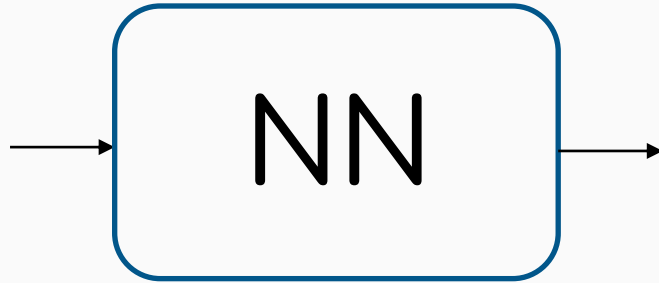


Lecture 6 : auto-encoders and semantic segmentation

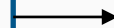
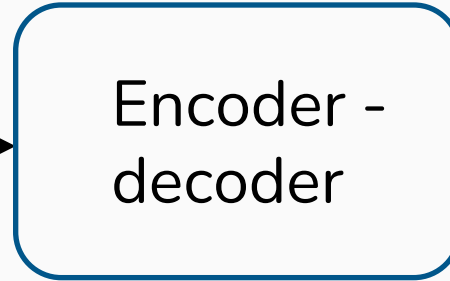
By : Khalil idrissi

The National
School of Applied
Sciences, Kenitra



Ecole Nationale
des Sciences
Appliquées de
Kénitra

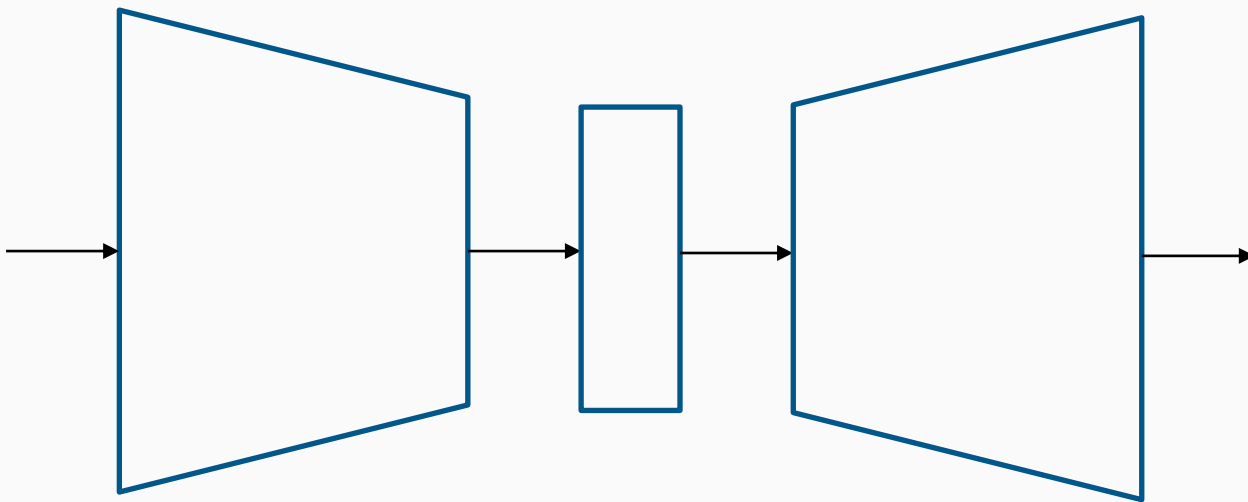
The National
School of Applied
Sciences, Kenitra



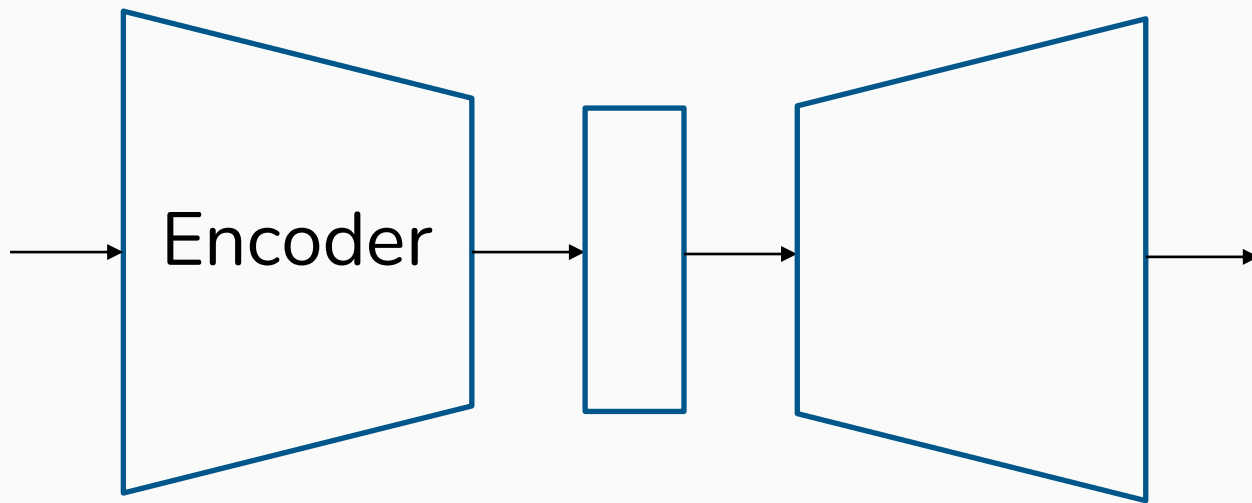
Ecole Nationale
des Sciences
Appliquées de
Kénitra

