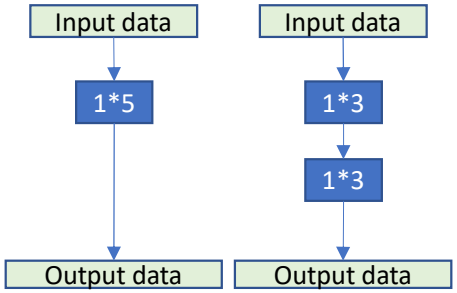
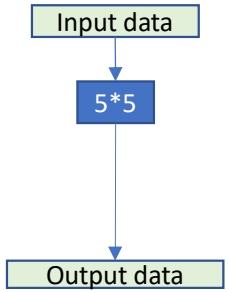
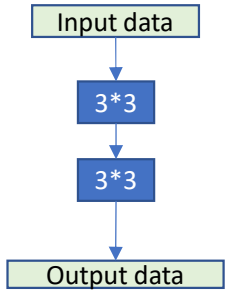
Difference between OS and Inception

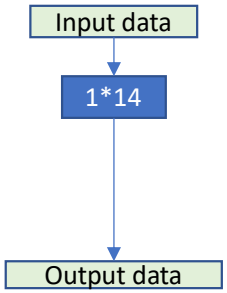
Start from a simple case

- In 1D-CNN kernel factorization cannot reduce mode size as 2D-CNN. It will increase the model size.

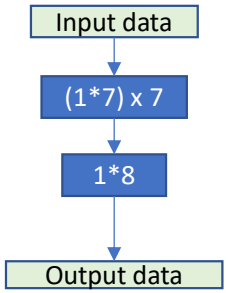
	1D		2D	
	A	B	C	D
				
Number of parameter	5	6	25	18

Can we do something to reduce model size?

- Before talk about the solution, we should talk about the channel requirements
- As the subsection "**No representation ability lose**" in the paper says
  - If we want to do kernel factorization we need to increase the channel number of first layer

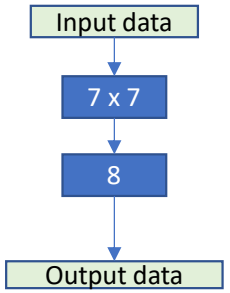


A



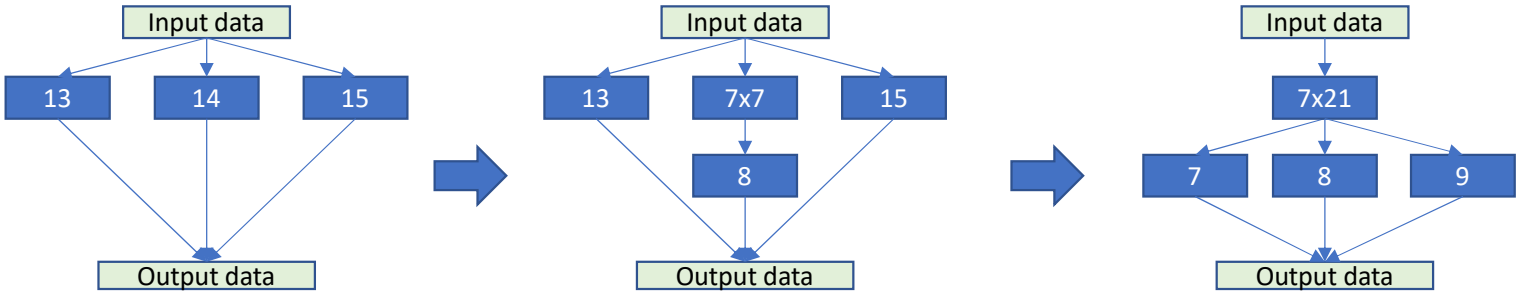
B

From A to B the channel requirement for 1*7 kernel is 7
Let's write it in and remove 1* for the following discussion is all based on 1D-CNN. Then we should write B as C



C

Simple factorization

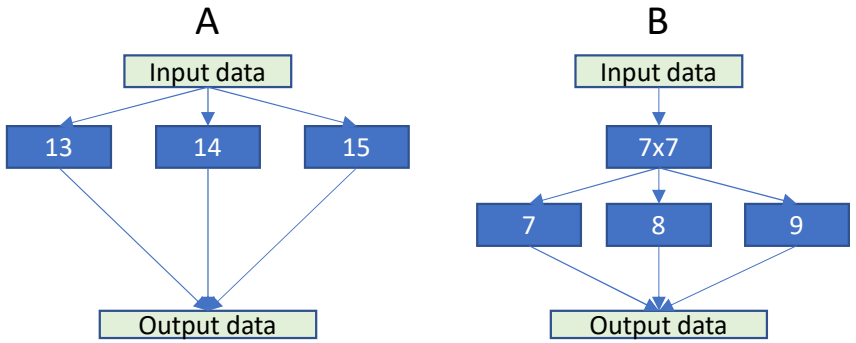


Number of parameter	42	13+8+15+7*7 = 85	7+8+9+7*21 = 171

However

- The objective that we want kernels of all size is:
  - we want to find the proper kernel.
  - Therefore, we don't really need 21 channels.

- For example, if we just want 1 kernel. Then, the model should selected from {13, 14, 15} during training time, and model B should be selected from {7, 8, 9}. And for network B, only 7 channels would be enough!



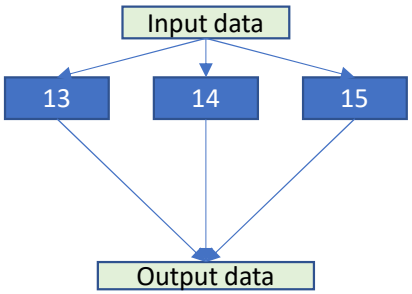
Number of parameter	42	7+8+9+7*7 = 73

It seems A is still of smaller number of parameter than B, But…..

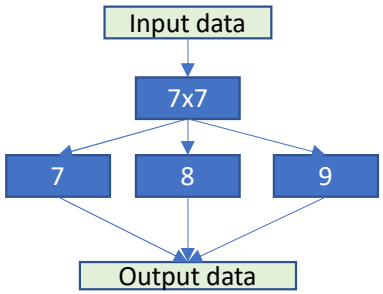
If we don’t want to select from {13, 14, 15}, if we want to select from 8 to 20

This is of smaller number of parameter!

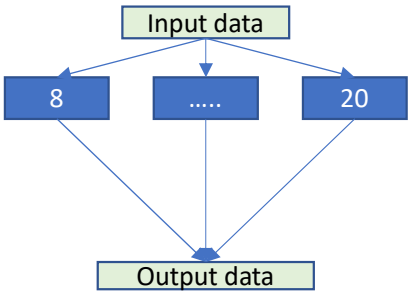
A



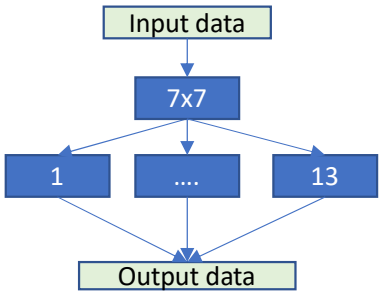
B



A’



B’



	A	B	A’	B’
Number of parameter	42	7+8+9+7*7 = 73	8+9+,…,+20 = 182	(1+2,…,+13)+49= 140

Differences

- Inception does not mention the kernel size selection in 1D-CNN. Here, OS's kernel selection is based on build receptive filed of all sizes.

- Inception-v1 does not talk about model size, and the model size reduction in inception-v2 does not work in 1D-CNN. In here, OS CNN reduces the model size differently.