

Reinforcement Learning

Lab #5: Policy Gradients and Attention

Daniil Arapov

Partially based on msu.ai lectures

Test questions for today:

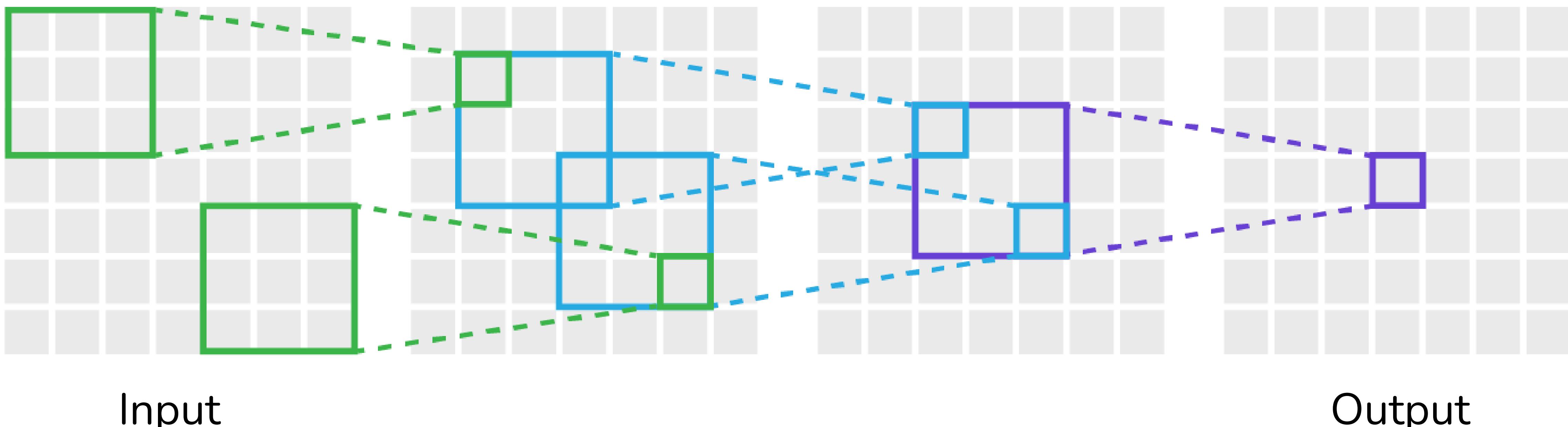
1. What are advantages of policy-based RL?
2. What is Gaussian policy, when one may use it?
3. Describe “roles” of Actor and Critic.
4. Compare Dyna with Model-free and Model-based RL.
5. Describe Monte-Carlo Tree Search.