Encoder - decoder

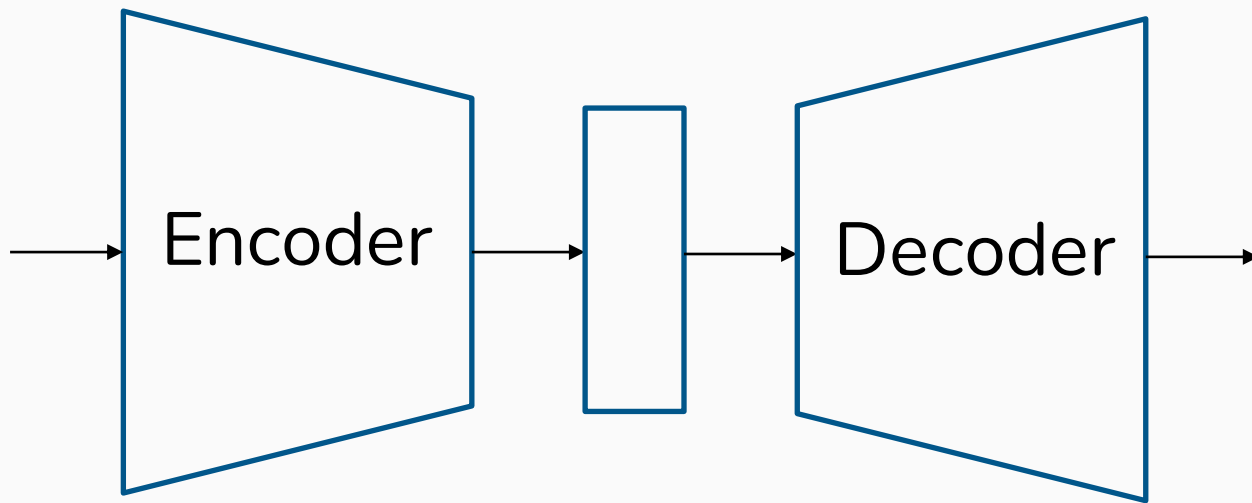
Encoder - decoder



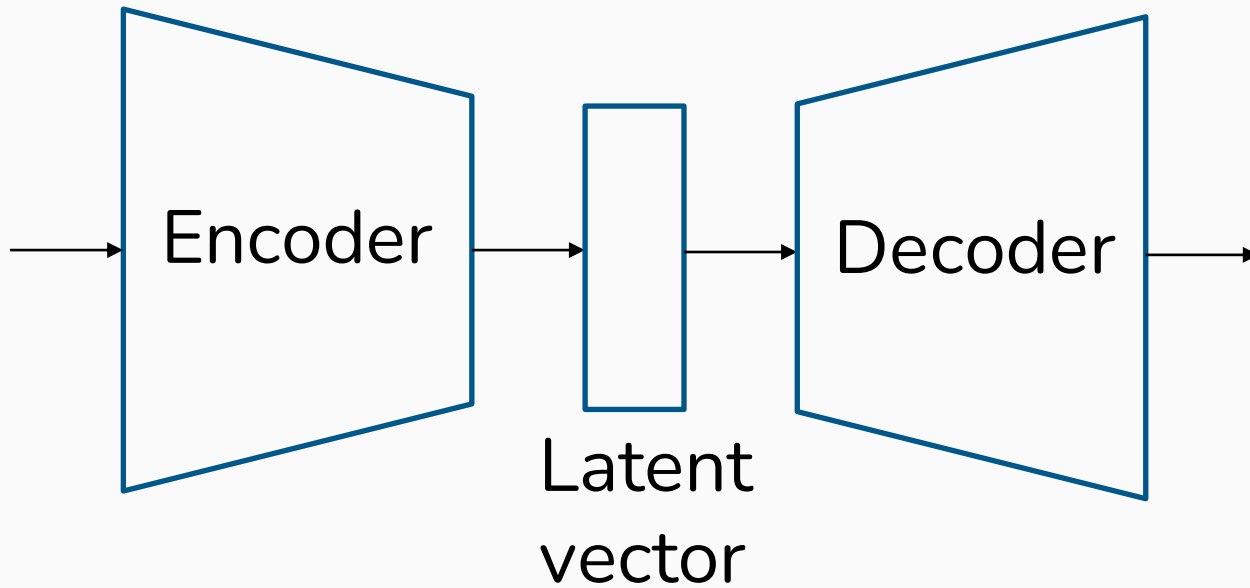
Encoder - decoder



Encoder - decoder

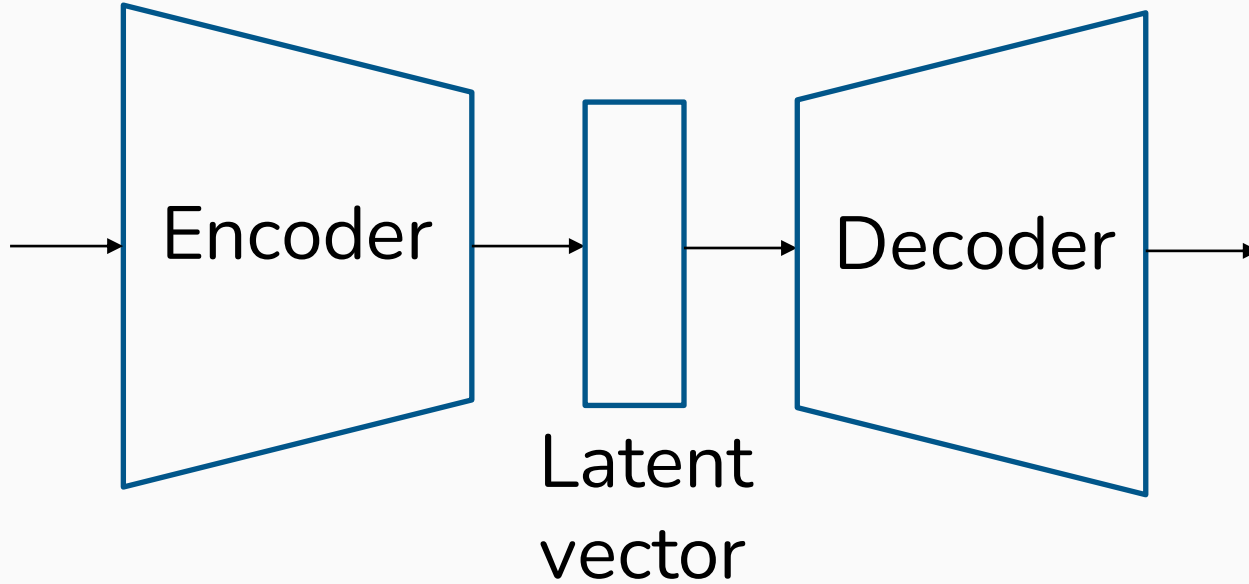


Encoder - decoder

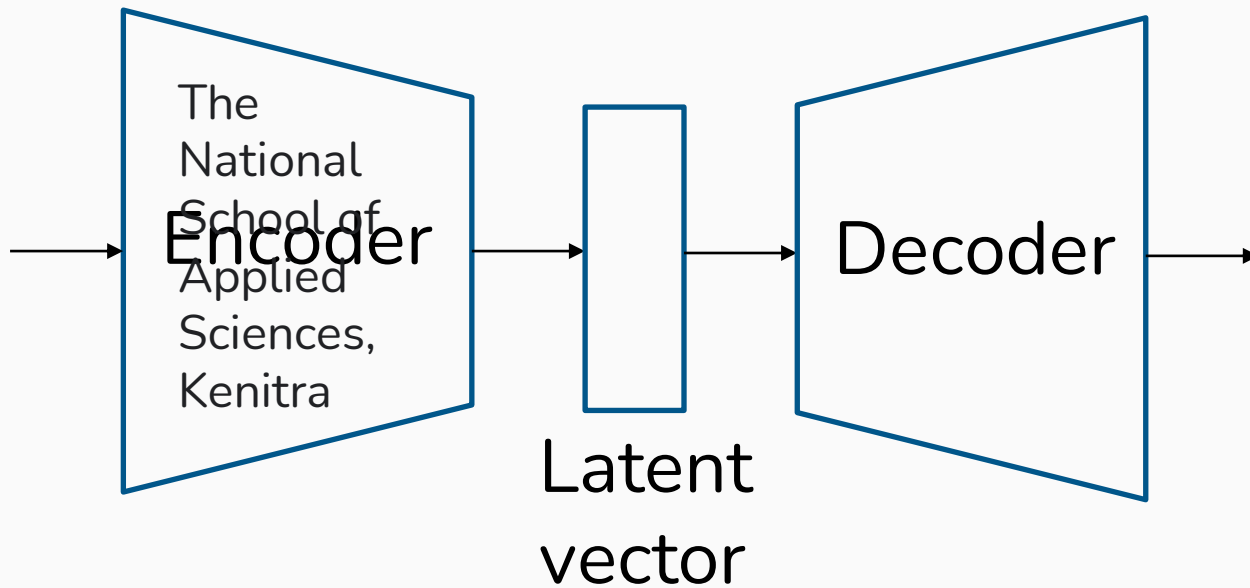


Encoder - decoder

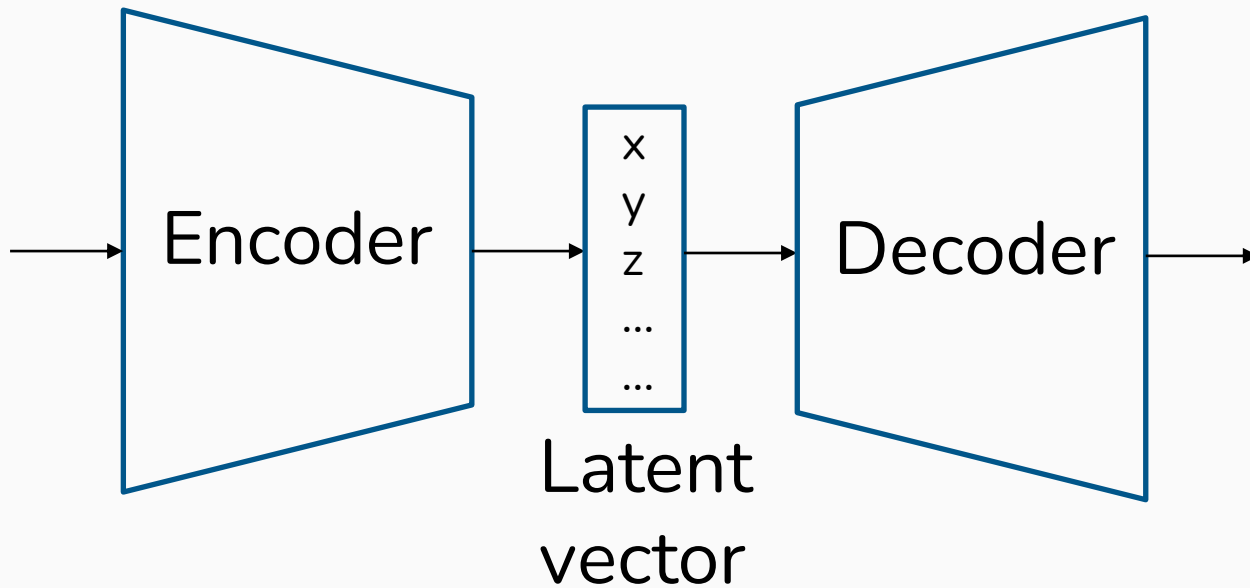
The
National
School of
Applied
Sciences,
Kenitra



