate them using max-pooling, will be activated when only one of the neurons in their receptive field spike. So, by setting the complex cell's threshold to 2, we are violating what is actually happening in the brain.

2.3 Both




Gabor Filter


$\theta = 90^{\circ}, Conv_{thresh} = -56mv, Pooling_{thresh} = -58mv$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter

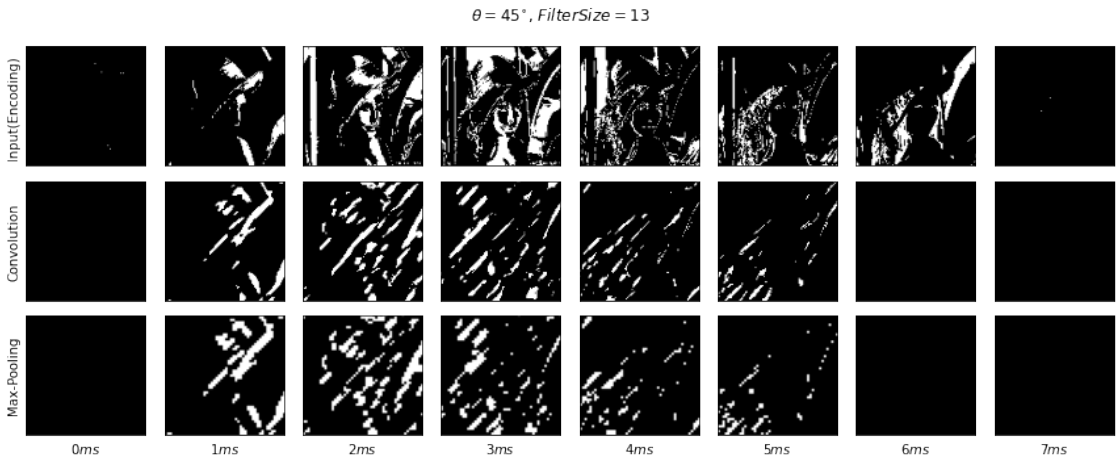
Chapter 3

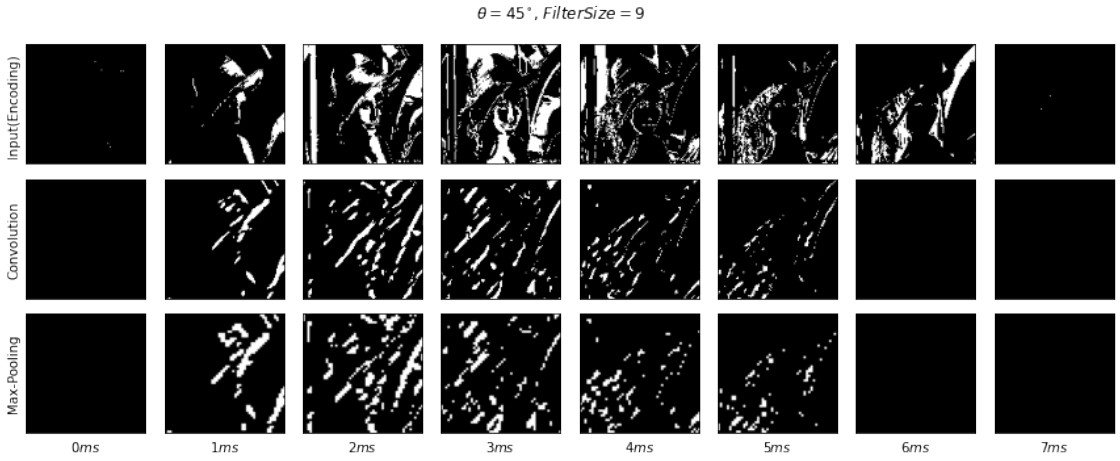
Part 2 (Convolution and Pooling Parameters)

3.1 Convolution Params

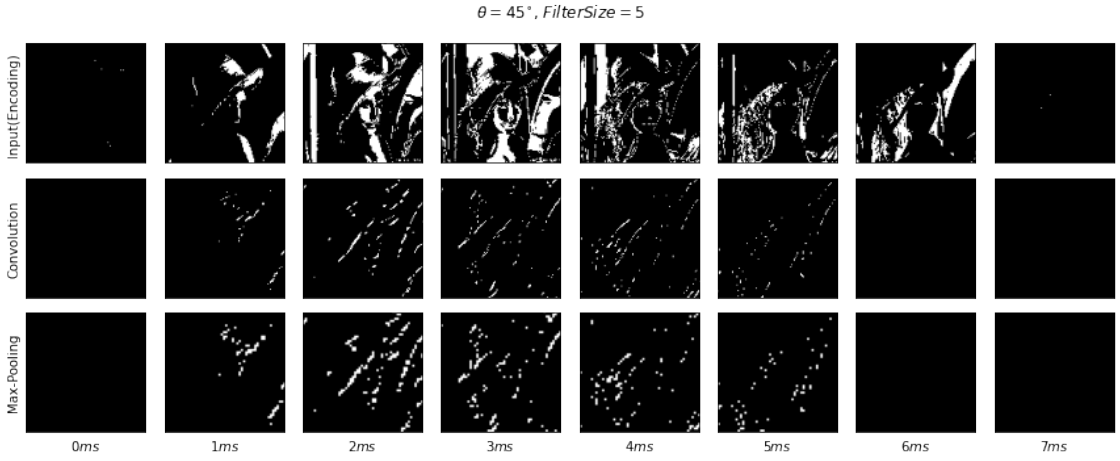
3.1.1 KernelSize

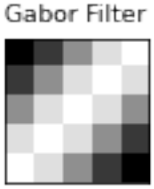
Note: When the receptive field is smaller, lower amount of activity accumulates in one convolution layer neuron, so the output will be more sparse. To reduce this effect we change $Convolution_{threshold}$ for $FilterSize = 5$ from $-57mv$ to $-58mv$.