Topic #3

Attention

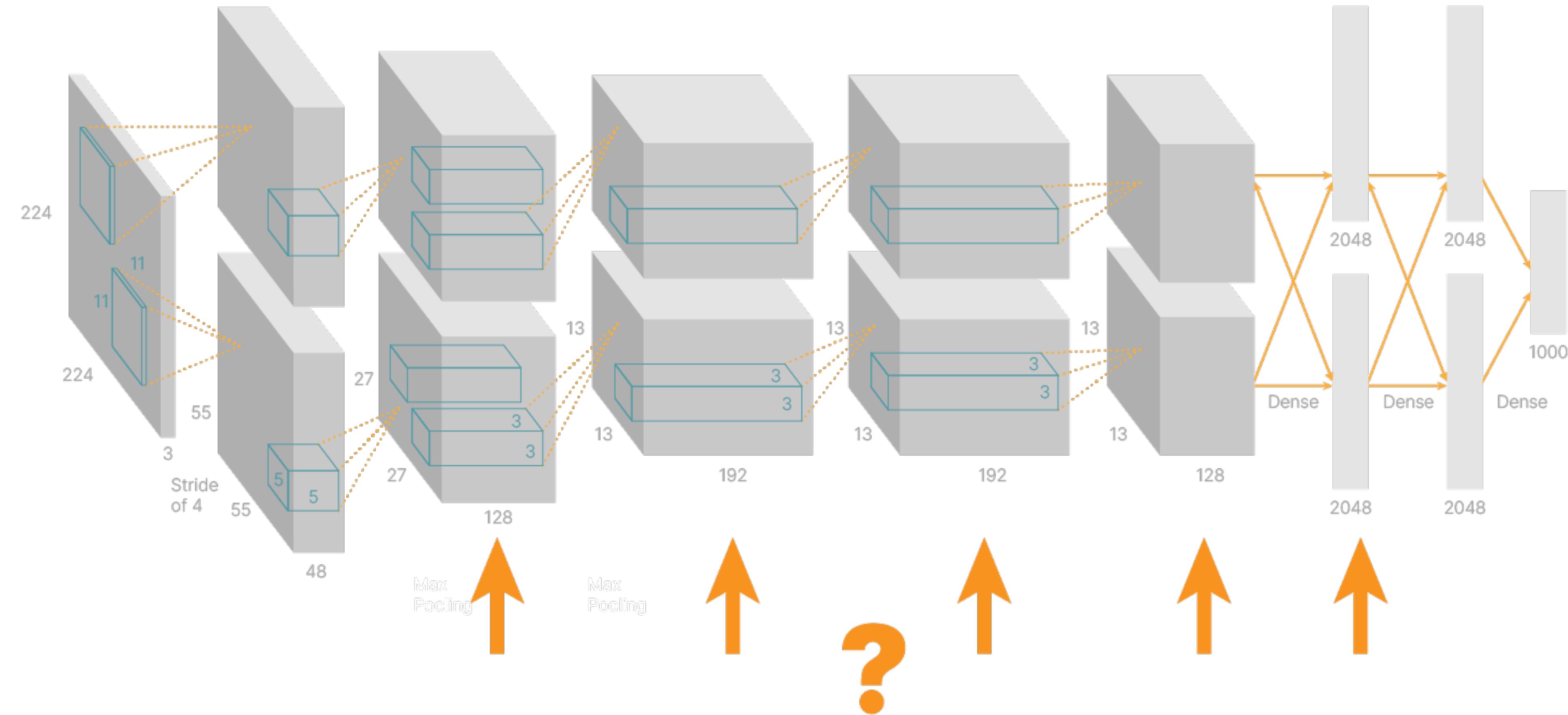


Convolution limitation



Receptive field

Convolution limitation

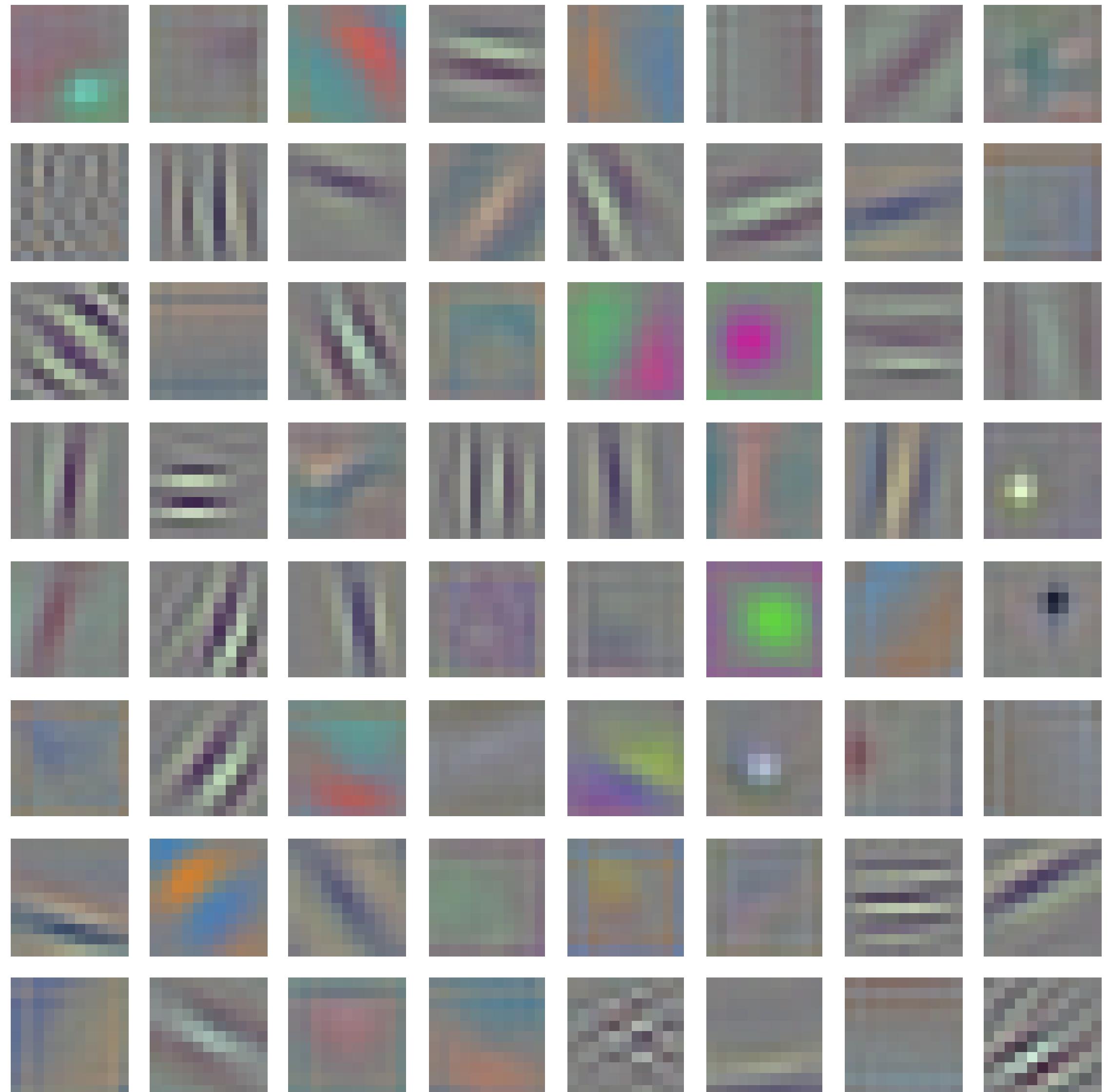


Receptive field

Kernel visualization

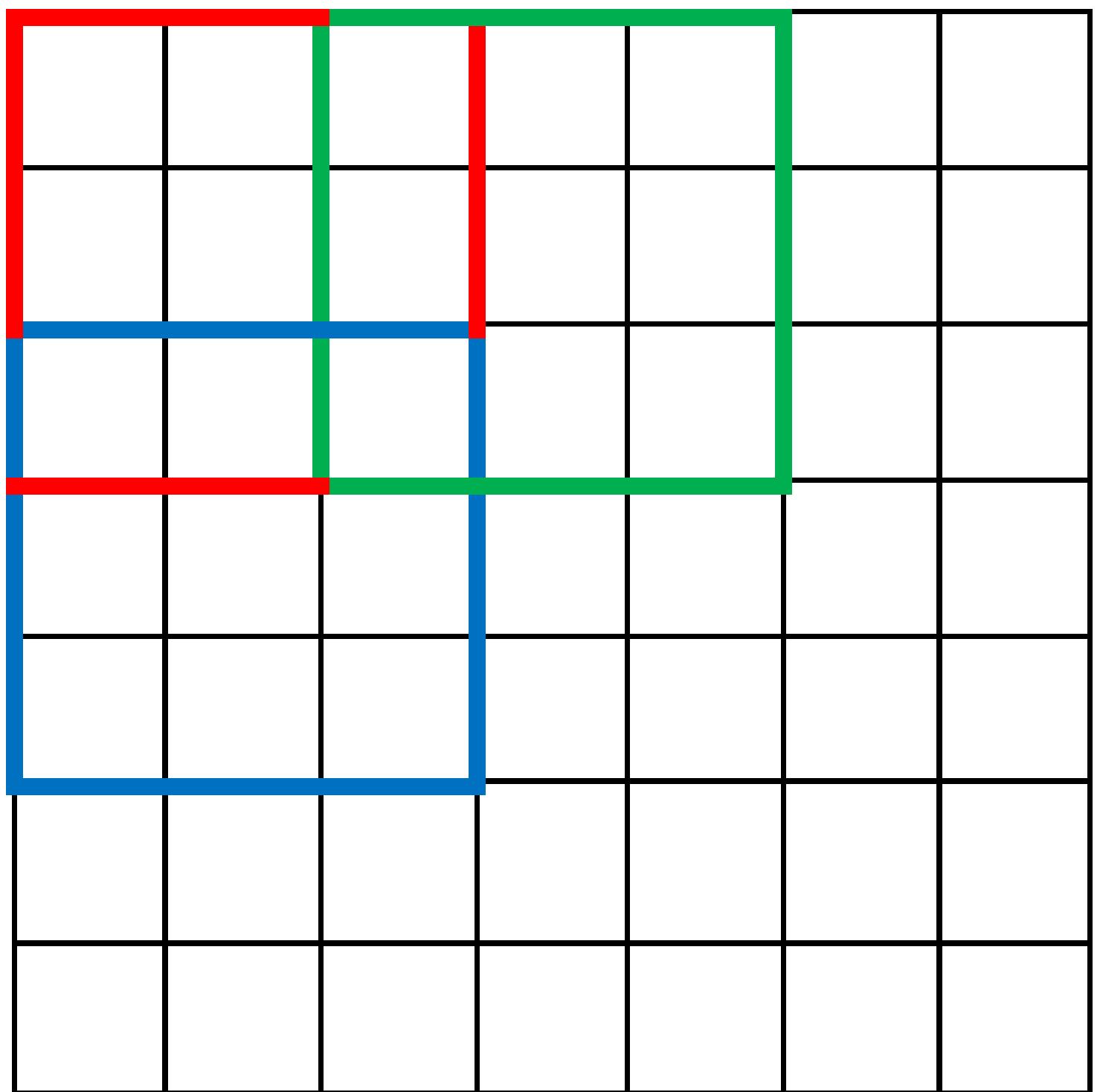
Kernels for the first layer of AlexNet

After each convolution operation with kernel size K , receptive field of a neuron increases by $K - 1$.

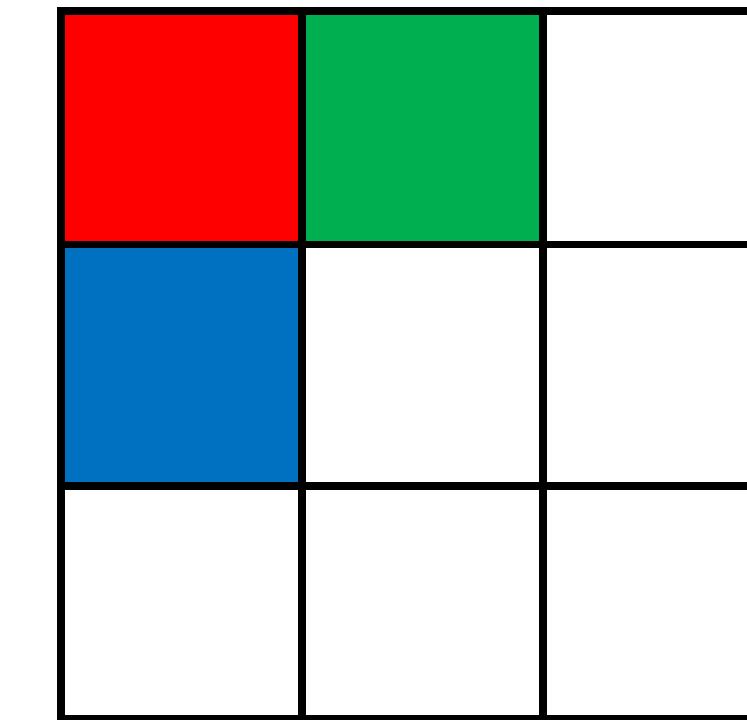


Problem solution (1)