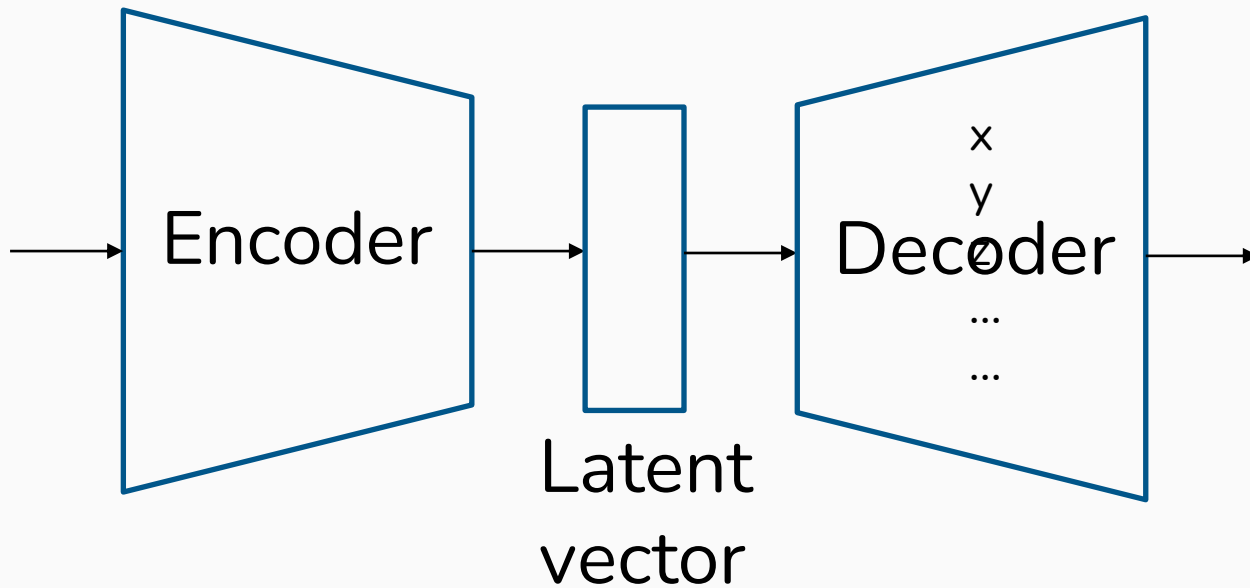
Encoder - decoder



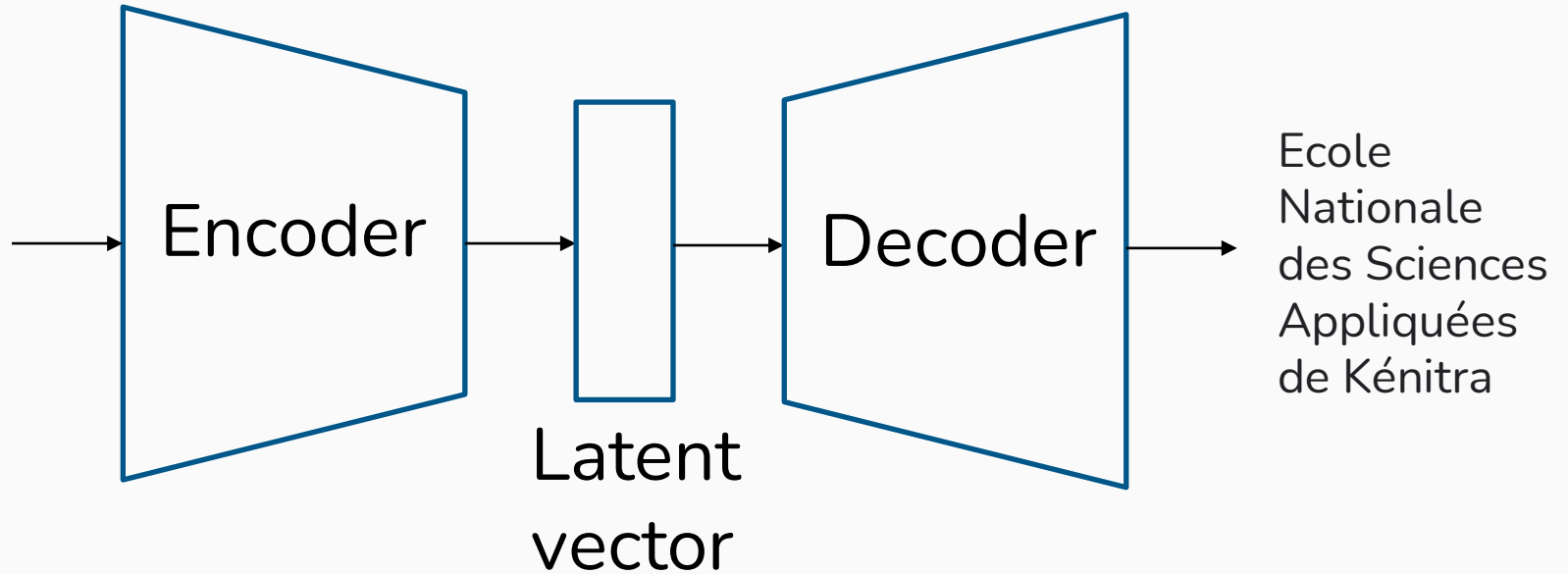
Encoder - decoder



Encoder - decoder



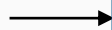
Encoder - decoder



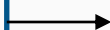
When
dim of input
=
dim of output

When
dim of input
=
dim of output

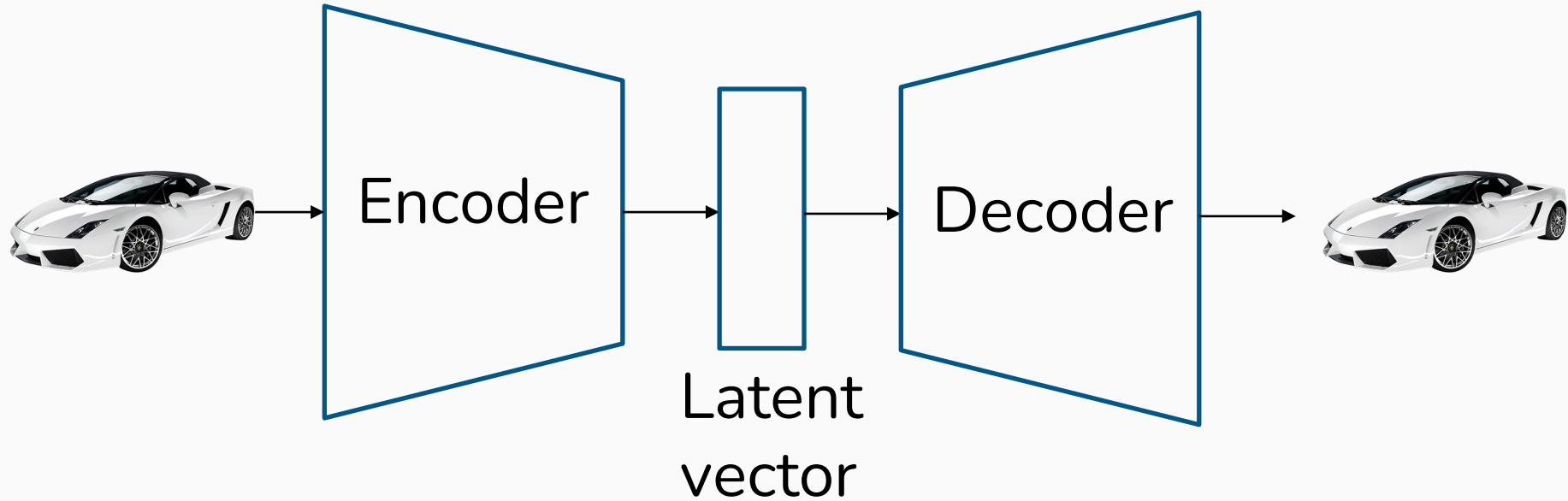
autoencoders



autoencoder

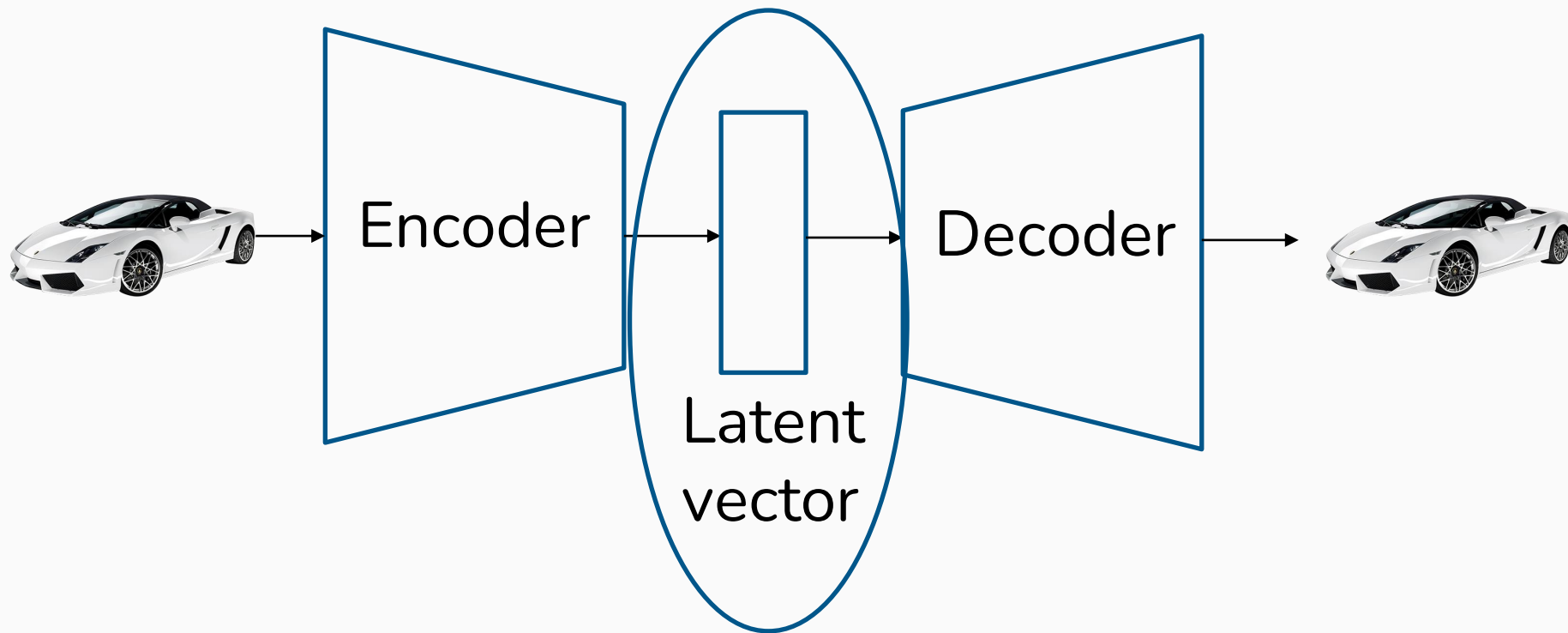


Encoder - decoder



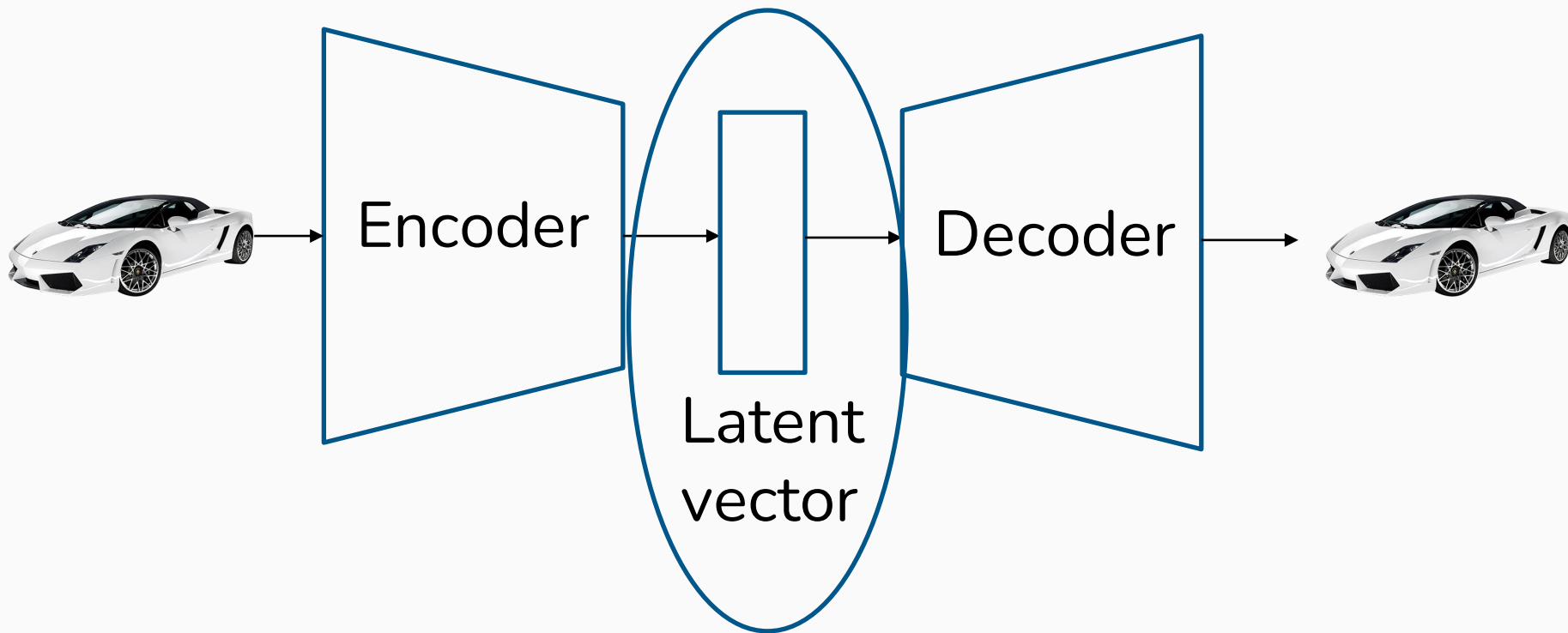
What's the
point

Encoder - decoder



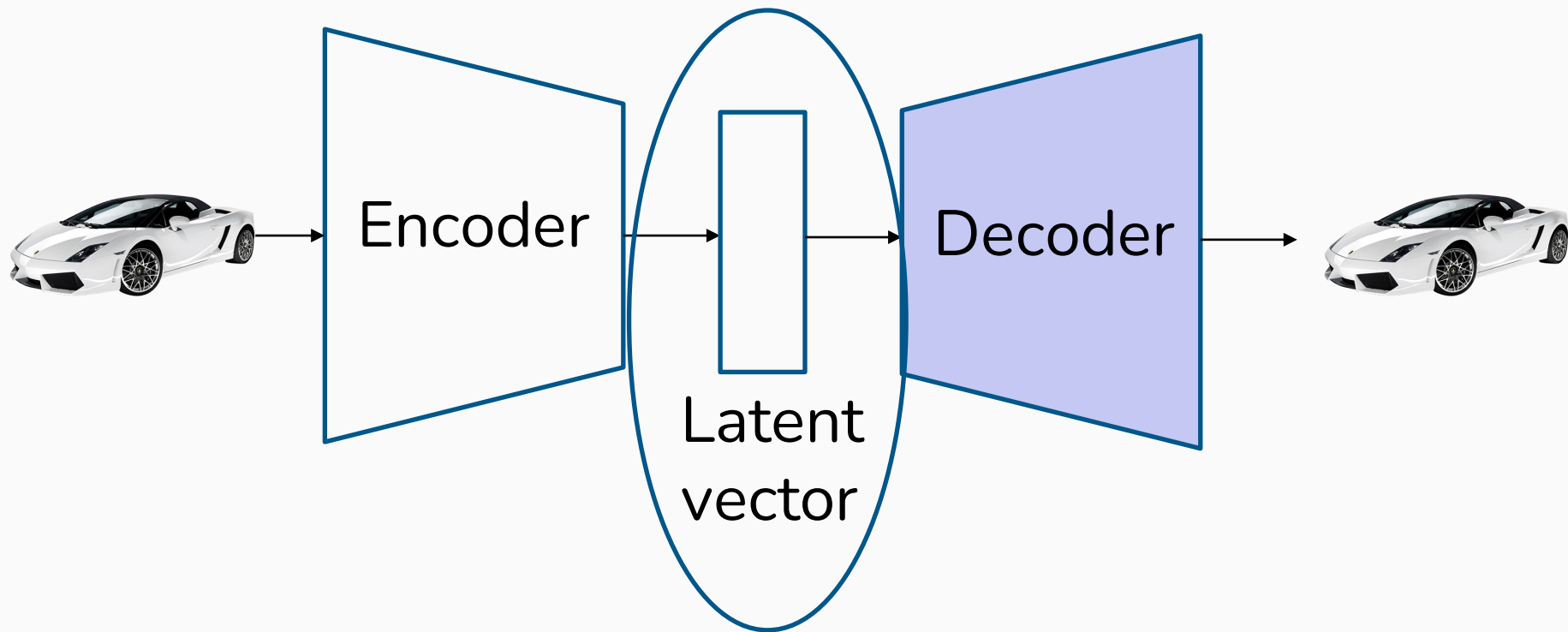
Captures the most relevant informations

Encoder - decoder



Captures the most relevant informations

Encoder - decoder



Captures the most relevant informations

