VINI CARIDÁ

Human-Centered Data & AI

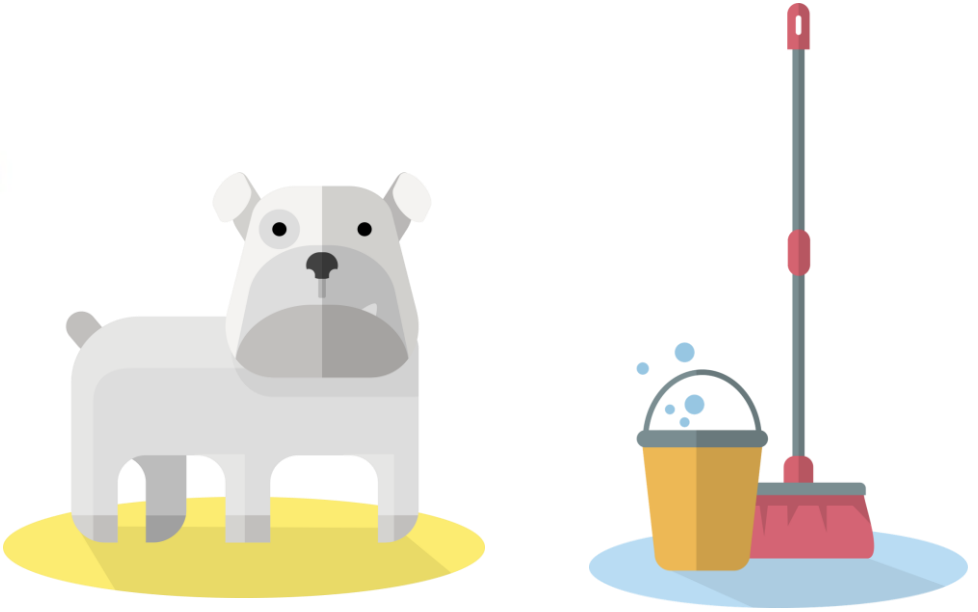
VINICARIDÁ
Human-Centered Data & AI

Vinicius Caridá, Ph.D.

- Executive Specialist, Artificial Intelligence and Data - Itaú
- MBA Professor – FIAP and ESPM

@vinicius caridá @vfcarida @vinicius caridá @vfcarida @vinicius caridá @vfcarida

What about a dog and a mop? Easy, right?

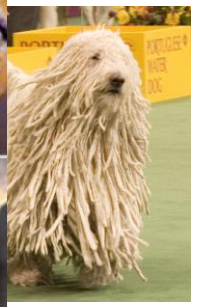


Not so fast...

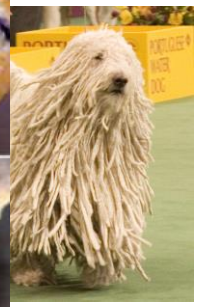




fast...



fast...



Make or Buy

API

Property	88%
House	87%
Architecture	85%
Home	81%

Landmark	94%
Sky	93%
House	89%
Building	89%

House	94%
Home	92%
Property	92%
Real Estate	80%

Property	92%
Building	84%
Town	84%
House	83%

Custom

Tudor	.5
Neoclassical	.1
Modern	**.89**
Ranch	.5

Tudor	.1
Victorian	**.94**
Modern	.3
Ranch	.1

Tudor	.2
Neoclassical	.3
Modern	.3
Ranch	**.93**

Tudor	**.92**
Neoclassical	.4
Modern	.2
Ranch	.2

VINICARIDÁ
Human-Centered Data & AI

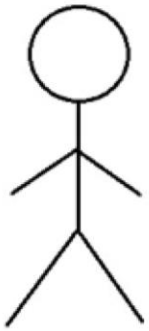
VINI CARIDÁ
Human-Centered Data & AI

“

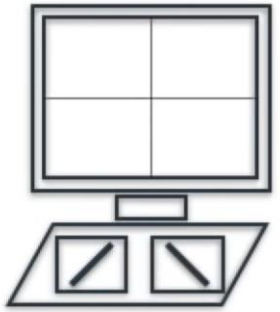
Convolution Neural Network (CNN)

Convolution Neural Network (CNN)

A simple world



Person

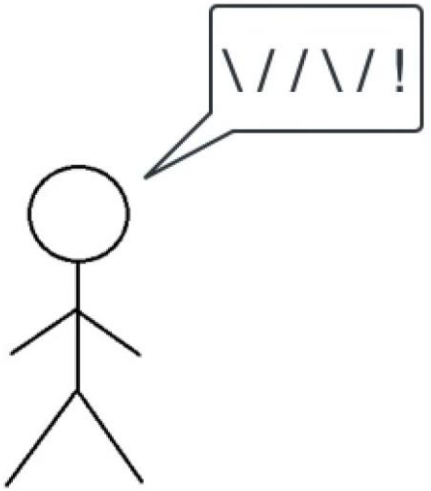


Computer

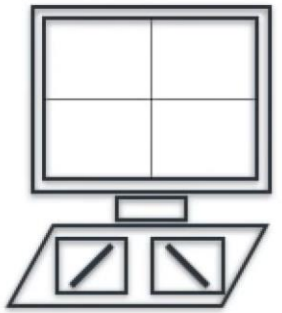
Alphabet

Convolution Neural Network (CNN)

A simple world



Person

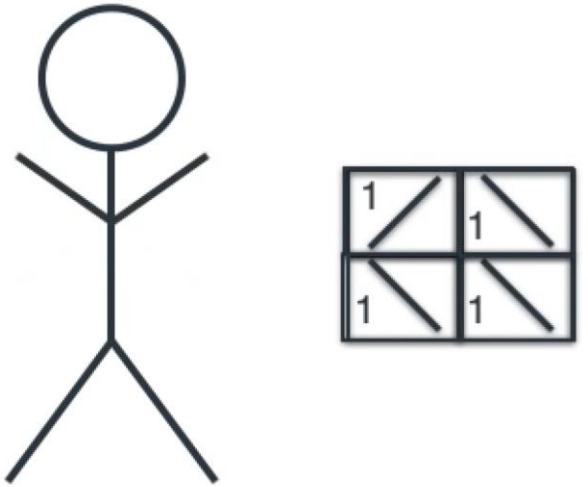


Computer

Alphabet

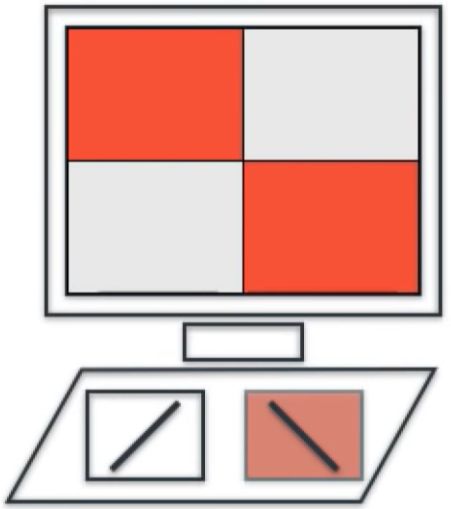
Convolution Neural Network (CNN)

A simple world



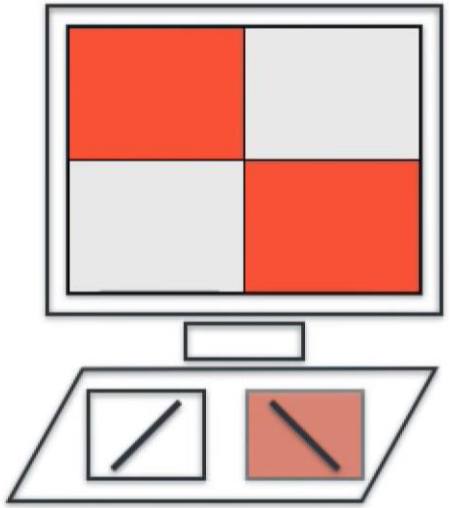
Convolution Neural Network (CNN)

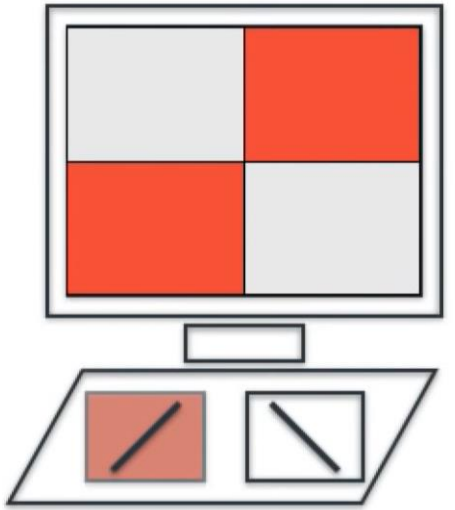
A simple world



Convolution Neural Network (CNN)

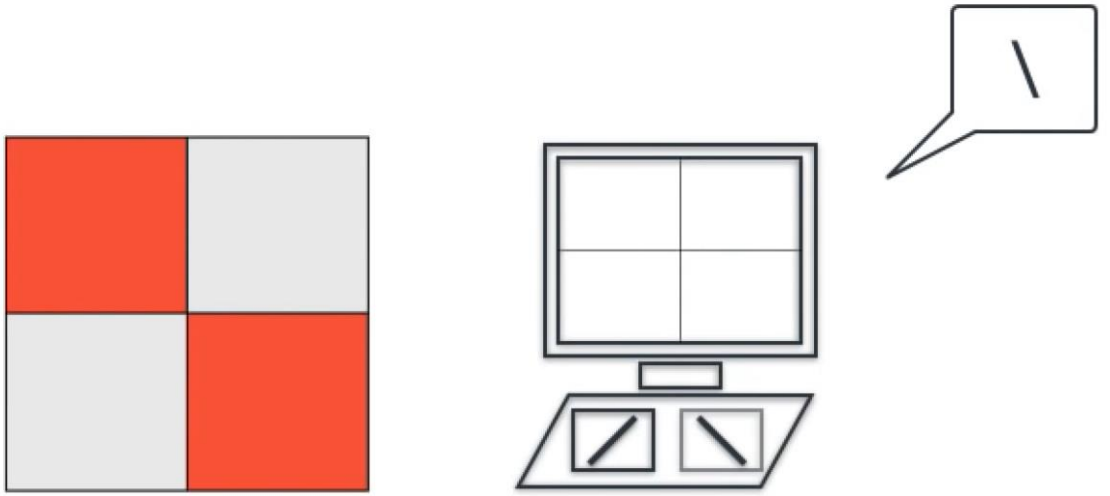
A simple world





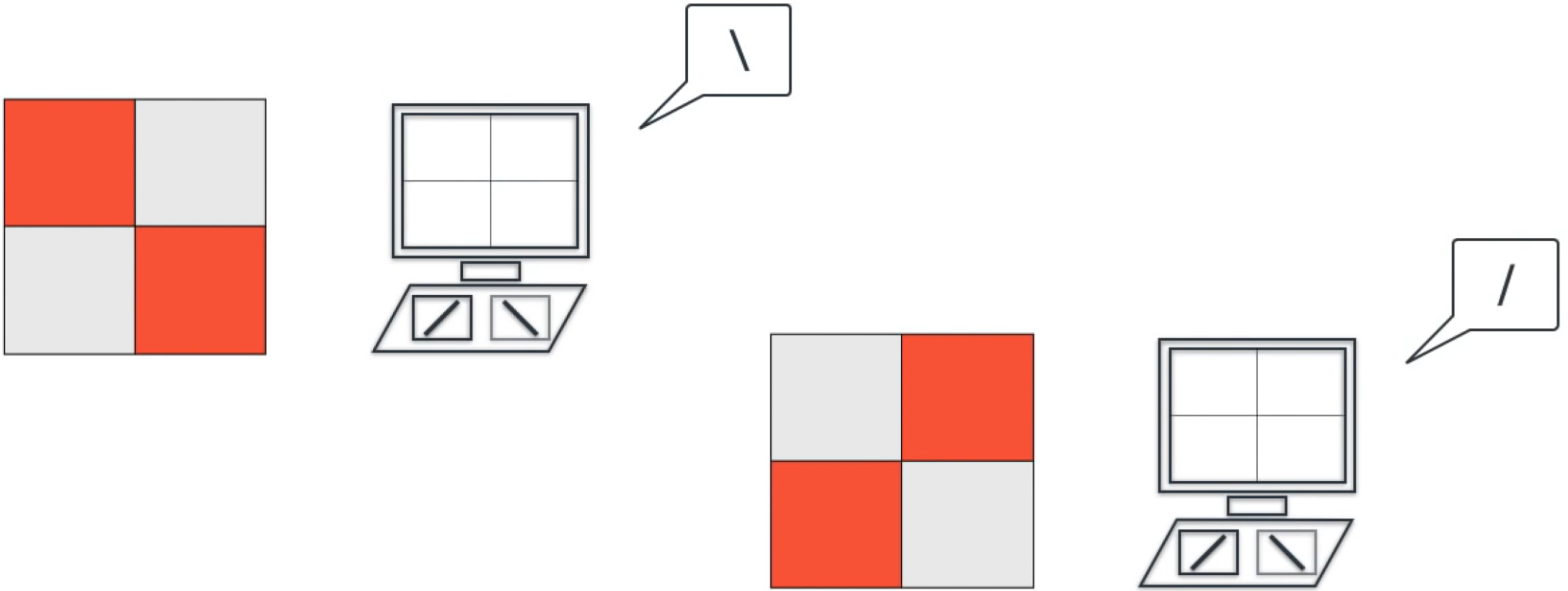
Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

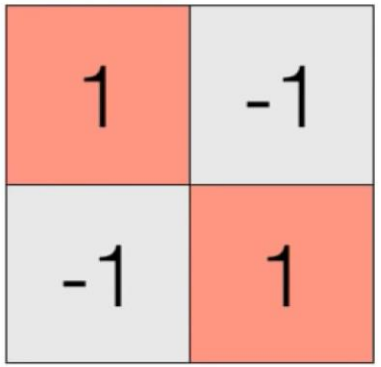
A simple world



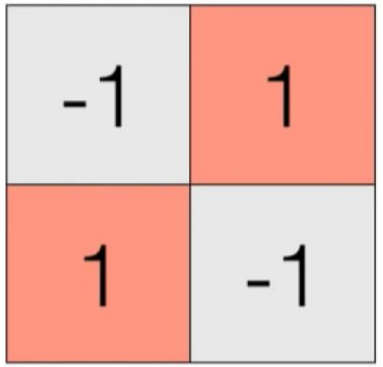
Convolution Neural Network (CNN)

A simple world

\



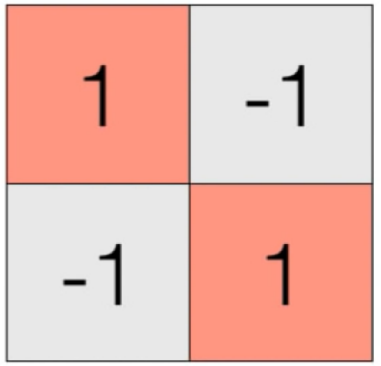
/



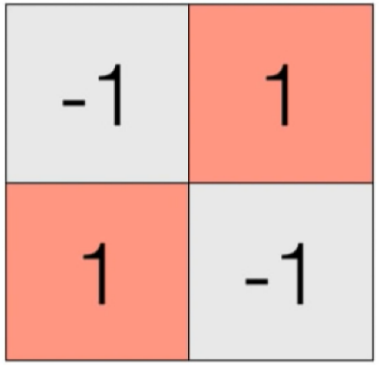
Convolution Neural Network (CNN)

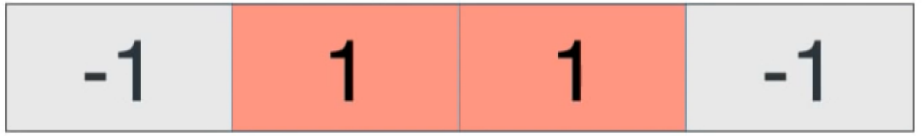
A simple world

\



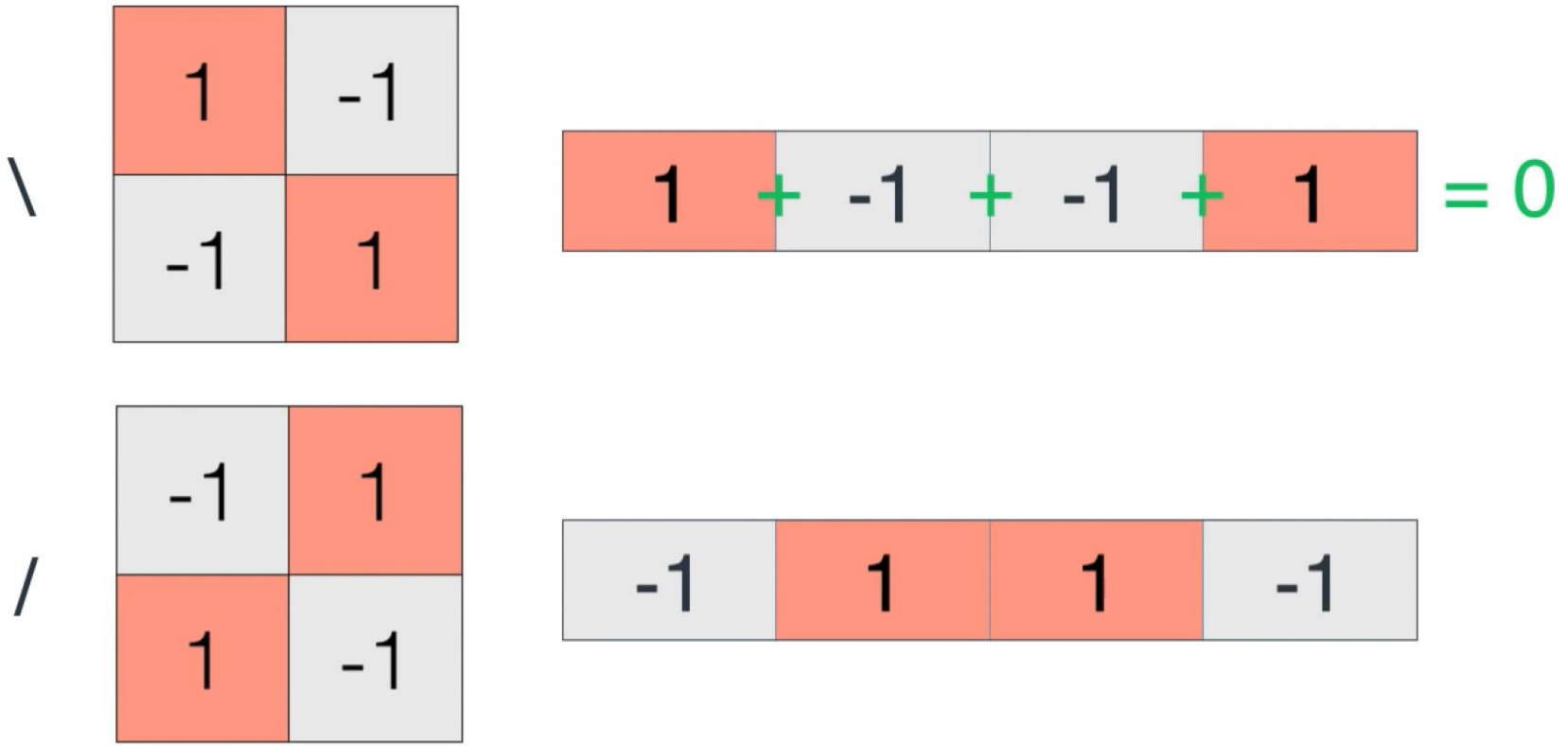
/





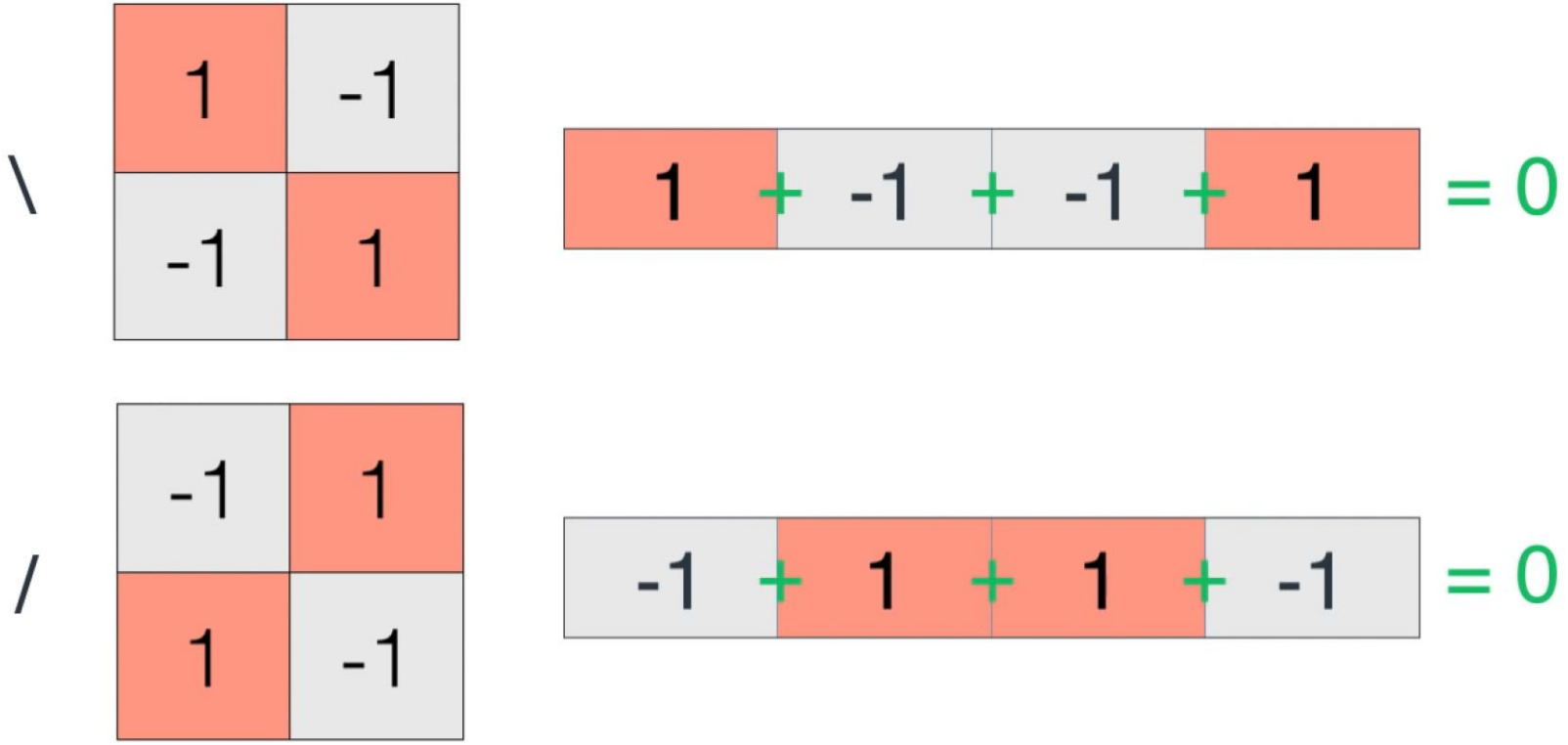
Convolution Neural Network (CNN)

A simple world



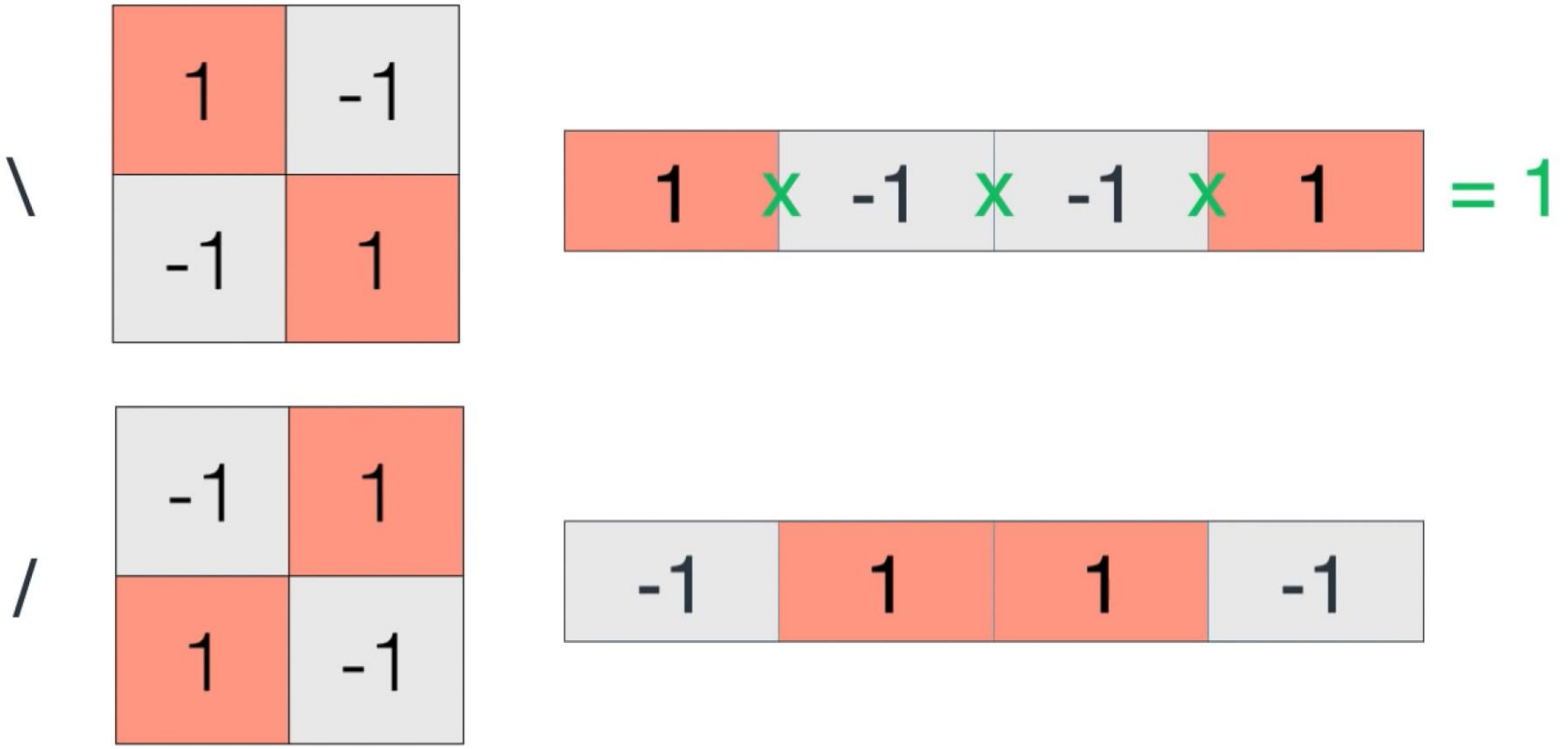
Convolution Neural Network (CNN)

A simple world



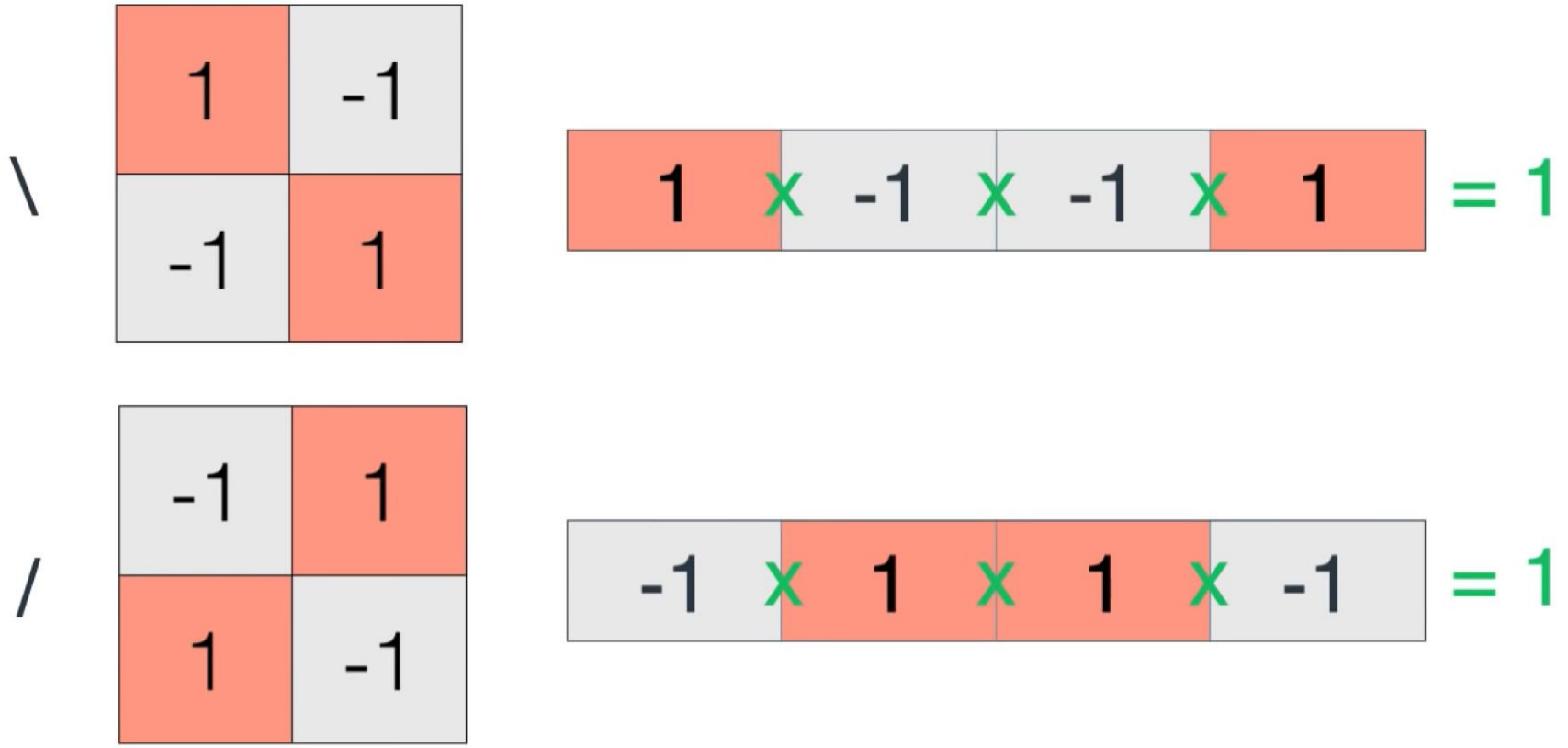
Convolution Neural Network (CNN)

A simple world



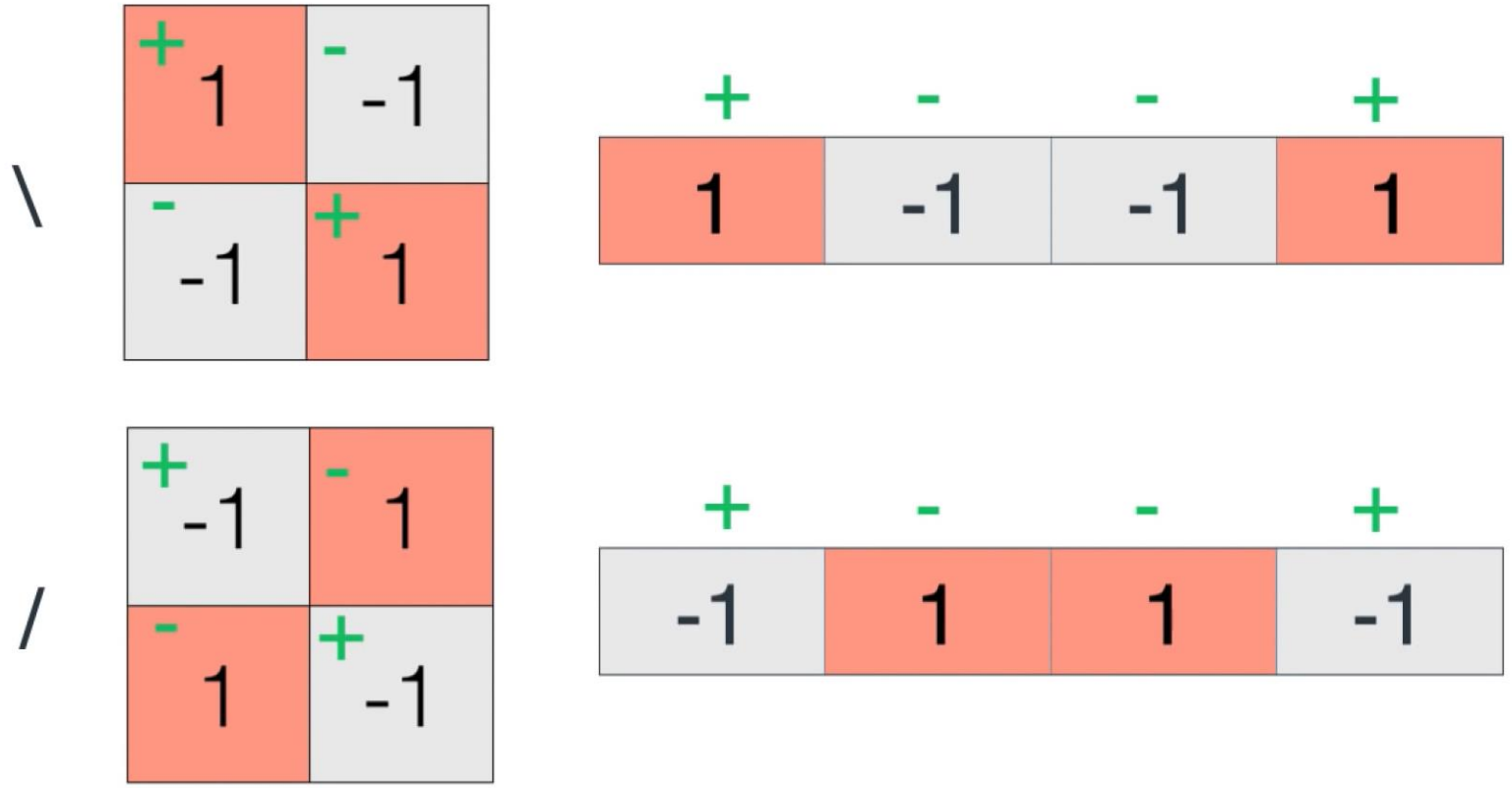
Convolution Neural Network (CNN)