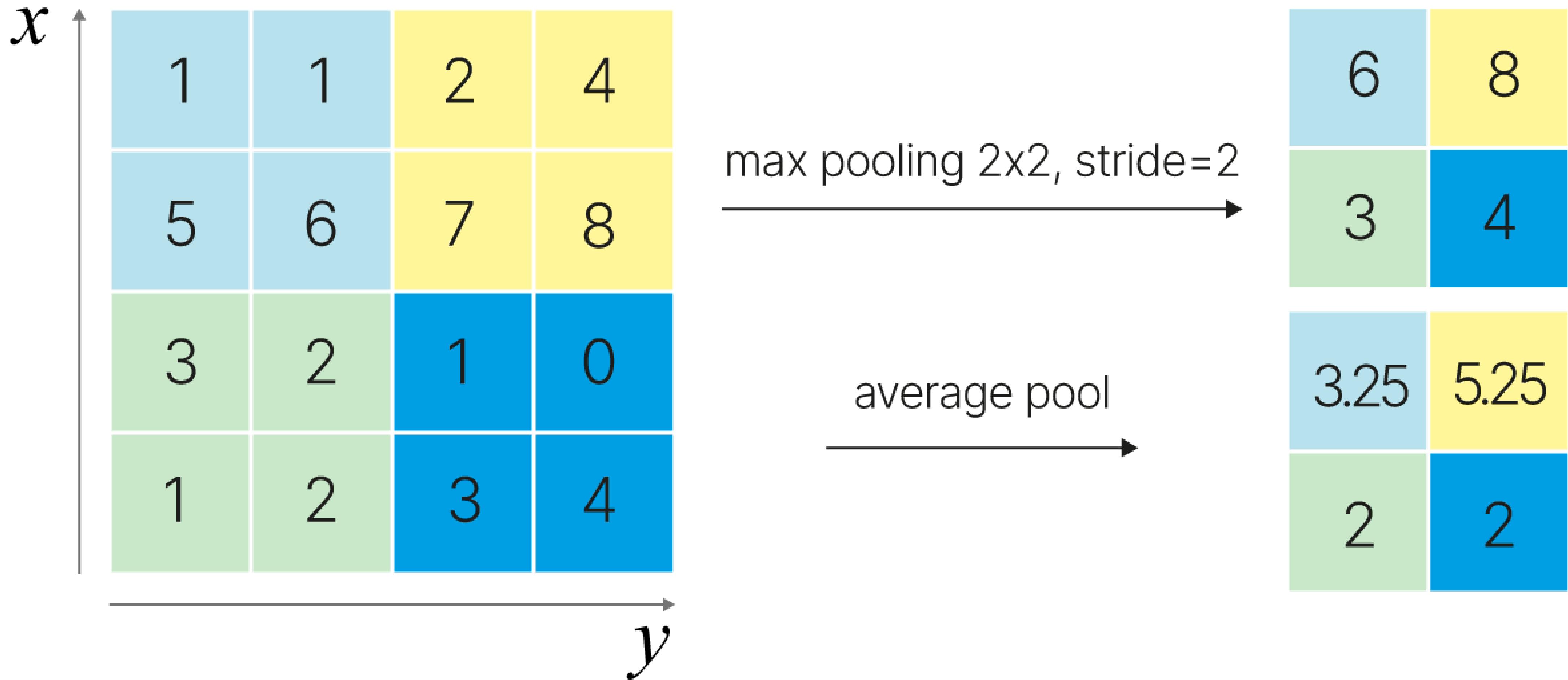
7 x 7 Input Volume



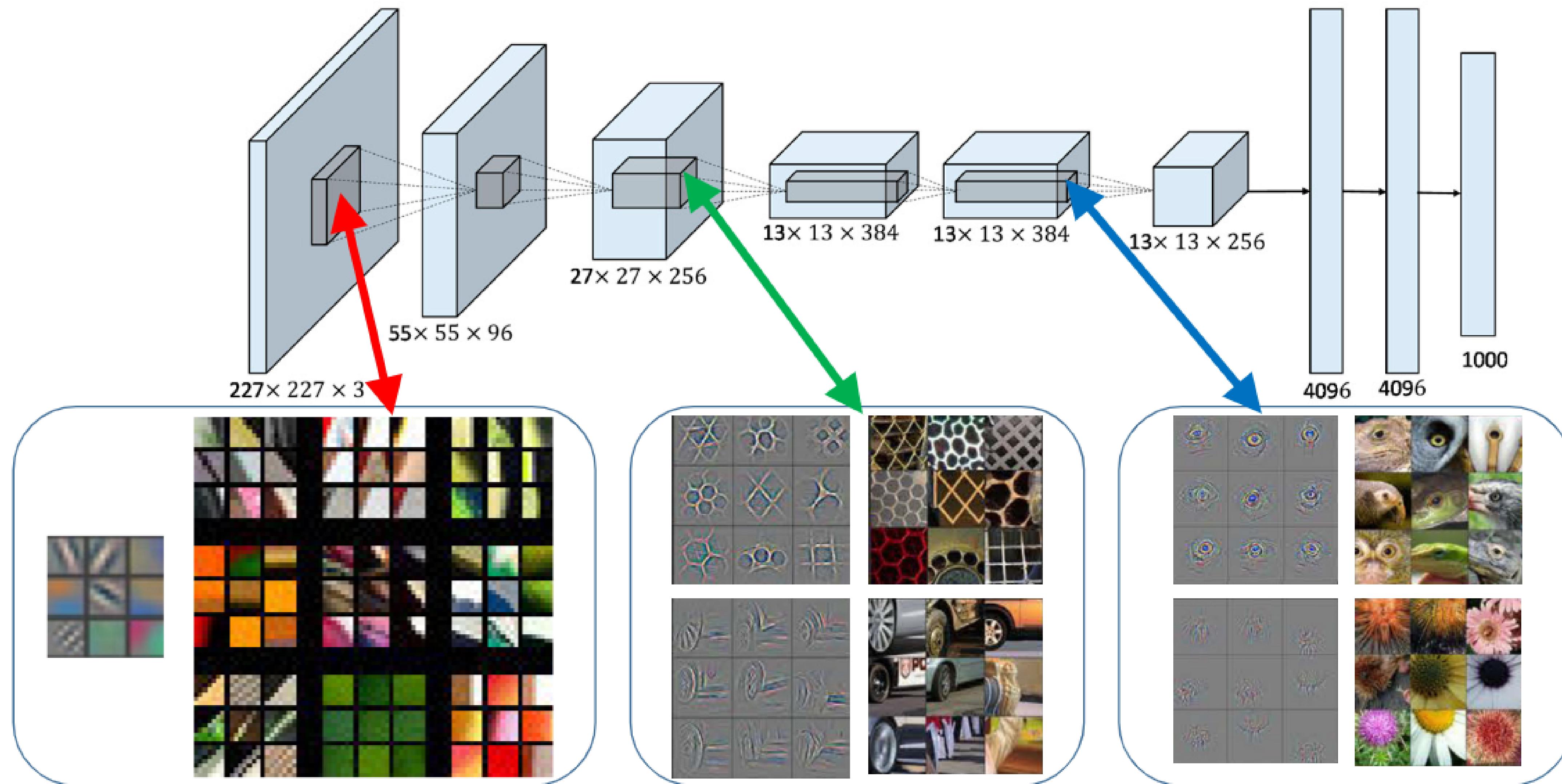
3 x 3 Output Volume



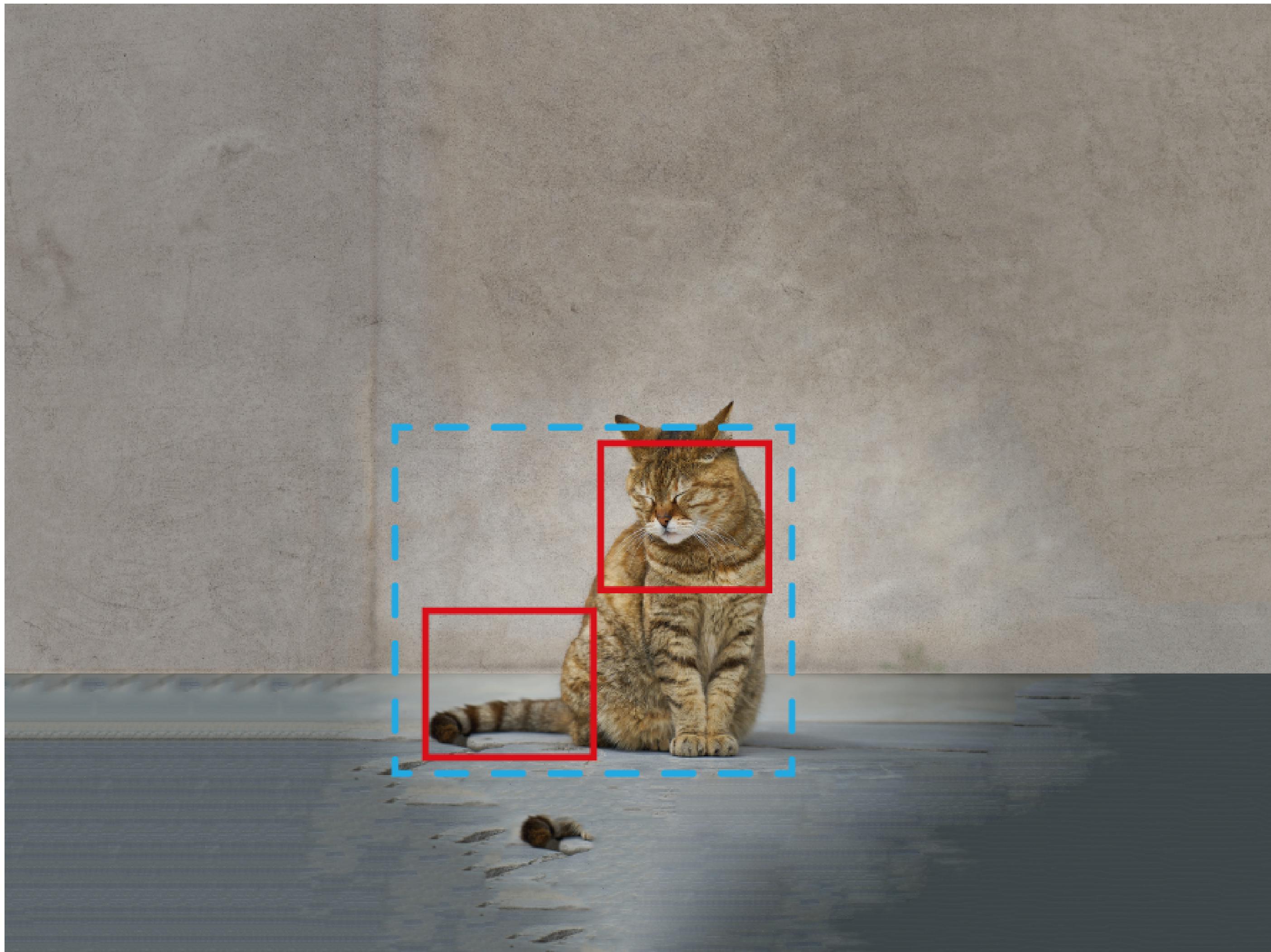
Problem solution (2)



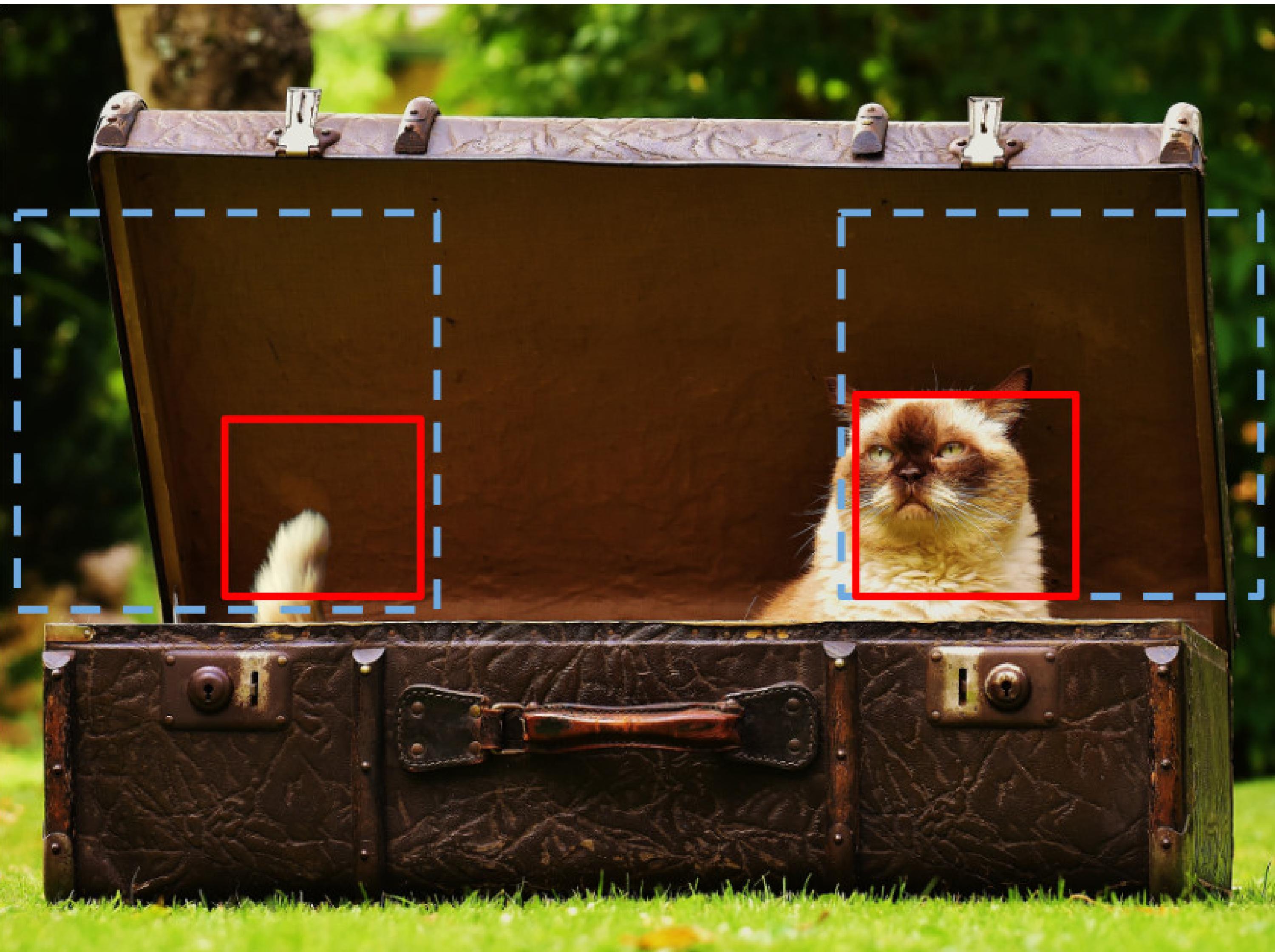
Deeper visualization



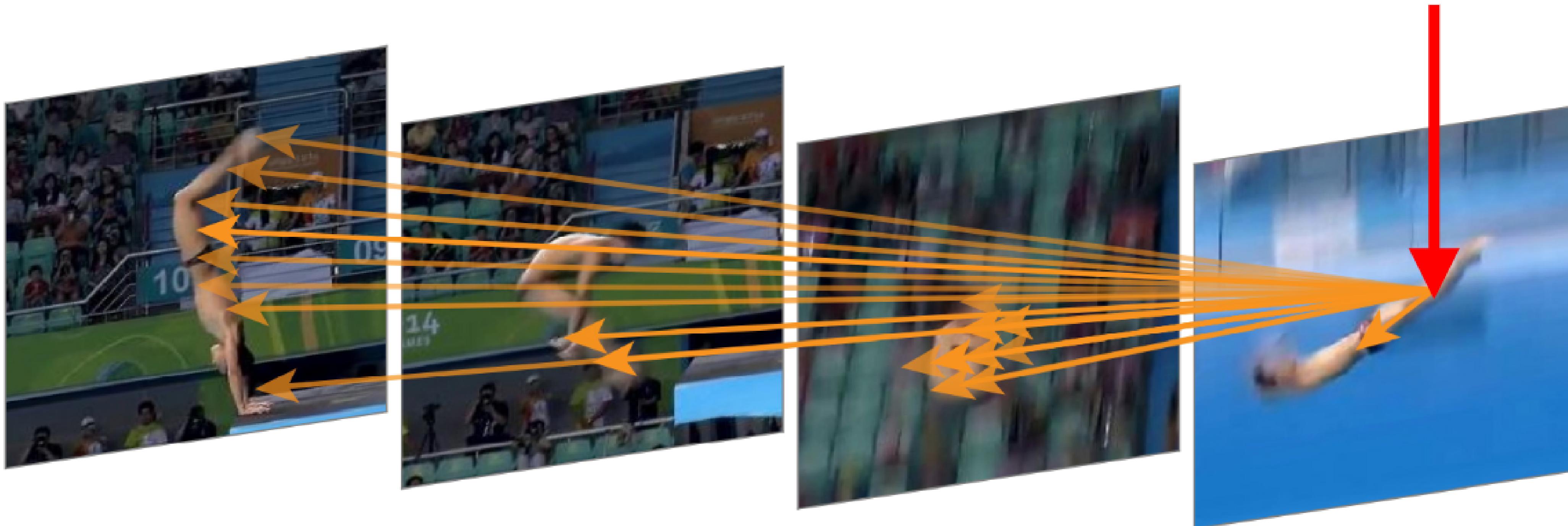
Do they solve the problem?