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

1	3	2
0	-1	2
1	-2	1

 $*$

1	0	0
-1	0	1
0	-1	1

 $=$

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

1	3	2
0	-1	2
1	-2	1

 $*$

1	0	0
-1	0	1
0	-1	1

 $=$

2	-5
-2	1

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

1	3	2
0	-1	2
1	-2	1

 $*$

1	0	0
-1	0	1
0	-1	1

 $=$

2	-5
-2	1

2	-5
-2	1

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

1	3	2
0	-1	2
1	-2	1

 \ast

1	0	0
-1	0	1
0	-1	1

 $=$

2	-5
-2	1

2	-5
-2	1

 \rightarrow

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

1	3	2
0	-1	2
1	-2	1

*

1	0	0
-1	0	1
0	-1	1

=

2	-5
-2	1

2	-5
-2	1

→

2	0	-5
0	0	0
-2	0	1

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

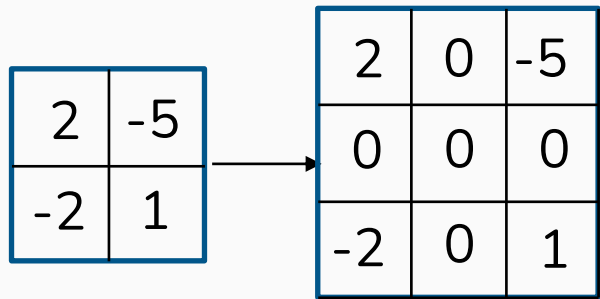