A simple world



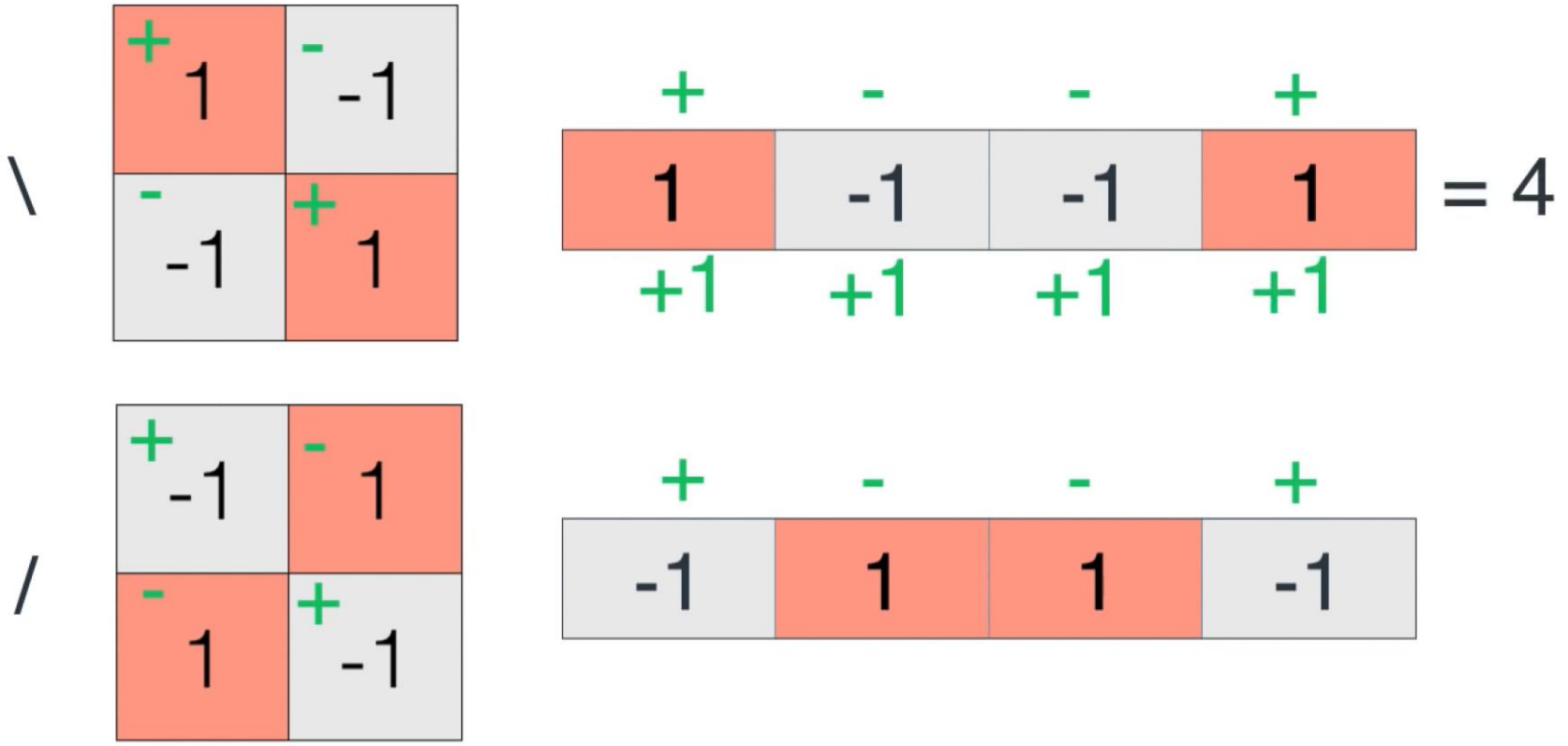
Convolution Neural Network (CNN)

A simple world



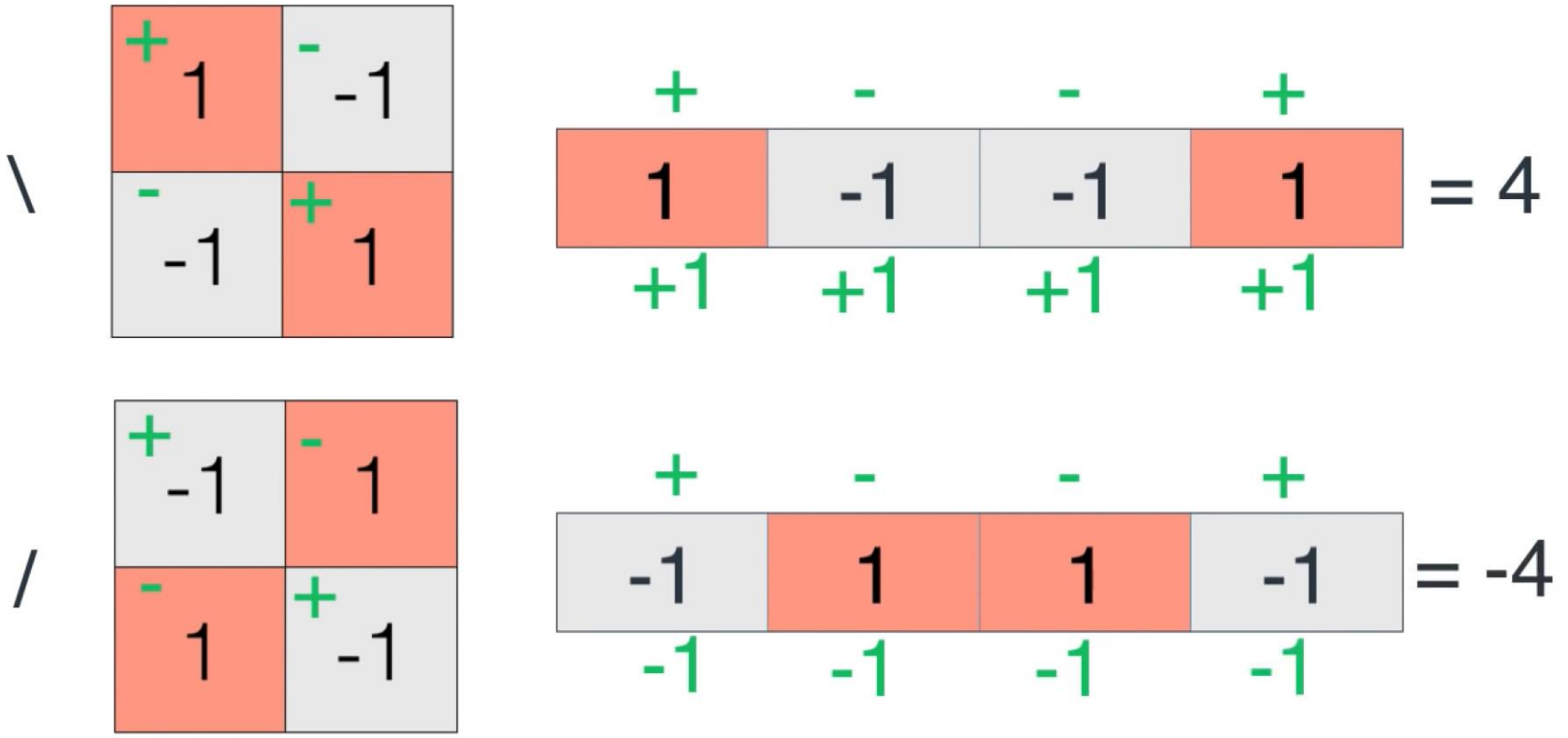
Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

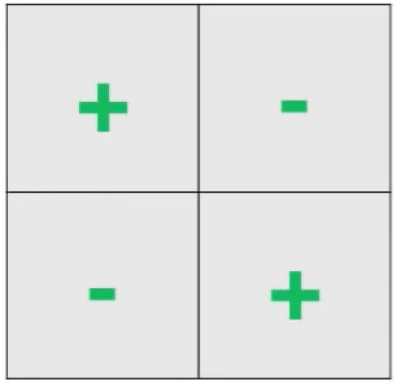
A simple world



Convolution Neural Network (CNN)

A simple world

Image Recognition Classifier

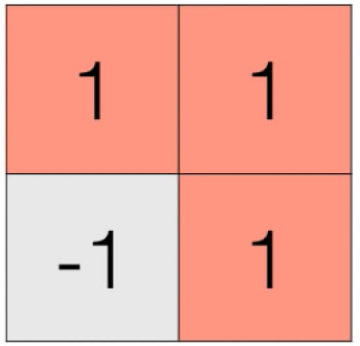


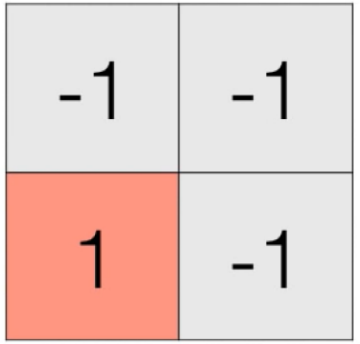
If positive, “\”

If negative, “/”

Convolution Neural Network (CNN)

A simple world

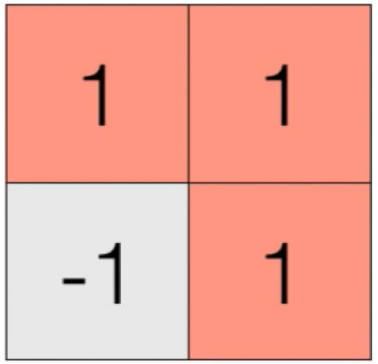




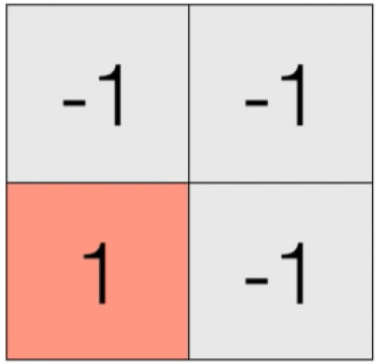
Convolution Neural Network (CNN)

A simple world

\

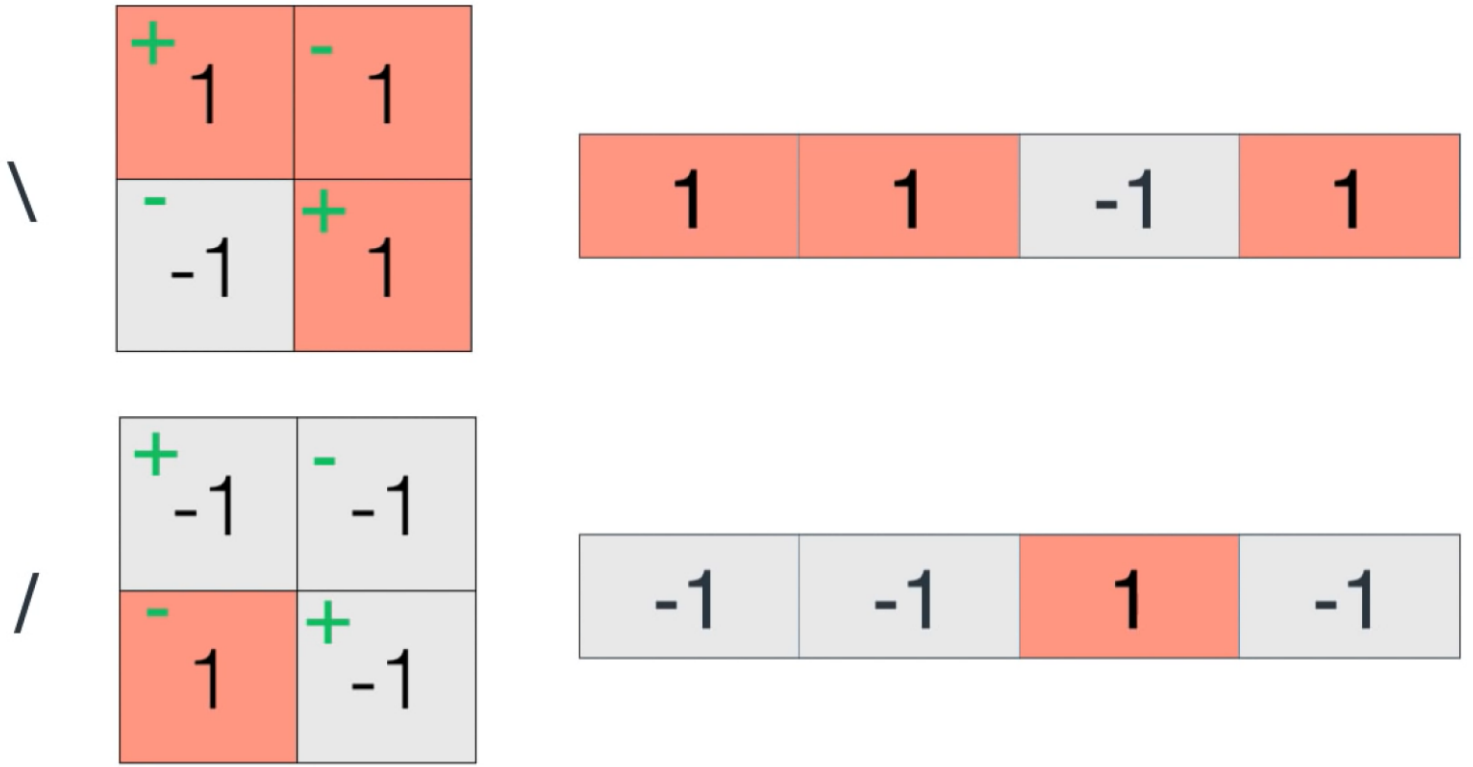


/



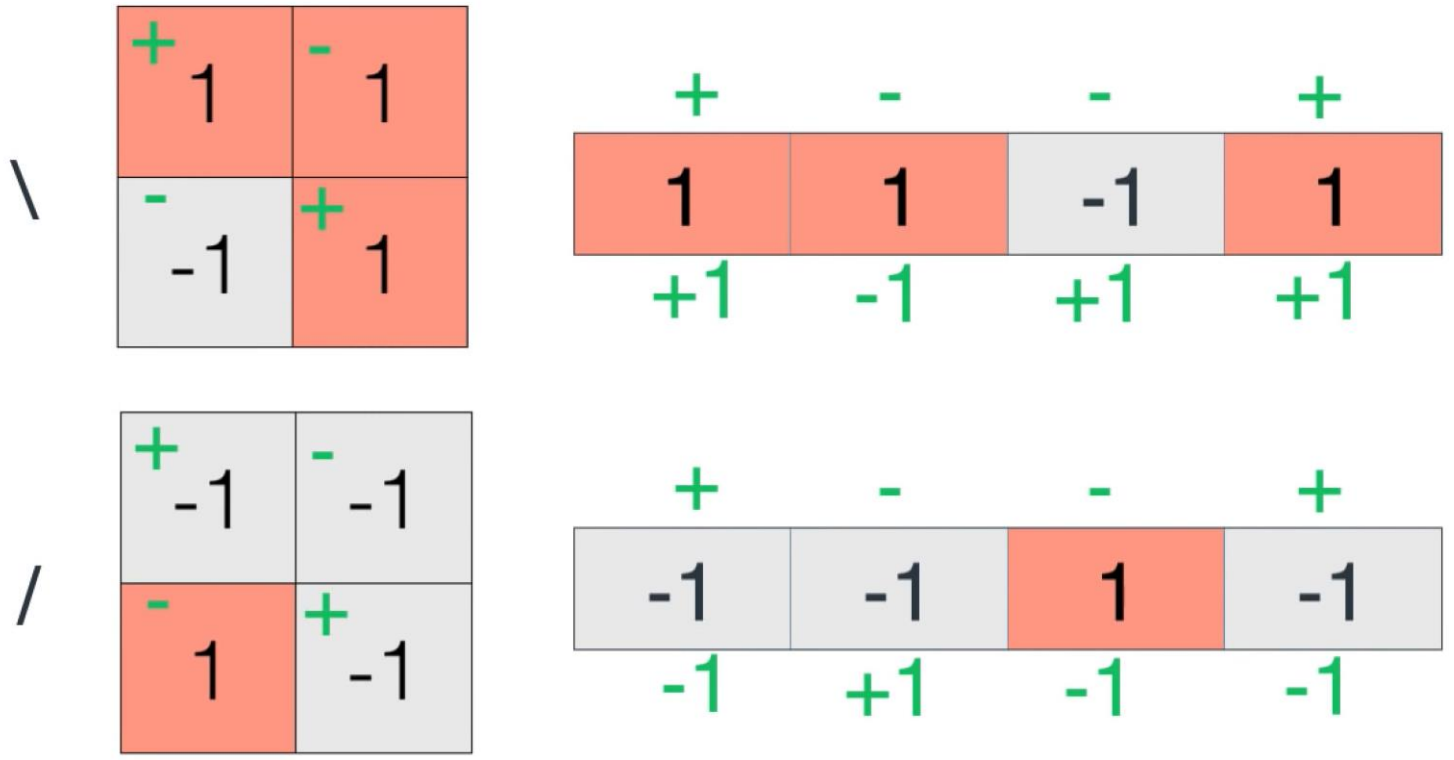
Convolution Neural Network (CNN)

A simple world



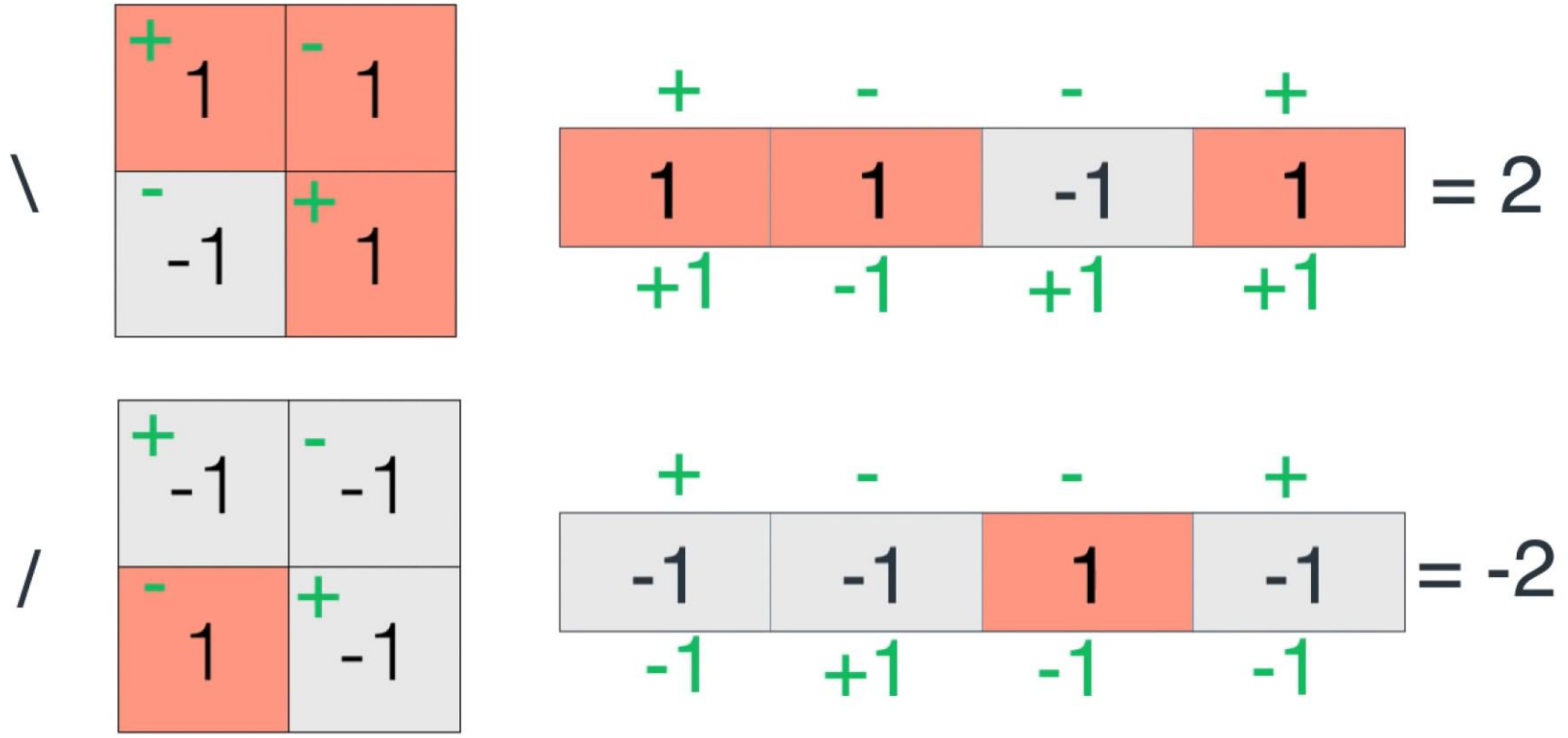
Convolution Neural Network (CNN)

A simple world



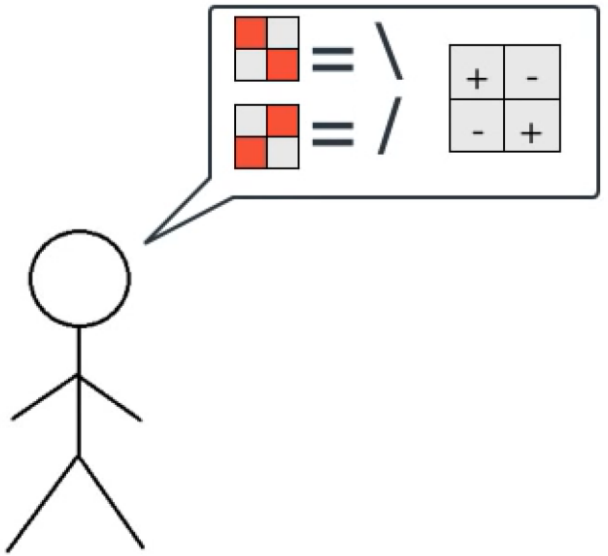
Convolution Neural Network (CNN)

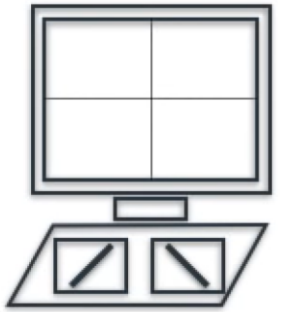
A simple world



Convolution Neural Network (CNN)

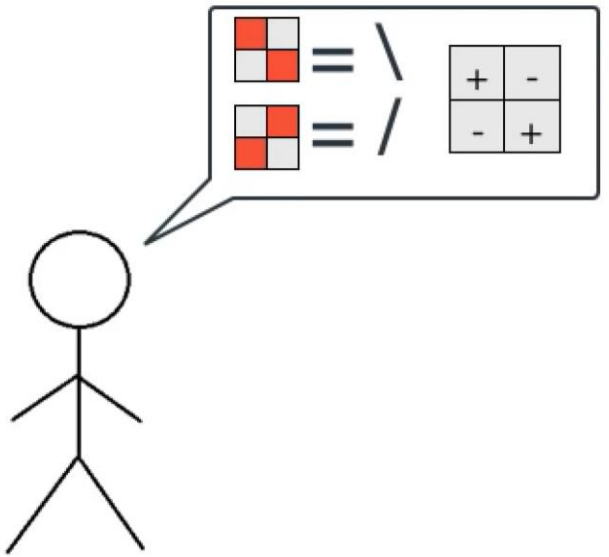
A simple world

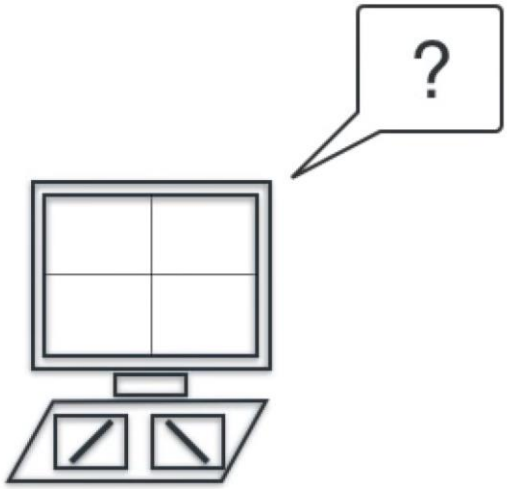




Convolution Neural Network (CNN)

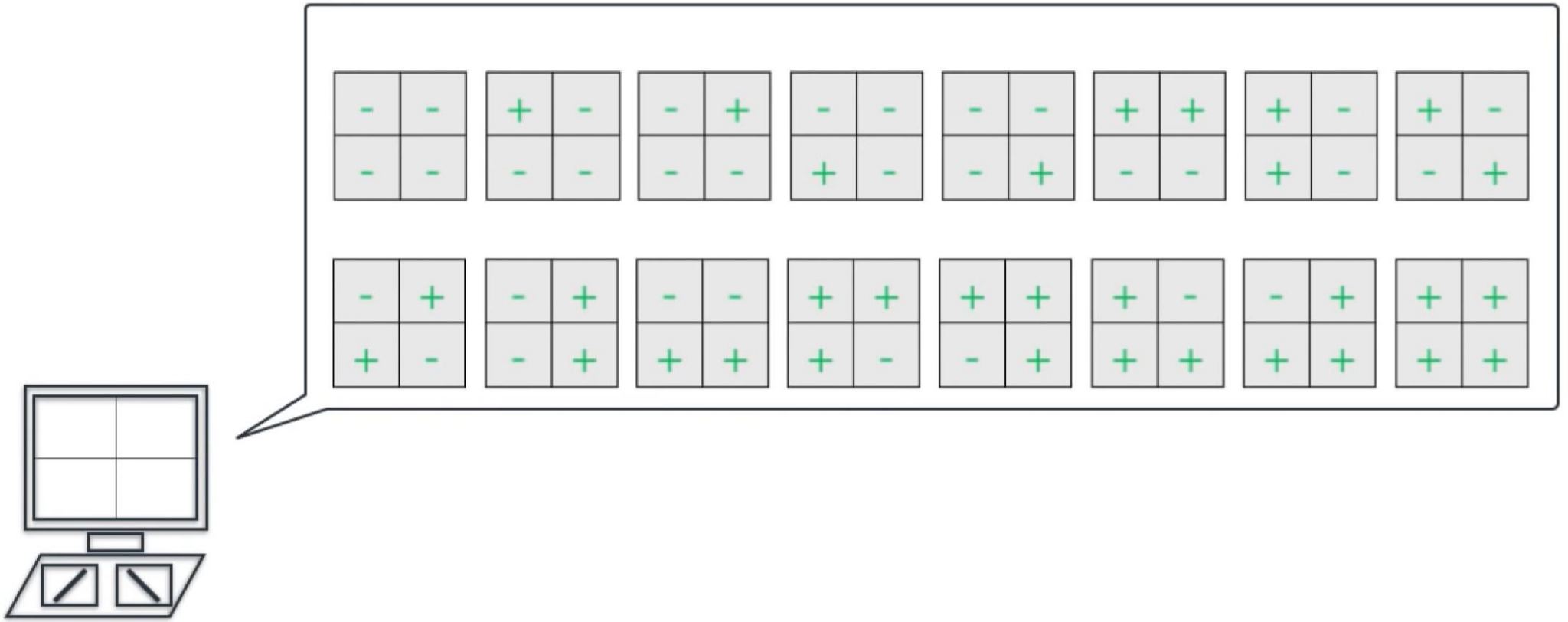
A simple world





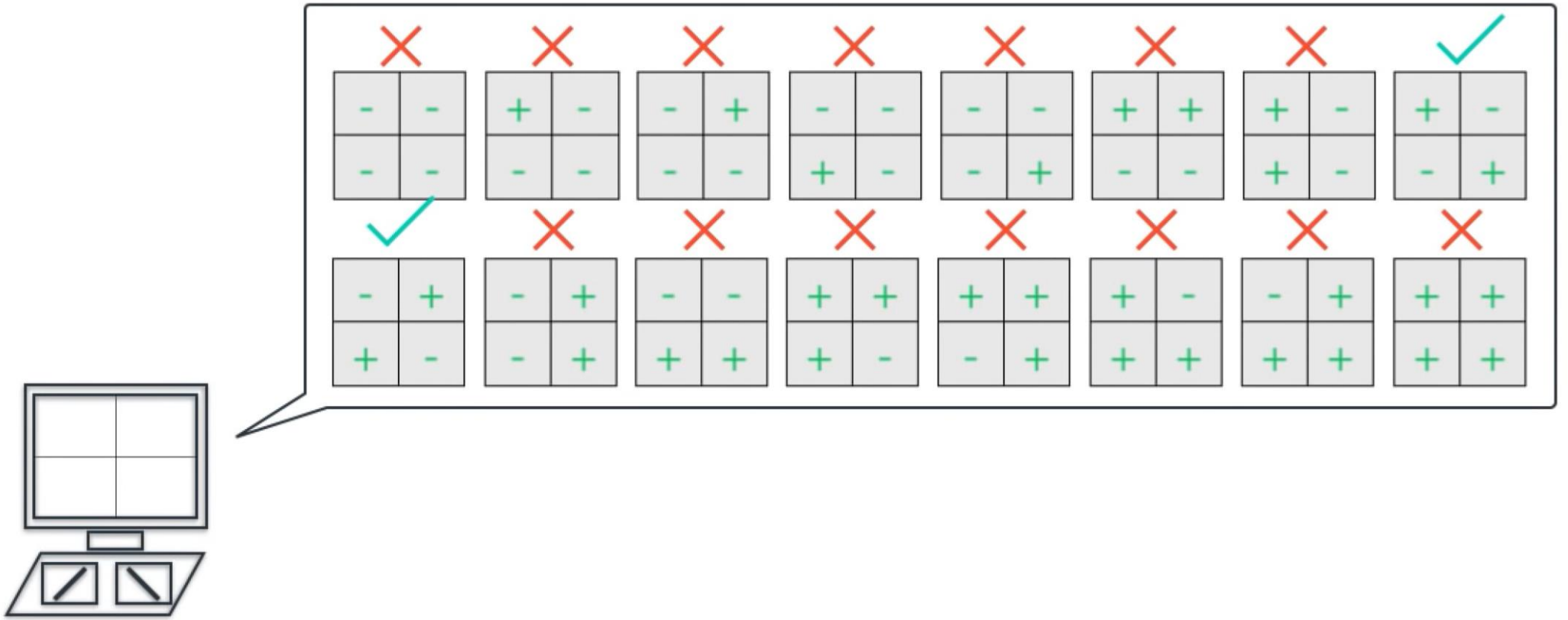
Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

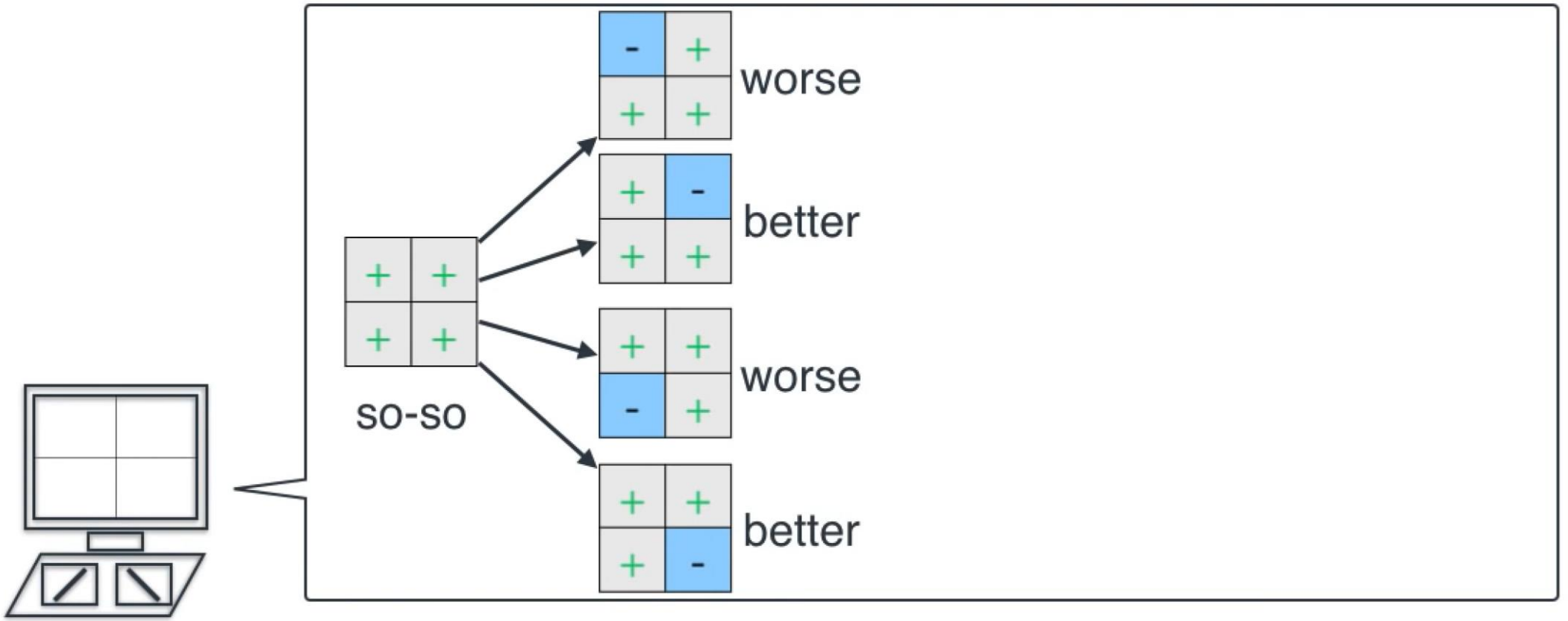
A simple world

Convolution Neural Network (CNN)