Do they solve the problem? (2)

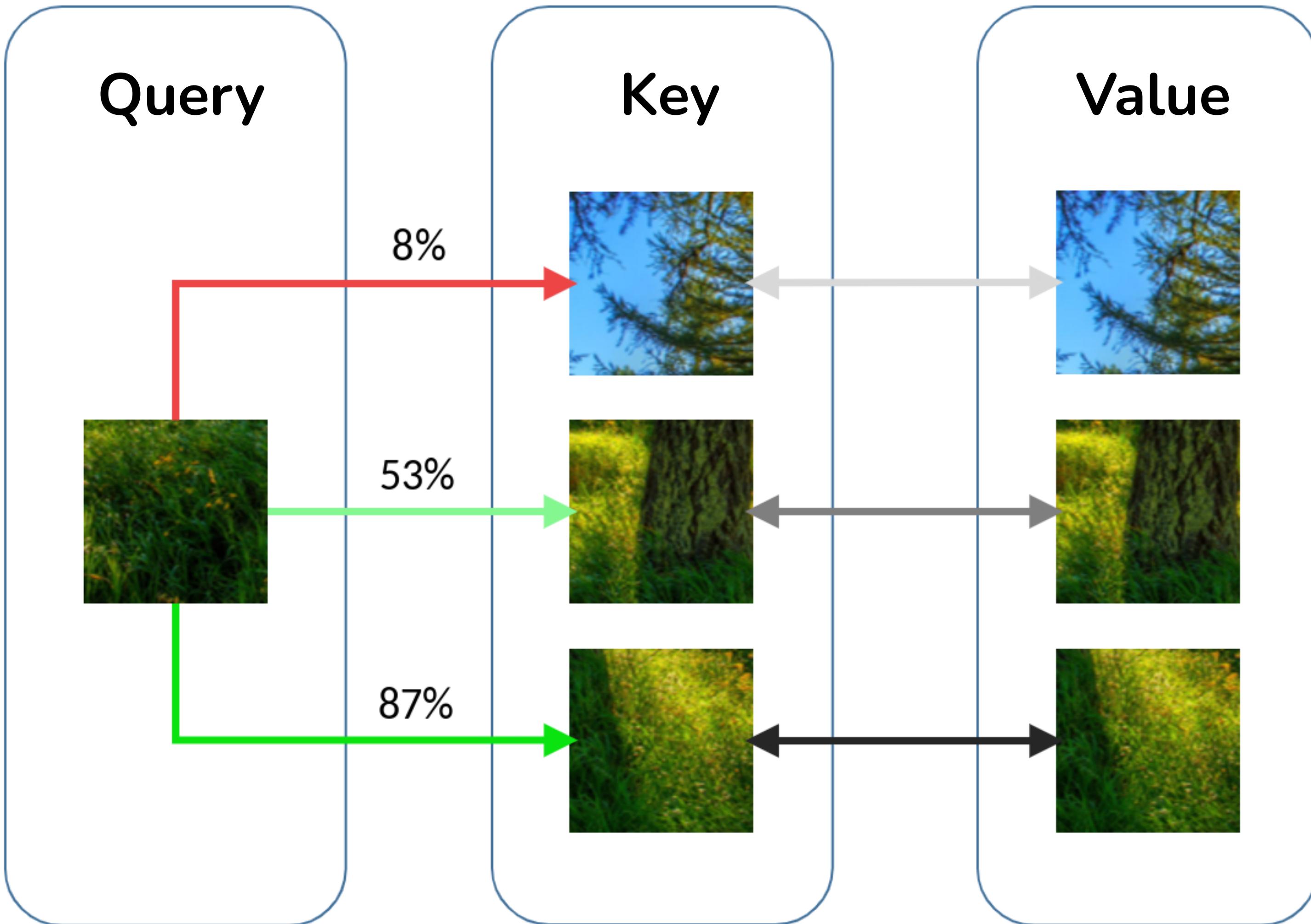
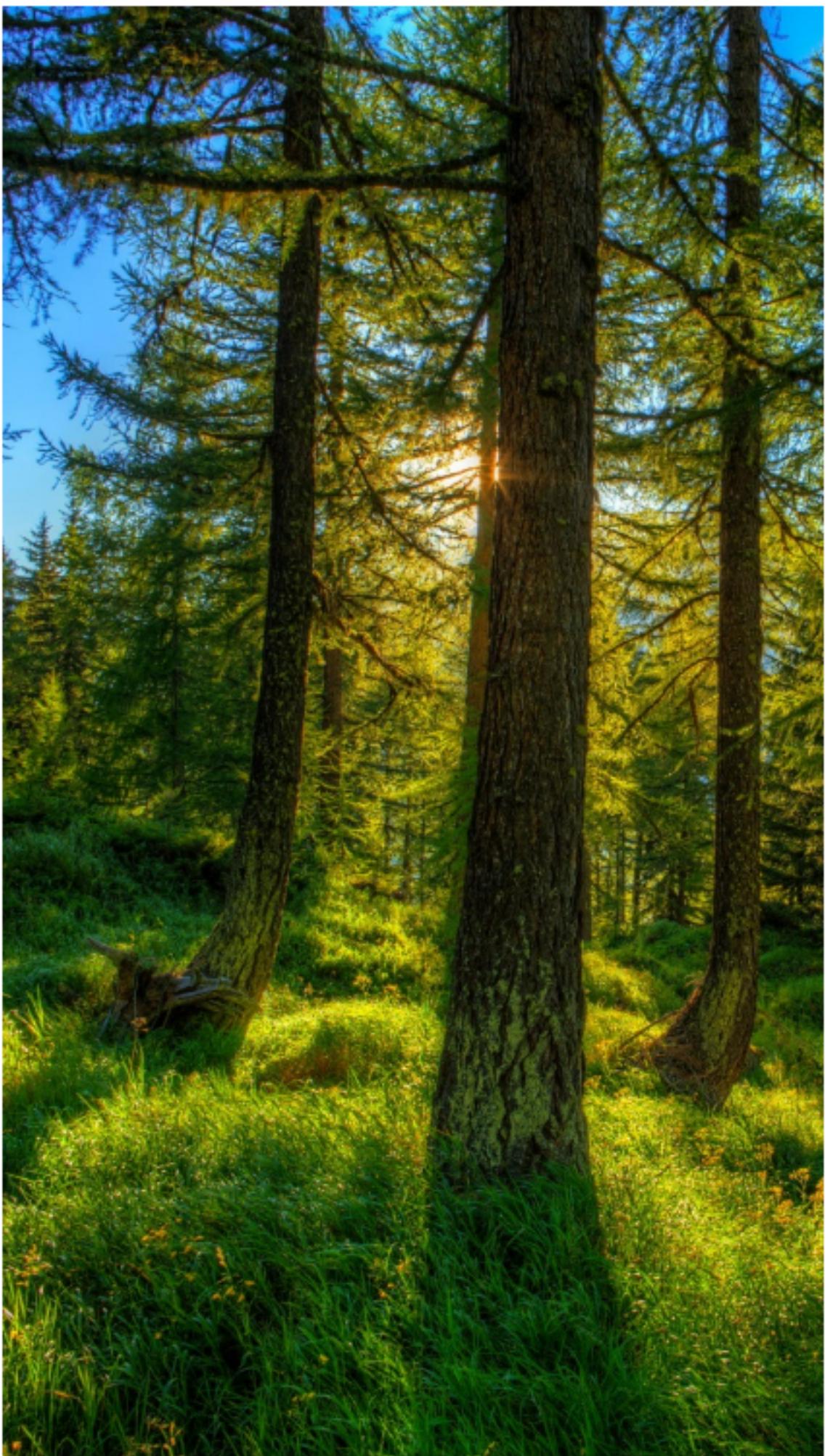


What do we want to achieve?



Considering whole image information “directly”