1	3	2
0	-1	2
1	-2	1

 \ast

1	0	0
-1	0	1
0	-1	1

 $=$

2	-5
-2	1



Atrous convolution/dilated

$$p = 1, s = 2, k = 3, H = 3, W = 3$$

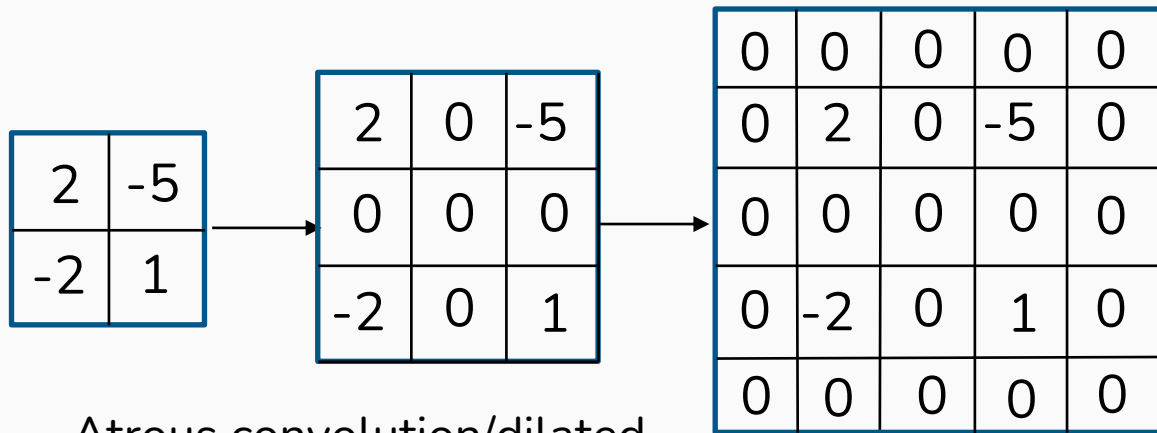
1	3	2
0	-1	2
1	-2	1

*

1	0	0
-1	0	1
0	-1	1

=

2	-5
-2	1



Atrous convolution/dilated

0	0	0	0	0
0	2	0	-5	0
0	0	0	0	0
0	-2	0	1	0
0	0	0	0	0

*

1	0	0
-1	0	1
0	-1	1

=

0	7	0
-2	3	5
0	-3	0

0	0	0	0	0
0	2	0	-5	0
0	0	0	0	0
0	-2	0	1	0
0	0	0	0	0

*

1	0	0
-1	0	1
0	-1	1

=

0	7	0
-2	3	5
0	-3	0

This is transposed convolution

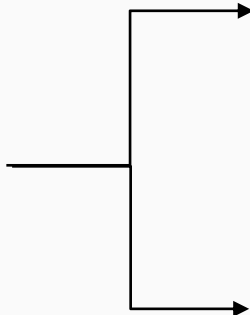
Max pooling :

Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1

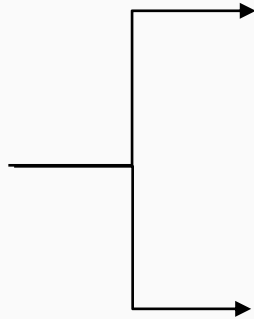
Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1



Max pooling :