A simple world

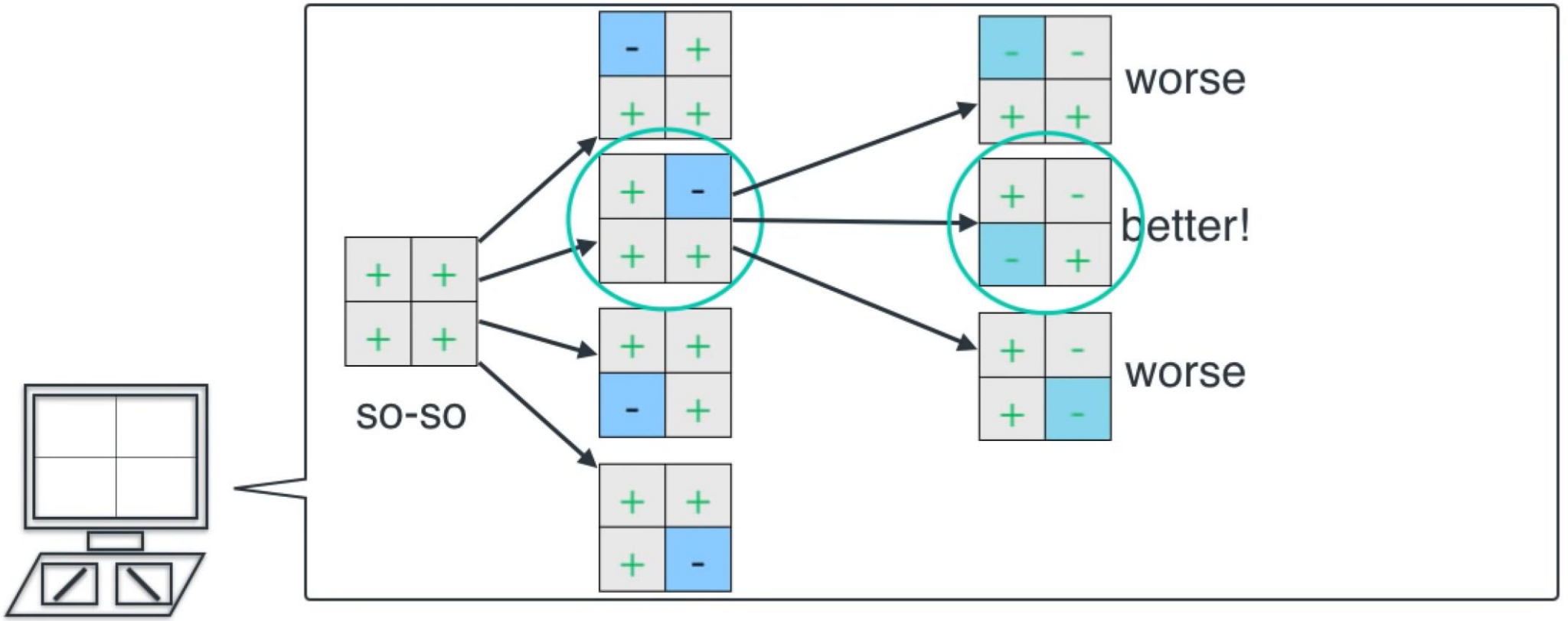
Convolution Neural Network (CNN)

A simple world



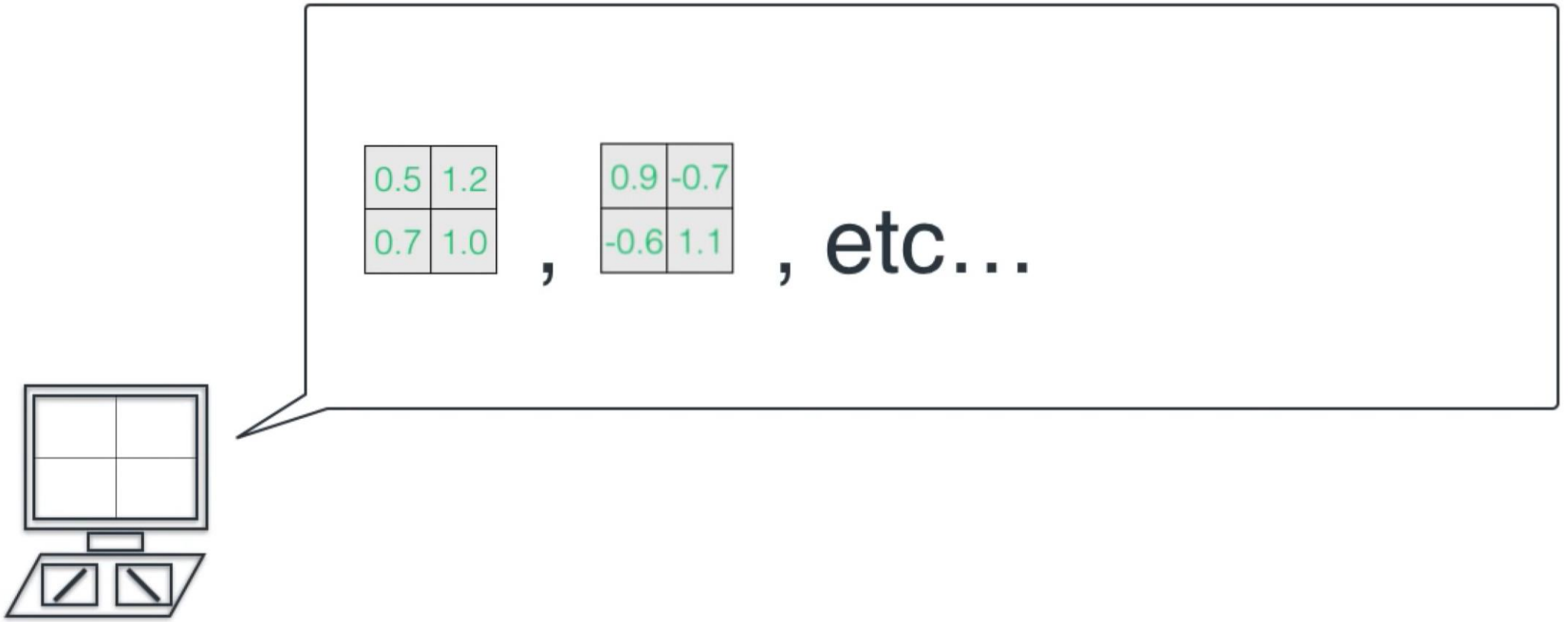
Convolution Neural Network (CNN)

A simple world



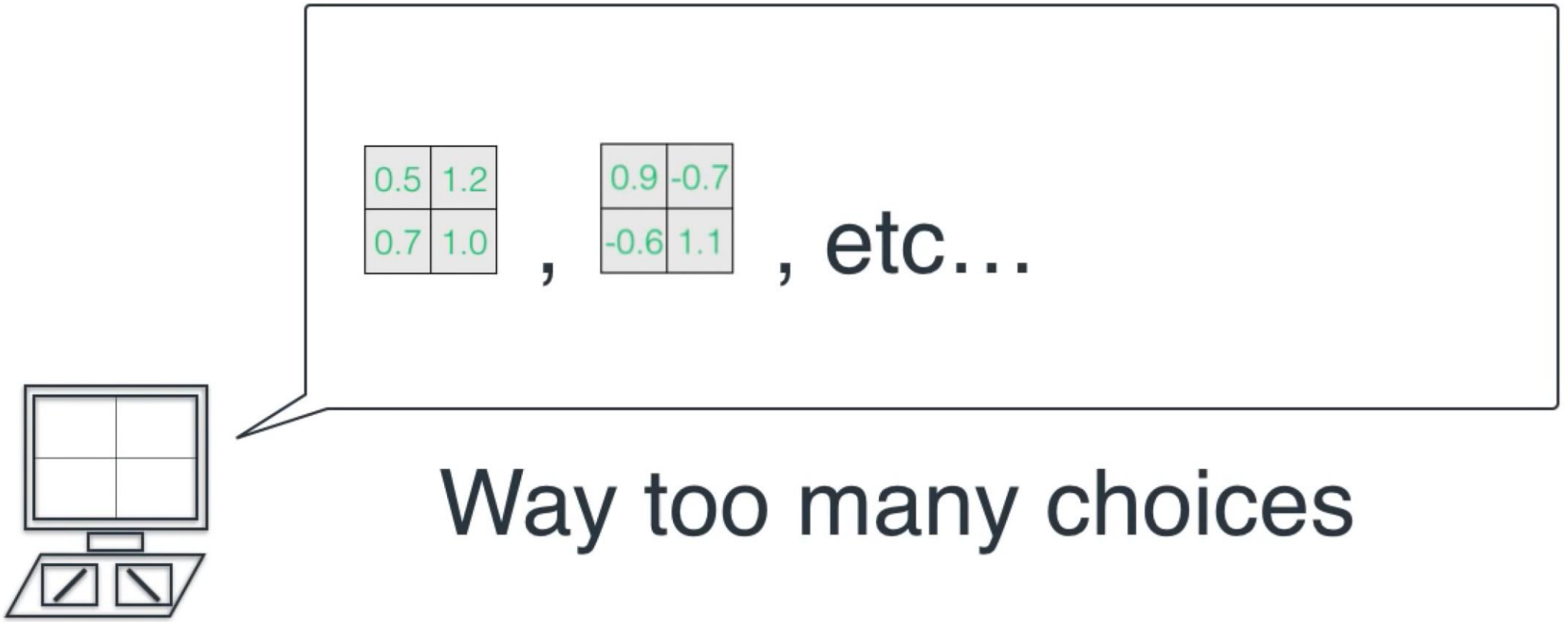
Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

A simple world

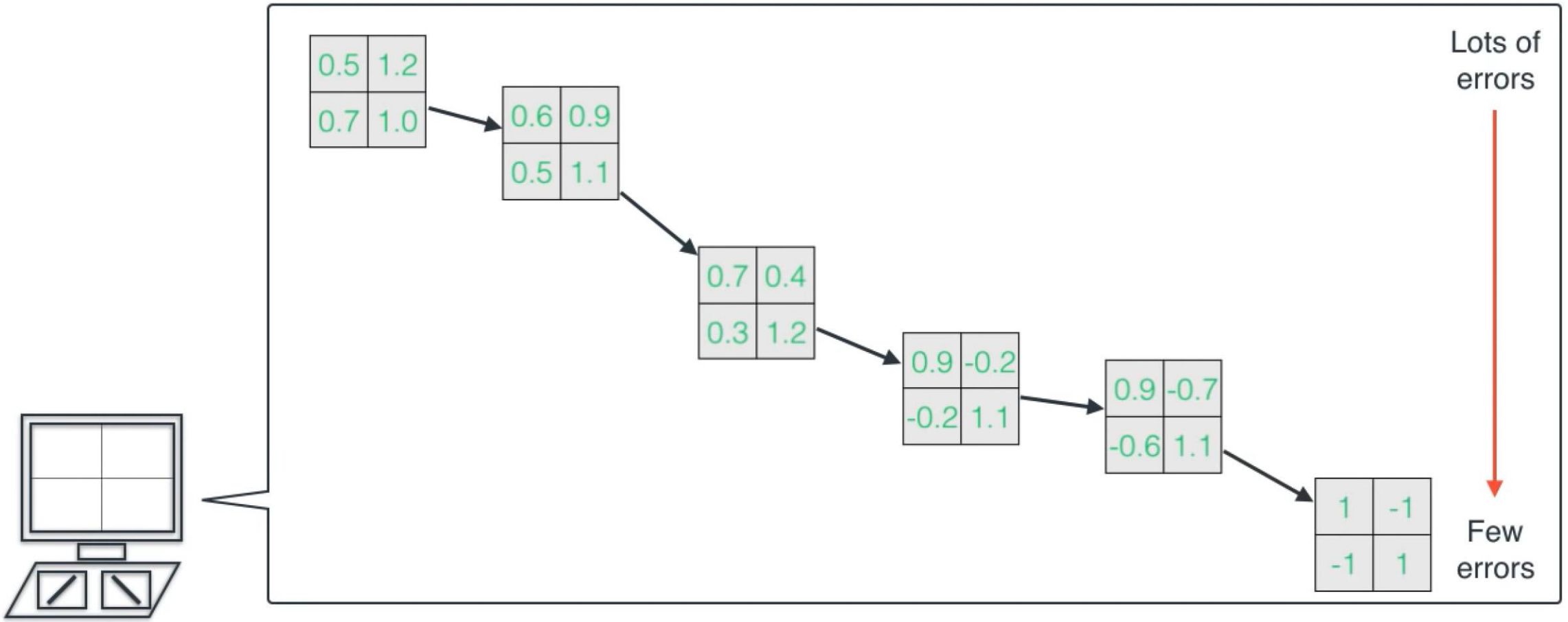


Convolution Neural Network (CNN)

A simple world

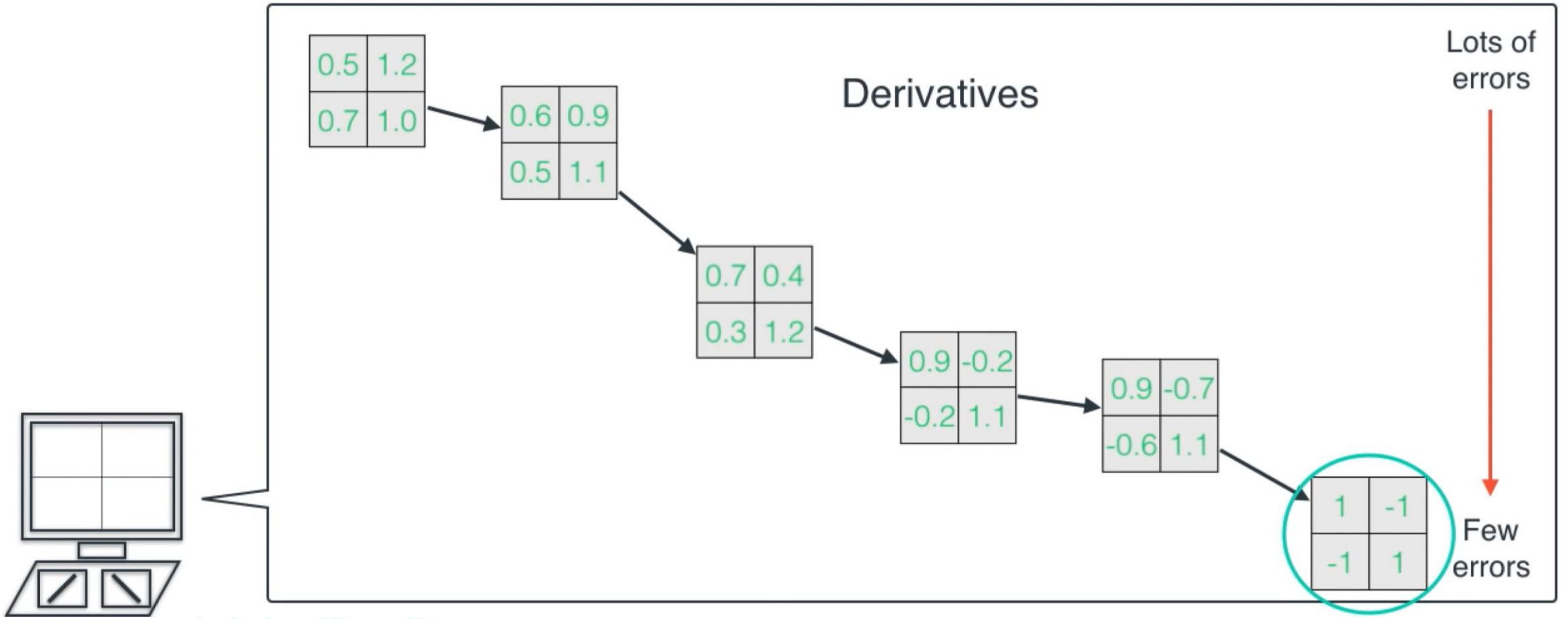
Convolution Neural Network (CNN)

A simple world



Convolution Neural Network (CNN)

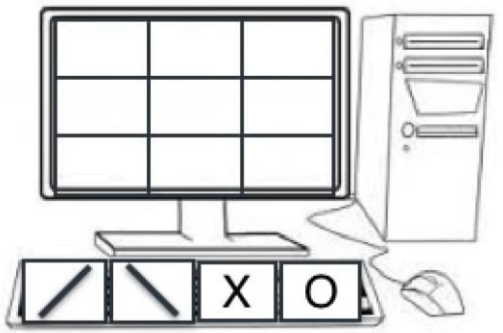
A simple world



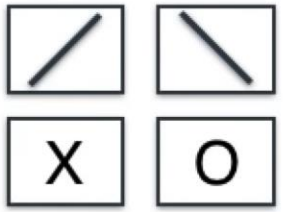
Convolution Neural Network (CNN)

A little more complex world

Person



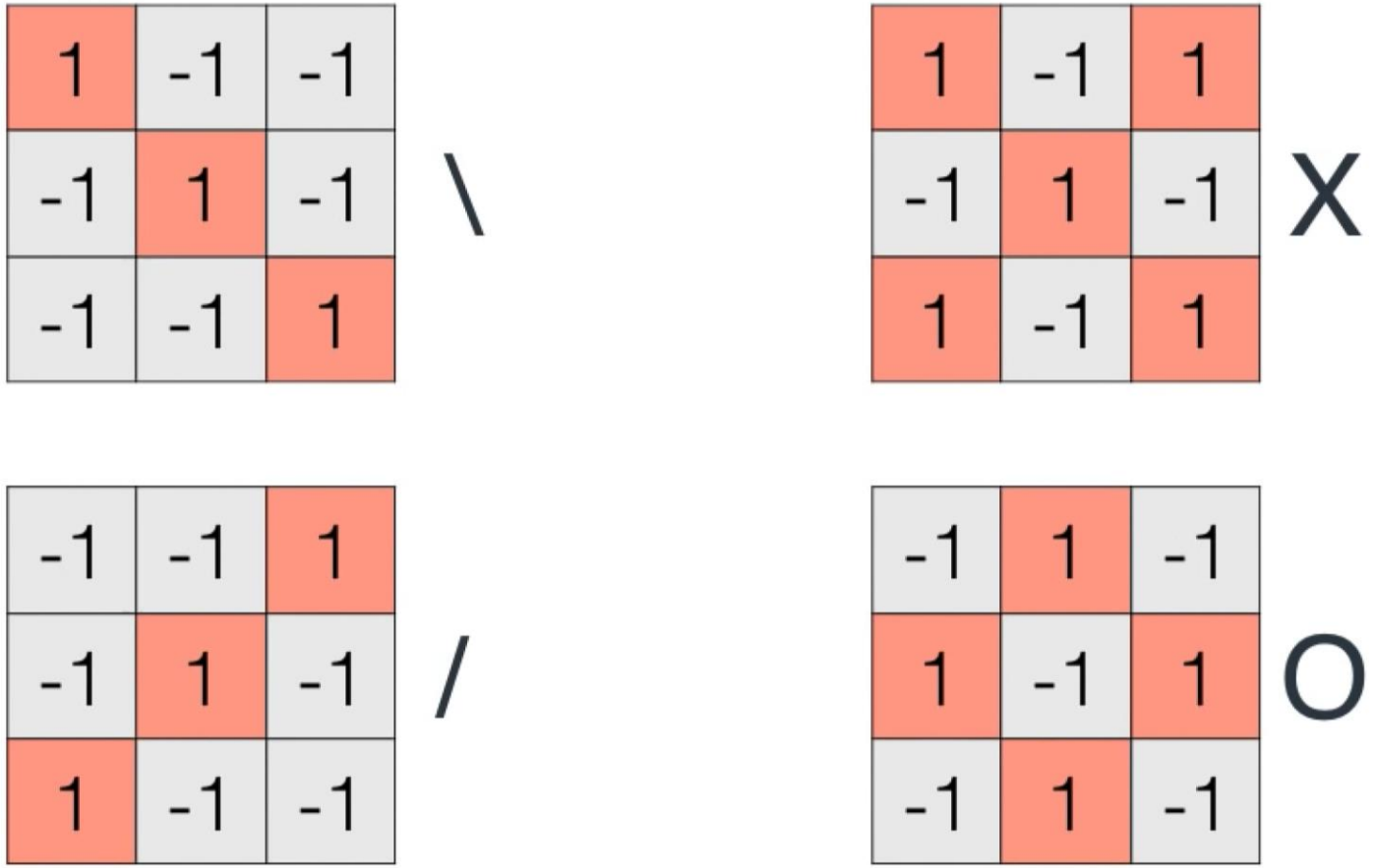
Computer



Alphabet

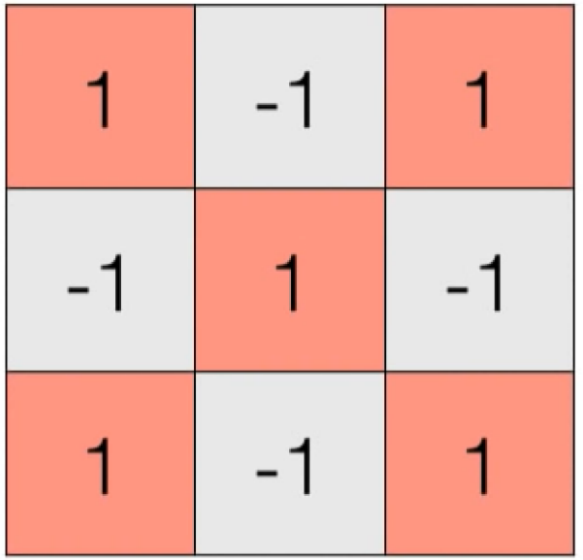
Convolution Neural Network (CNN)

A little more complex world



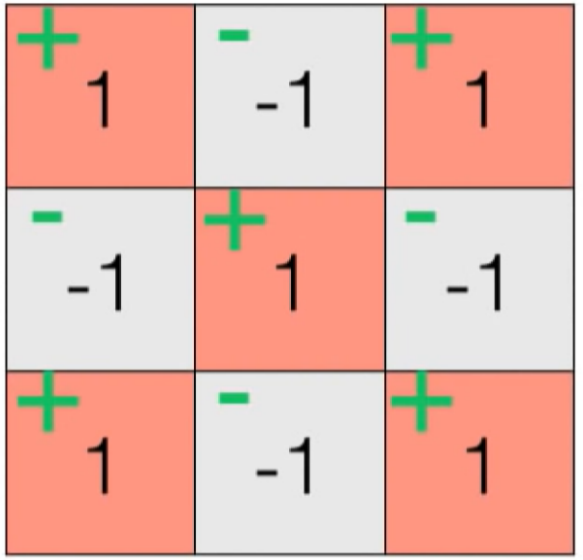
Convolution Neural Network (CNN)

A little more complex world



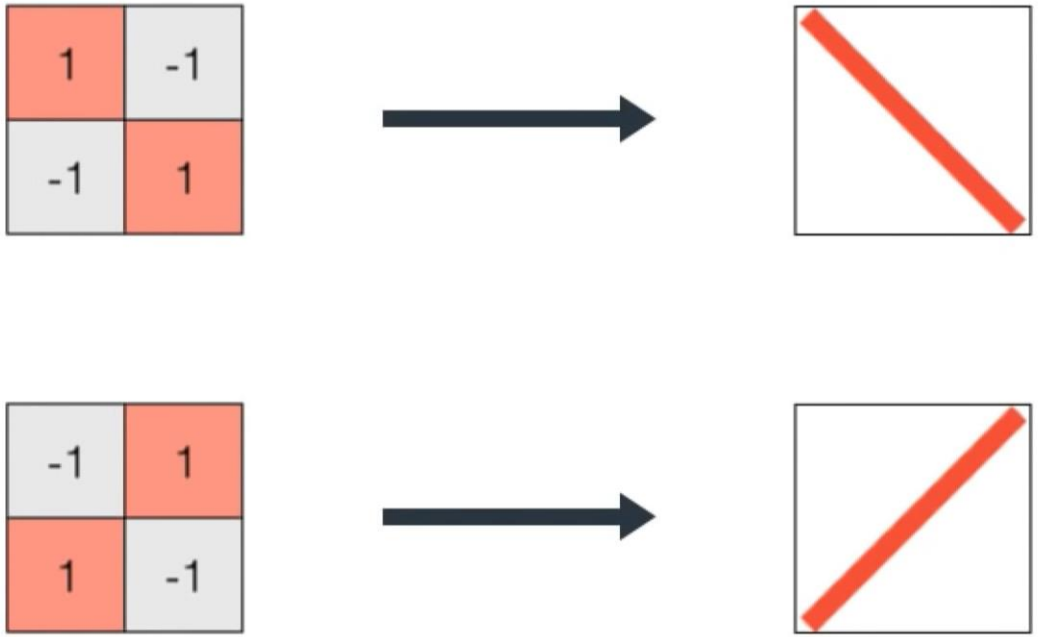
Convolution Neural Network (CNN)

A little more complex world



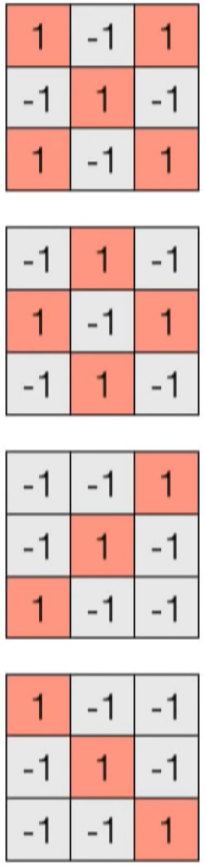
Convolution Neural Network (CNN)

A little more complex world



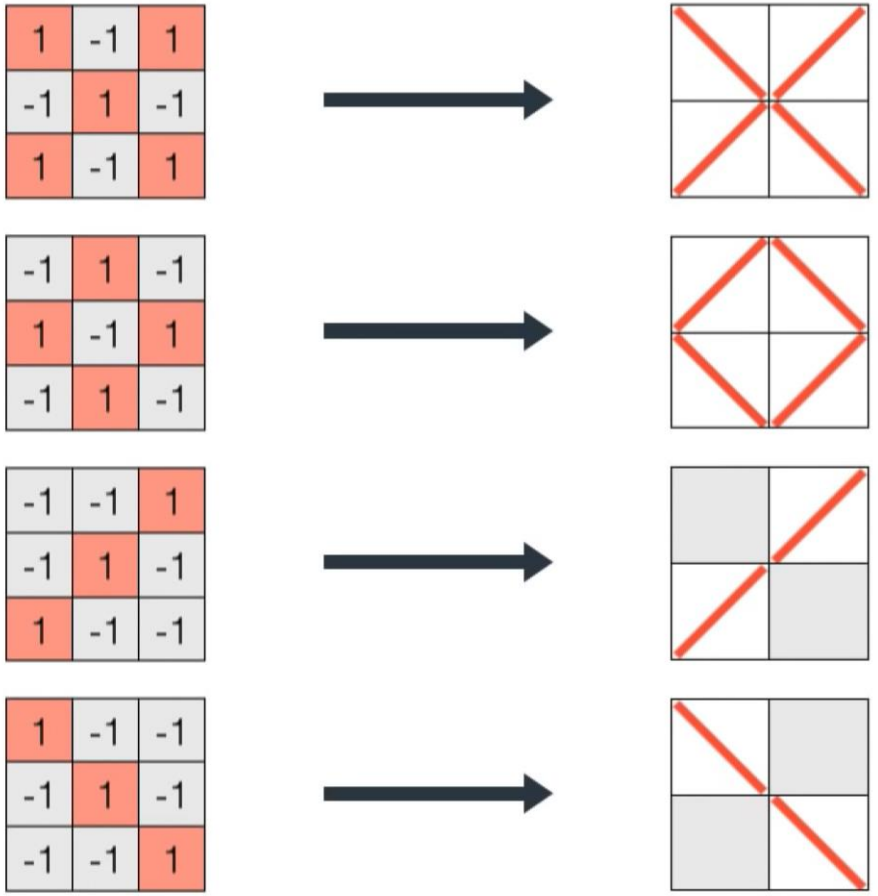
Convolution Neural Network (CNN)

A little more complex world



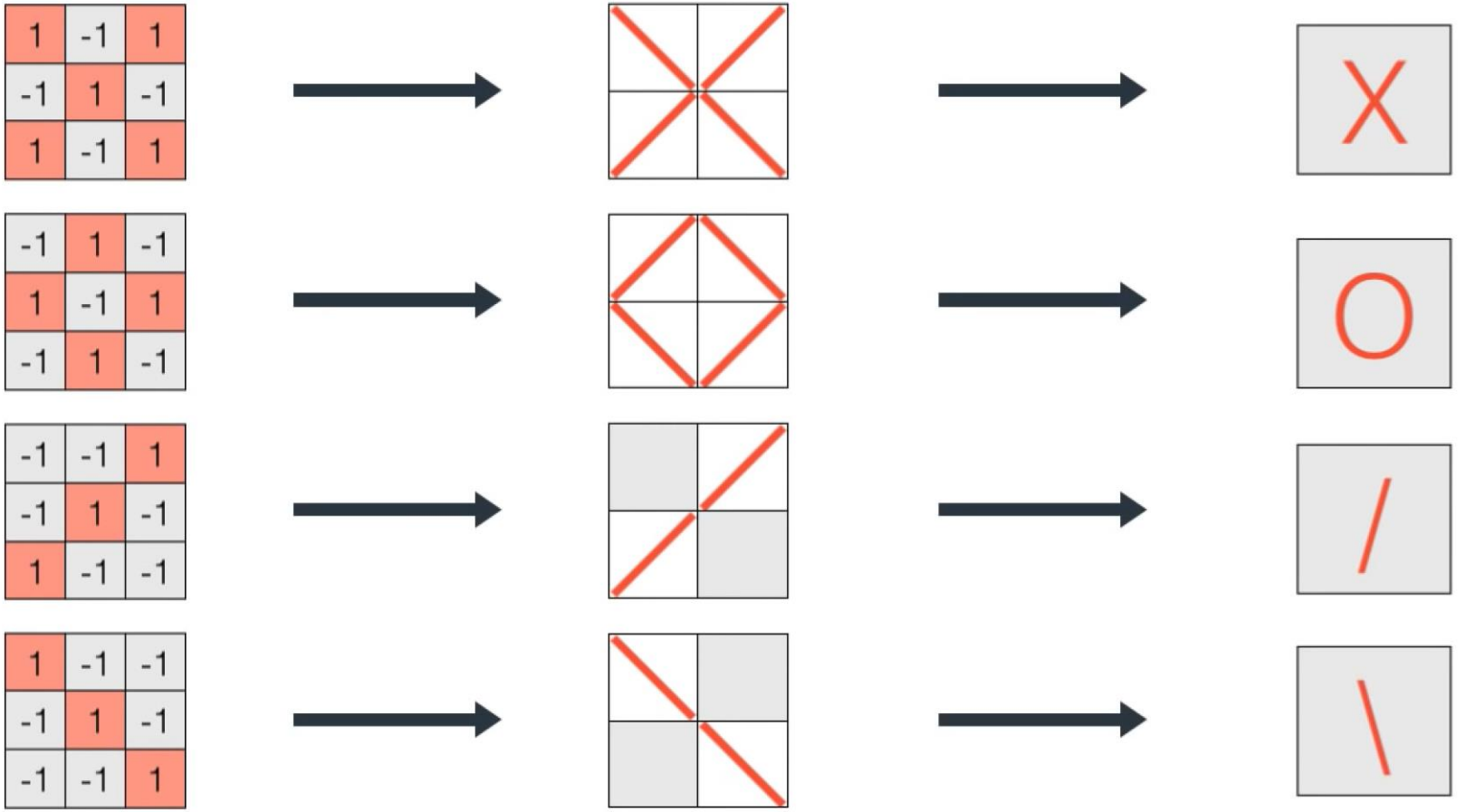
Convolution Neural Network (CNN)

A little more complex world



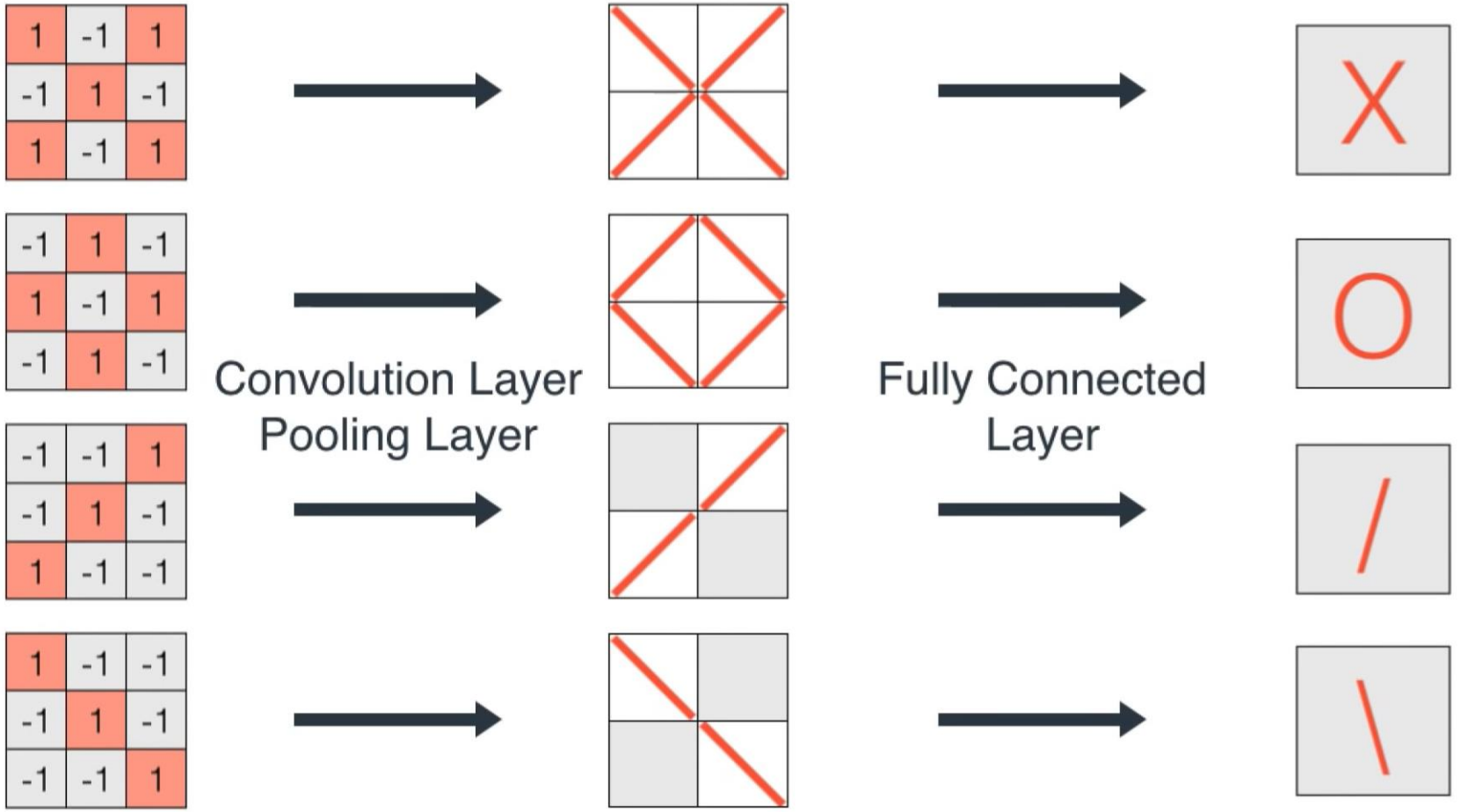
Convolution Neural Network (CNN)