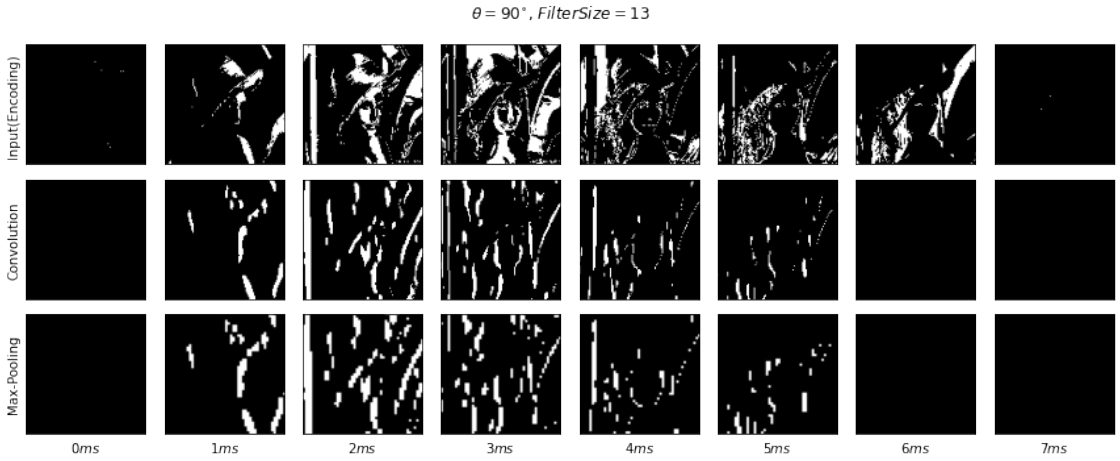



$\theta = 45^\circ, FilterSize = 9$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms

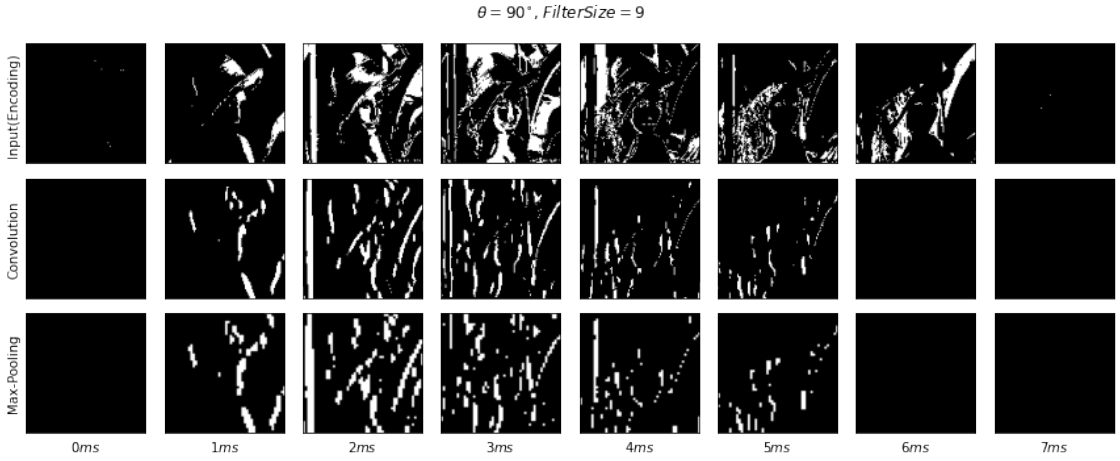

Gabor Filter

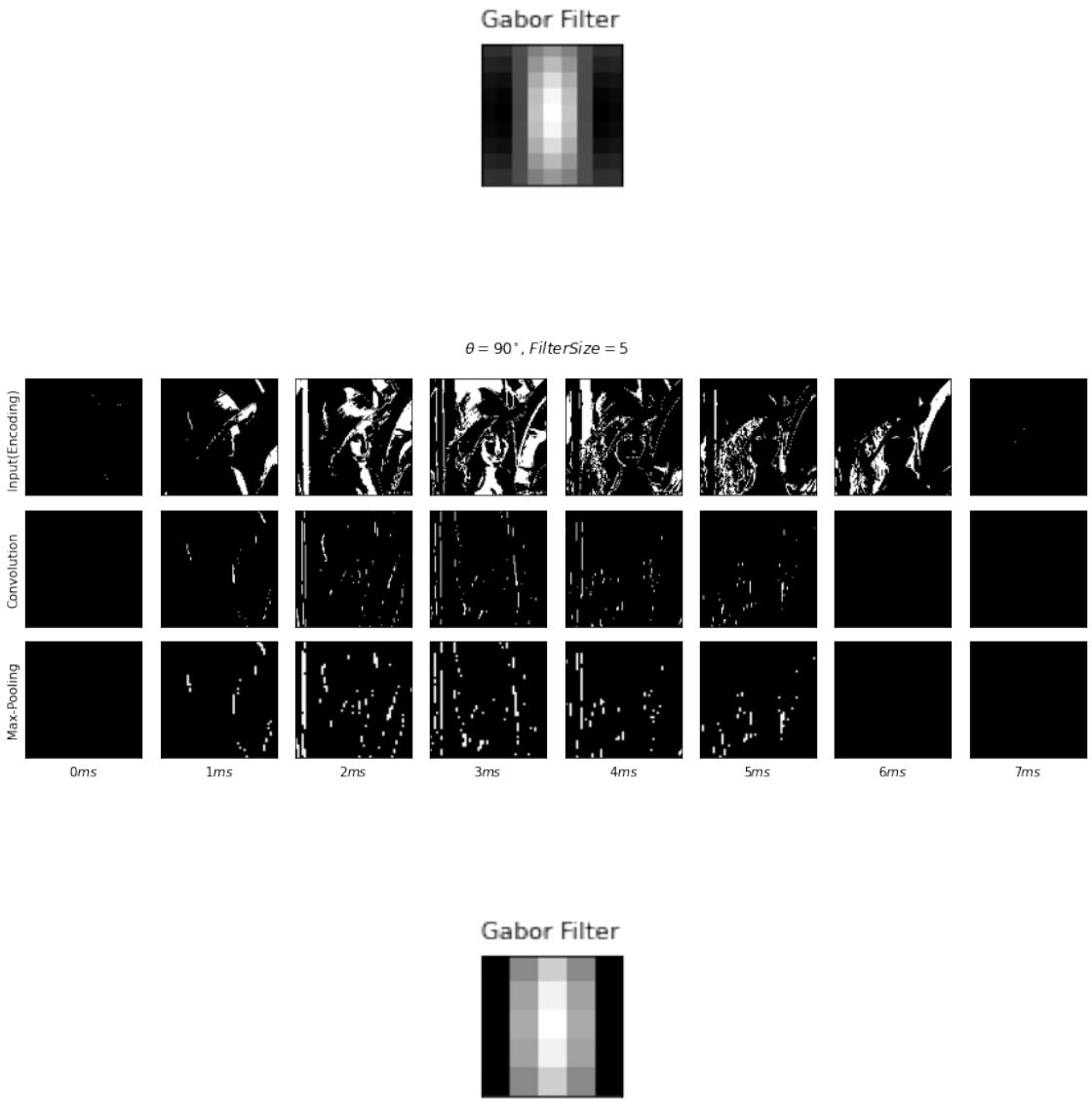

$\theta = 45^\circ, FilterSize = 5$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter


$\theta = 90^{\circ}, FilterSize = 13$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms

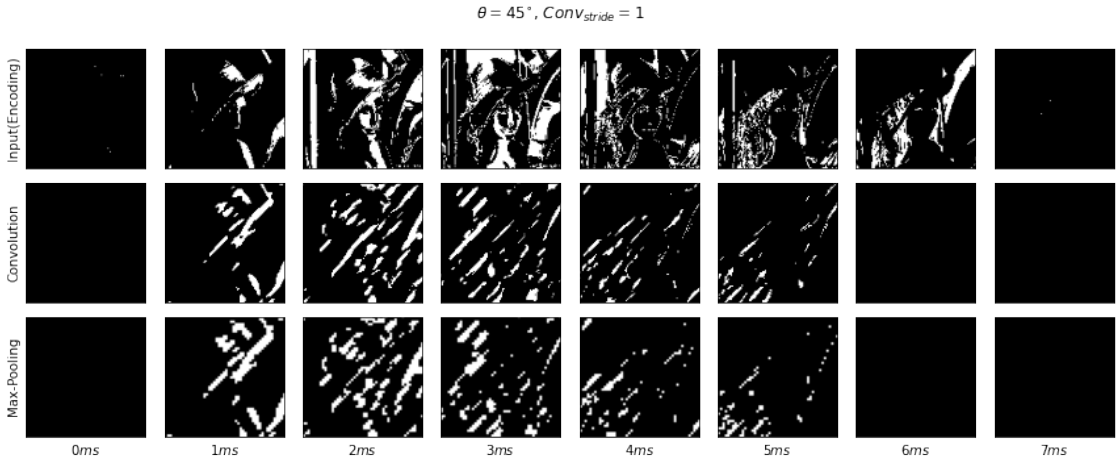

Gabor Filter

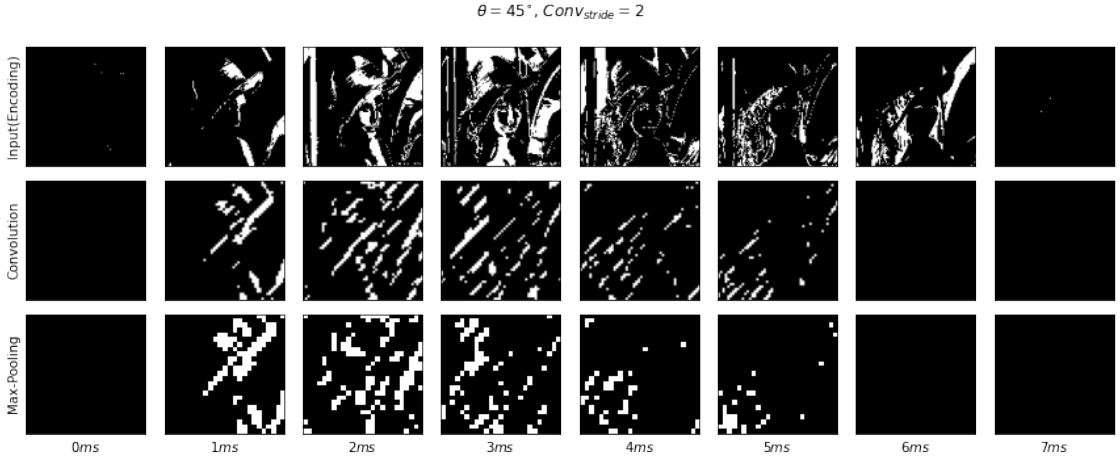

$\theta = 90^{\circ}, FilterSize = 9$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms

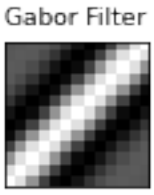


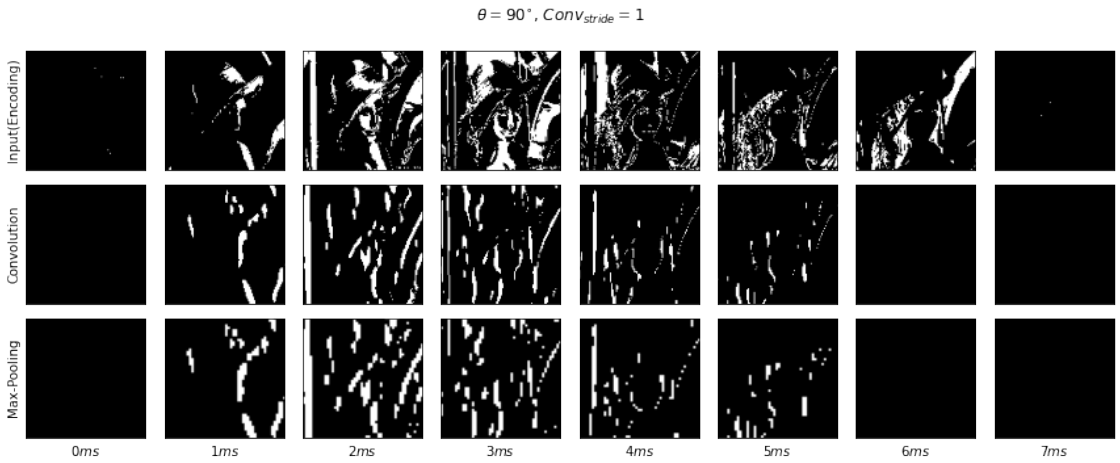
Smaller kernels detect smaller lines and larger kernels detect larger lines.