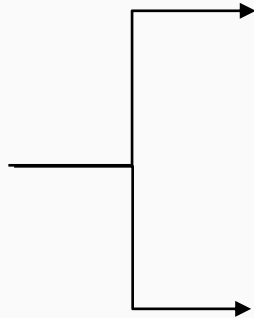
0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1



2	-5
6	3

Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1

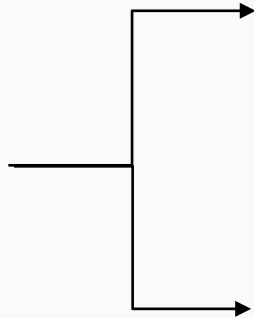


0	1	0	0
0	0	0	1
1	0	1	0
0	0	0	0

2	-5
6	3

Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1

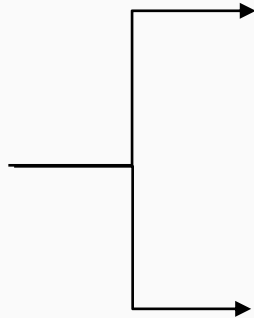


0	1	0	0
0	0	0	1
1	0	1	0
0	0	0	0

2	-5
6	3

Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1



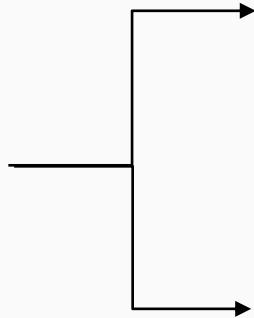
0	1	0	0
0	0	0	1
1	0	1	0
0	0	0	0

2	-5
6	3

Unpooling :

Max pooling :

0	2	3	0
0	0	0	5
6	0	3	0
0	-2	2	1



0	1	0	0
0	0	0	1
1	0	1	0
0	0	0	0

2	-5
6	3

Unpooling :

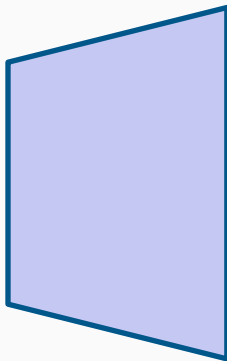
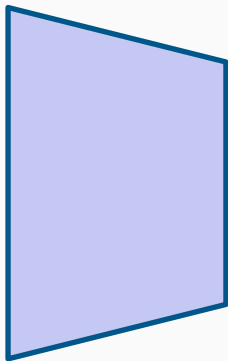
0	2	0	0
0	0	0	-5
6	0	3	0
0	0	0	0

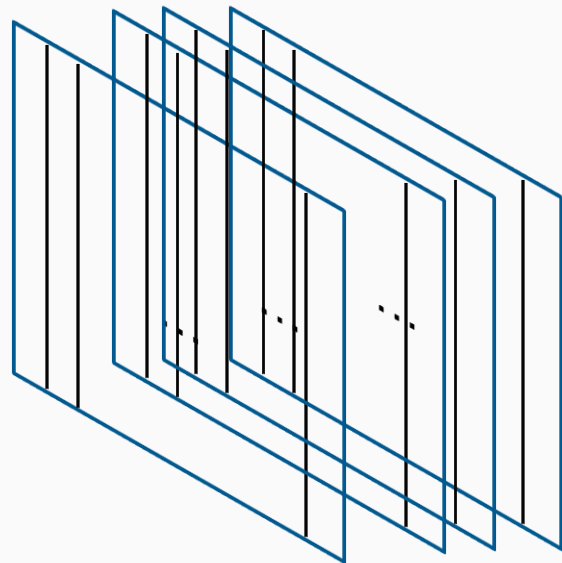
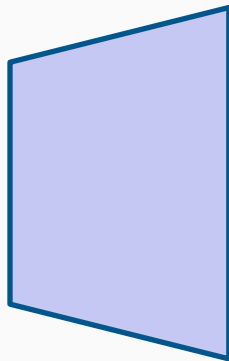
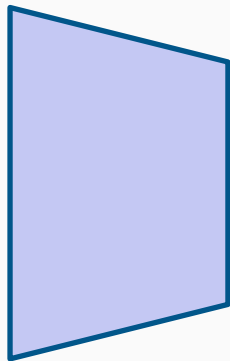
Semantic segmentation

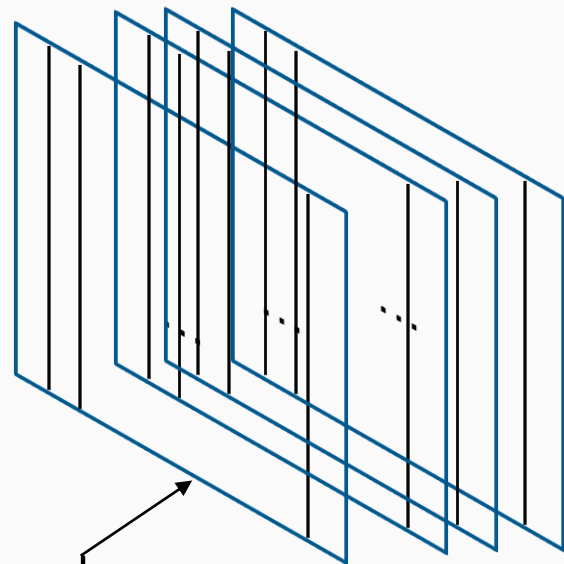
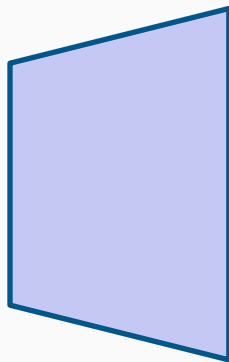
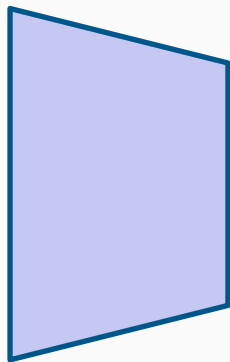
Labeling each pixel into a category



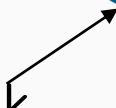


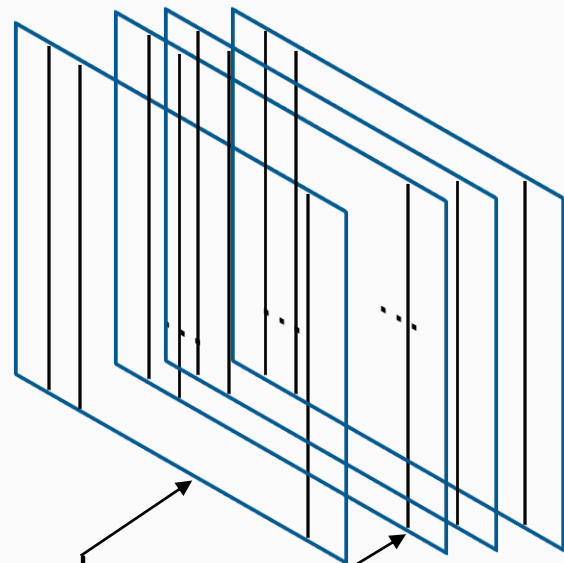
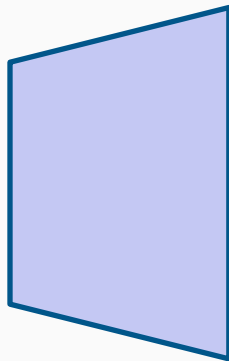
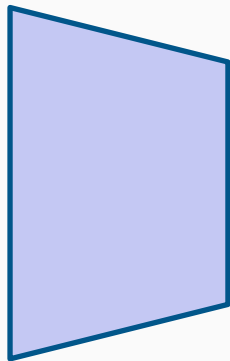






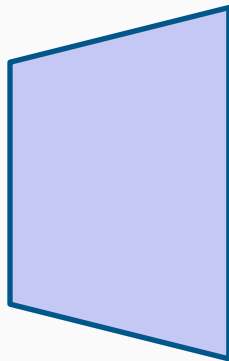
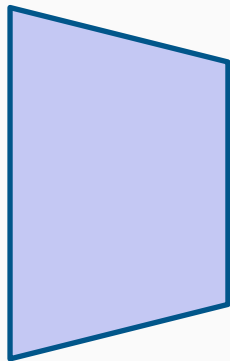
Cat mask