A little more complex world



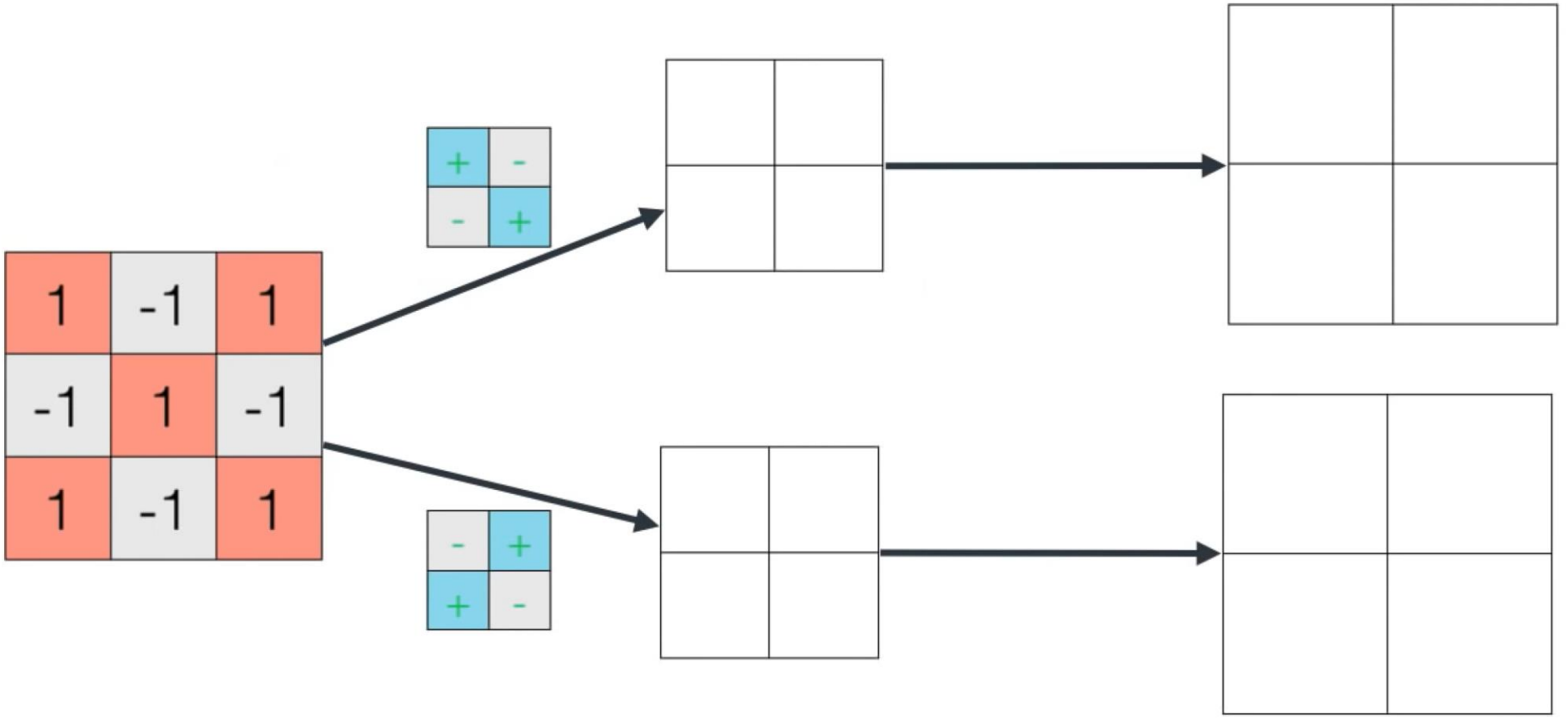
Convolution Neural Network (CNN)

A little more complex world



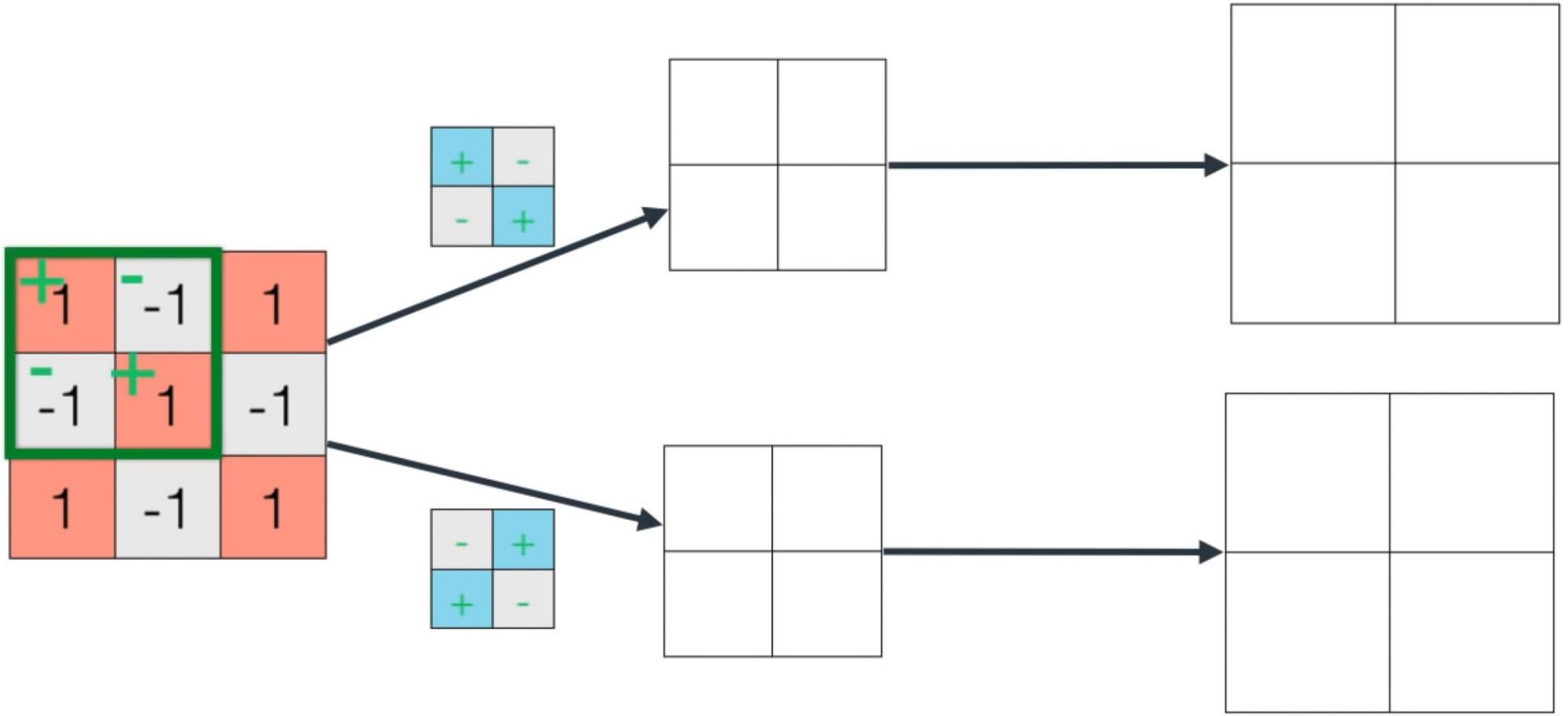
Convolution Neural Network (CNN)

A little more complex world



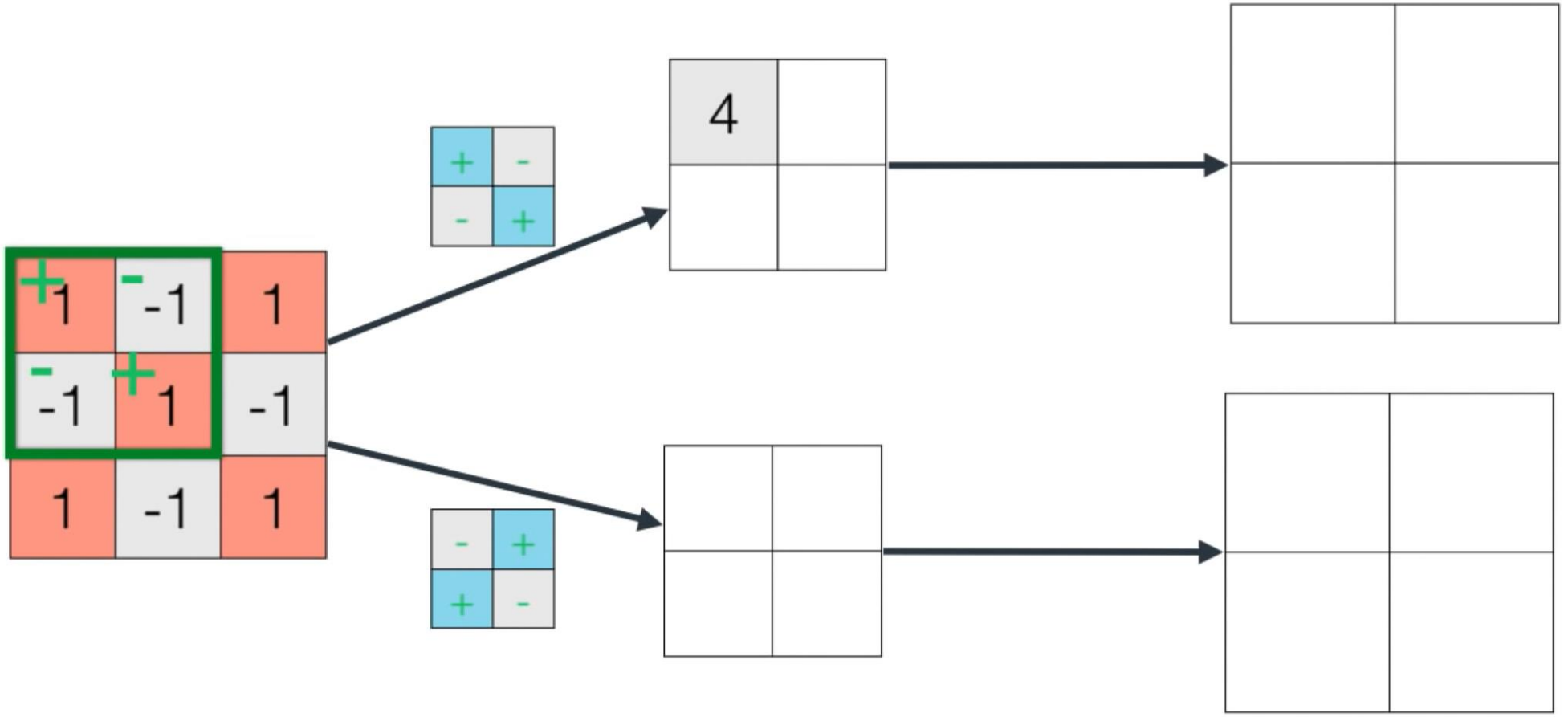
Convolution Neural Network (CNN)

A little more complex world



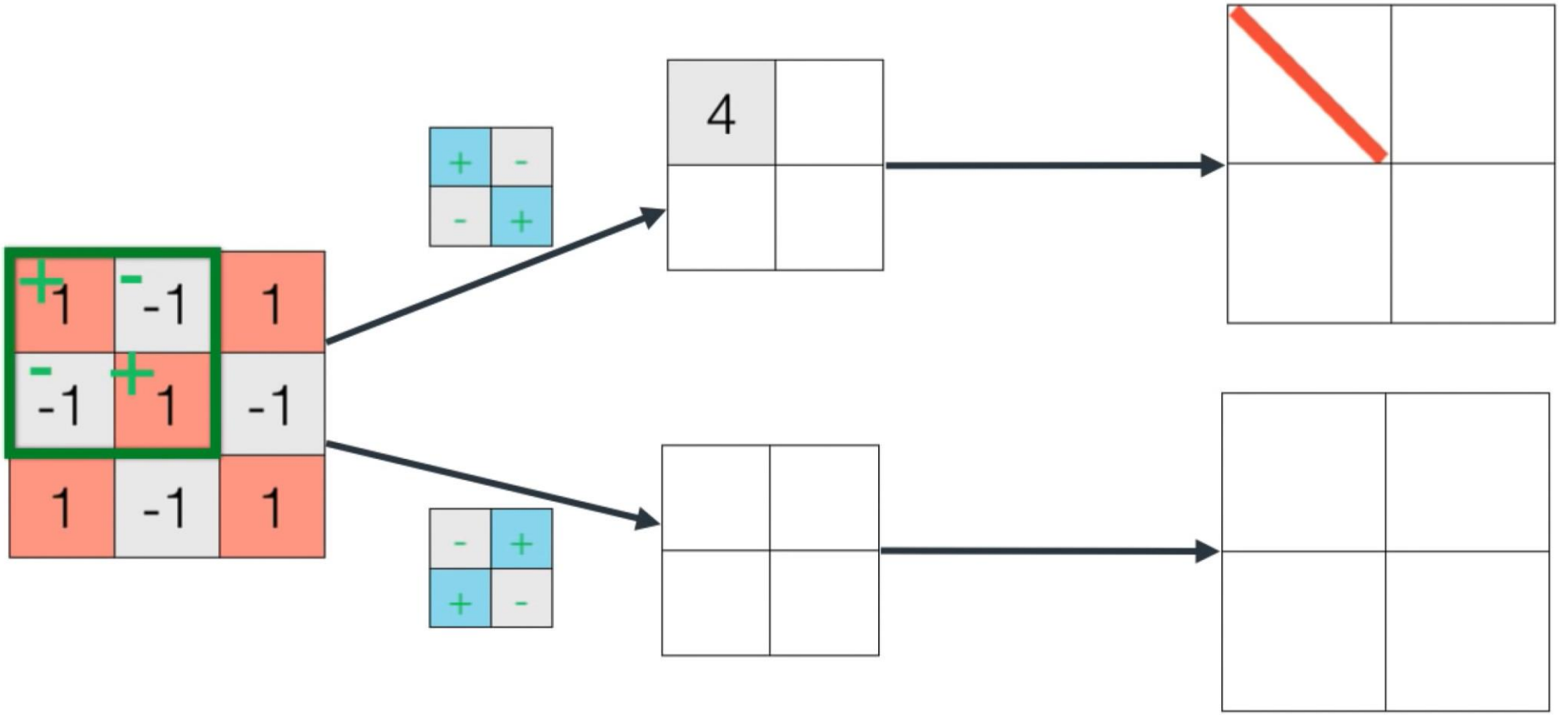
Convolution Neural Network (CNN)

A little more complex world



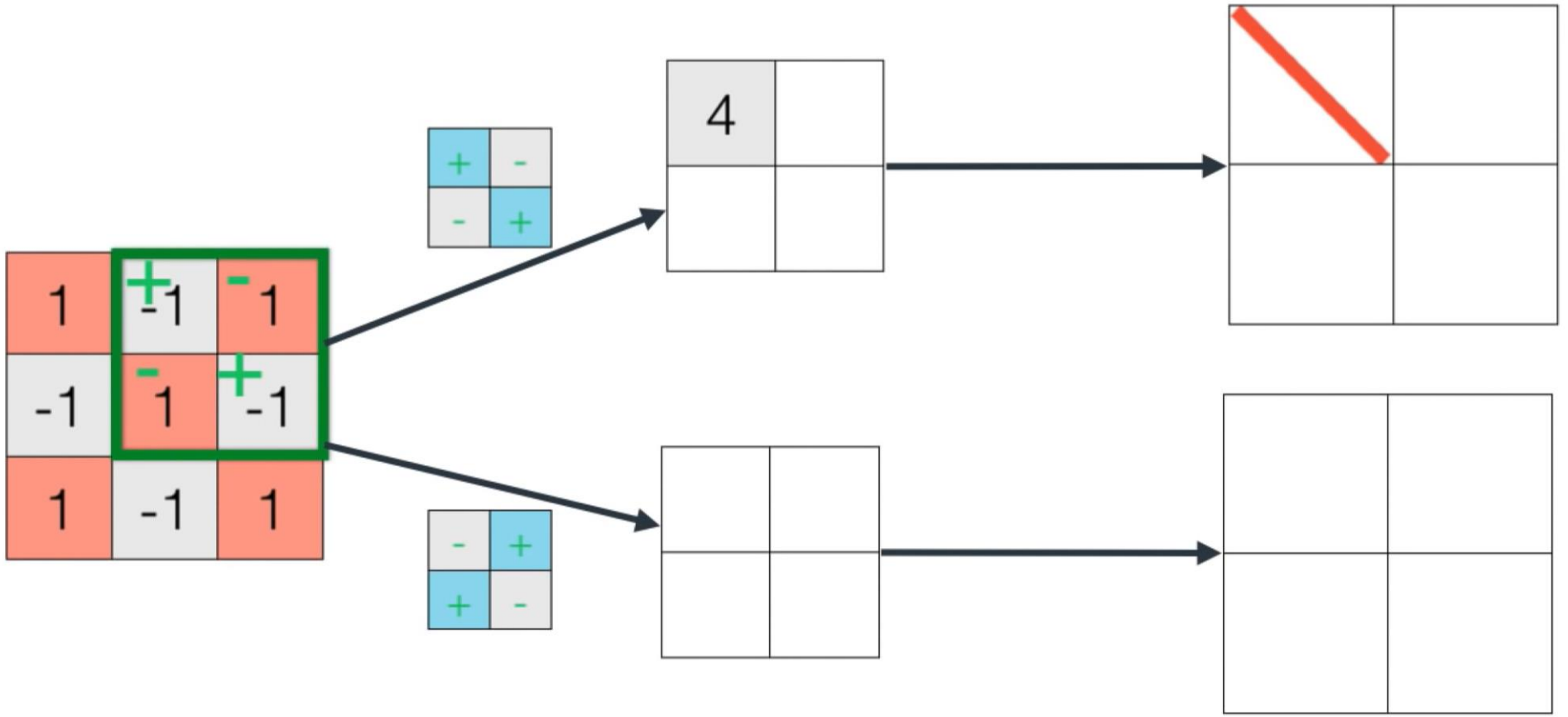
Convolution Neural Network (CNN)

A little more complex world



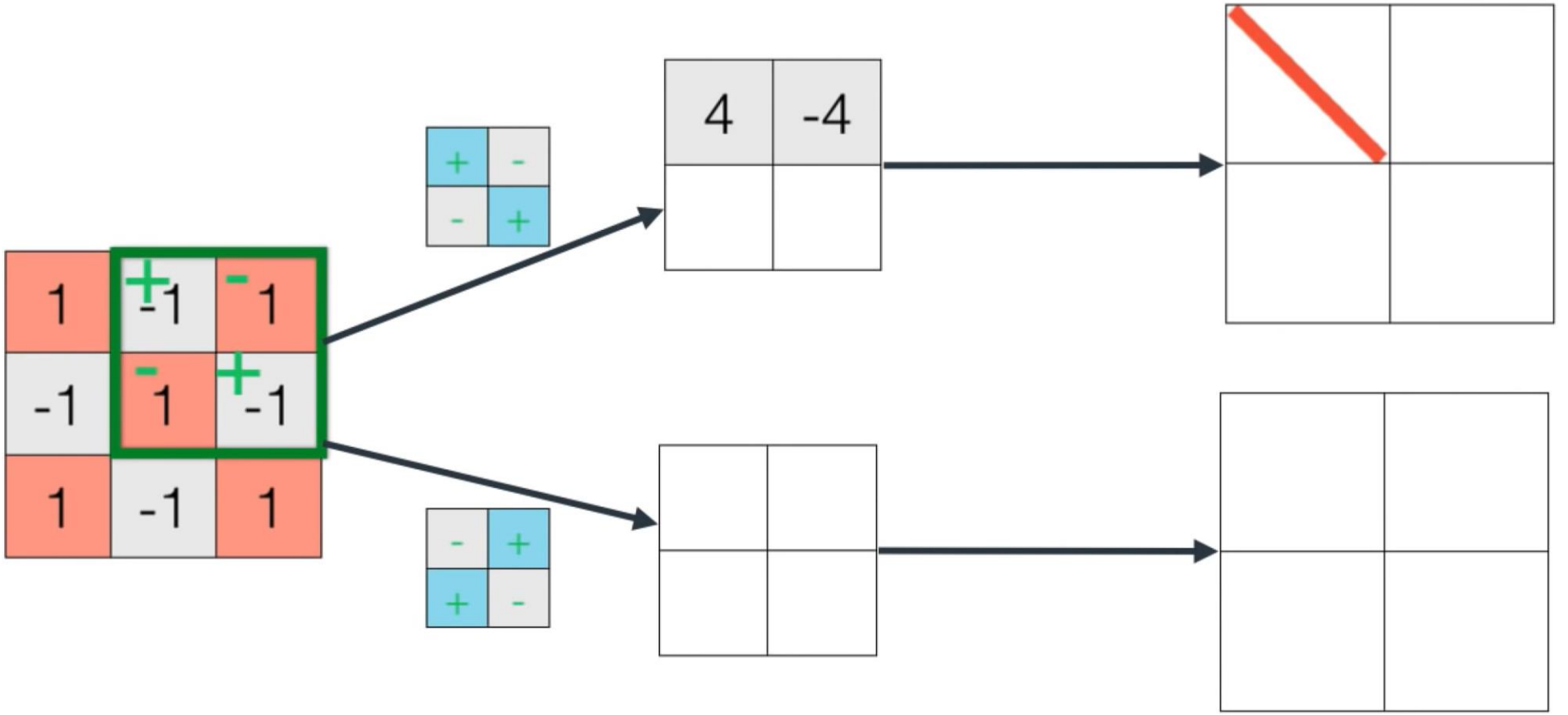
Convolution Neural Network (CNN)

A little more complex world



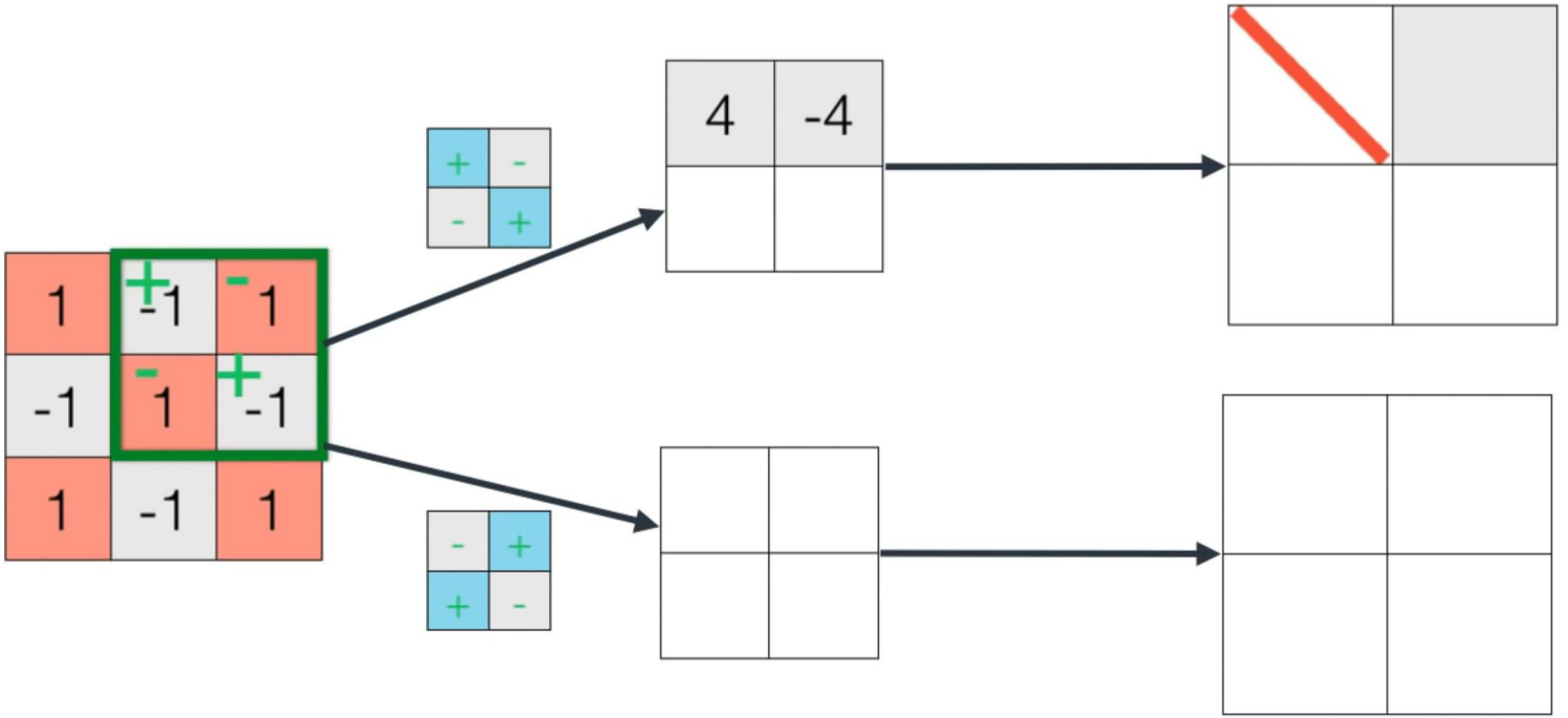
Convolution Neural Network (CNN)

A little more complex world



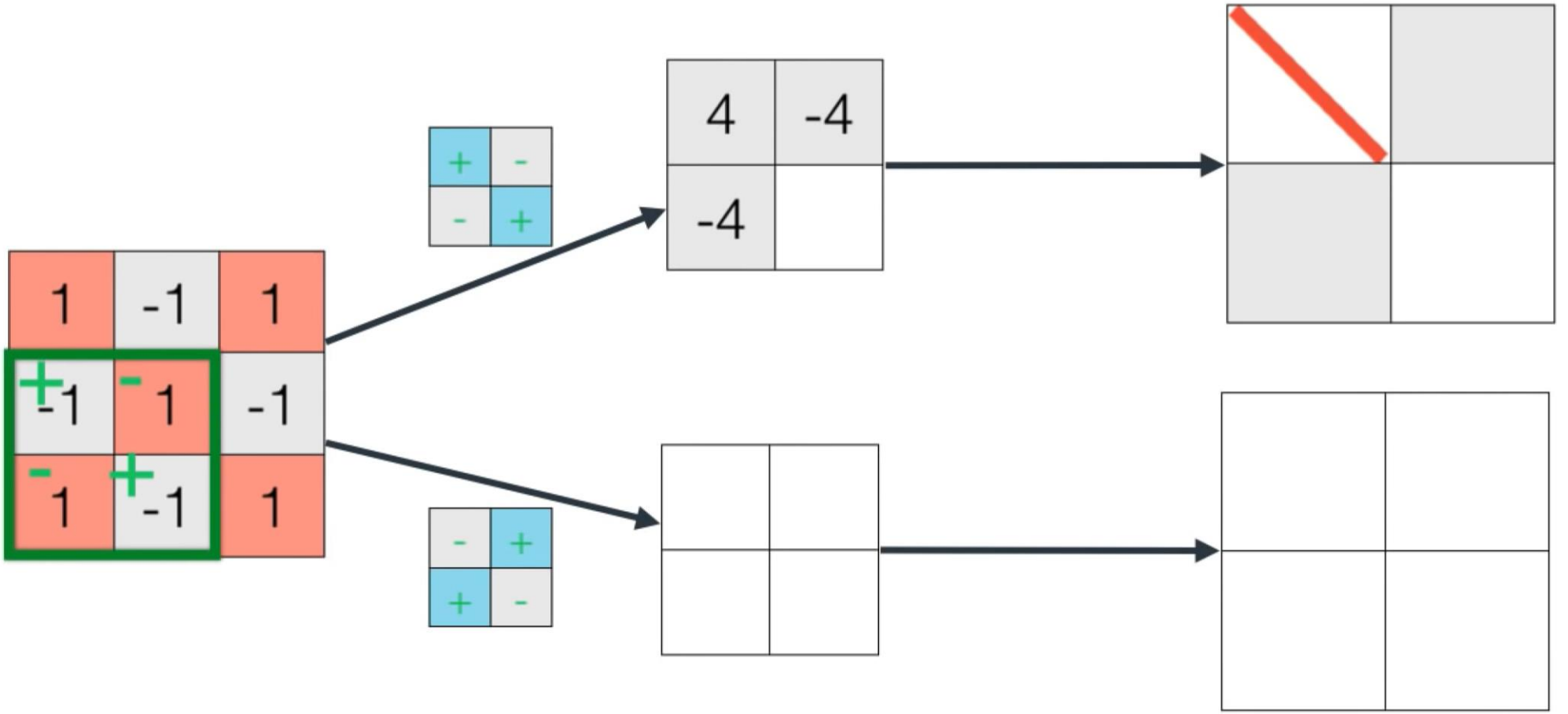
Convolution Neural Network (CNN)

A little more complex world



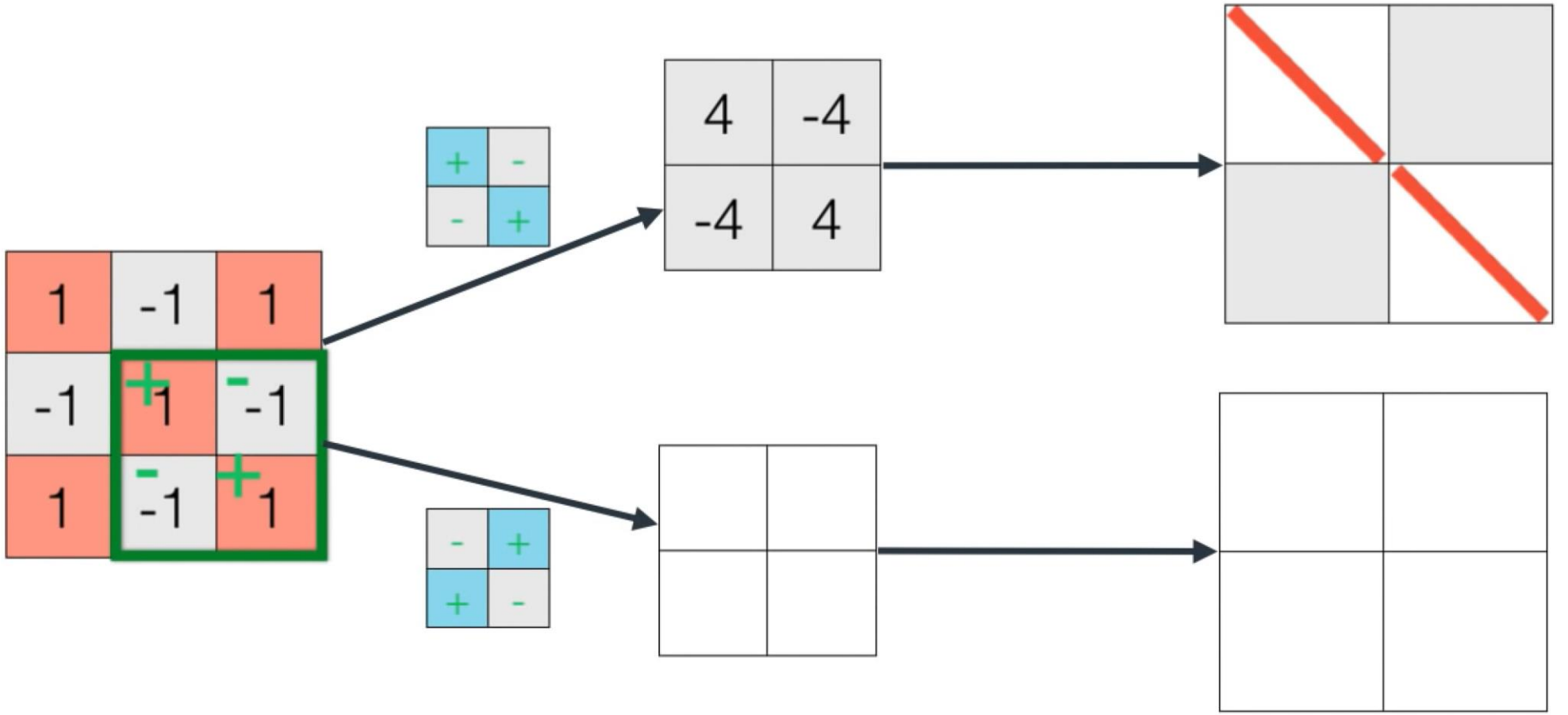
Convolution Neural Network (CNN)