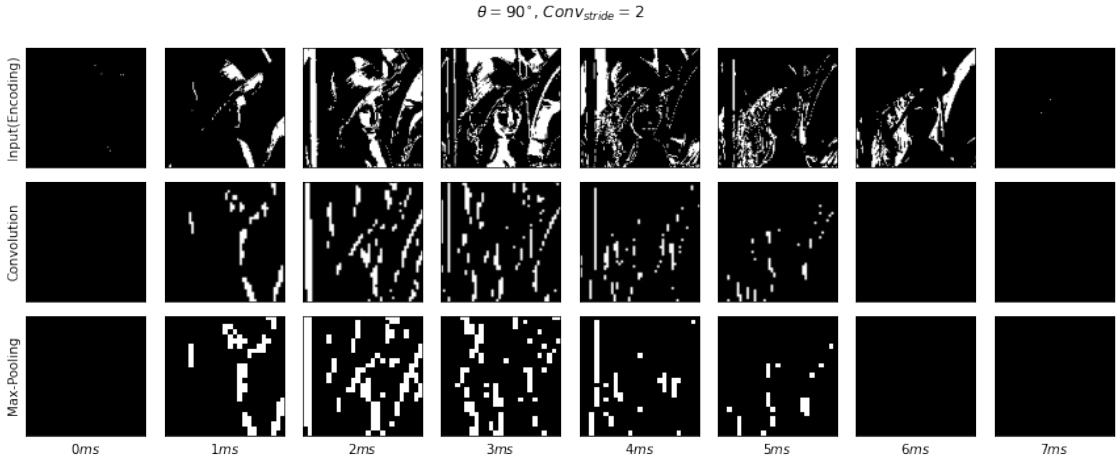
3.1.2 Stride


$\theta = 45^\circ, Conv_{stride} = 1$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


$\theta = 45^\circ, Conv_{stride} = 2$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter





We see that by increasing the stride of convolution, an operation similar to max-pooling is taking place, and the convolution output is also downsampled. By setting $Conv_{stride} = 2 \times 2$, the output shape of the convolution layer will be halved in each direction.

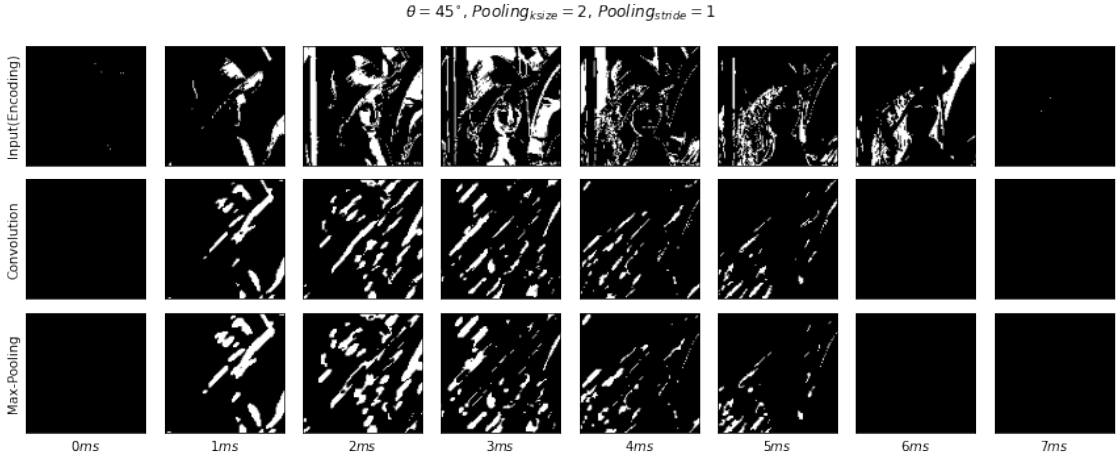
If we compare the output of max-pooling layer with $Conv_{stride} = 1 \times 1$ and the output of convolution layer with $Conv_{stride} = 2 \times 2$, the similarity between the two will be obvious. Note that the two outputs also have the same

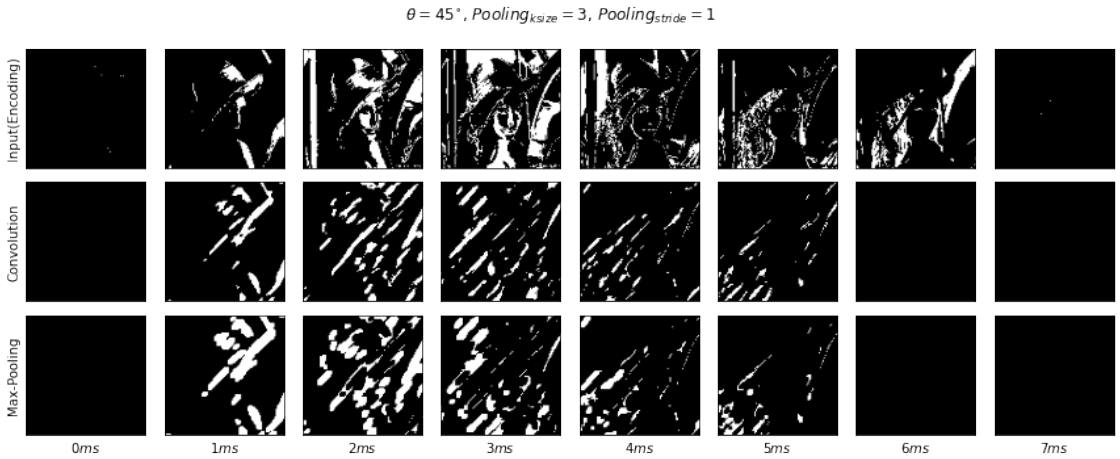
Note: Stride of convolution is modeling the `density` of simple cells, so it's better to keep the stride as low as possible (=1). This also makes our model as biologically plausible as we can (In V1 layer simple cells are very dense).