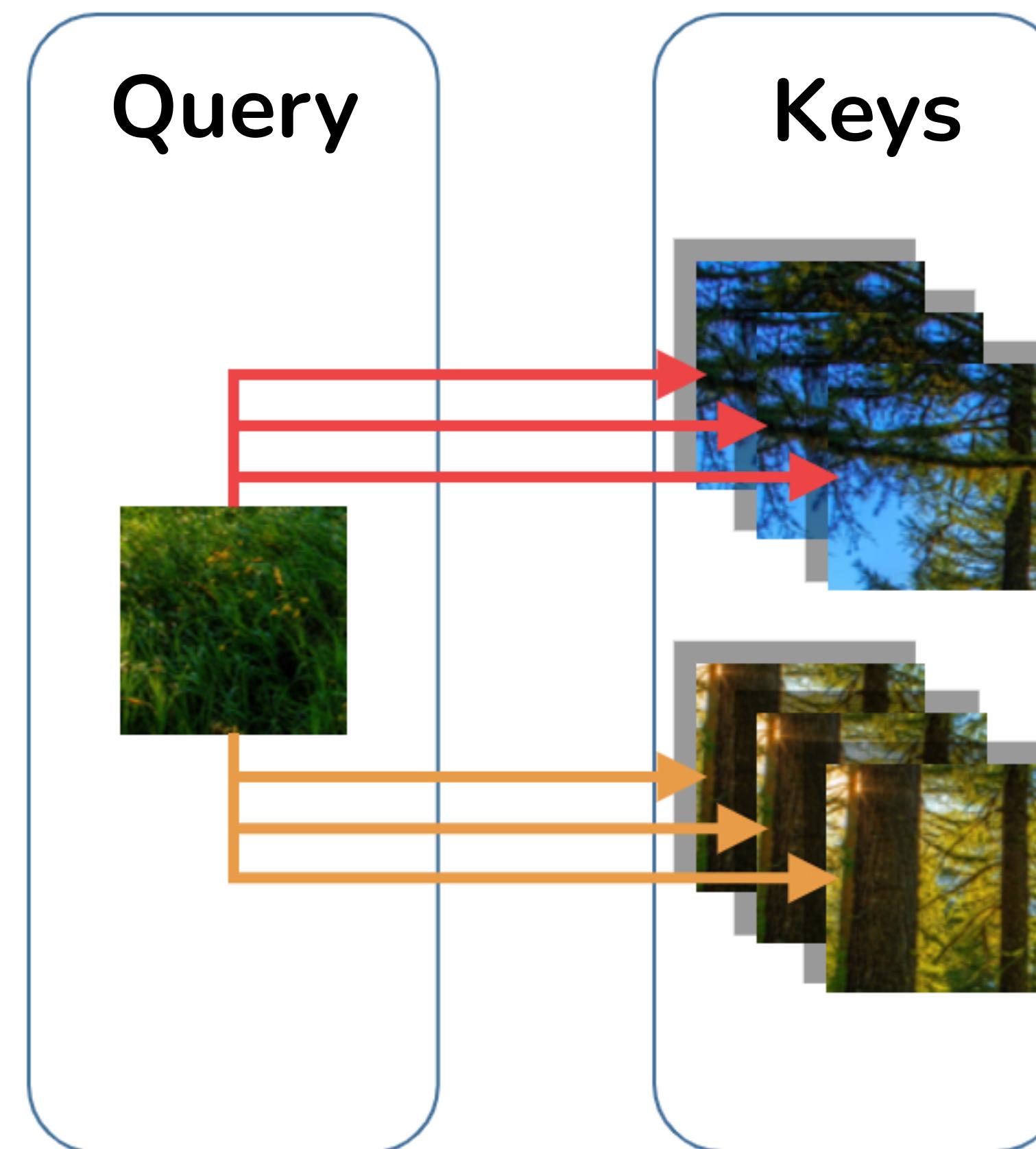
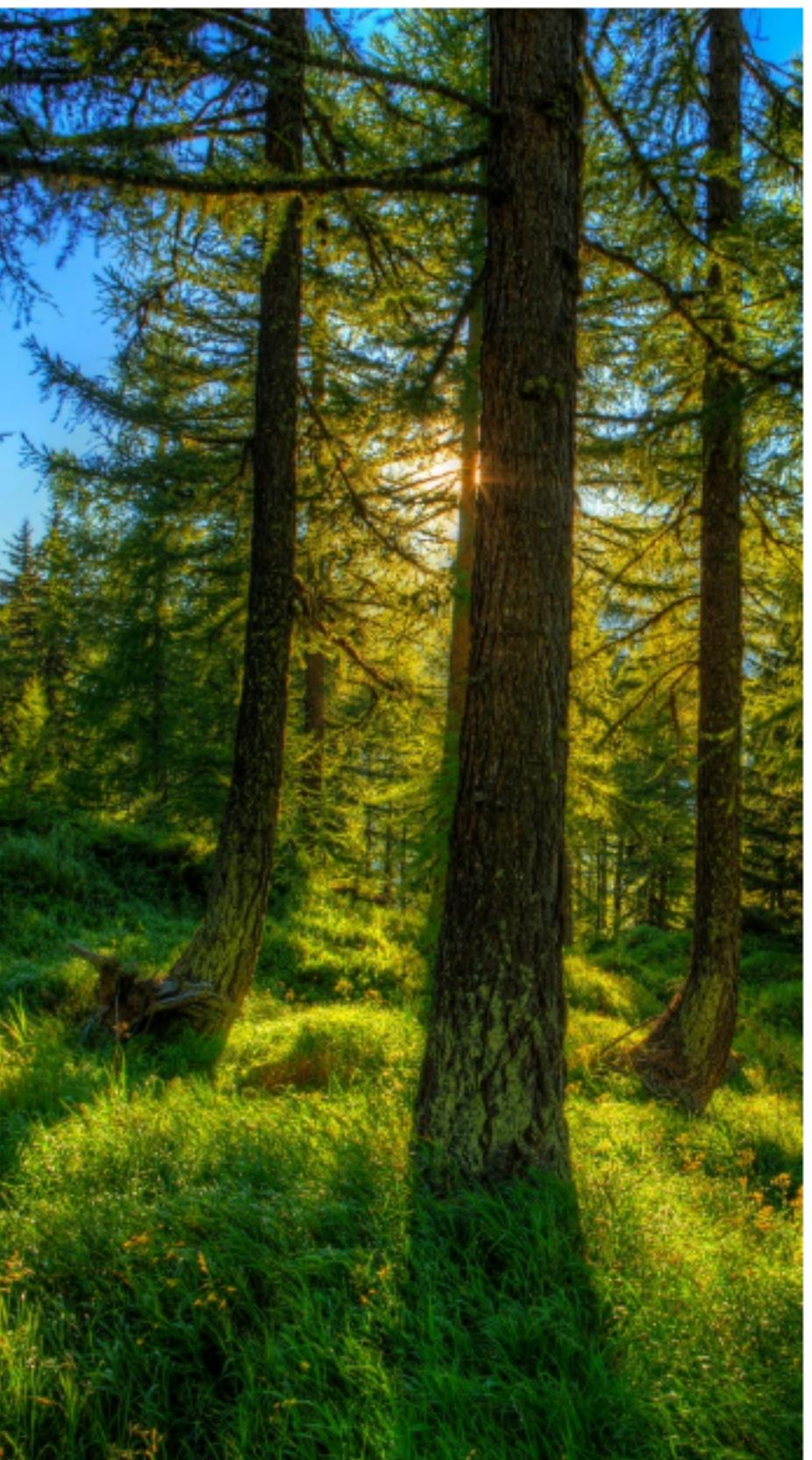
How do we want to achieve that?



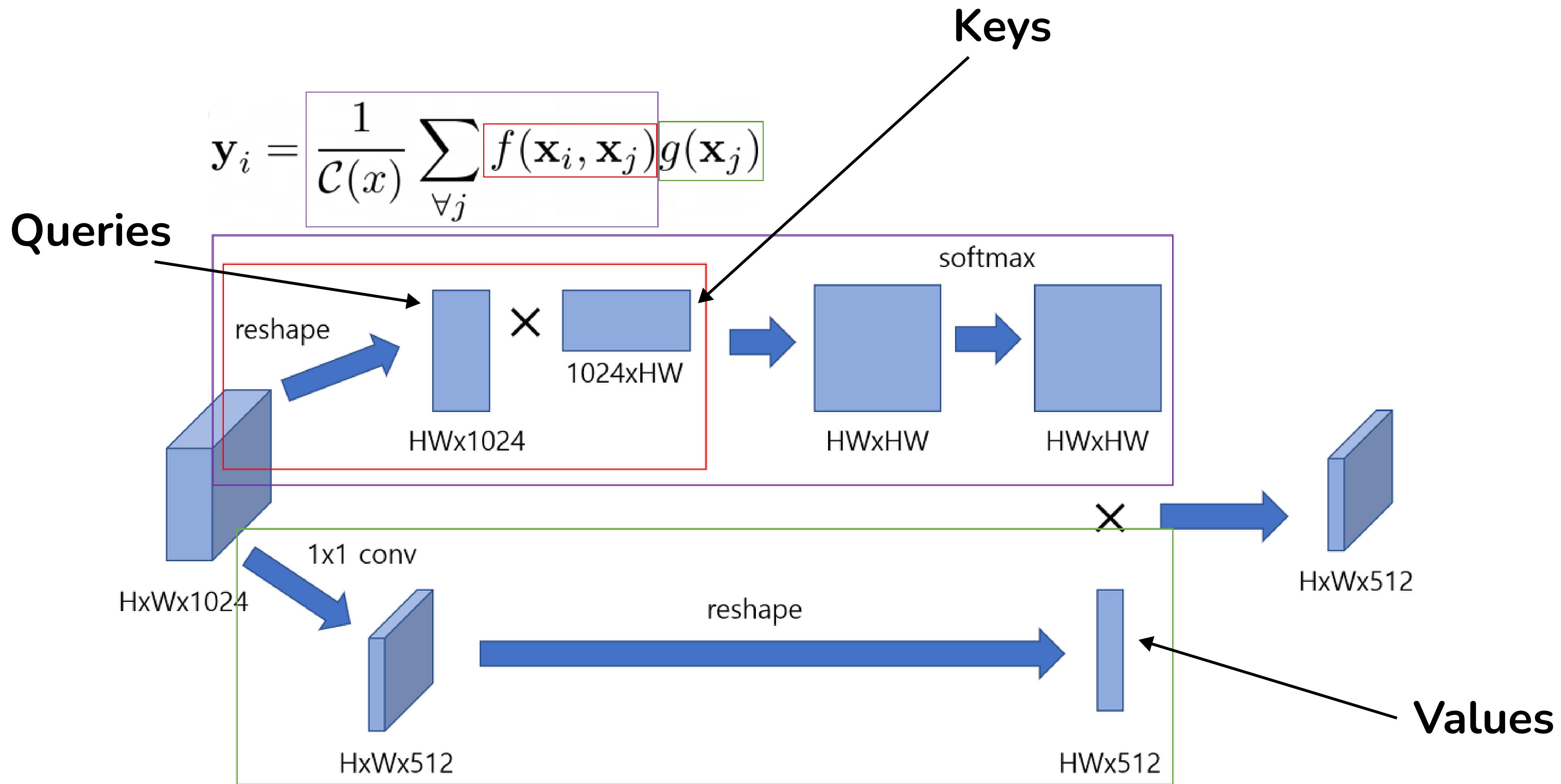
Non-local Networks (2018)