A little more complex world



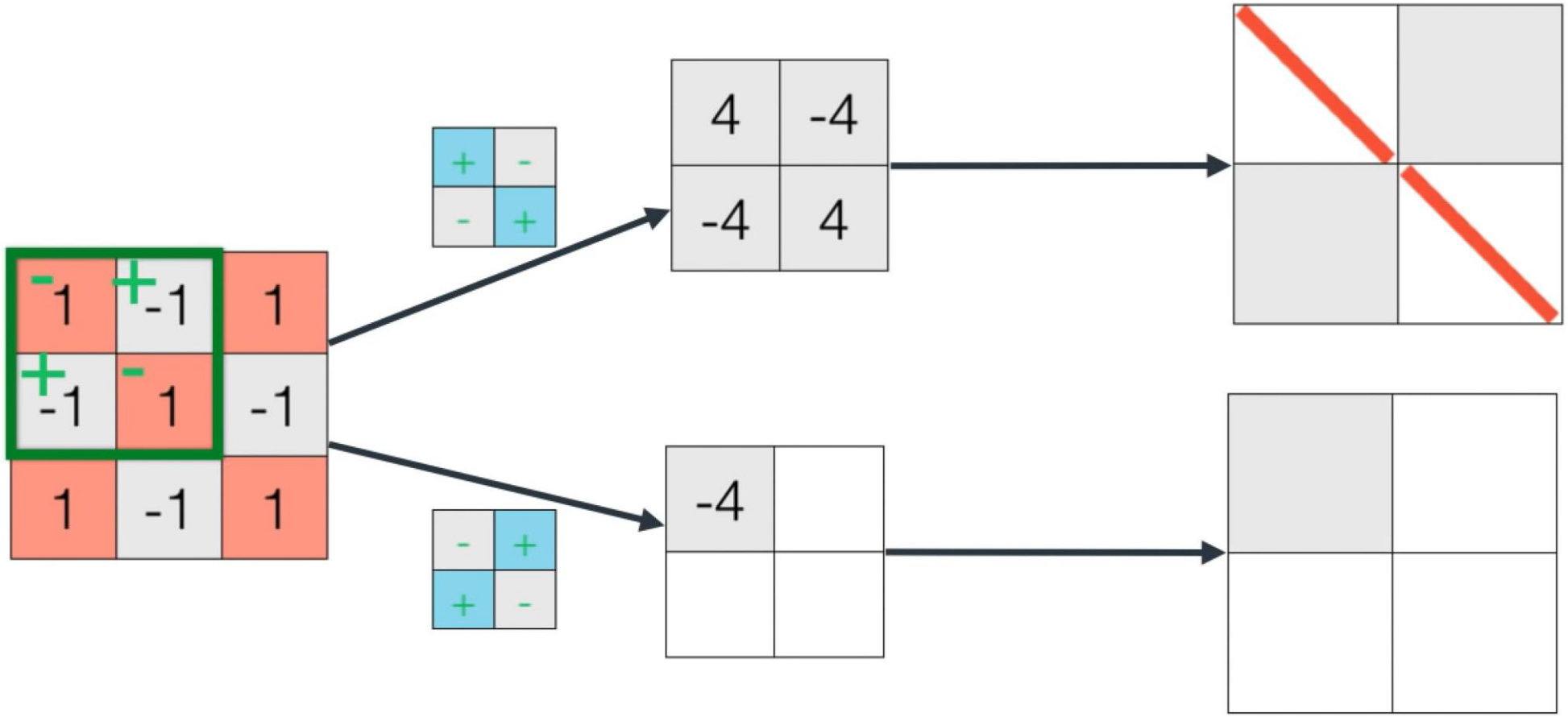
Convolution Neural Network (CNN)

A little more complex world



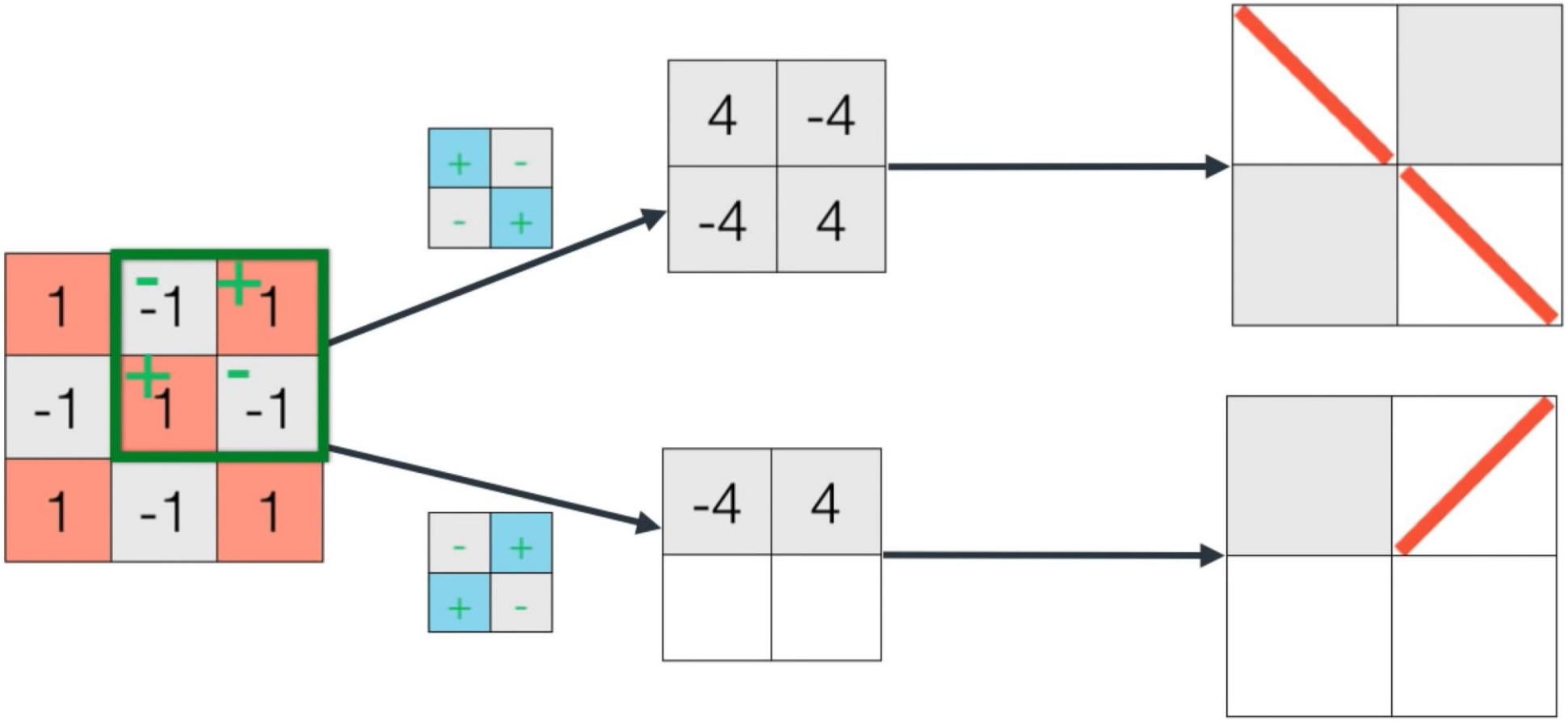
Convolution Neural Network (CNN)

A little more complex world



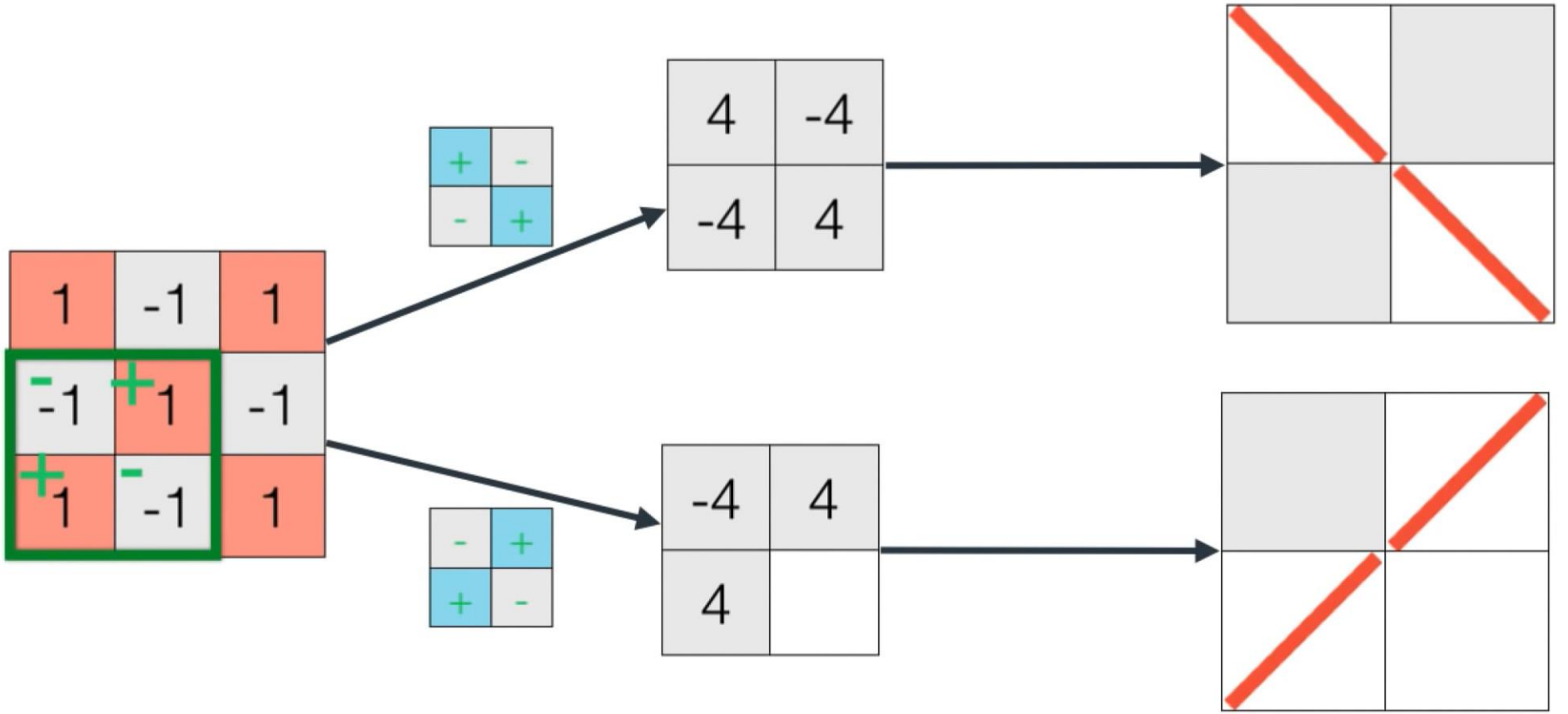
Convolution Neural Network (CNN)

A little more complex world



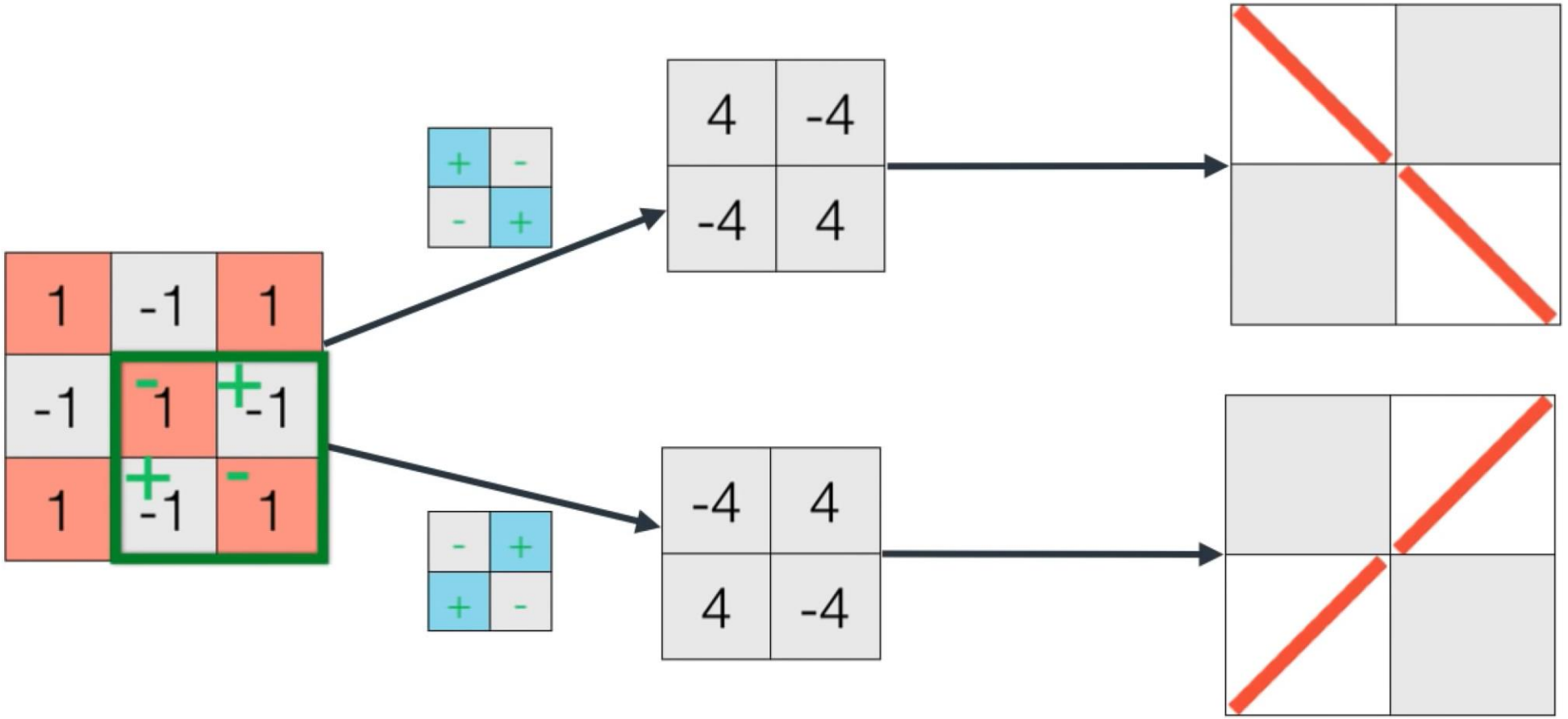
Convolution Neural Network (CNN)

A little more complex world



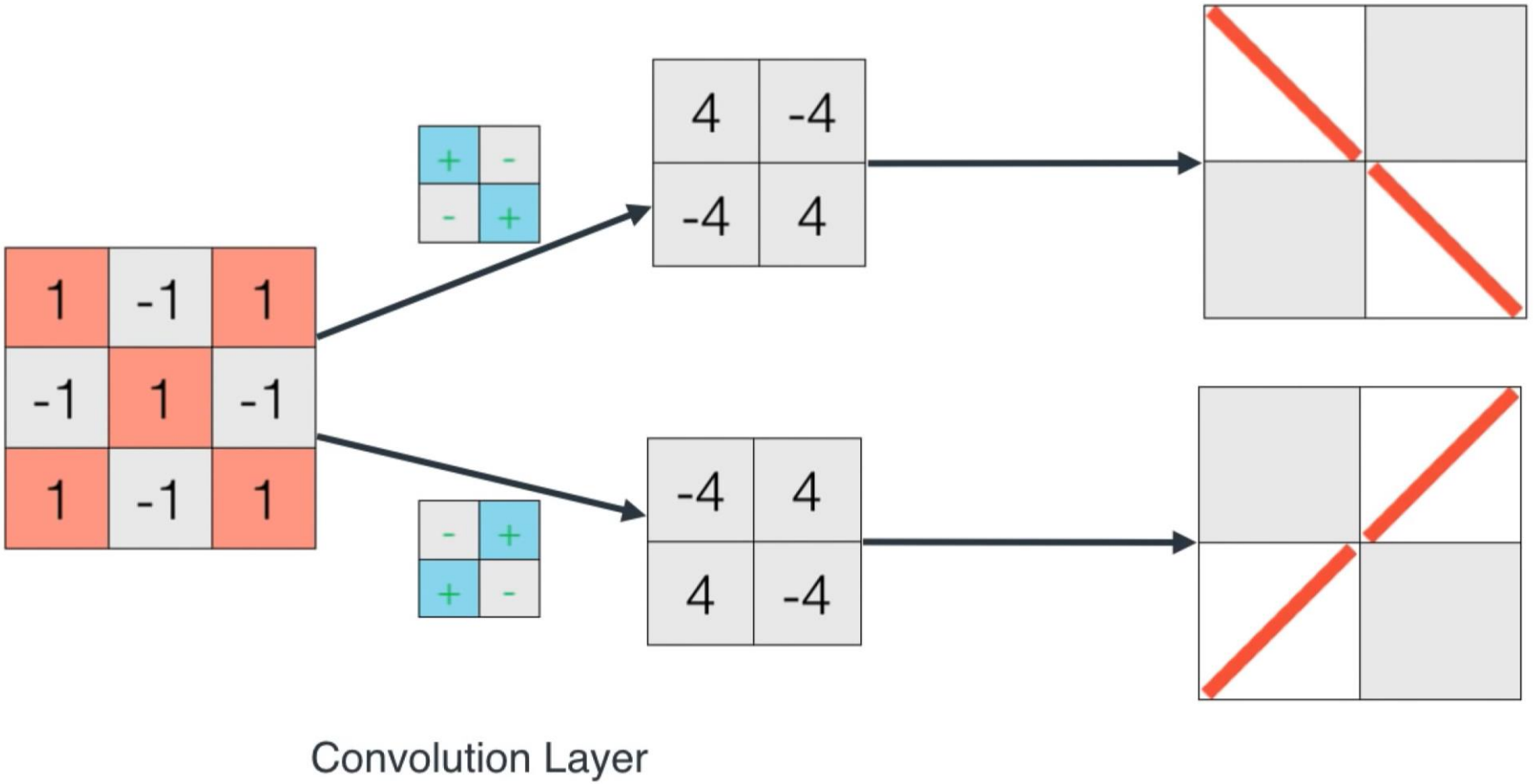
Convolution Neural Network (CNN)

A little more complex world



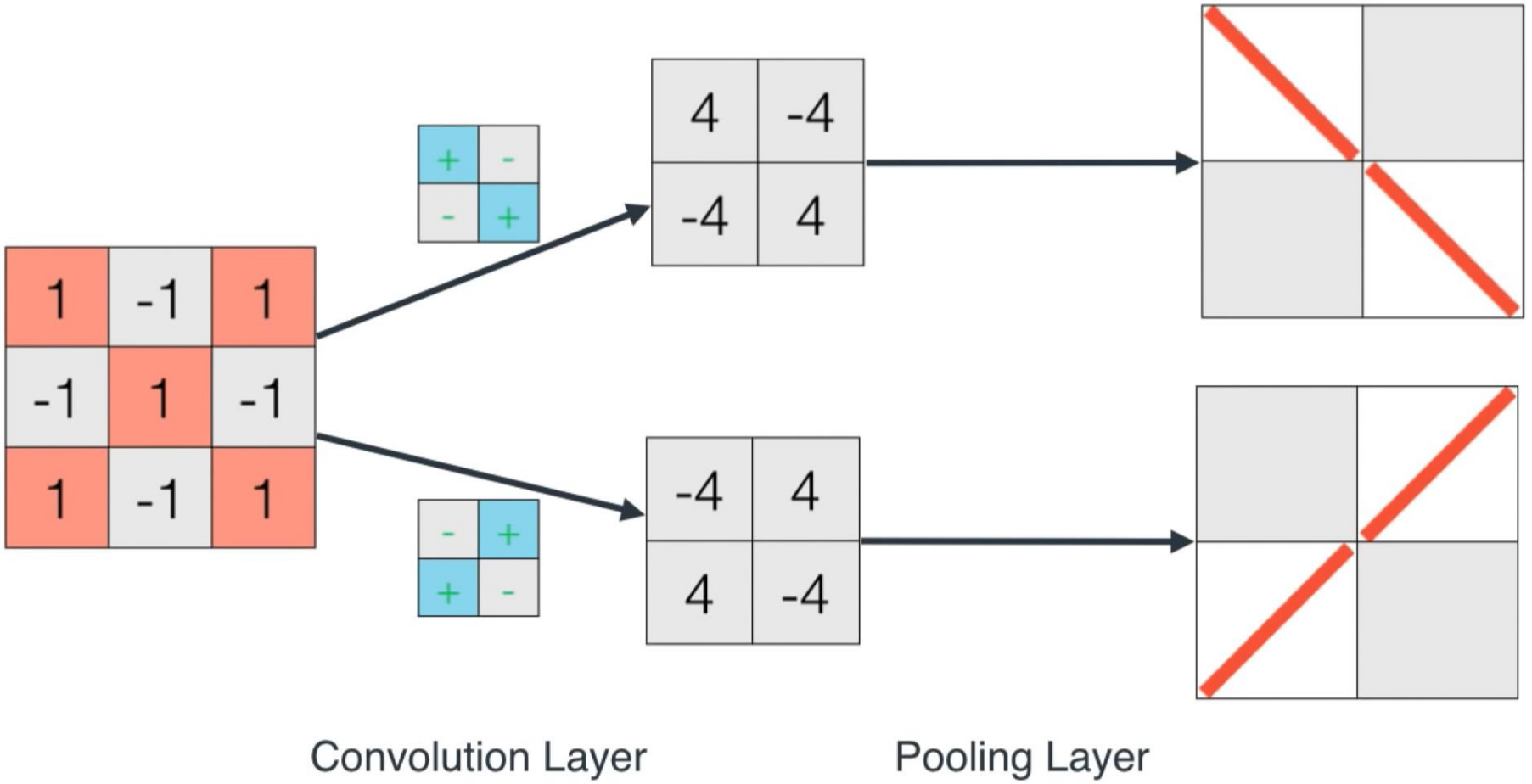
Convolution Neural Network (CNN)

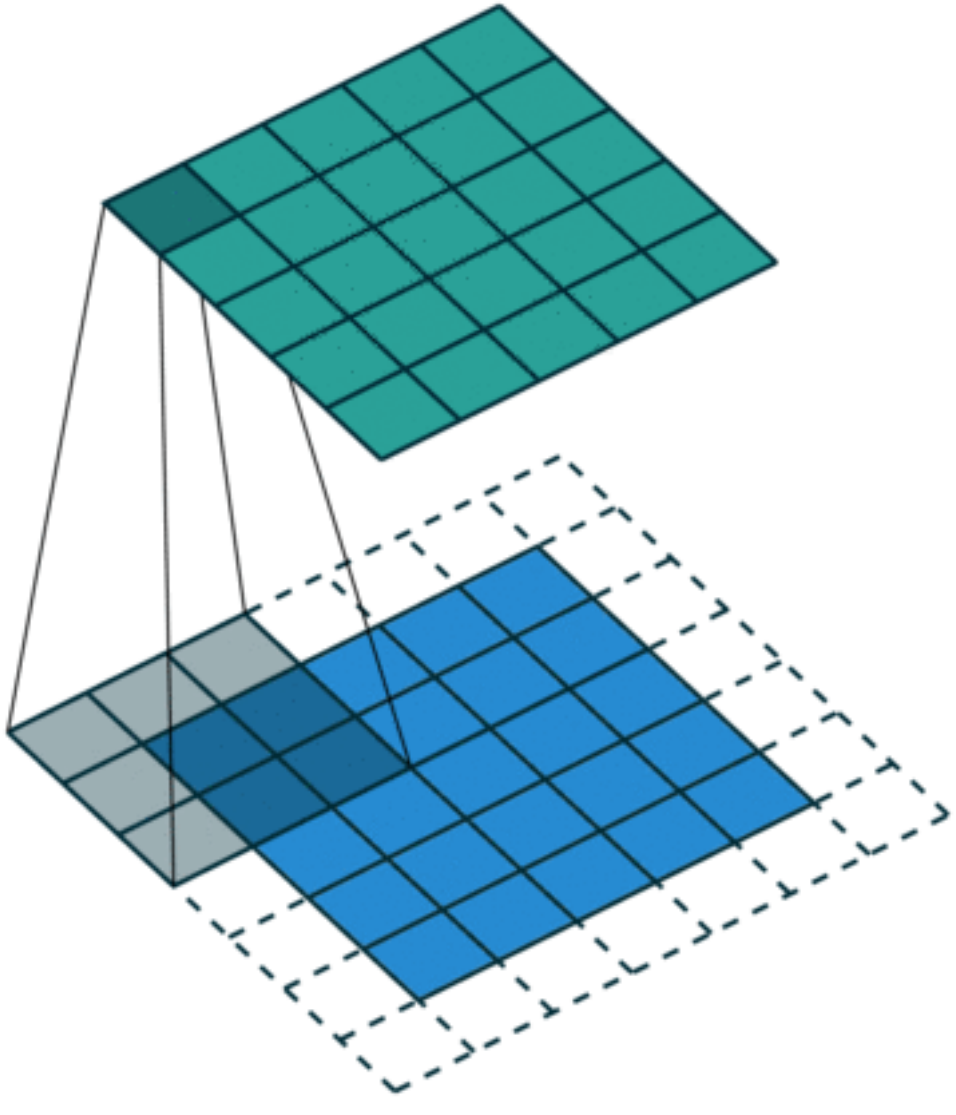
A little more complex world



Convolution Neural Network (CNN)

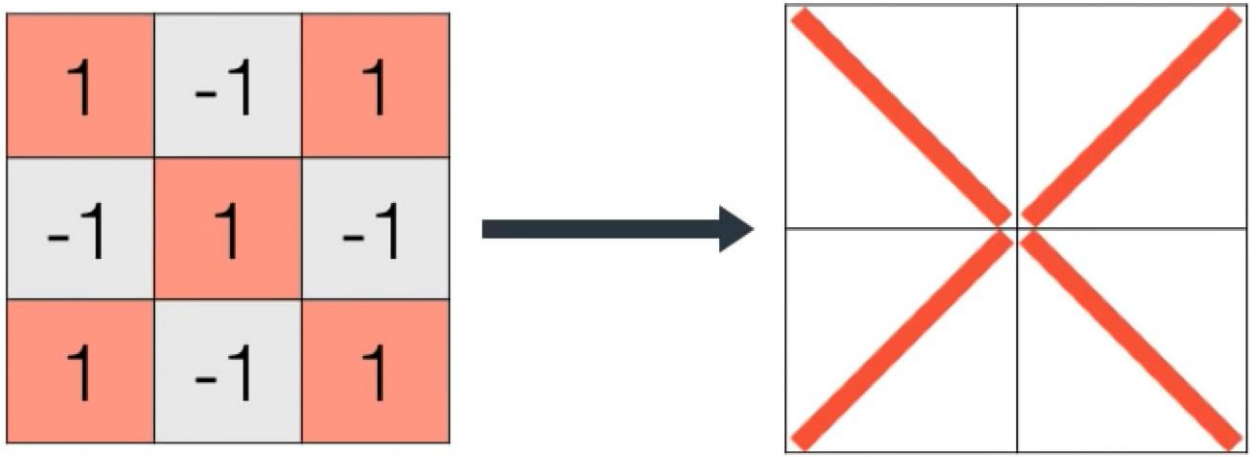
A little more complex world





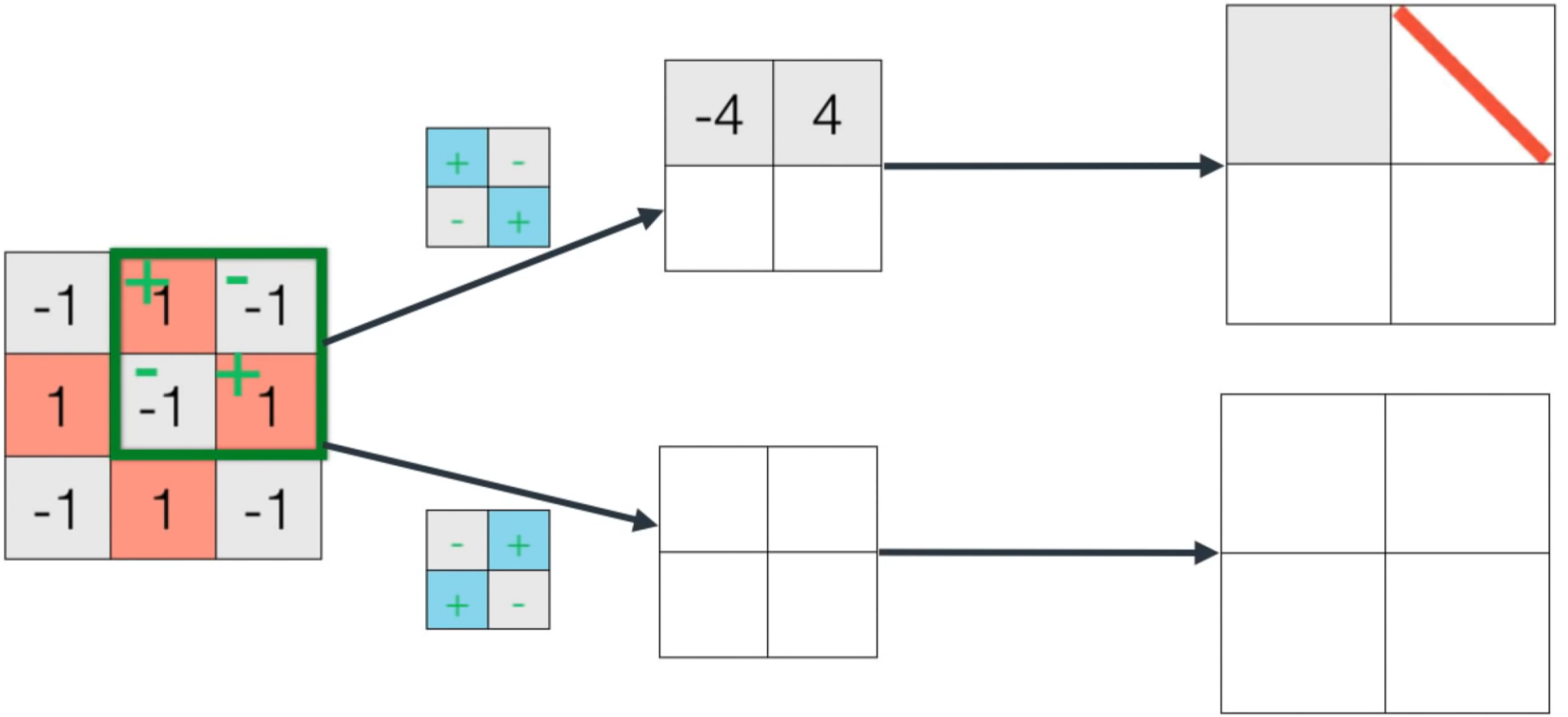
Convolution Neural Network (CNN)

A little more complex world



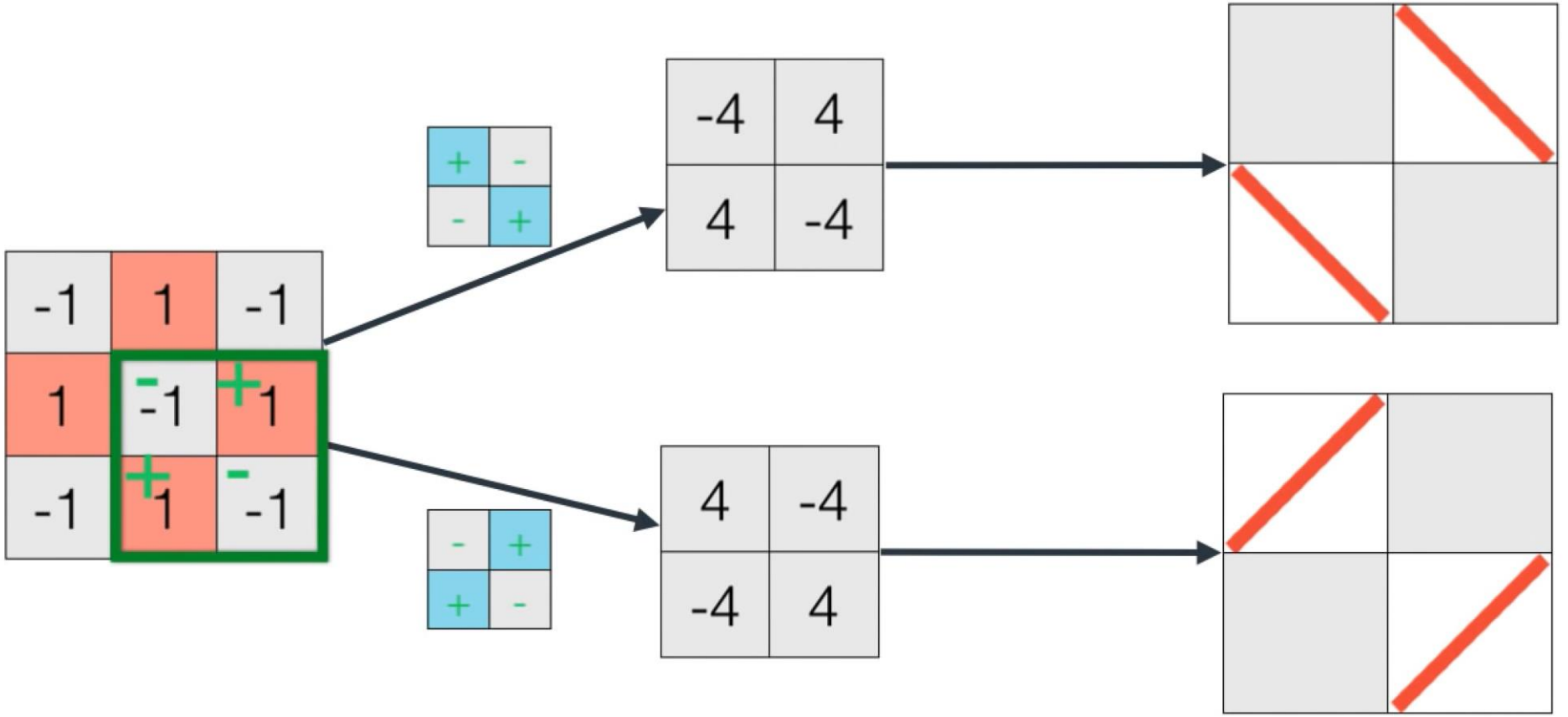
Convolution Neural Network (CNN)

A little more complex world



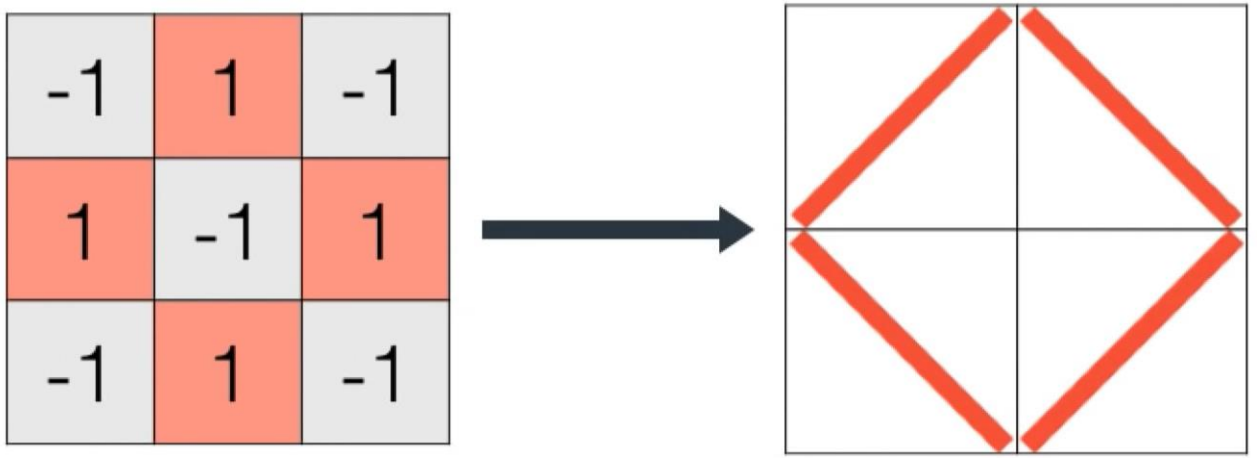
Convolution Neural Network (CNN)