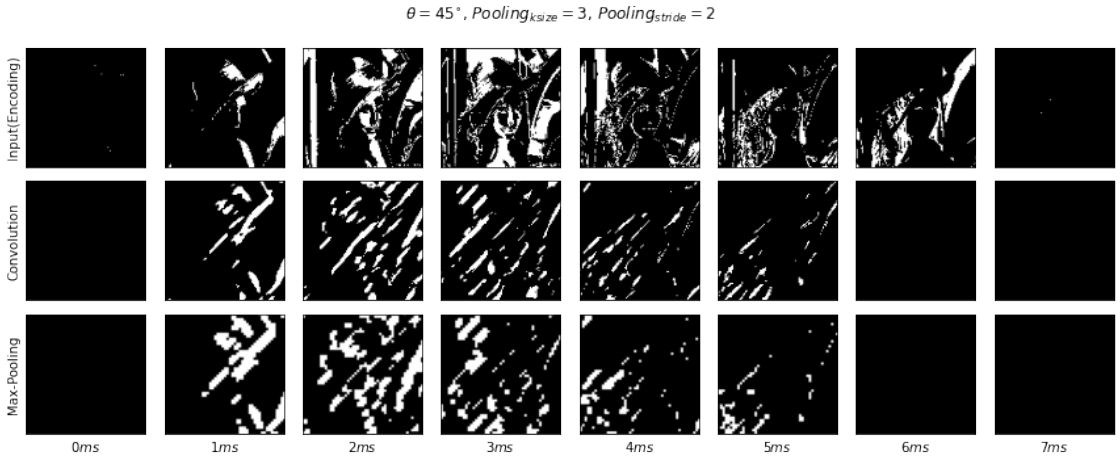
3.2 Pooling Params

In pooling layer KernelSize models the `receptive field` of `Complex cells` and Stride models the `density` of these neurons.