Compare all the queries
with all the keys

Too much memory
is required

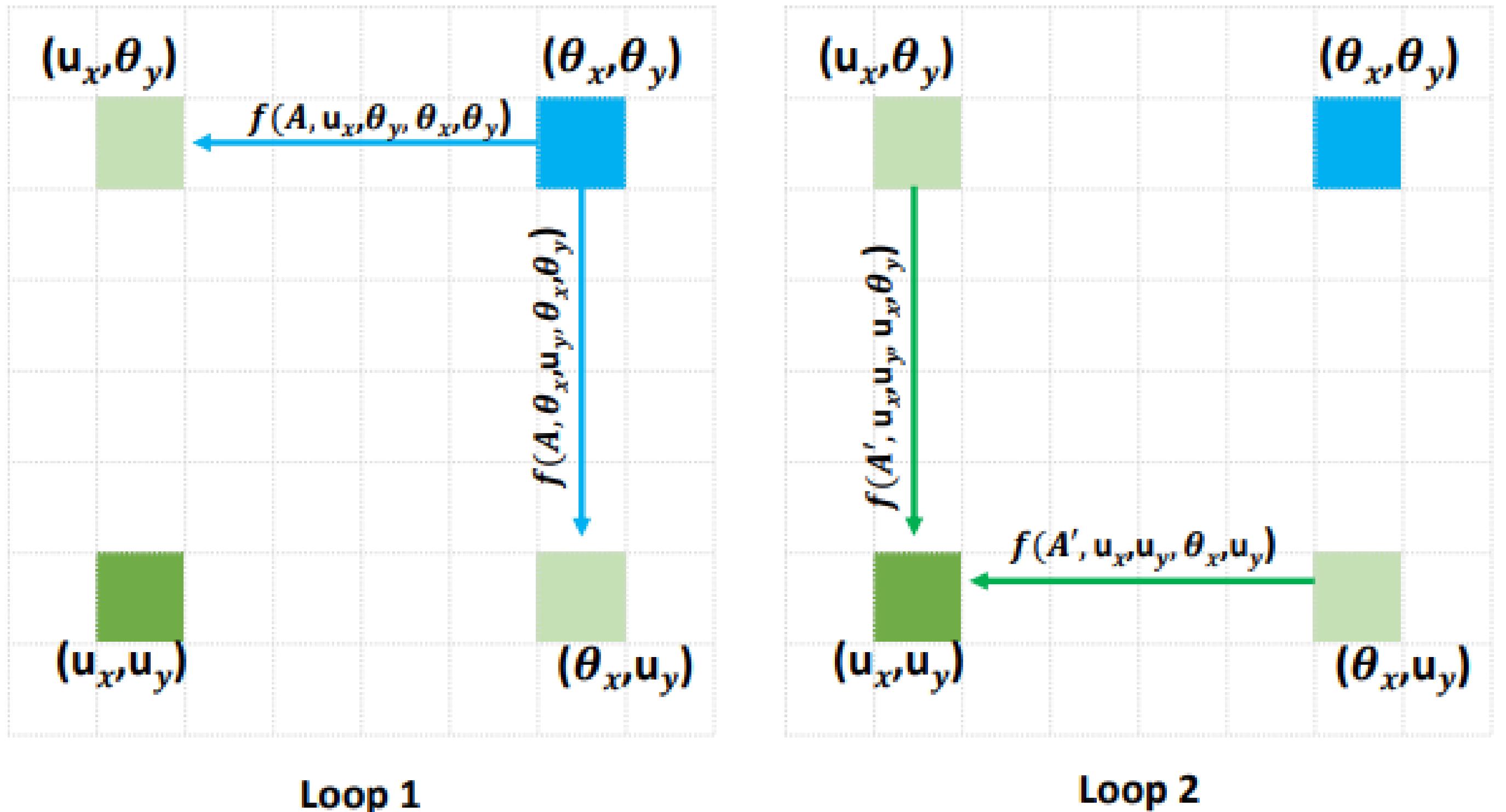


Non-local Networks (2018)

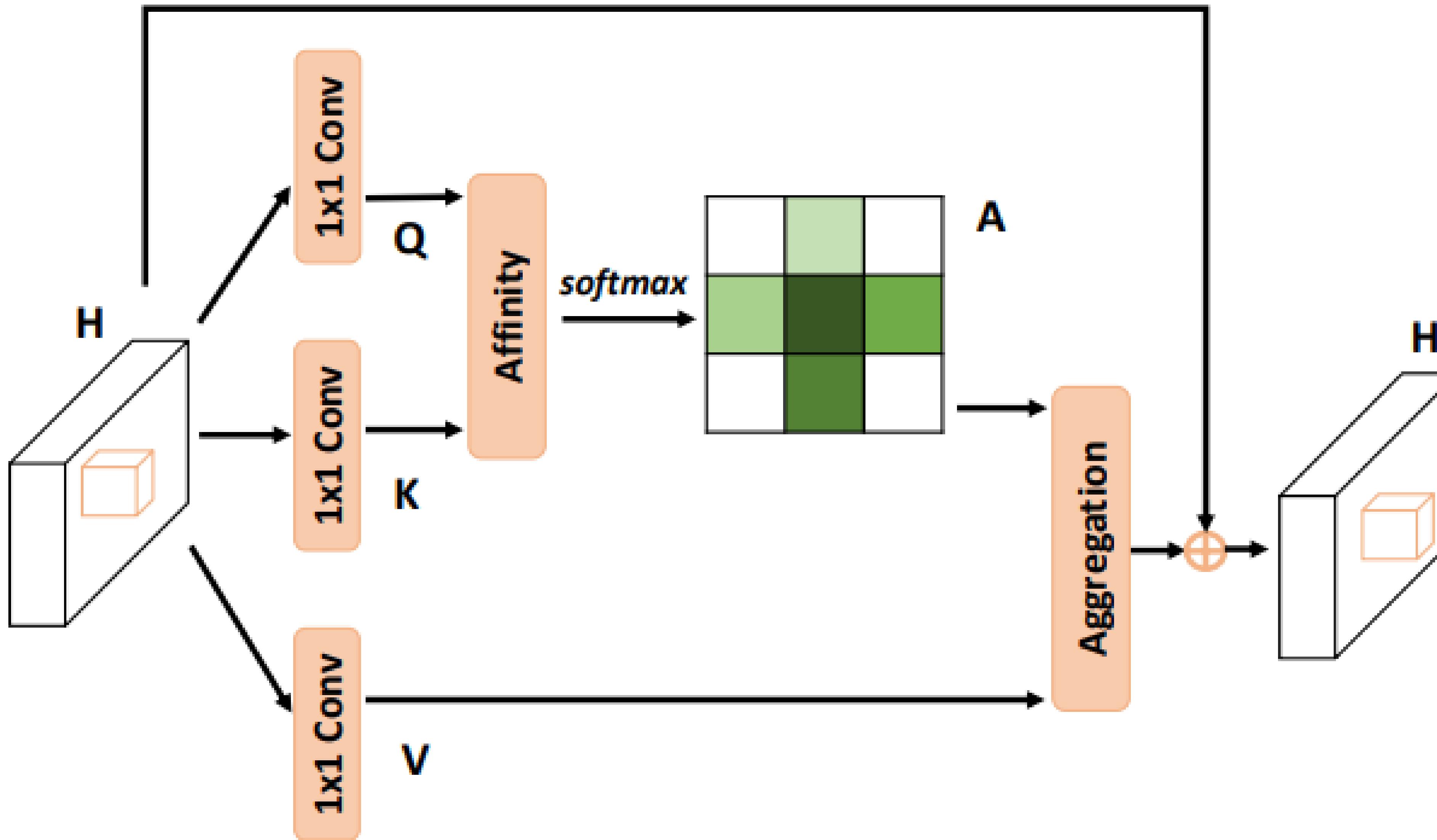


Criss-Cross Network (2019)