A little more complex world



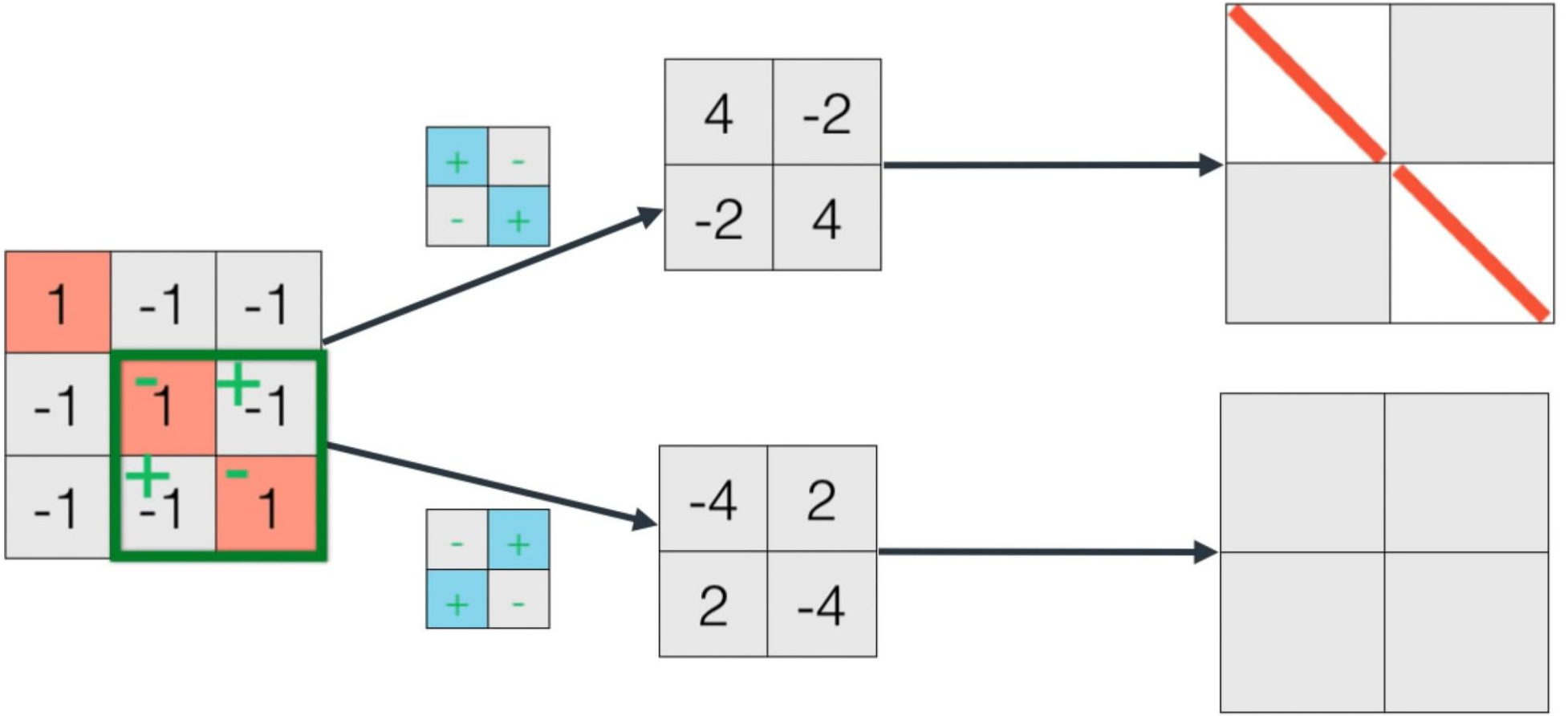
Convolution Neural Network (CNN)

A little more complex world



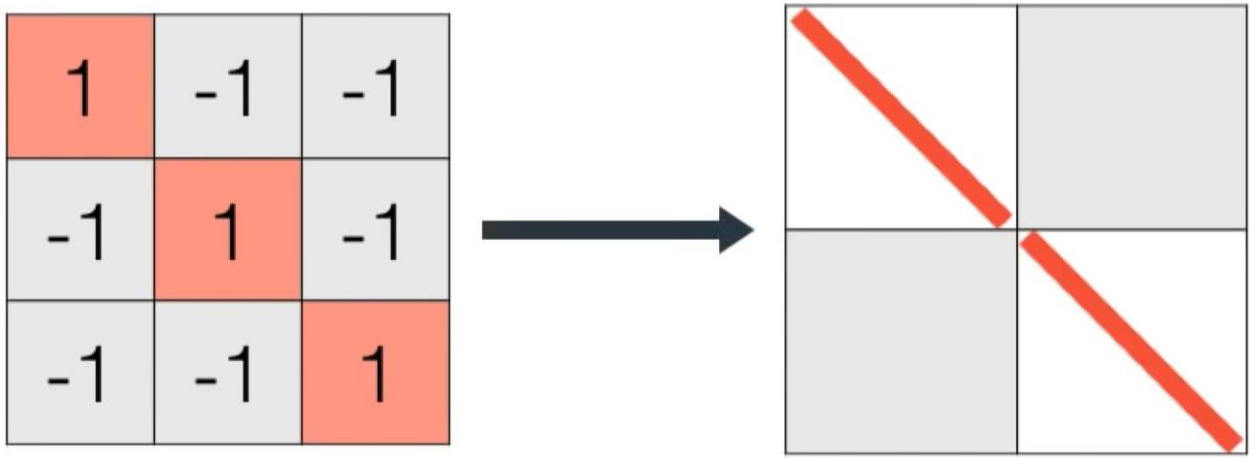
Convolution Neural Network (CNN)

A little more complex world



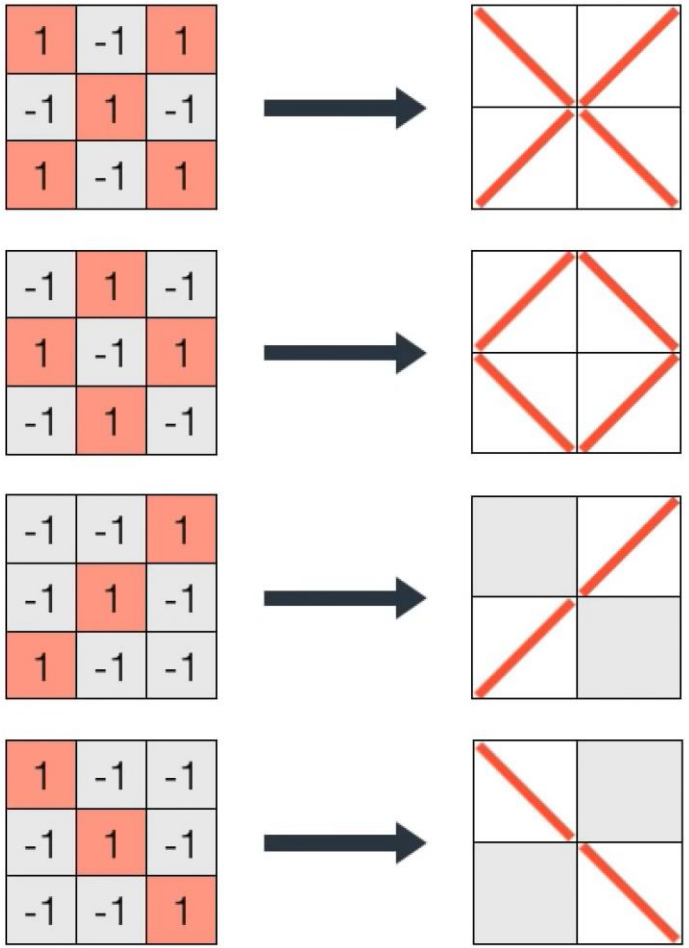
Convolution Neural Network (CNN)

A little more complex world



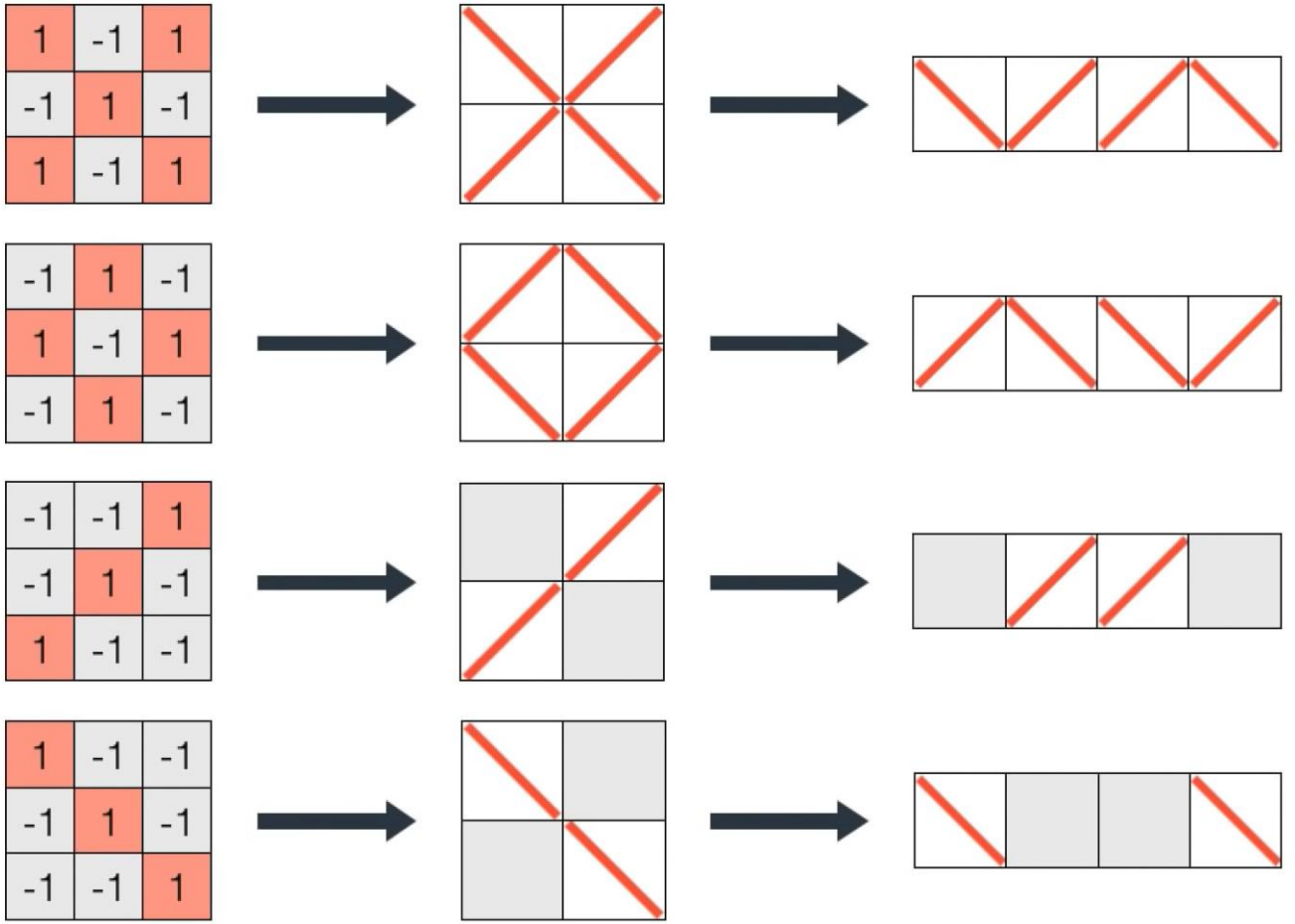
Convolution Neural Network (CNN)

A little more complex world



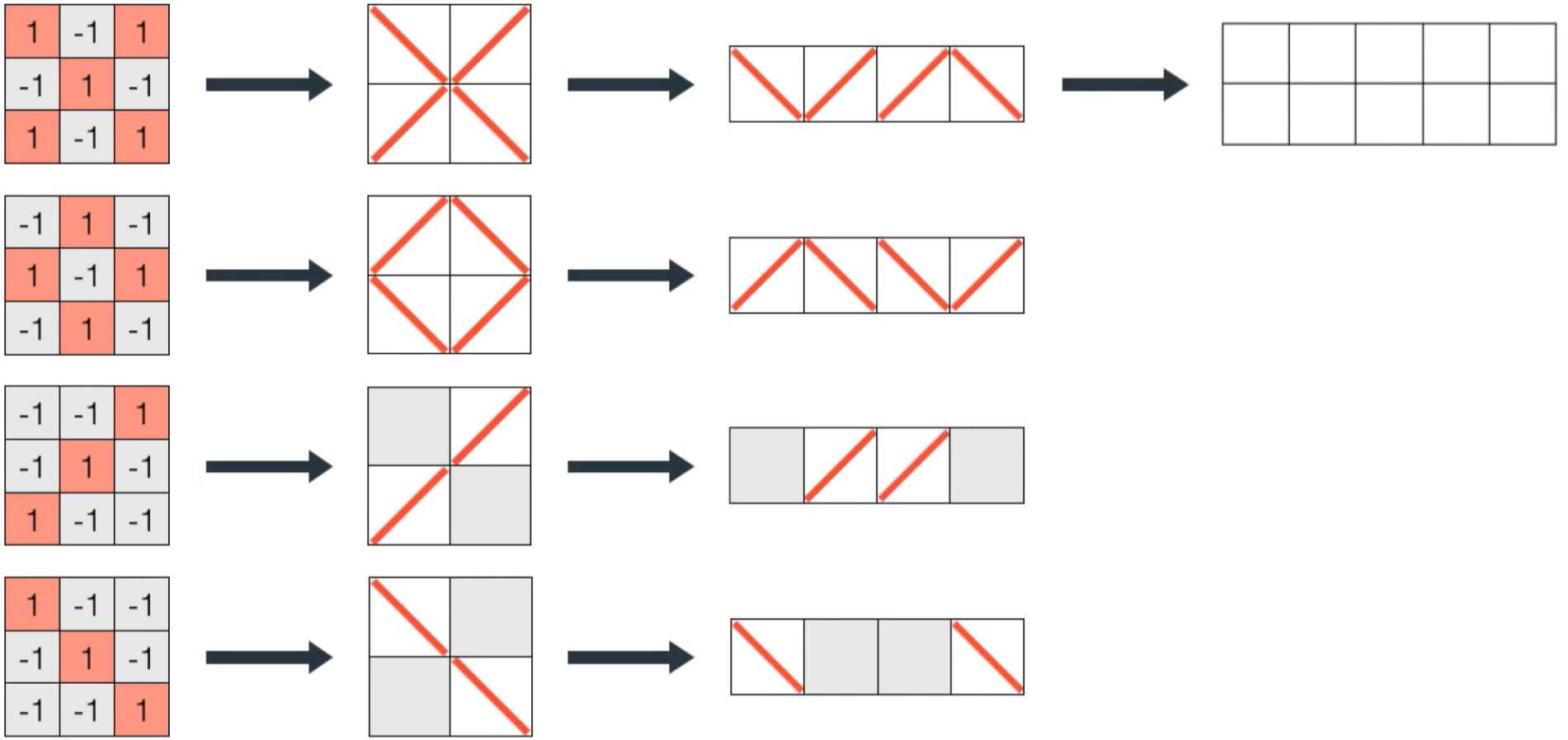
Convolution Neural Network (CNN)

A little more complex world



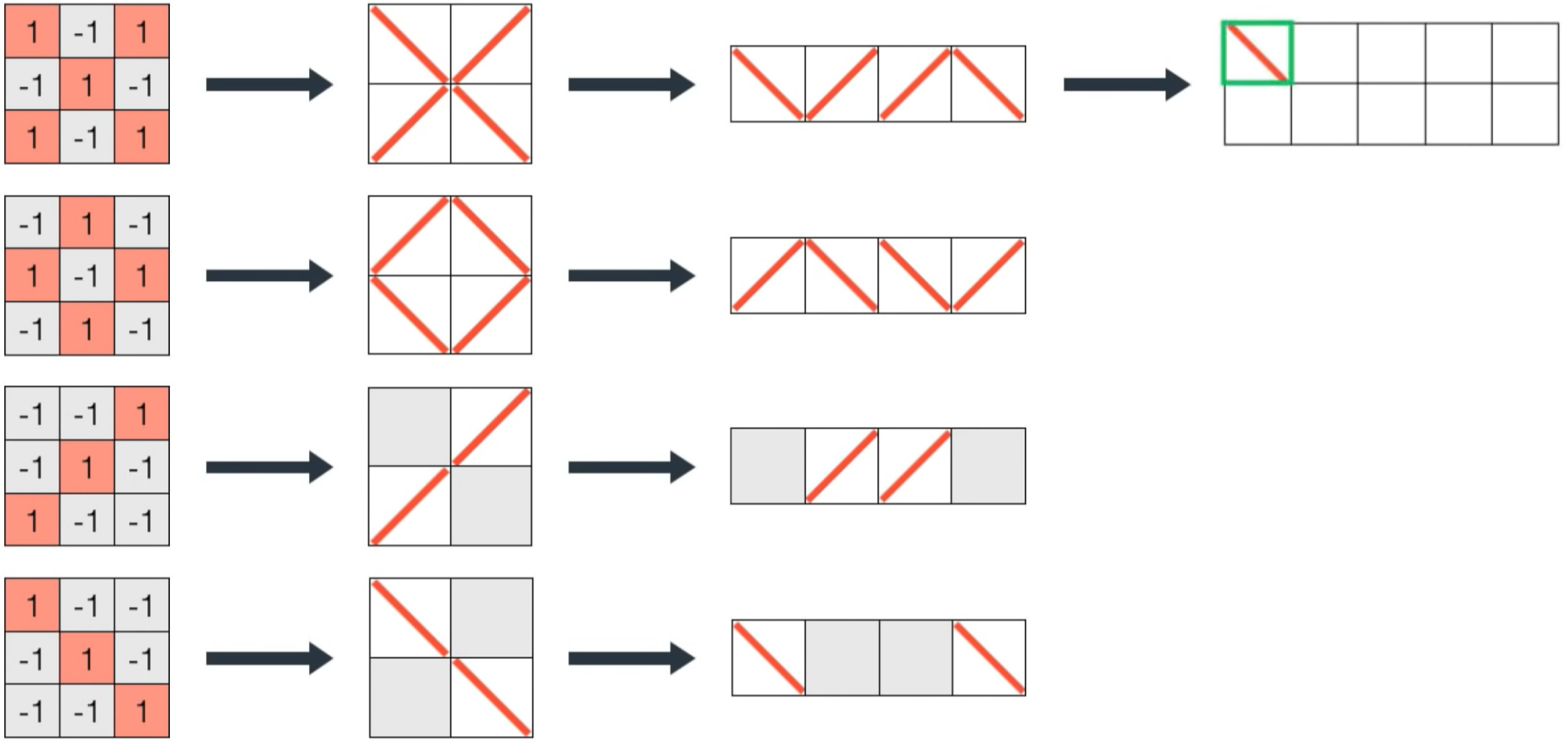
Convolution Neural Network (CNN)

A little more complex world



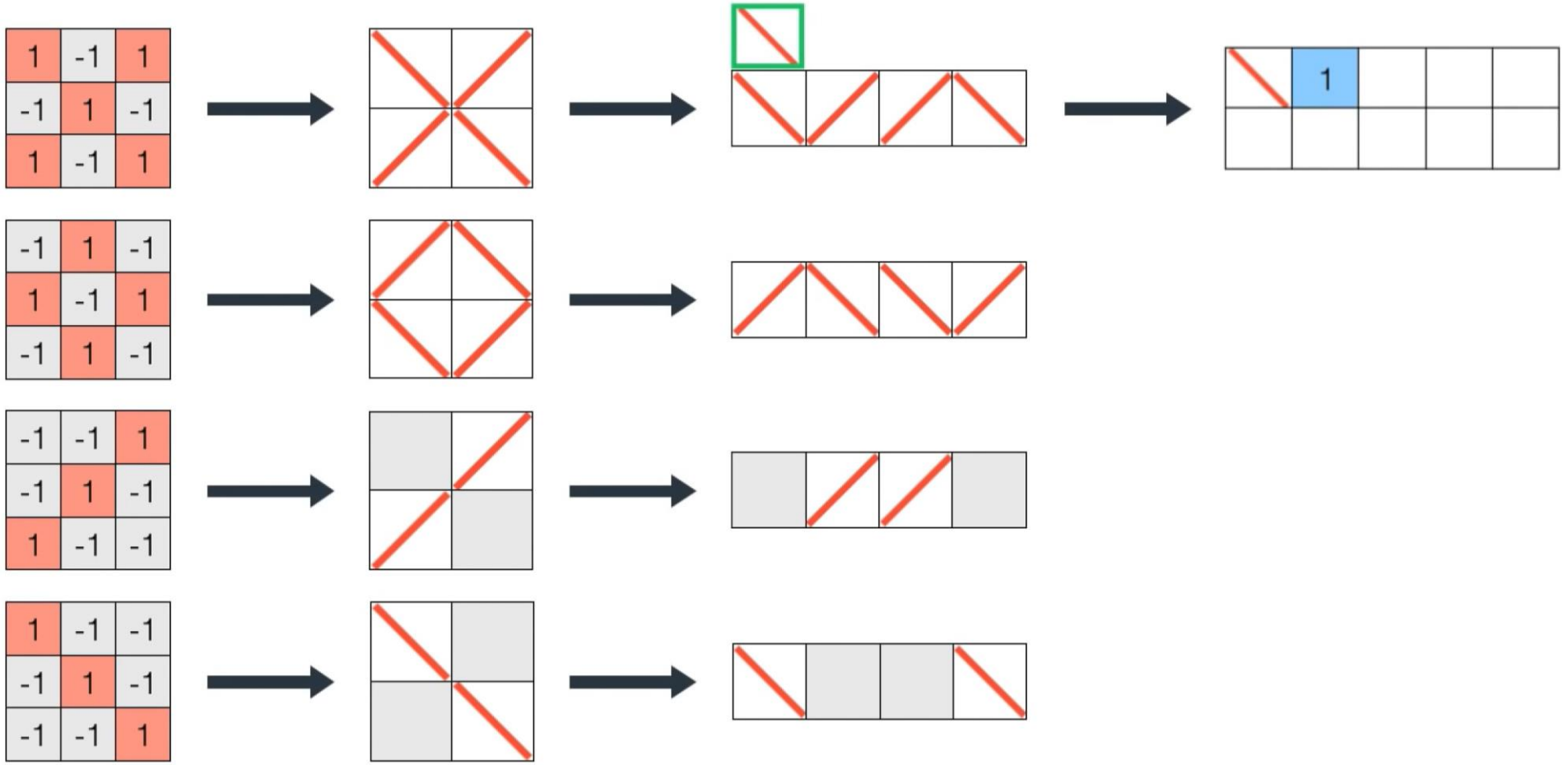
Convolution Neural Network (CNN)

A little more complex world



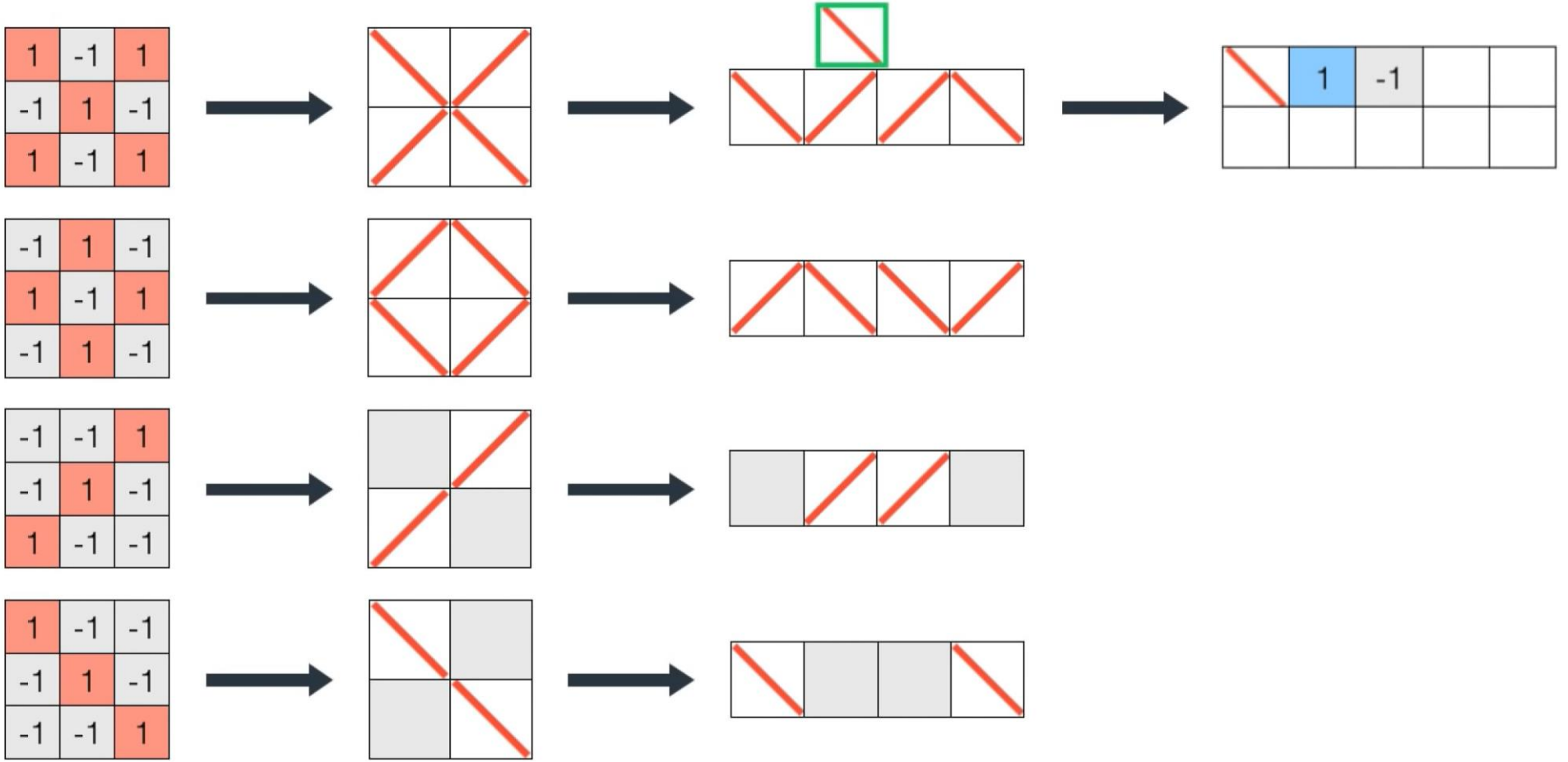
Convolution Neural Network (CNN)

A little more complex world



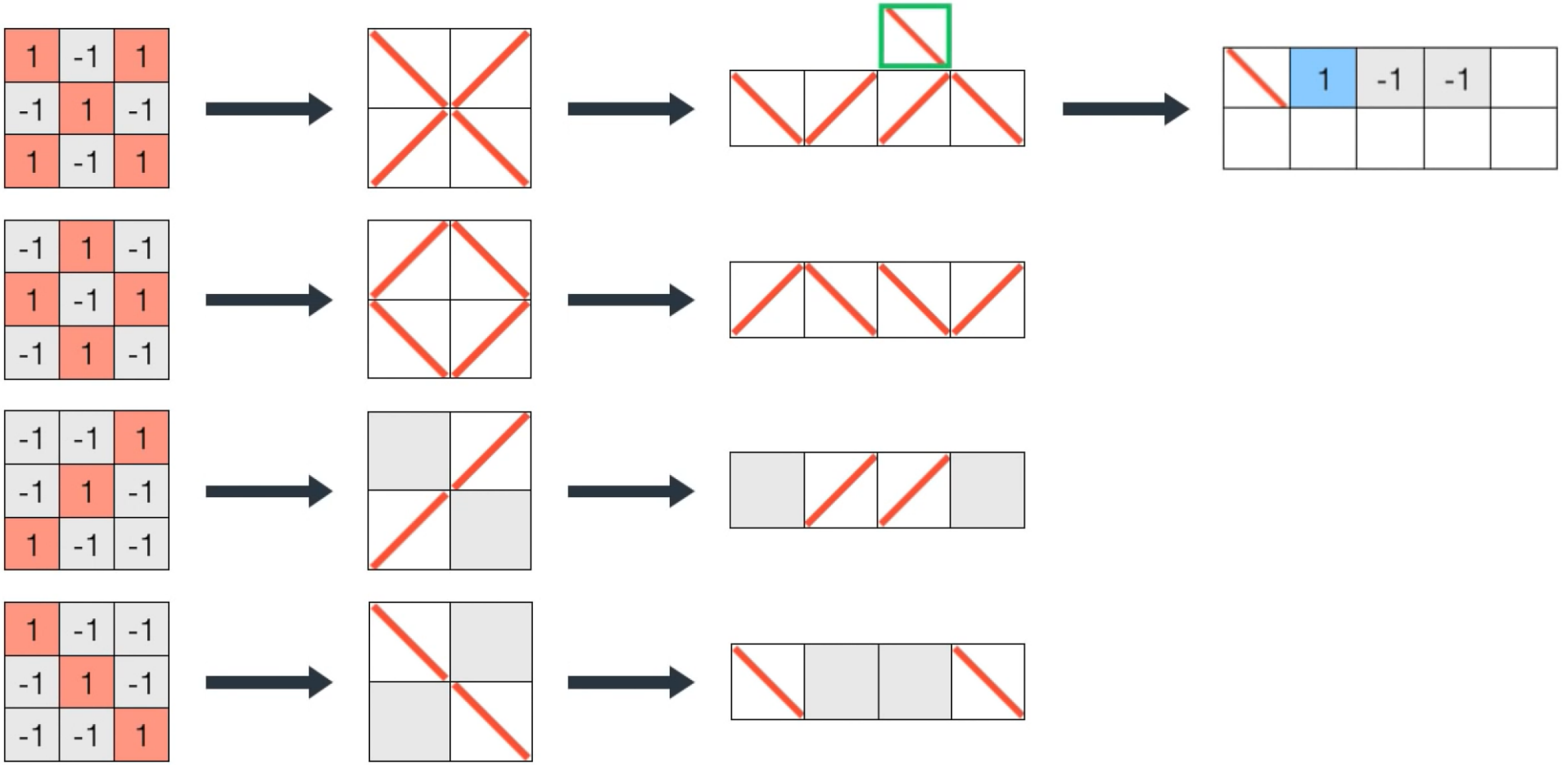
Convolution Neural Network (CNN)

A little more complex world



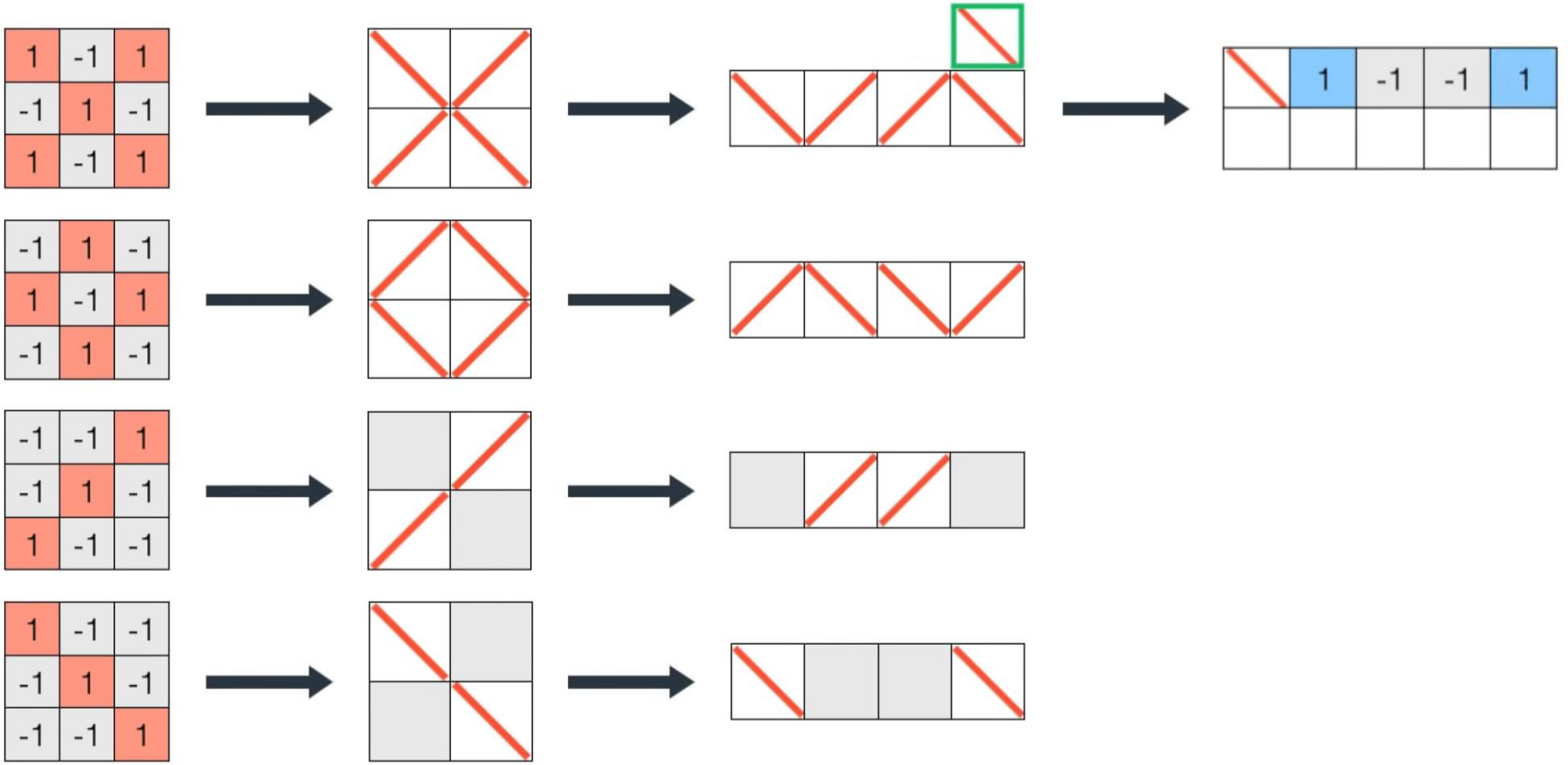
Convolution Neural Network (CNN)