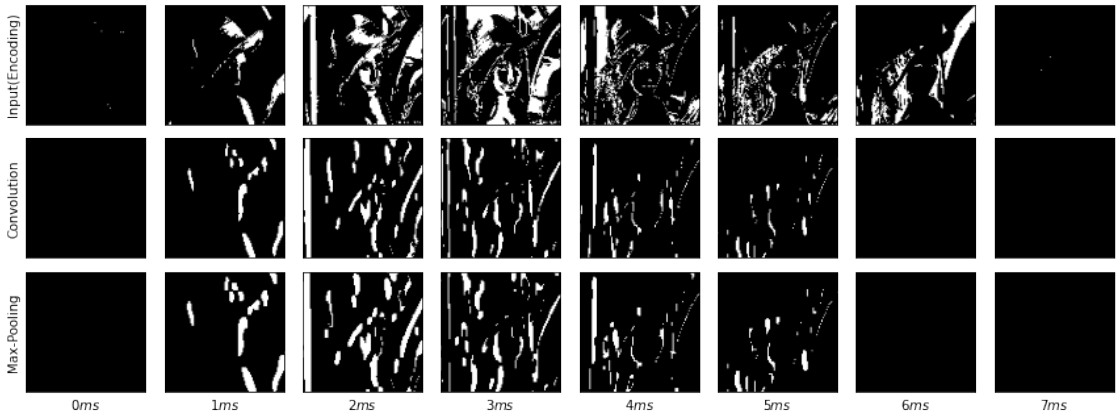
3.2.1 $KernelSize > Stride$

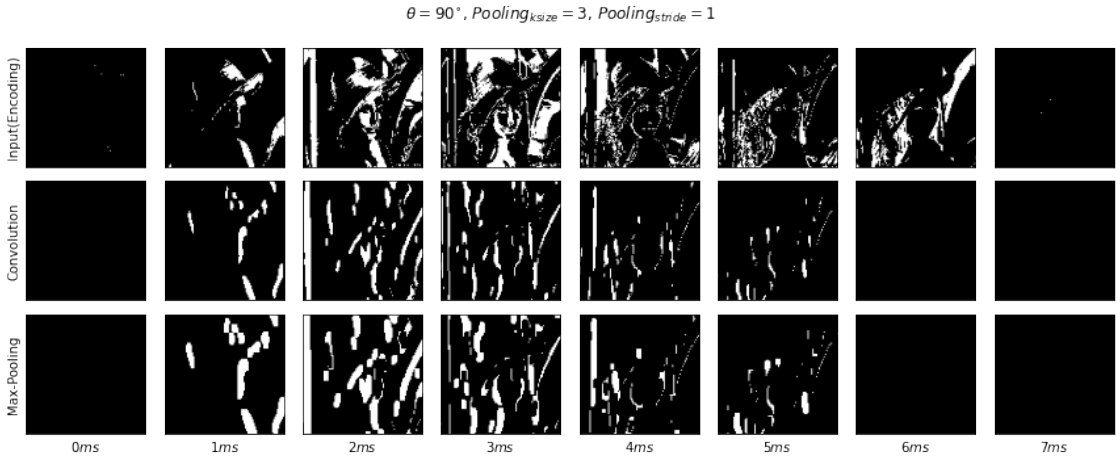


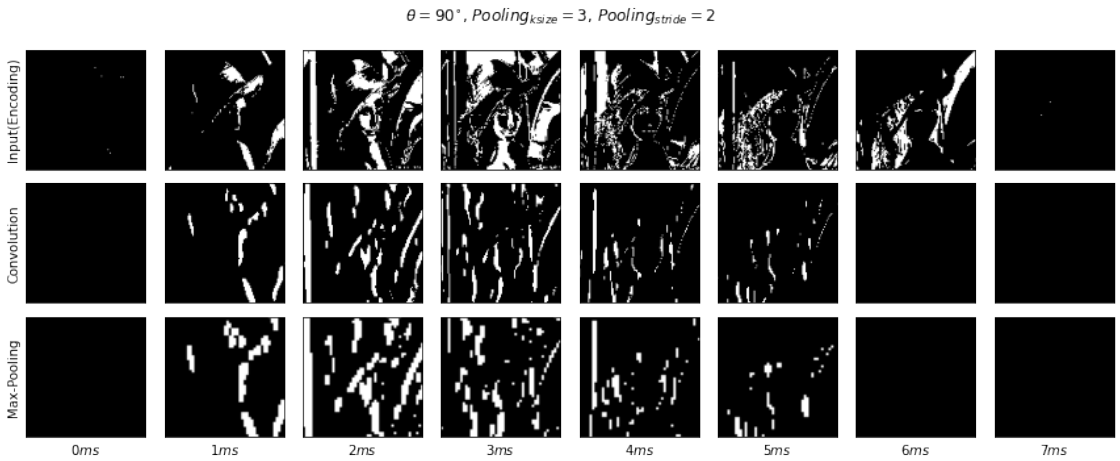



$\theta = 45^\circ$, $Pooling_{ksize} = 3$, $Pooling_{stride} = 2$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter


$\theta = 90^\circ$, $Pooling_{ksize} = 2$, $Pooling_{stride} = 1$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms



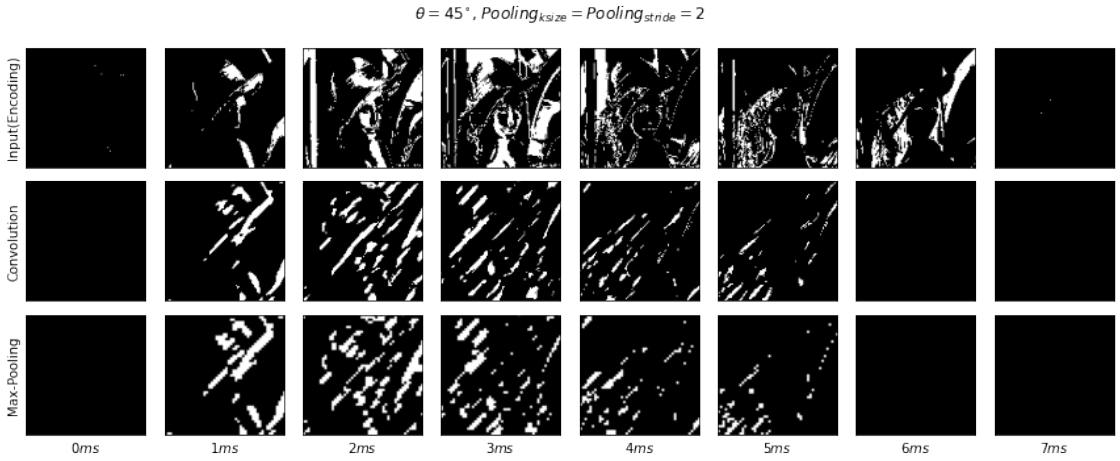


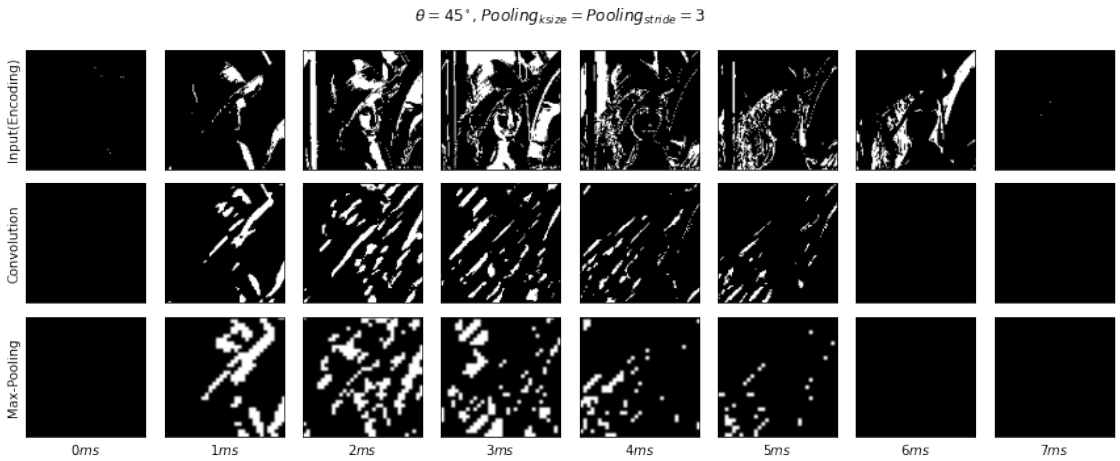
$Pooling_{stride} = 1$, adds a negligible amount of invariance to the convolution output and does not change the shape. By increasing the $Pooling_{ksize}$ and fixing the $Pooling_{stride} = 1$, we do not gain anything as the shape will be preserved and no redundant information will be neglected.

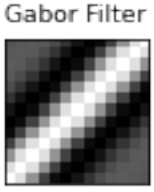
If we think of complex cells and simple cells in visual cortex, we may better understand why $stride = 1$ is not doing anything useful. This is like having some simple cells connected to more than just one complex cell. The goal of complex cells is to add invariance to simple cells' output, but with $stride = 1$, and more generally with $KernelSize > Stride$, this goal is not achieved.