Cat mask

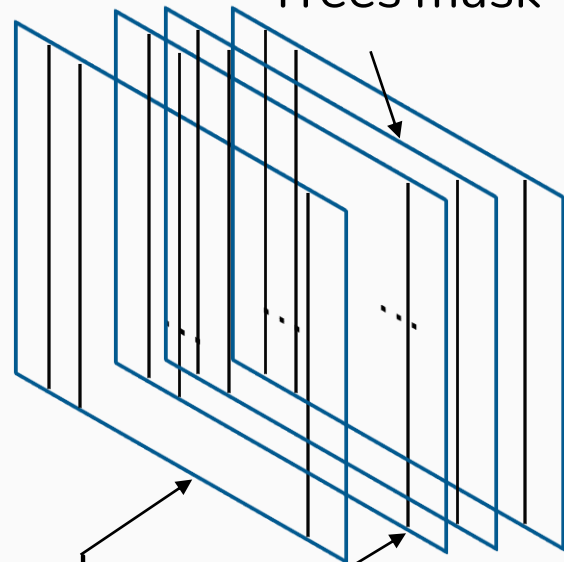
Sky mask

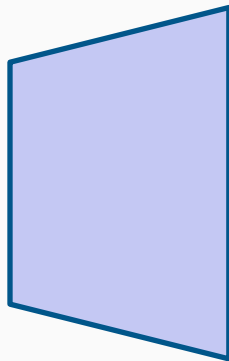
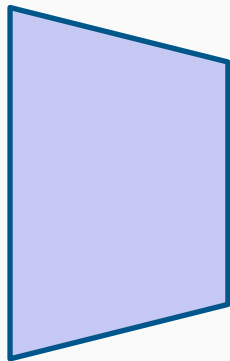


Cat mask

Sky mask

Trees mask



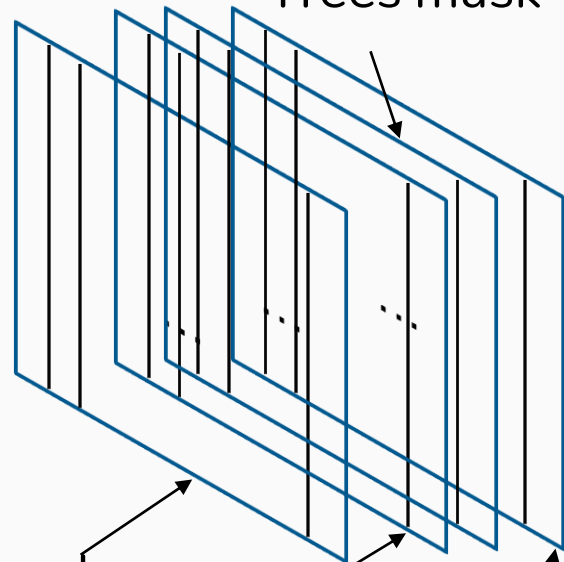


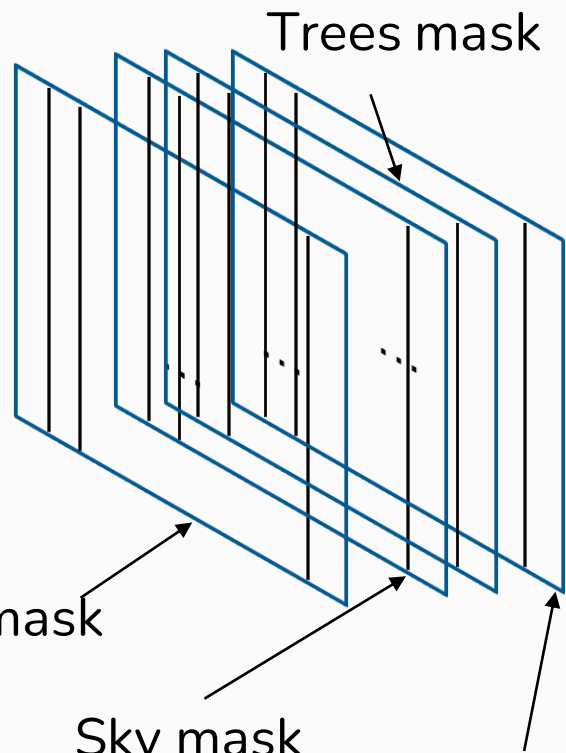
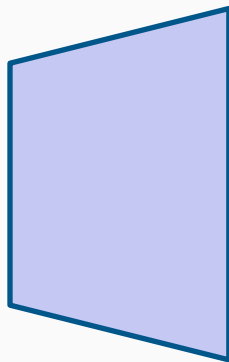
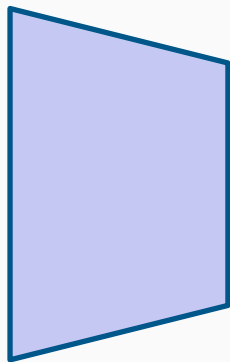
Cat mask

Sky mask

grass mask

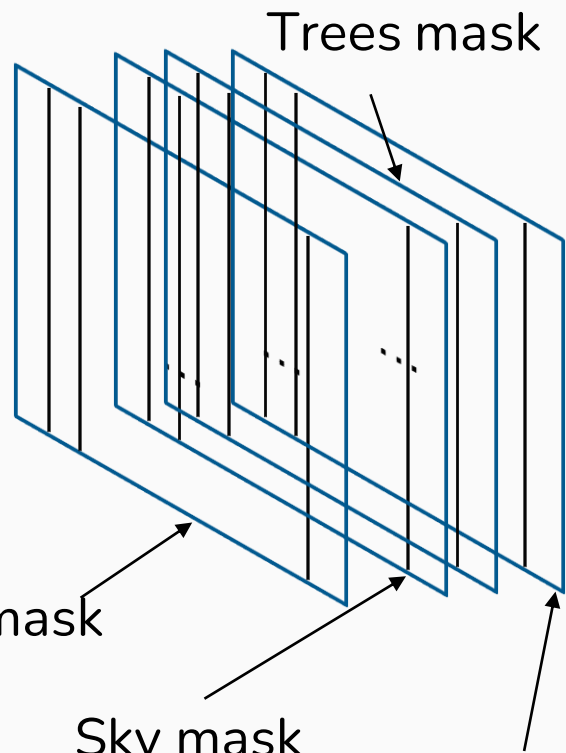
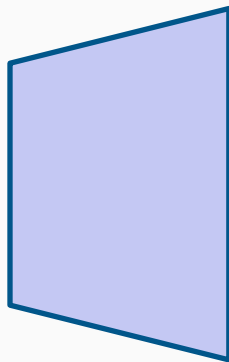
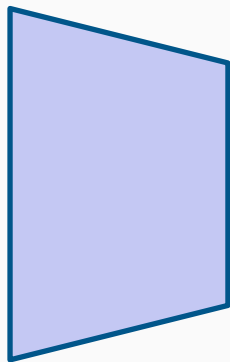
Trees mask



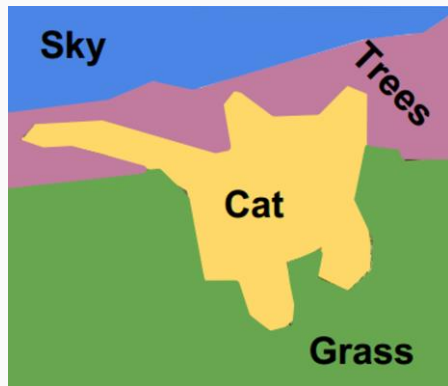


argmax





argmax



Cost function ?