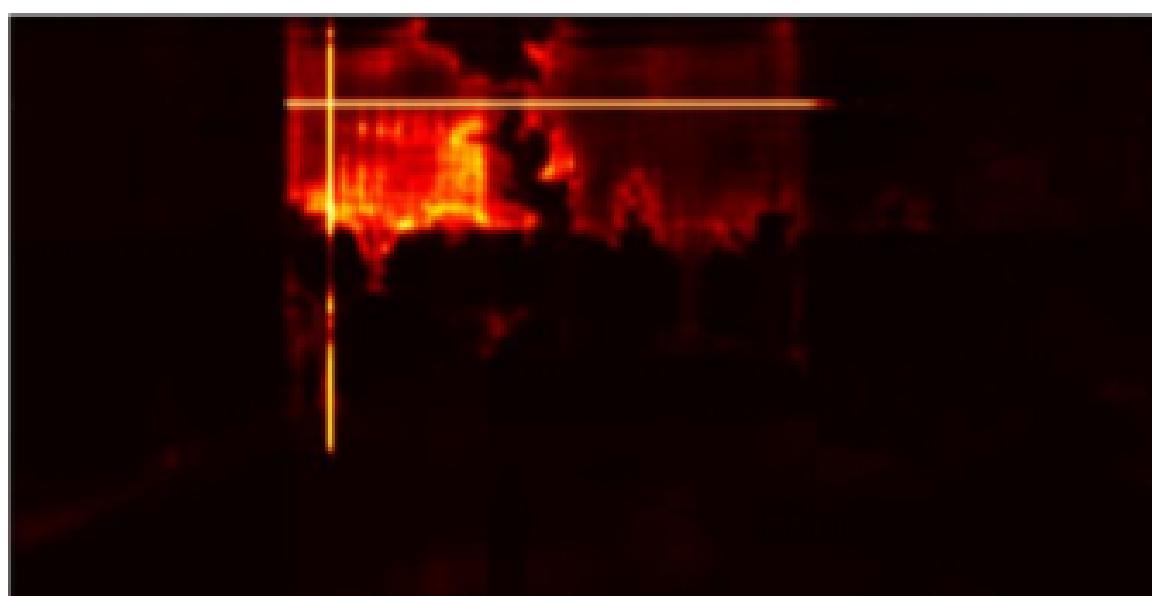
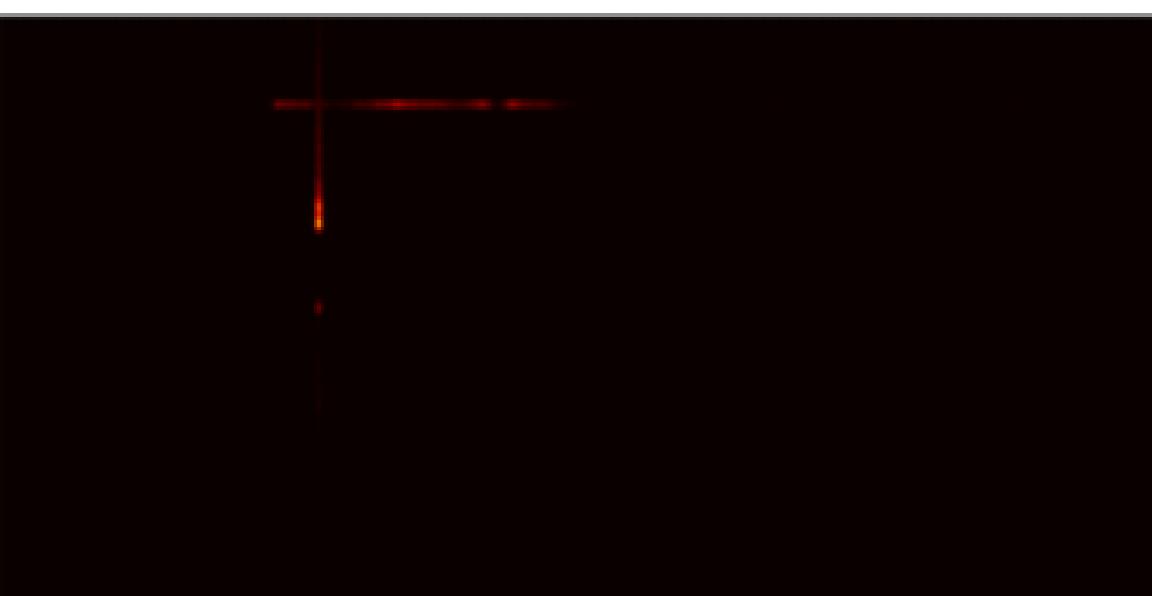
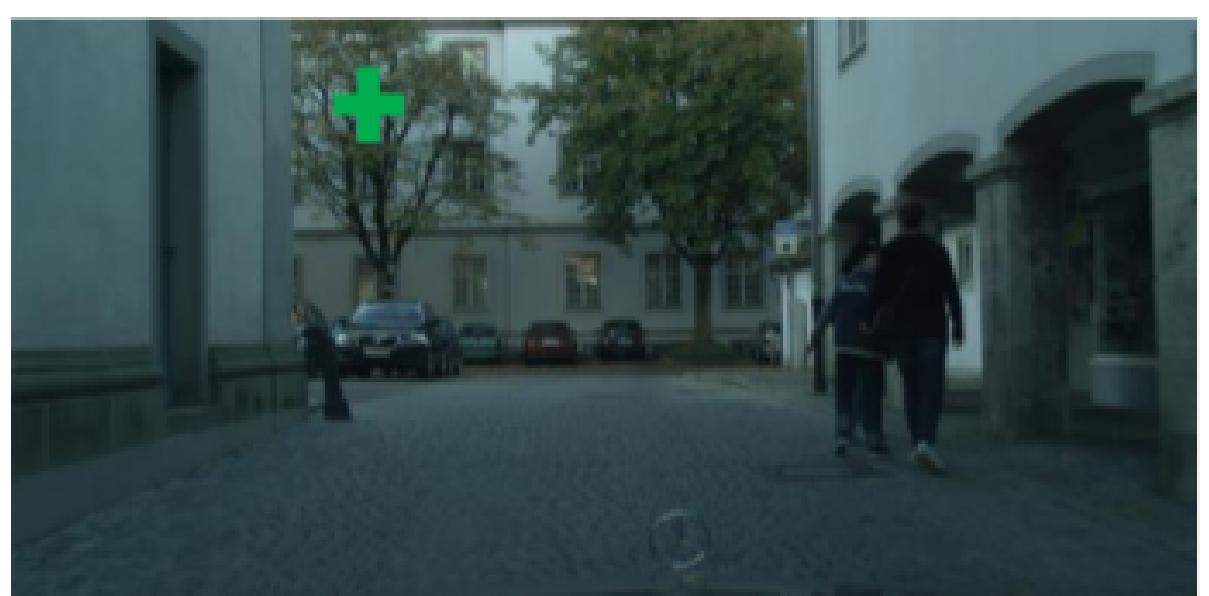
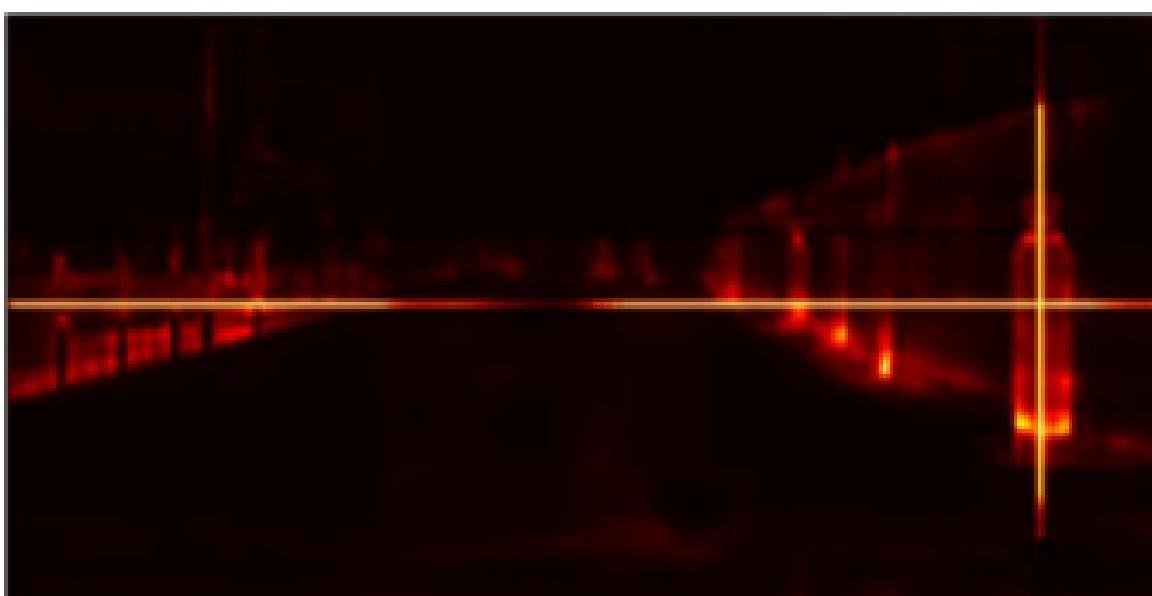
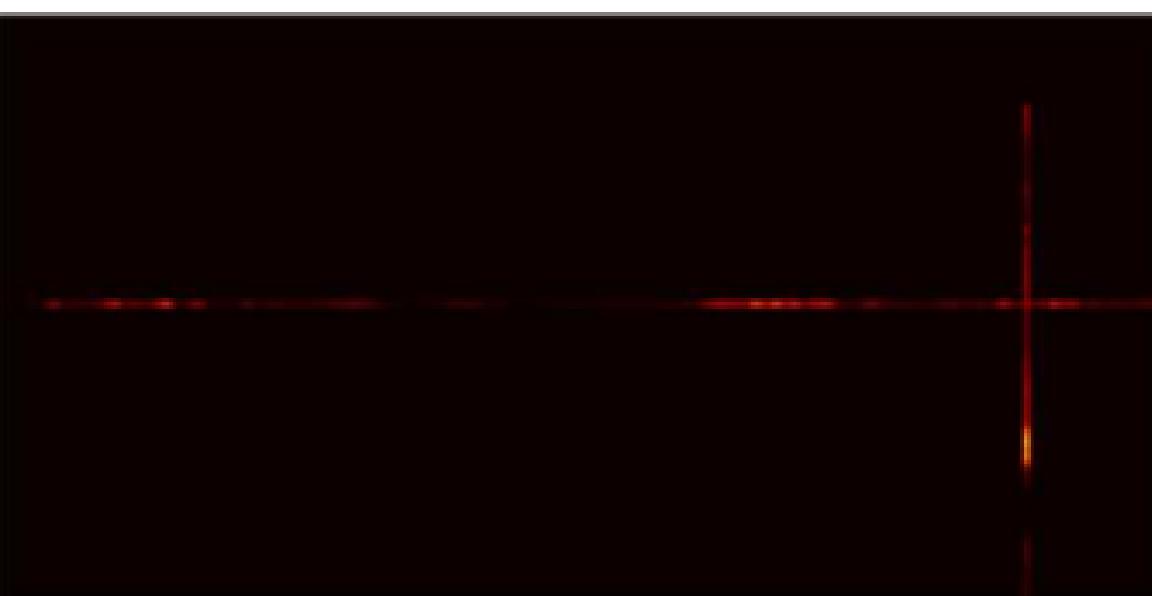
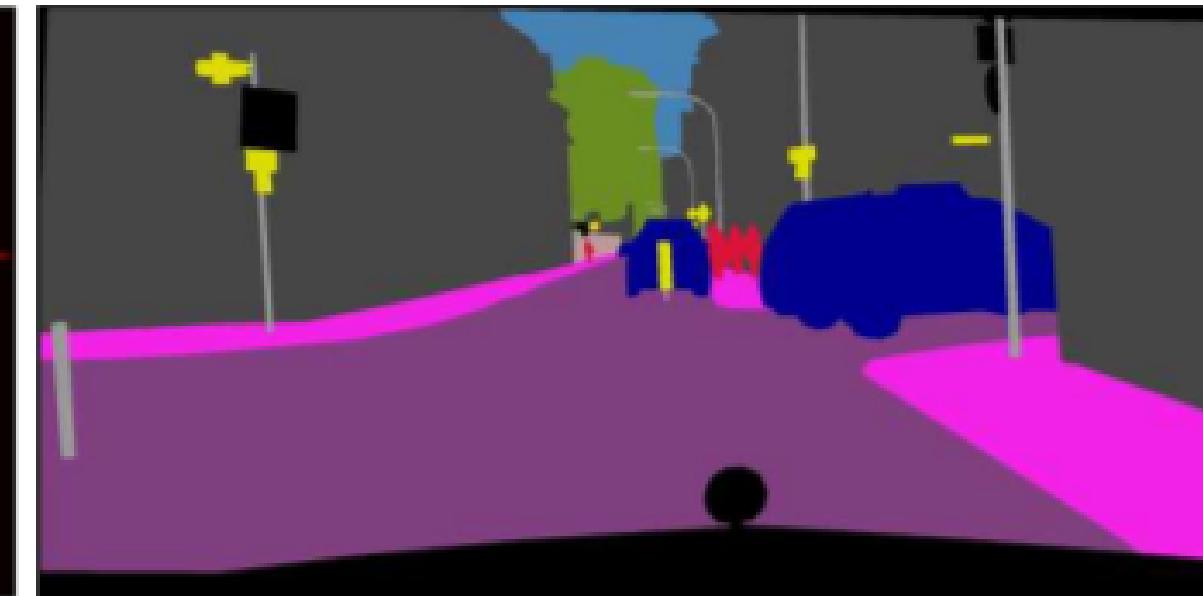
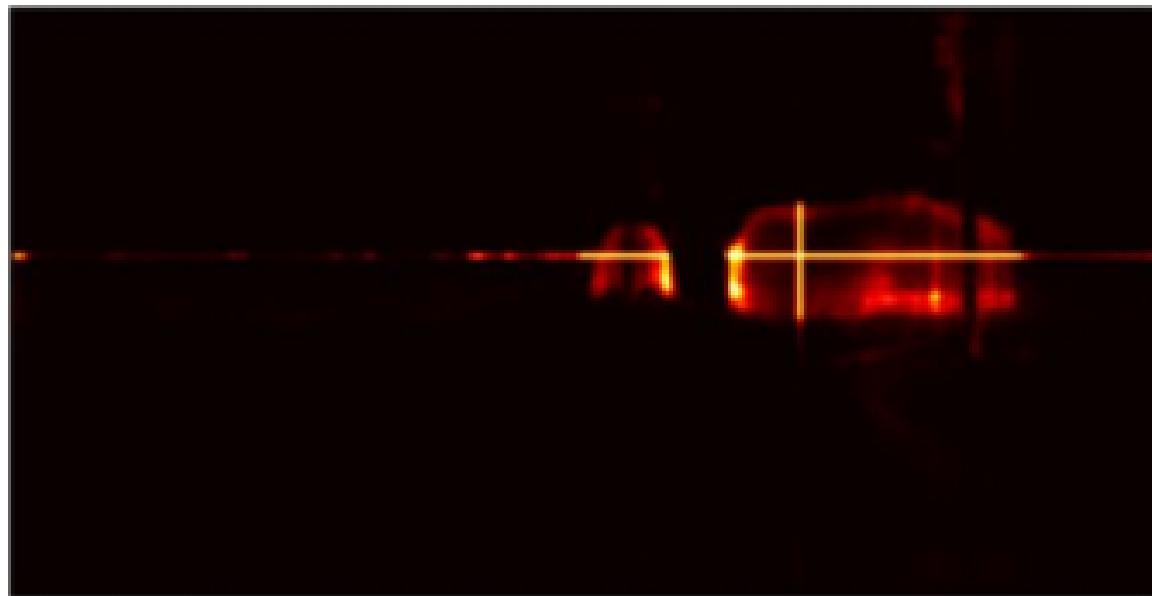
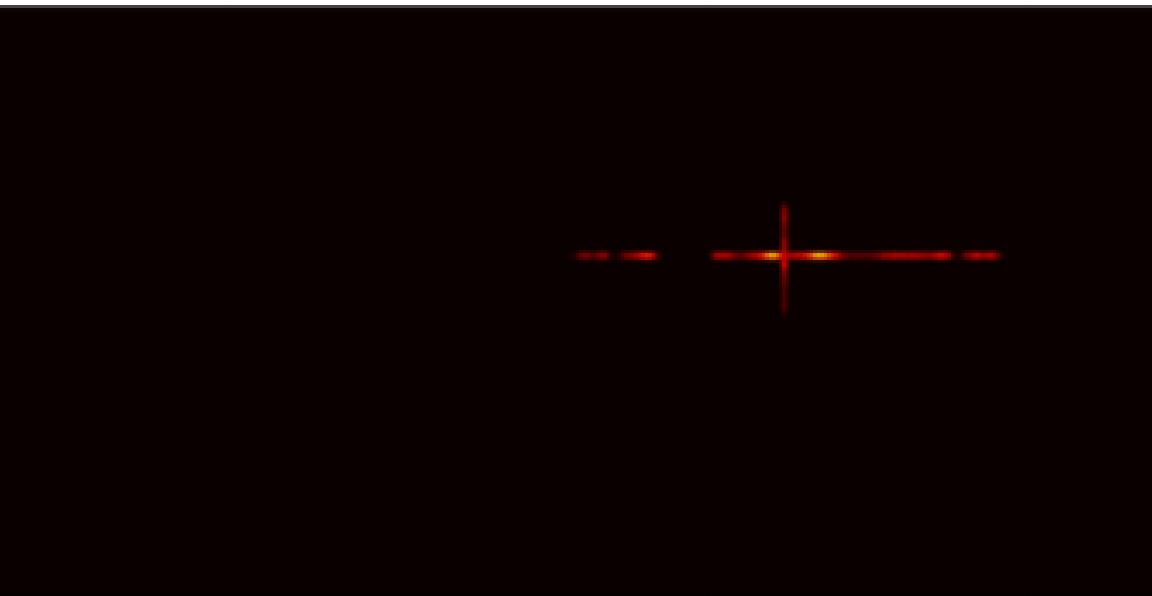
Compare all the queries
with all the keys having
same x or y