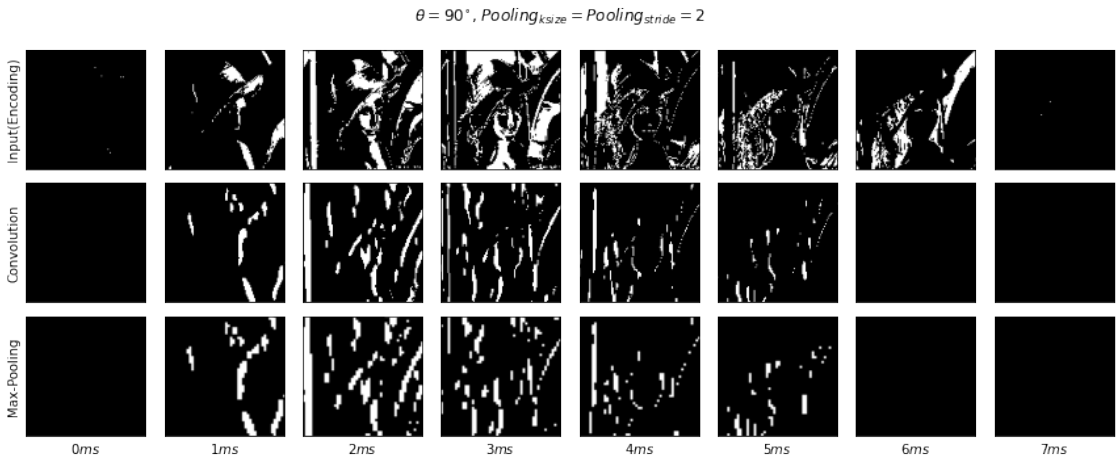
3.2.2 $Stride = KernelSize$

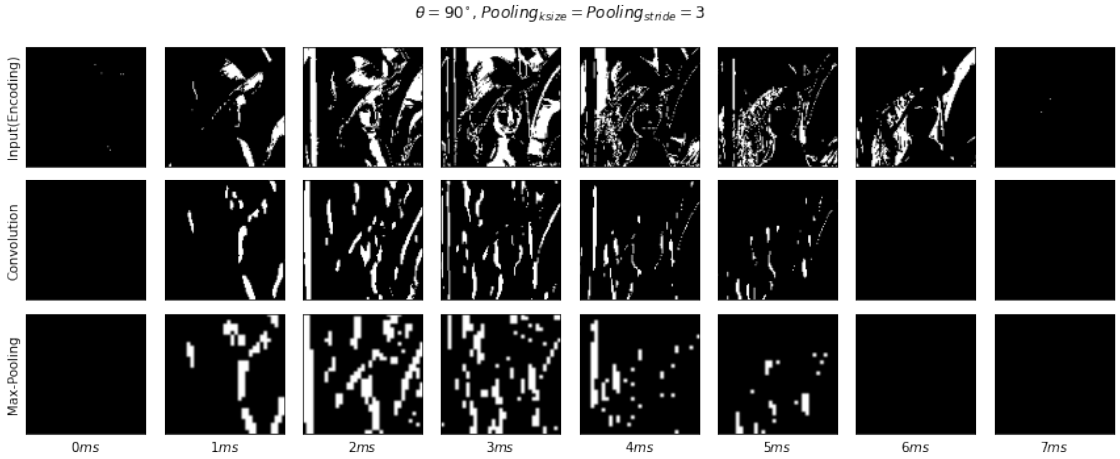
From the previous section we saw that using $KernelSize > Stride$ is not doing what we expect.







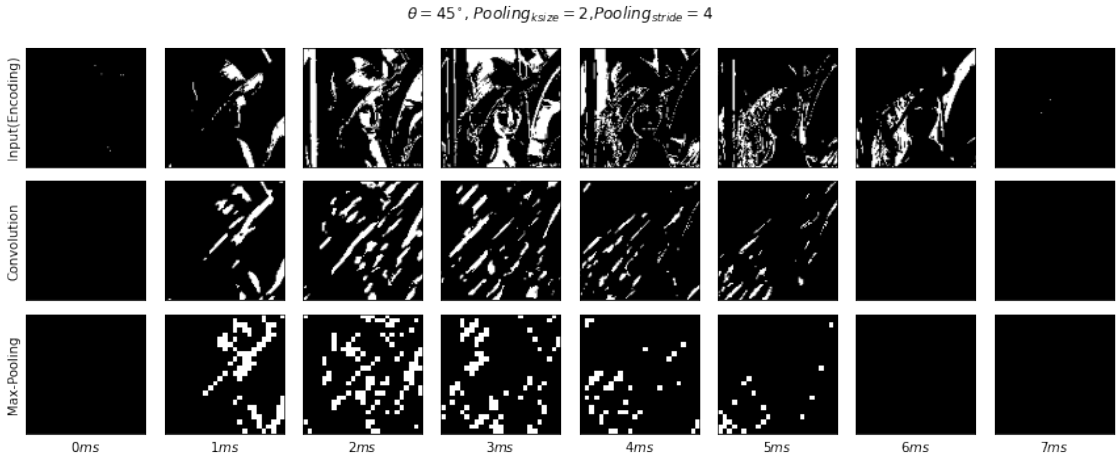


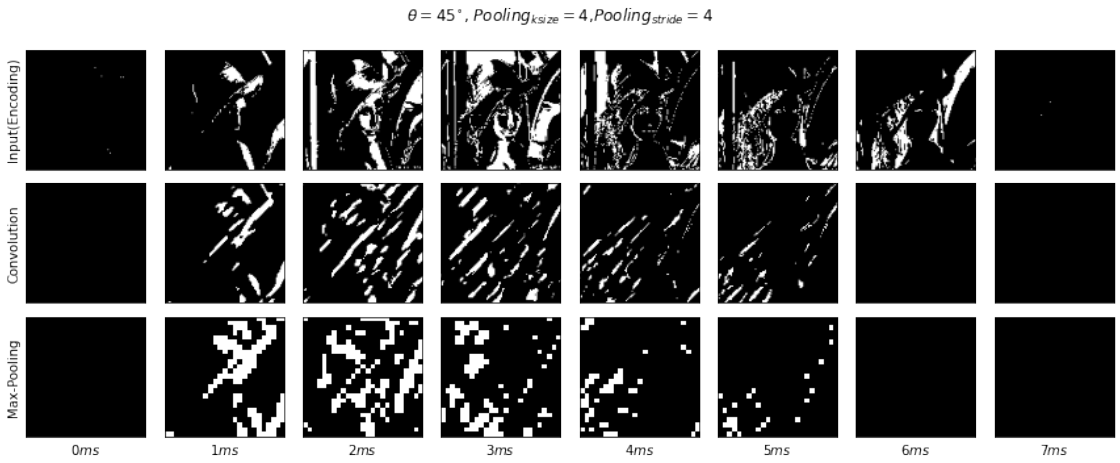


We see that by increasing $Pooling_{ksize}$ and $Pooling_{stride}$ together, the amount of invariance will be increased, and the shape will be reduced. This is what we expect from a model that mimics the complex cells. $Pooling_{ksize} = Pooling_{stride} = 2$ seems to be a good choice here.