A little more complex world



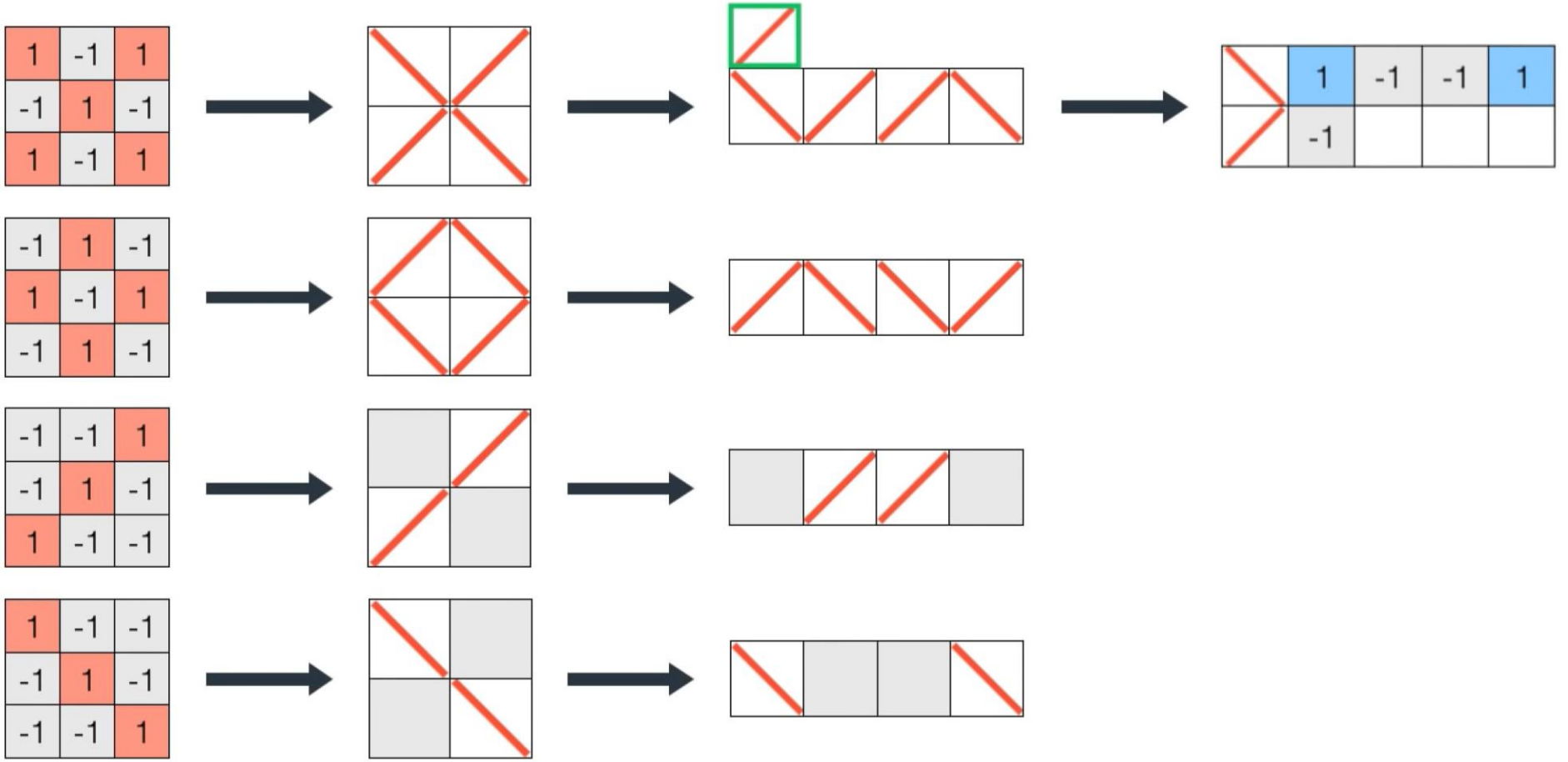
Convolution Neural Network (CNN)

A little more complex world



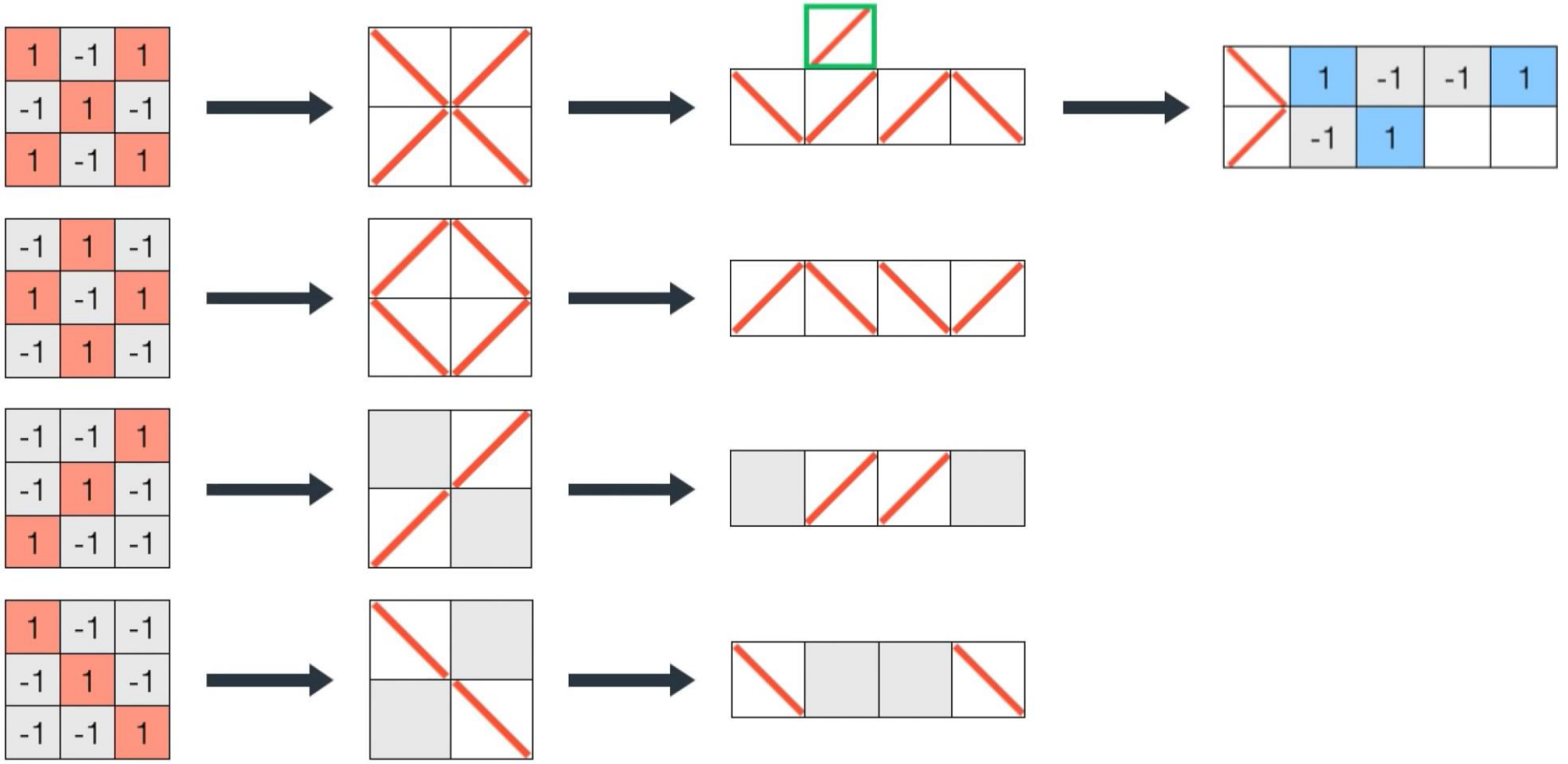
Convolution Neural Network (CNN)

A little more complex world



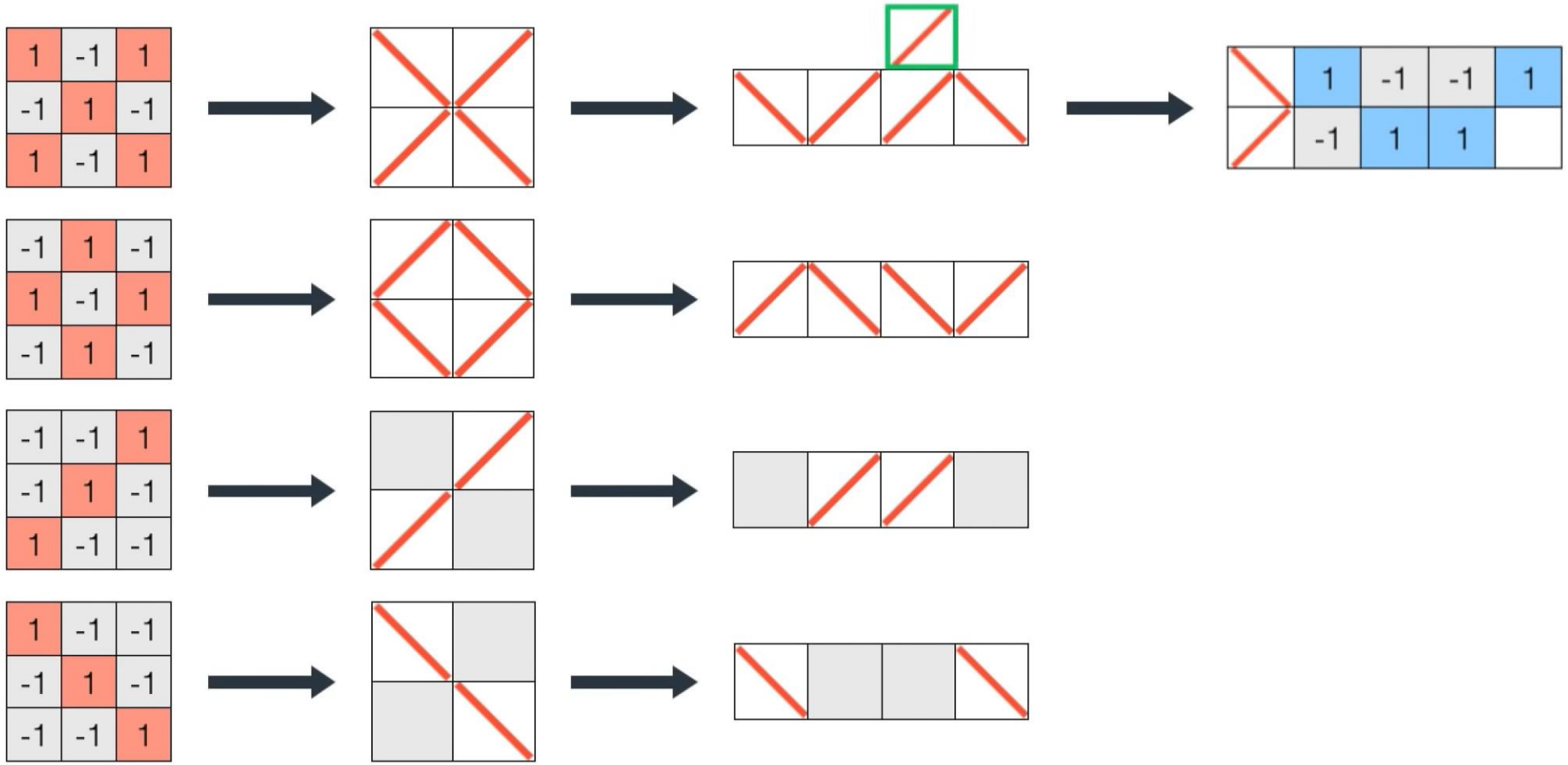
Convolution Neural Network (CNN)

A little more complex world



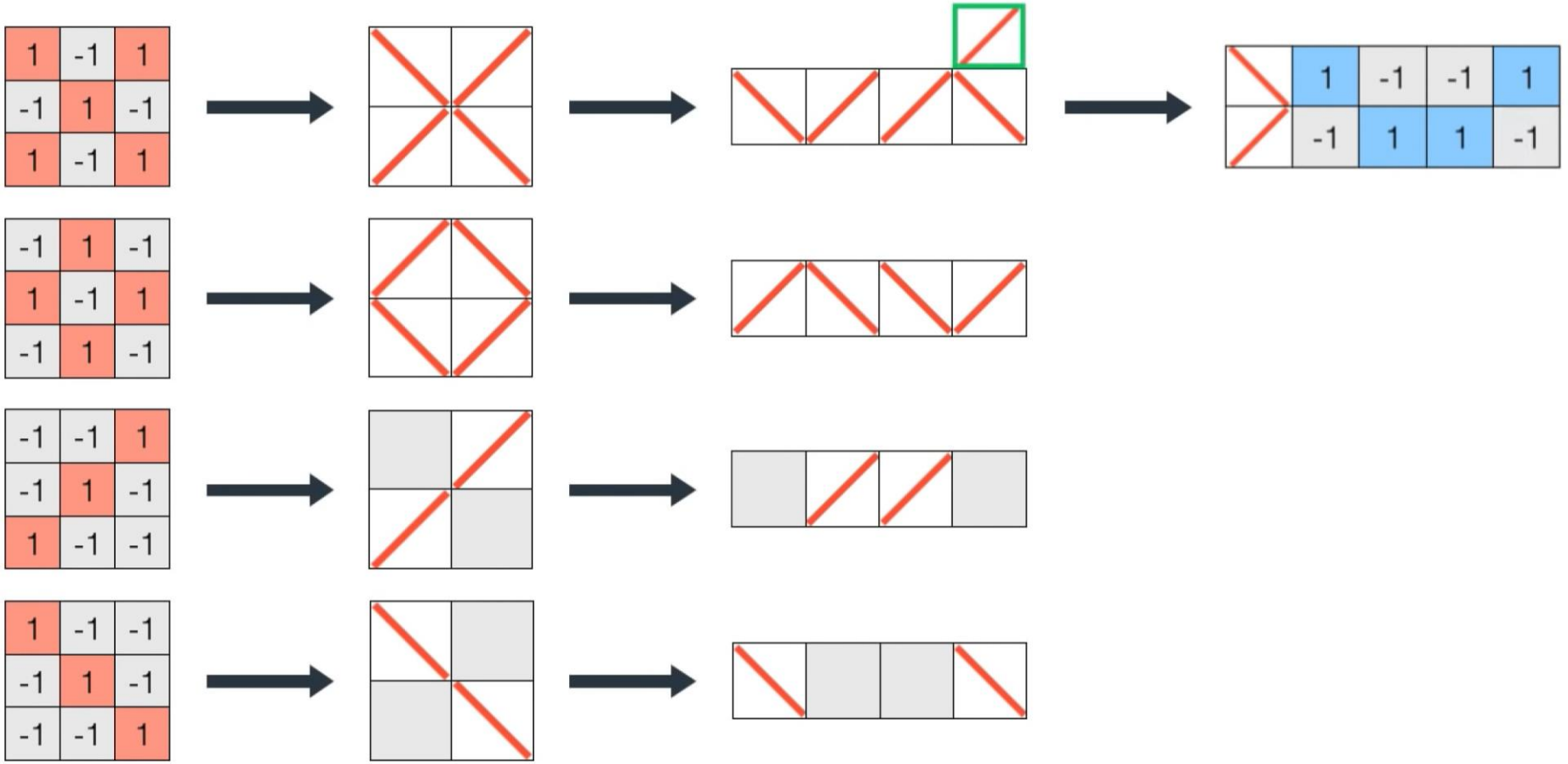
Convolution Neural Network (CNN)

A little more complex world



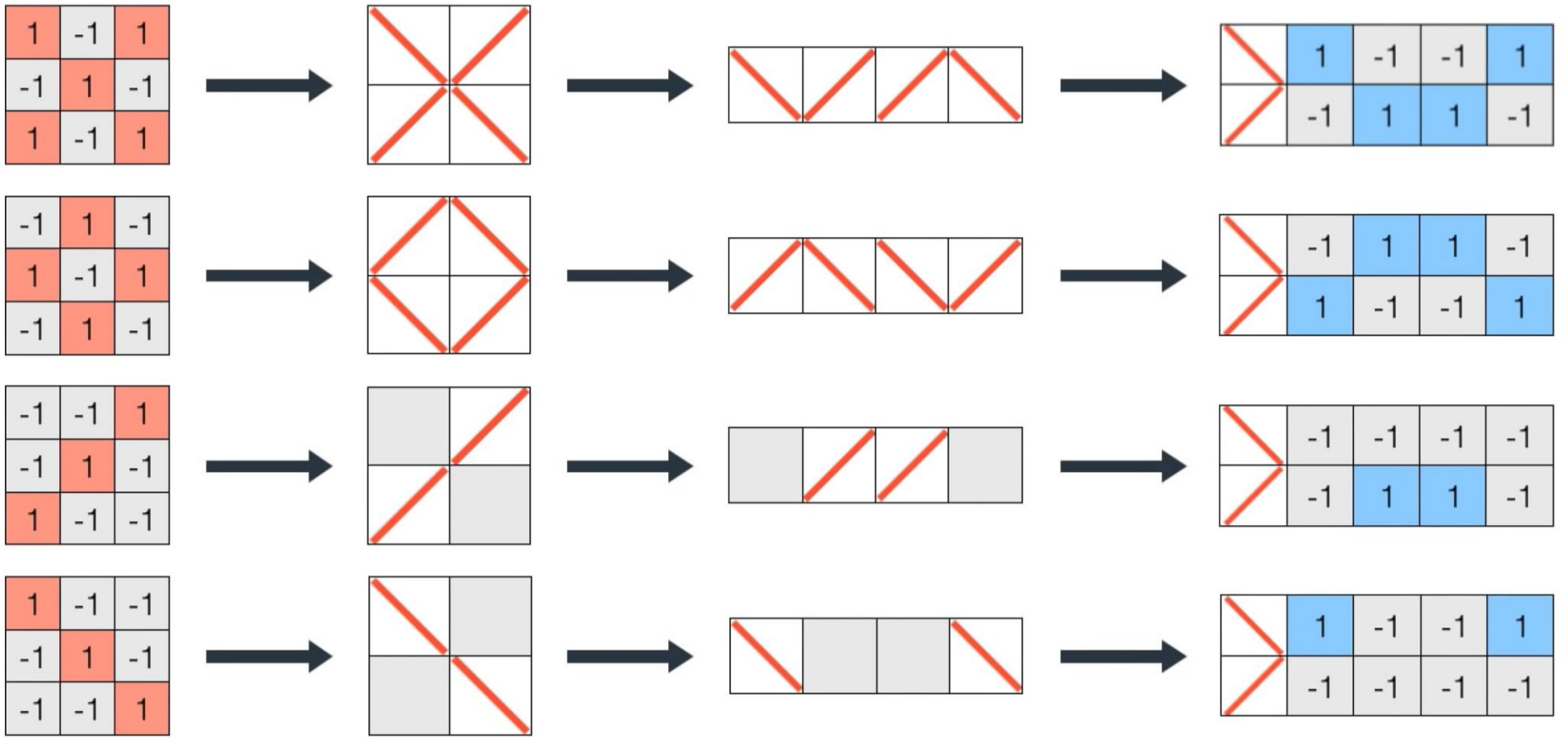
Convolution Neural Network (CNN)

A little more complex world



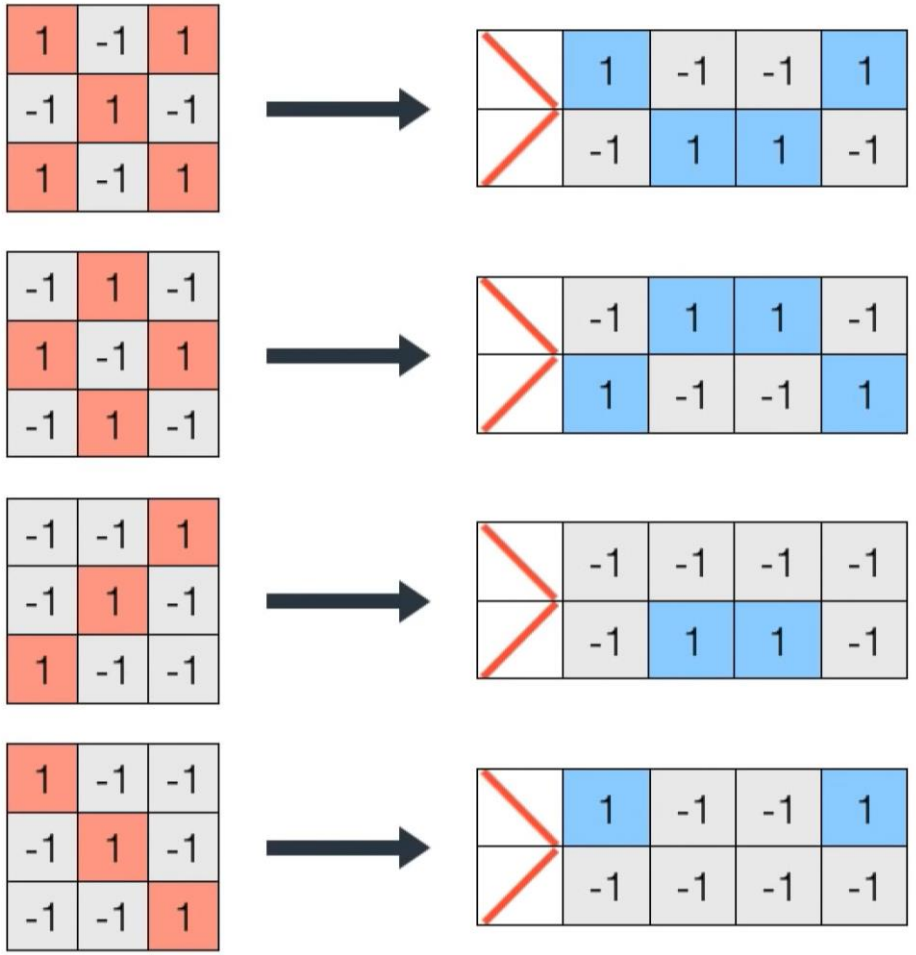
Convolution Neural Network (CNN)

A little more complex world



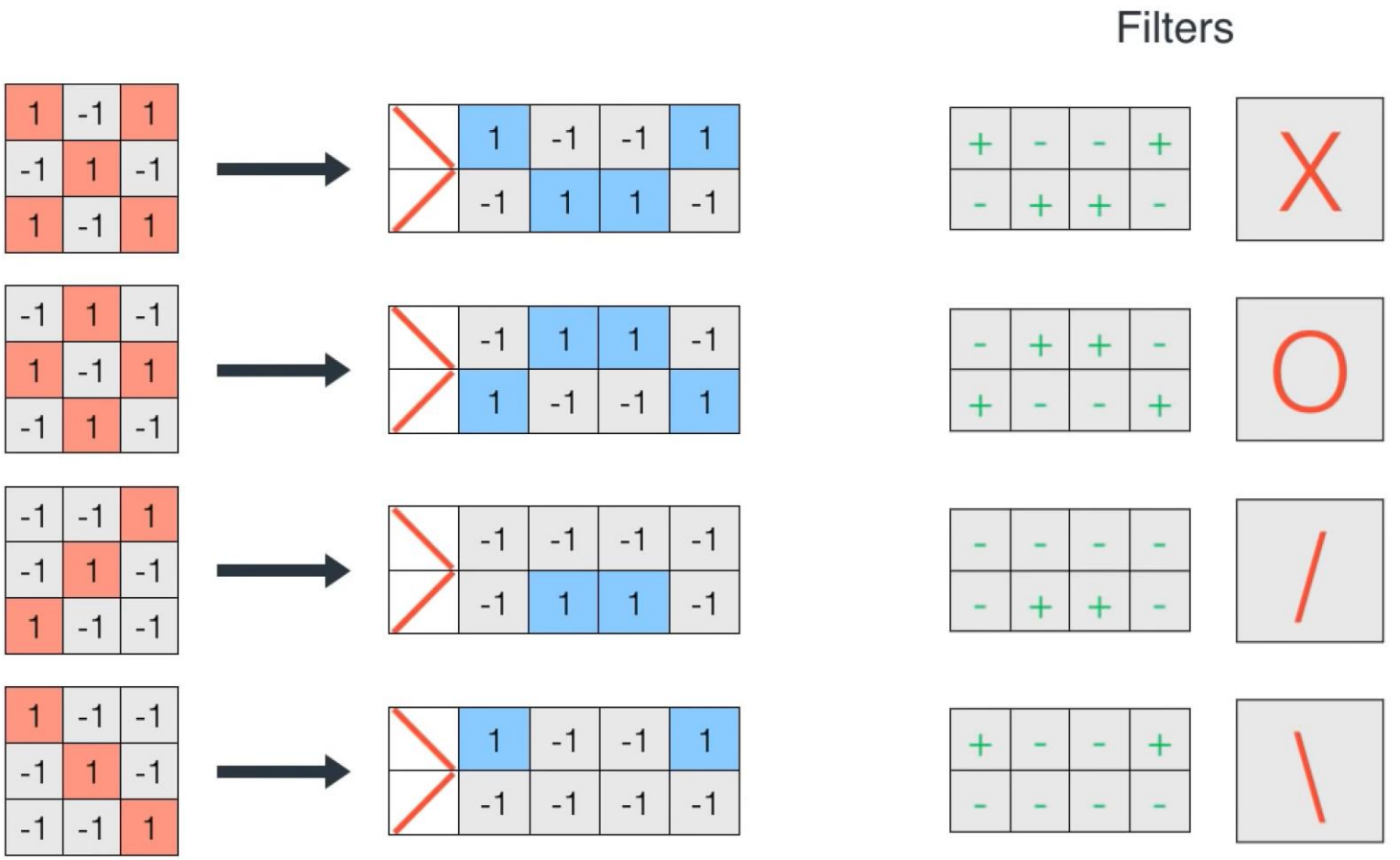
Convolution Neural Network (CNN)

A little more complex world



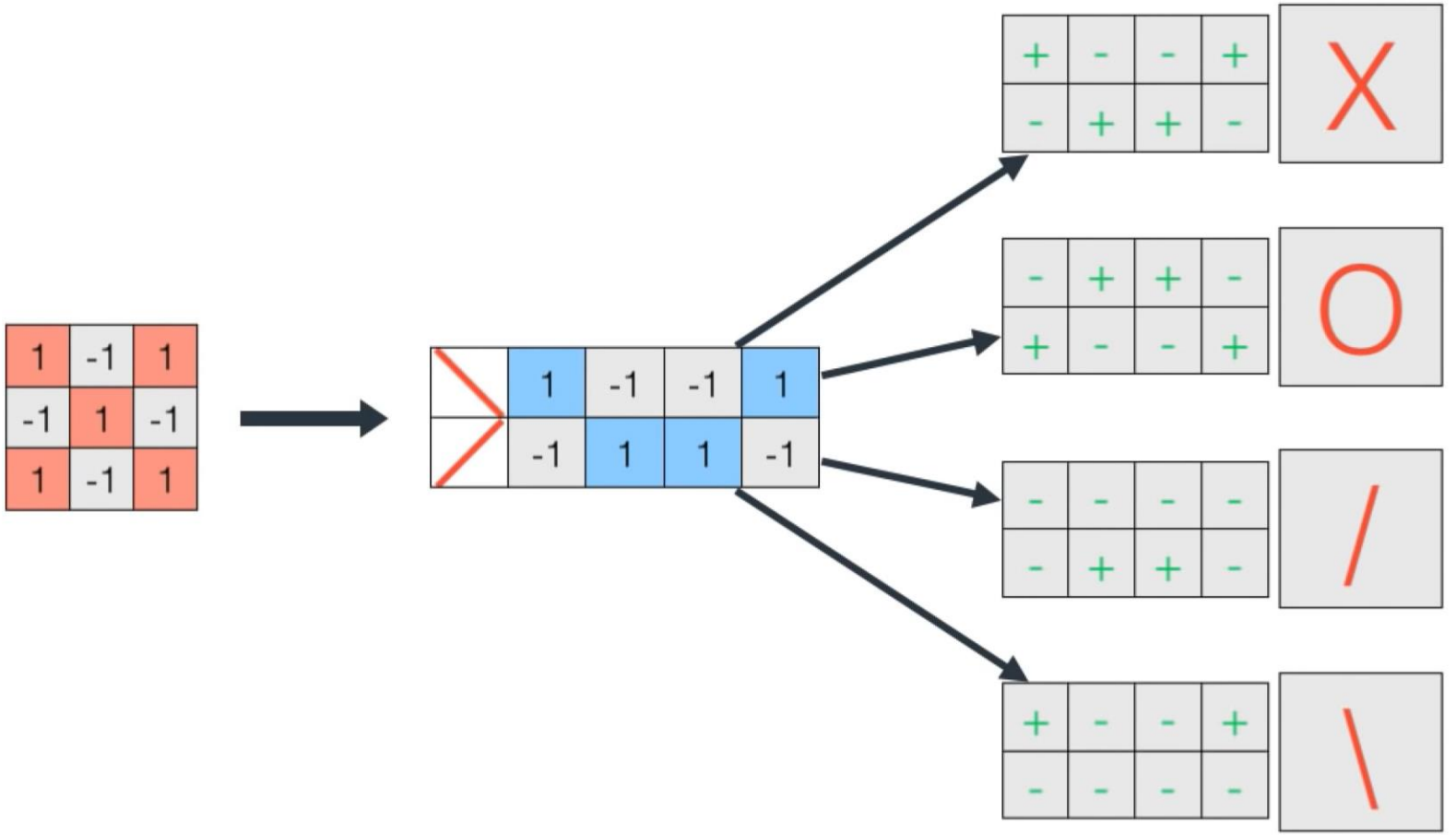
Convolution Neural Network (CNN)

A little more complex world



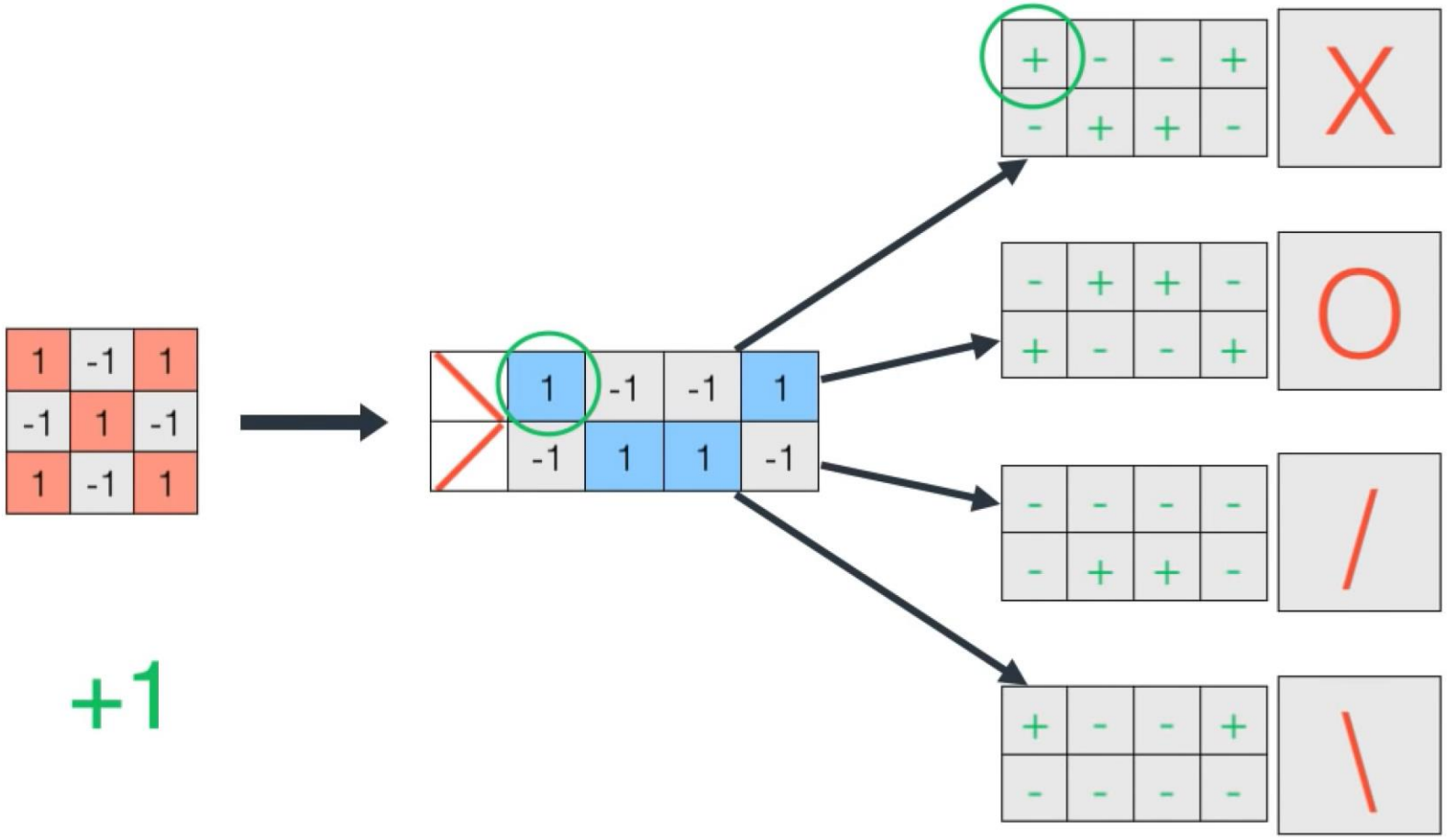
Convolution Neural Network (CNN)