Cross entropy

Cross entropy

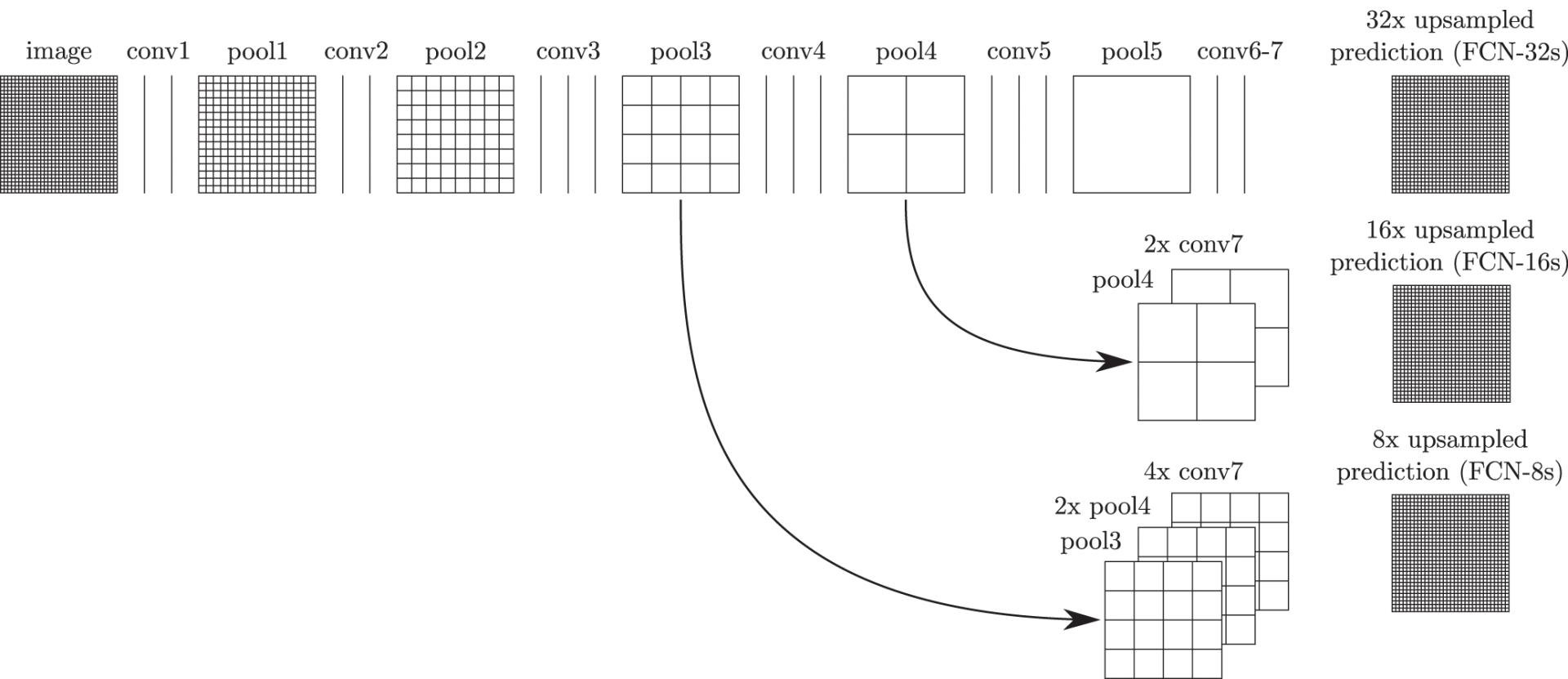
$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

Cross entropy

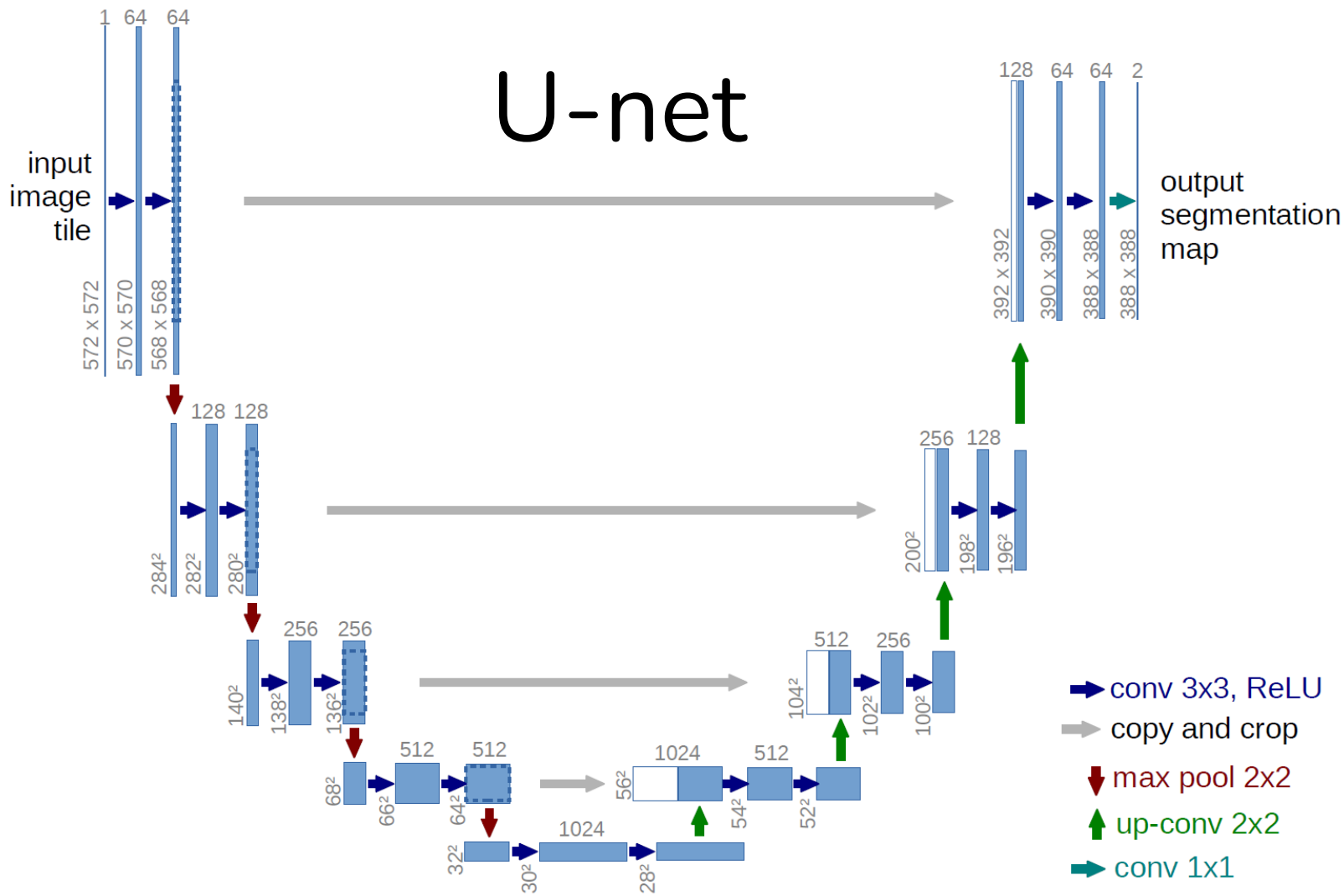
$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

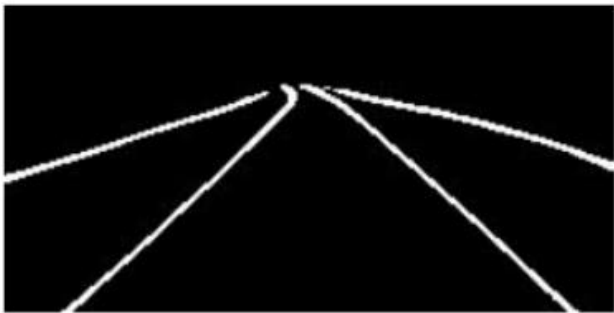
$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|}$$

FCN



U-net





end