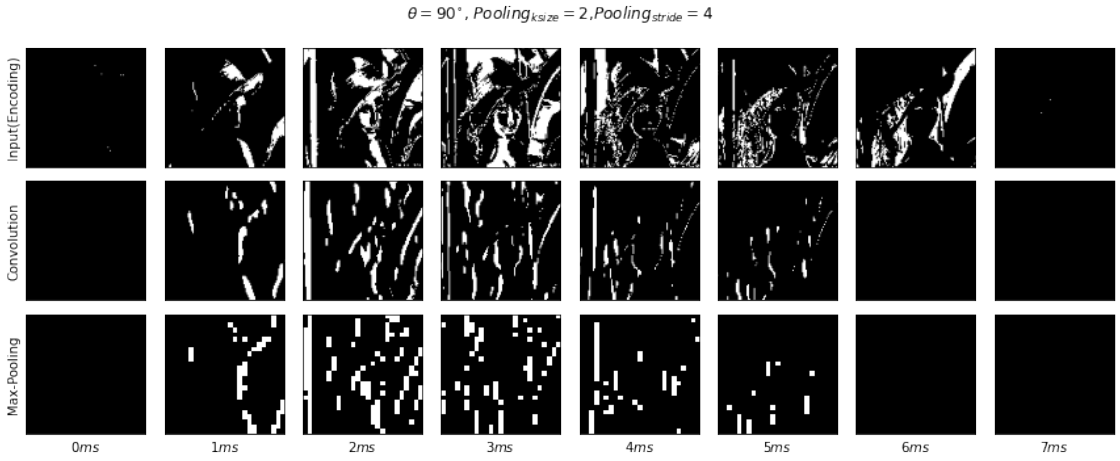
3.2.3 $KernelSize > Stride$

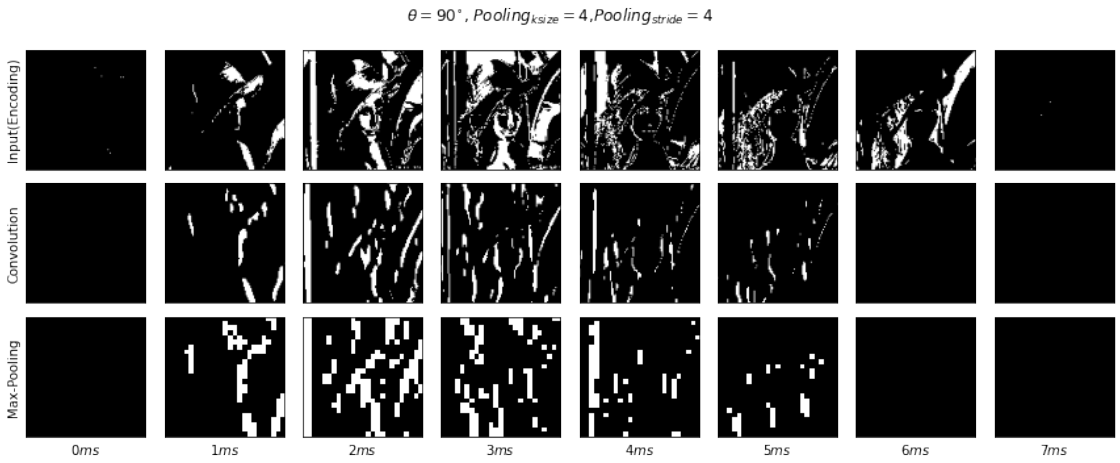
In this case, some input pixels will be neglected altogether. Although the outputs are not very different from previous section, this setting is **not biologically plausible** because it means that some simple cells are not connected to any complex cells. Therefore, we do not use this setting for our simulations.


$\theta = 45^\circ$, $Pooling_{ksize} = 2$,$Pooling_{stride} = 4$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


$\theta = 45^\circ$, $Pooling_{ksize} = 4$,$Pooling_{stride} = 4$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter


$\theta = 90^\circ$, $Pooling_{ksize} = 2$,$Pooling_{stride} = 4$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


$\theta = 90^\circ$, $Pooling_{ksize} = 4$,$Pooling_{stride} = 4$
Input(Encoding)
Convolution
Max-Pooling
0ms
1ms
2ms
3ms
4ms
5ms
6ms
7ms


Gabor Filter

Chapter 4

Summary

1. Model's sensitivity to the input could be controlled by changing convolution layer neurons (simple cells) threshold. The more the threshold, the model is more robust to noise. However, too high thresholds make the model to not respond to some useful information.

2. It's better to keep pooling layer neurons' threshold to 1 unit above their u_{rest}. This is because in brain, complex cells mostly spike when only one of the simple cells in their receptive field spike. Also, in the simulations, we saw that setting this parameter to 2, will cause the model to neglect much useful information.

3. Smaller convolution kernels detect smaller lines and larger kernels detect larger lines. Neurons in convolution layer with smaller receptive field should have lower threshold and vice versa.

4. By increasing the stride of convolution, we are lowering the density of `simple cells` and the output is similar to the output of applying max pooling. It's better to keep the $Convolution_{stride}$ as low as possible to have a model with more bio plausibility.

5. $Stride$ and $KernelSize$ are better to be equal in the pooling layer. If $Stride < KernelSize$, the effect of pooling is not clearly visible. If $Stride > KernelSize$, we are neglecting some neurons in the previous layer completely, which means that we are loosing useful information. Also, if we consider complex cells in the visual cortex, $Stride = KernelSize$ is more biologically plausible because it means that each `simple cell` is connected to only 1 `complex cell`.