Criss-Cross Network (2019)



Criss-Cross Network (2019)



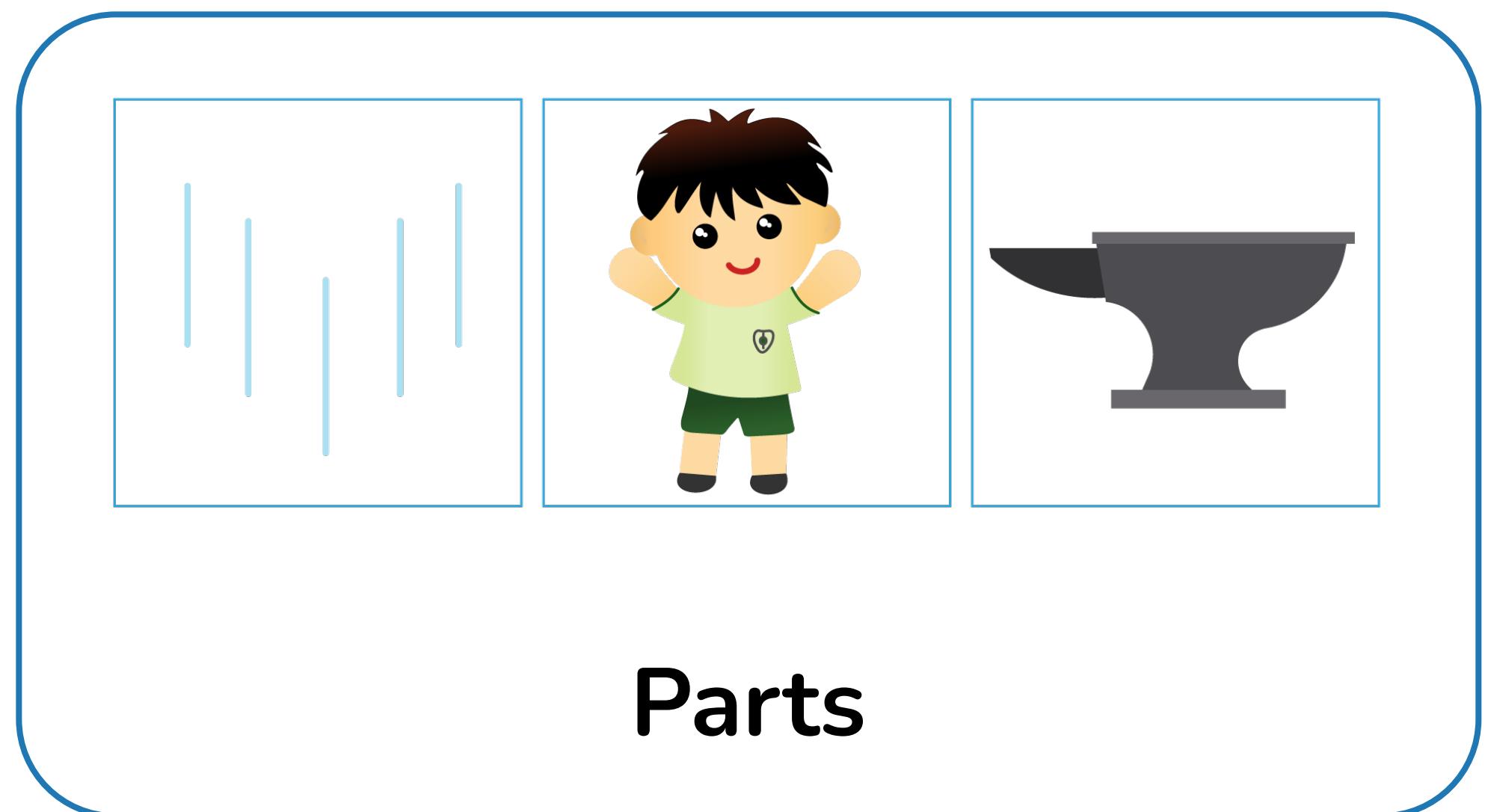
Image

R=1

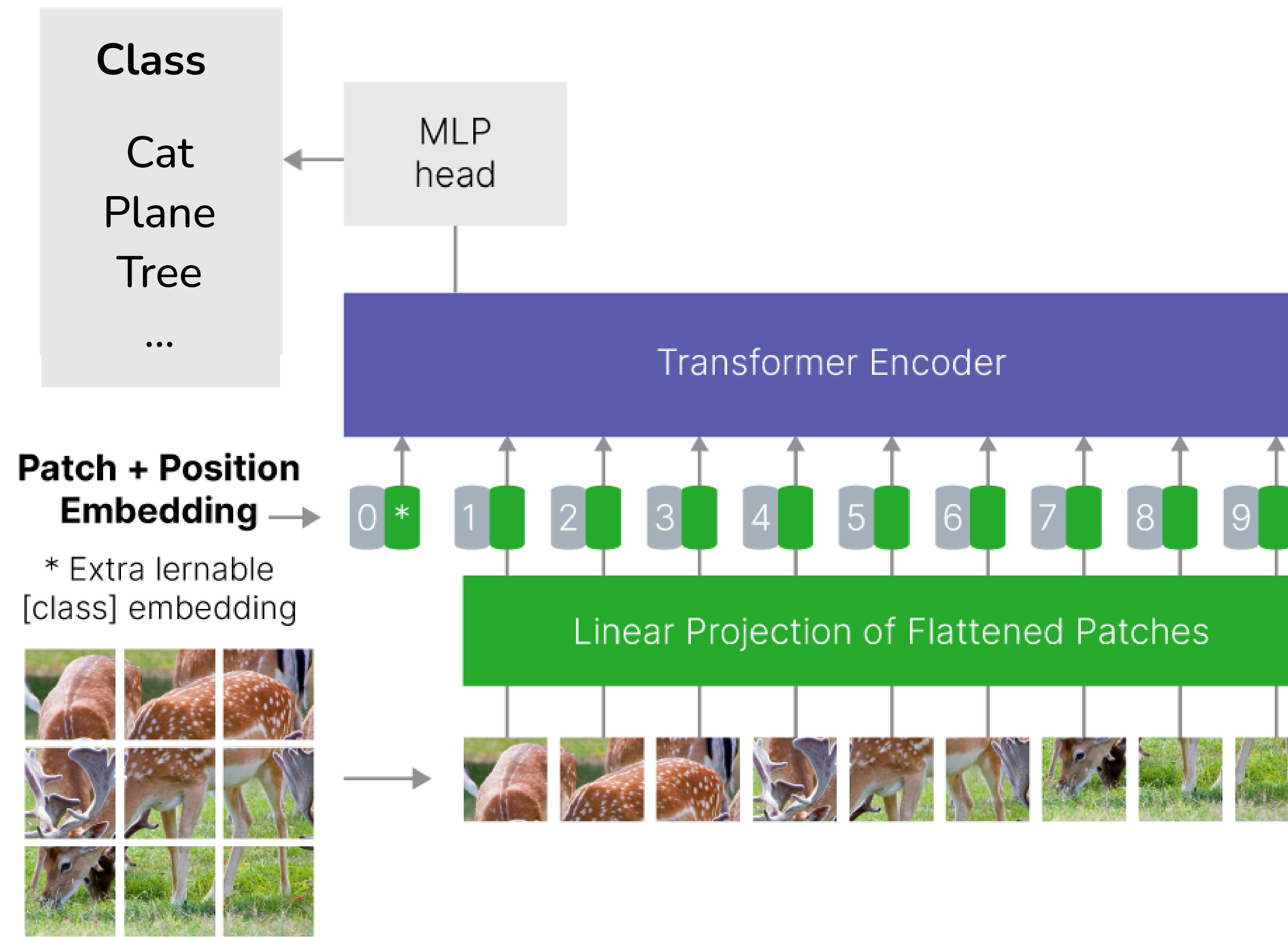
R=2

Ground Truth

Positional Encoding



Vision Transformer (2020)



Convolution-free model?

Positional Encoding for sequences