A little more complex world



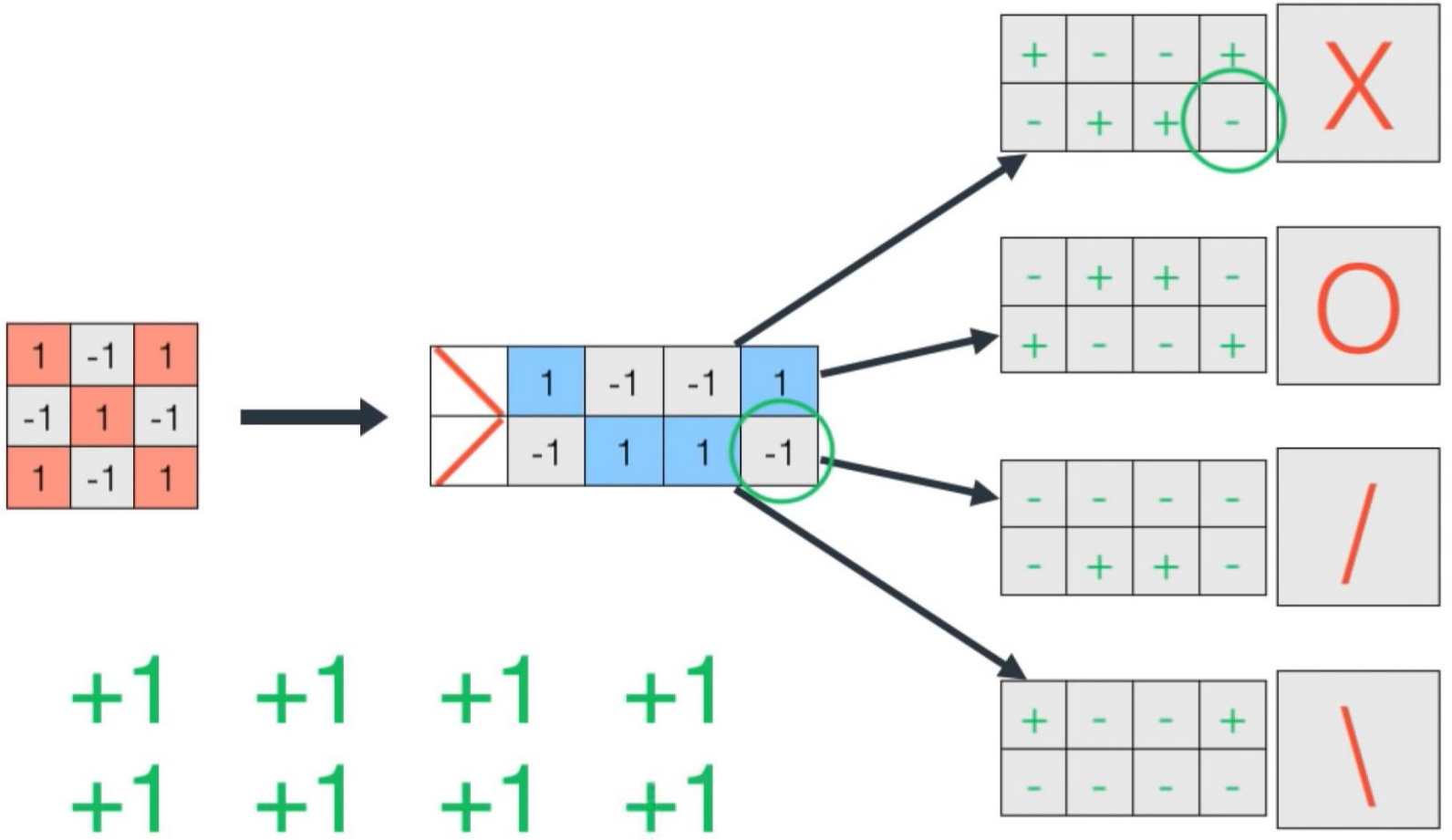
Convolution Neural Network (CNN)

A little more complex world



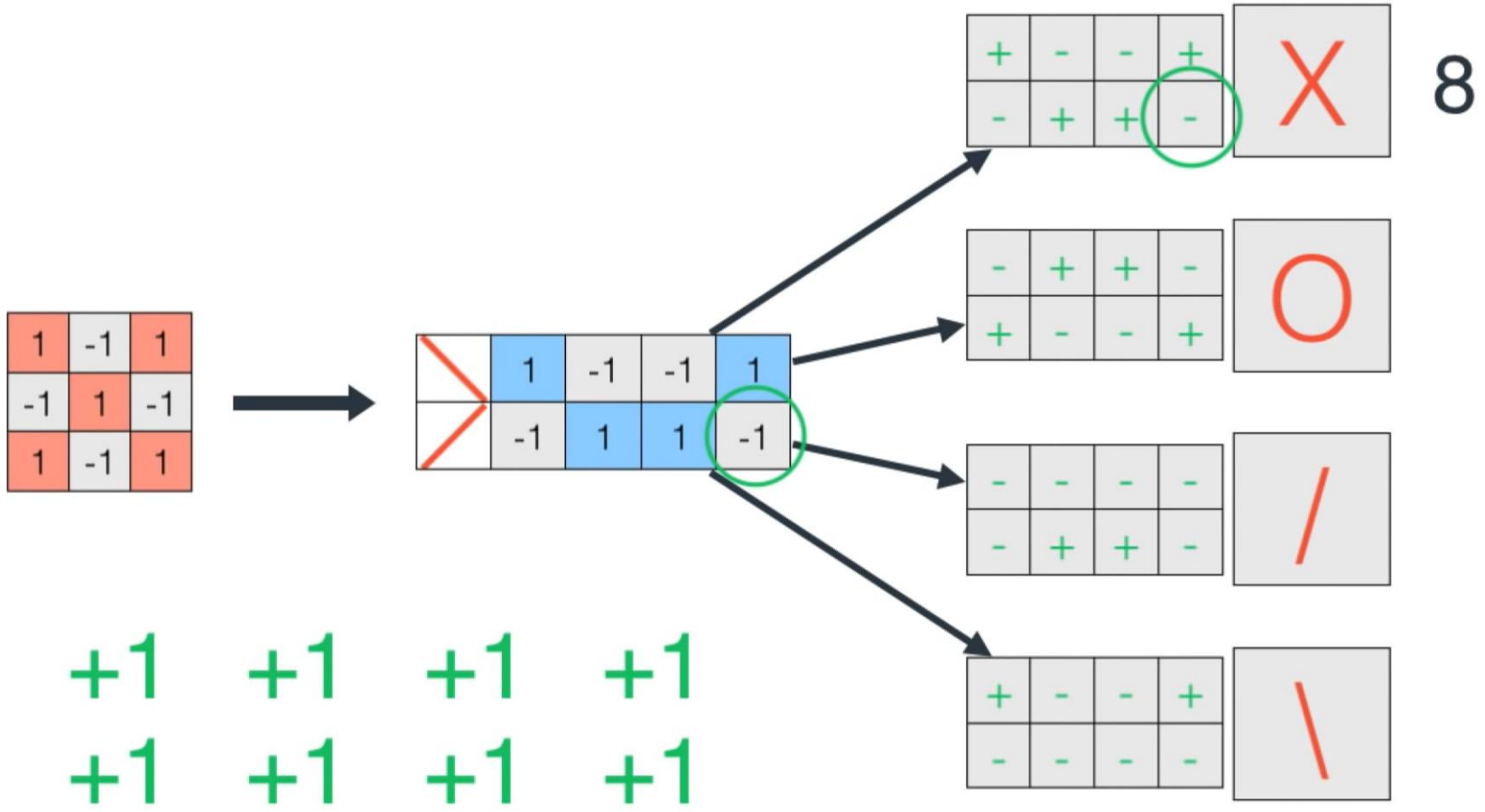
Convolution Neural Network (CNN)

A little more complex world



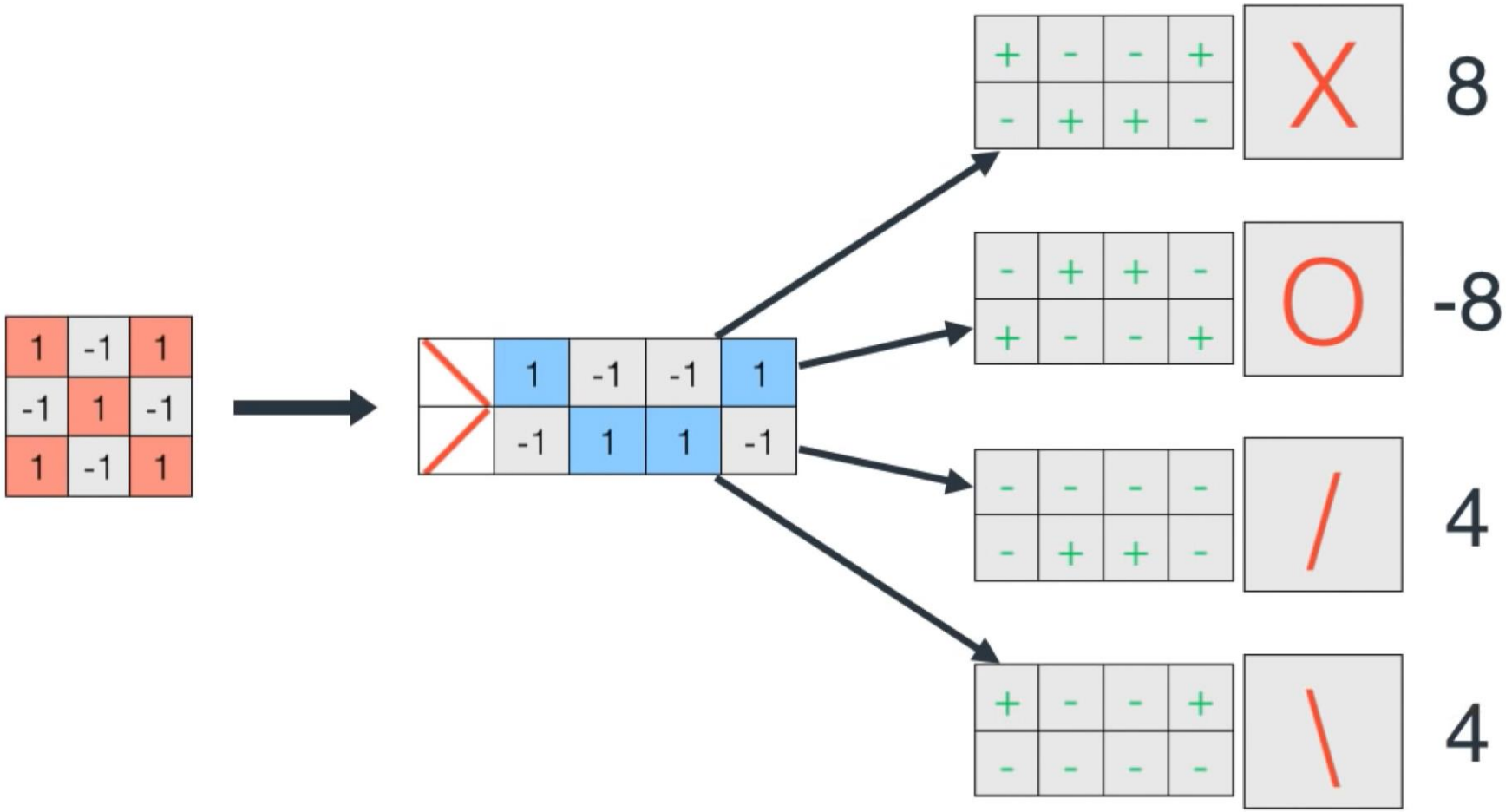
Convolution Neural Network (CNN)

A little more complex world



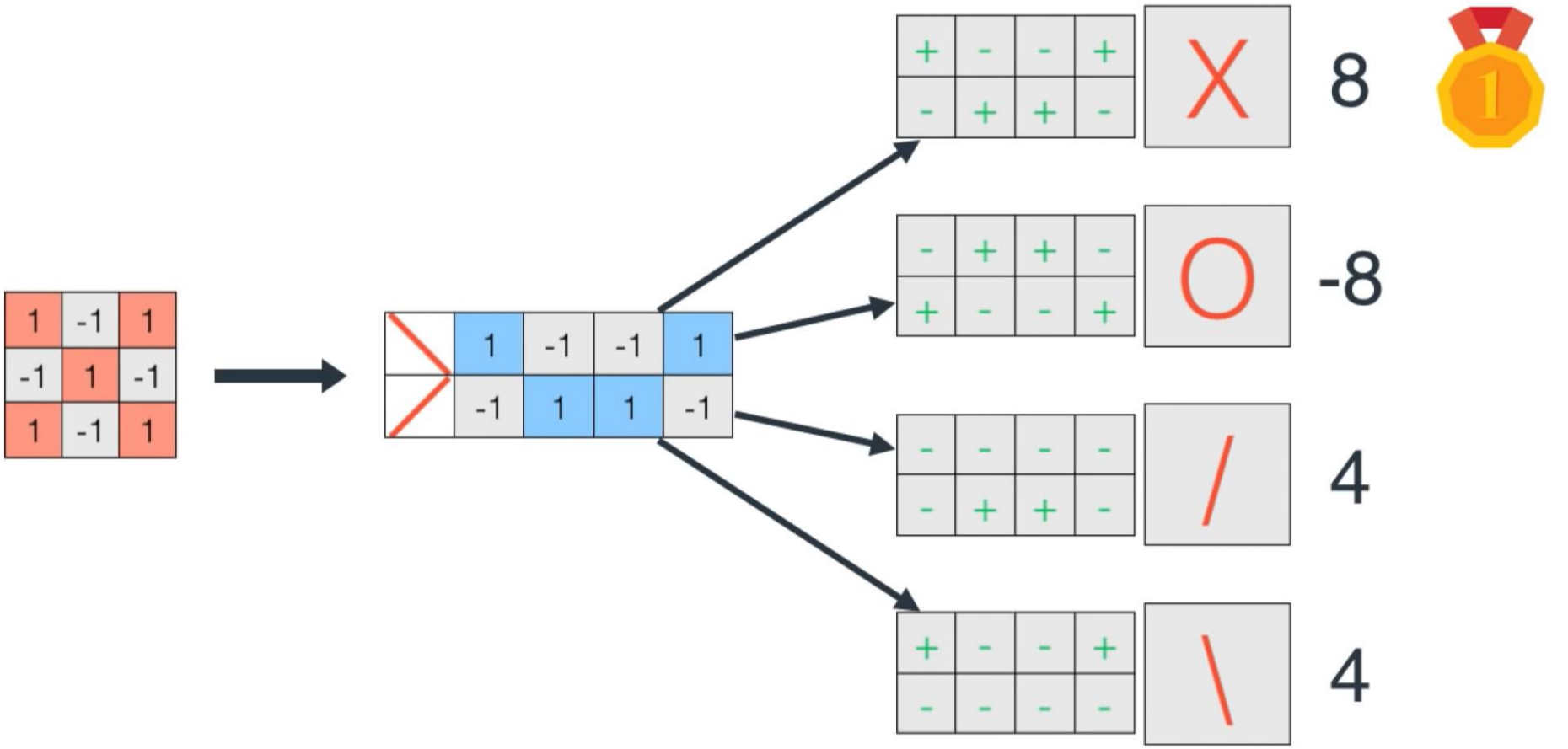
Convolution Neural Network (CNN)

A little more complex world



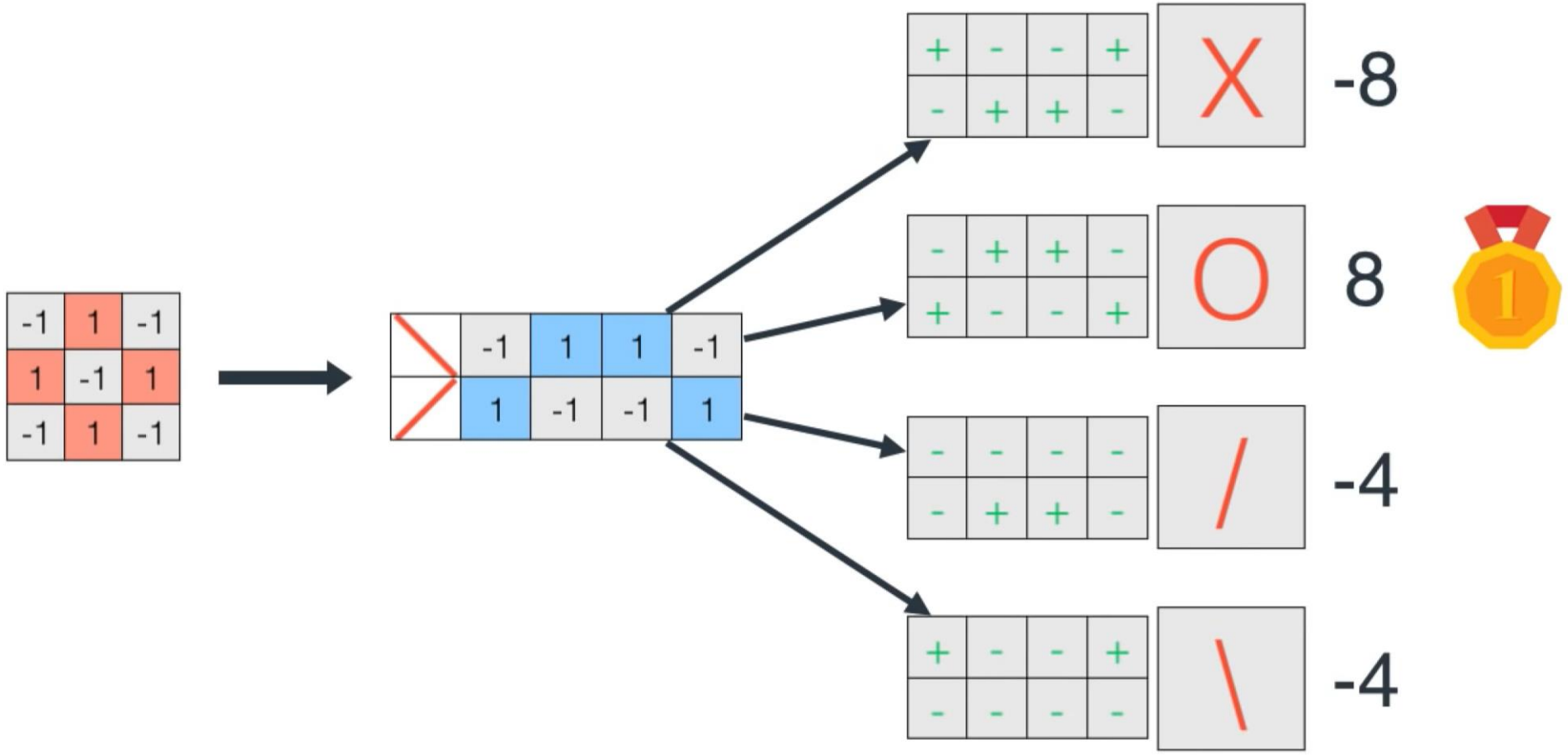
Convolution Neural Network (CNN)

A little more complex world



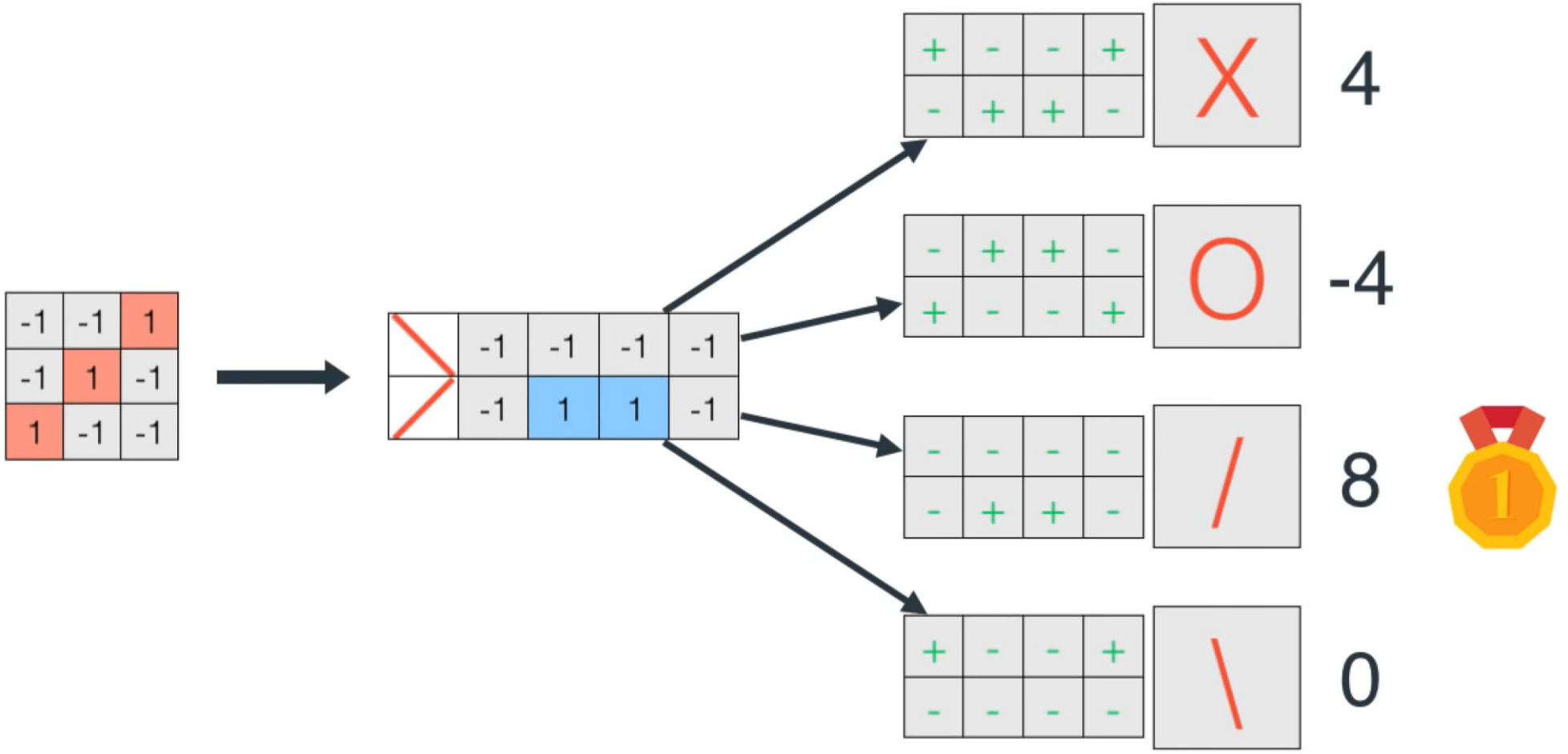
Convolution Neural Network (CNN)

A little more complex world



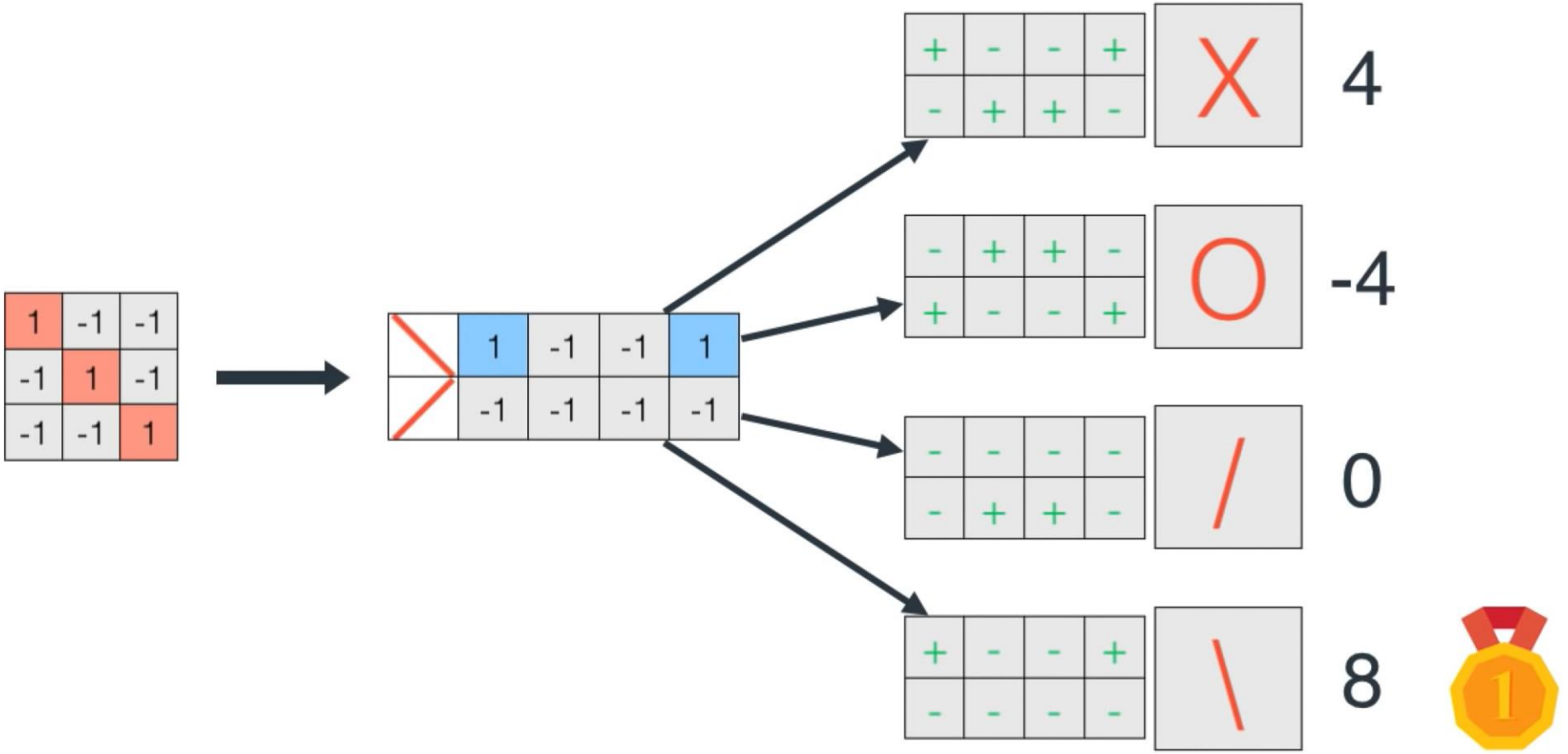
Convolution Neural Network (CNN)

A little more complex world



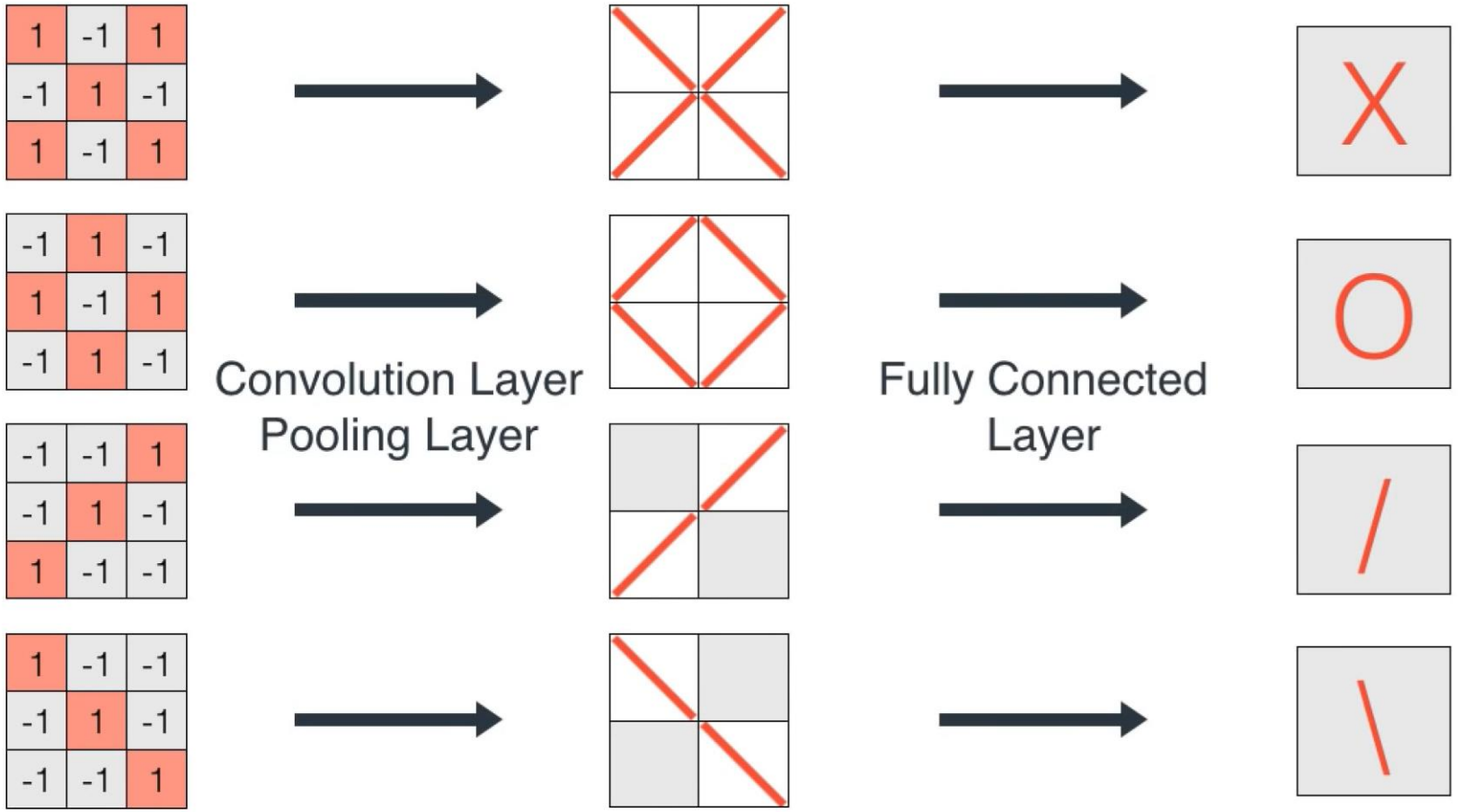
Convolution Neural Network (CNN)

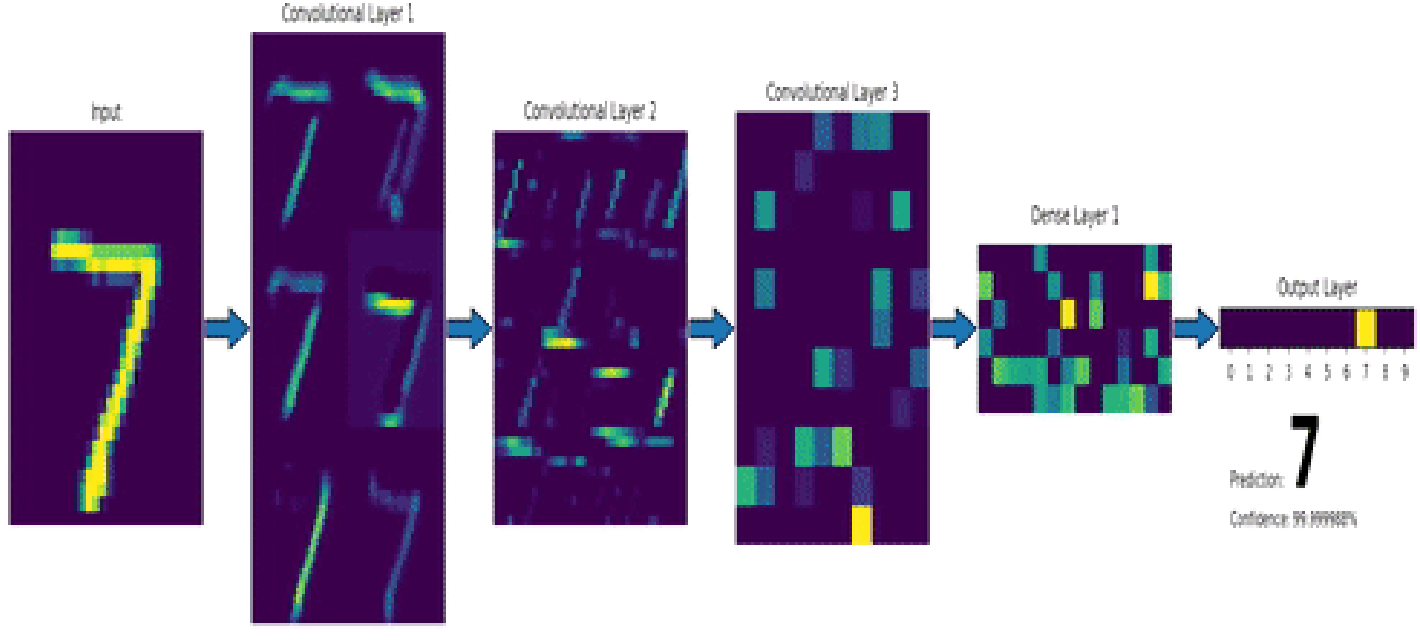
A little more complex world



Convolution Neural Network (CNN)

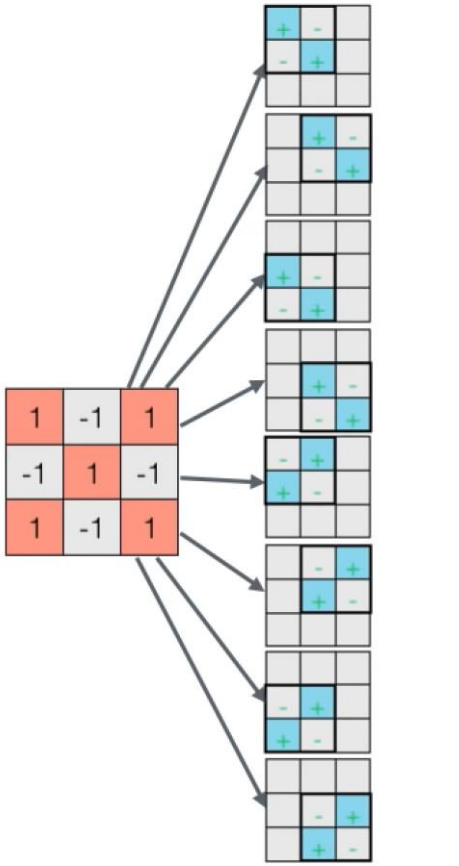
A little more complex world




Input
Convolutional Layer 1
Convolutional Layer 2
Convolutional Layer 3
Dense Layer 1
Output Layer
0 1 2 3 4 5 6 7 8 9
Prediction:
7
Confidence: 99.999988%

Convolution Neural Network (CNN)

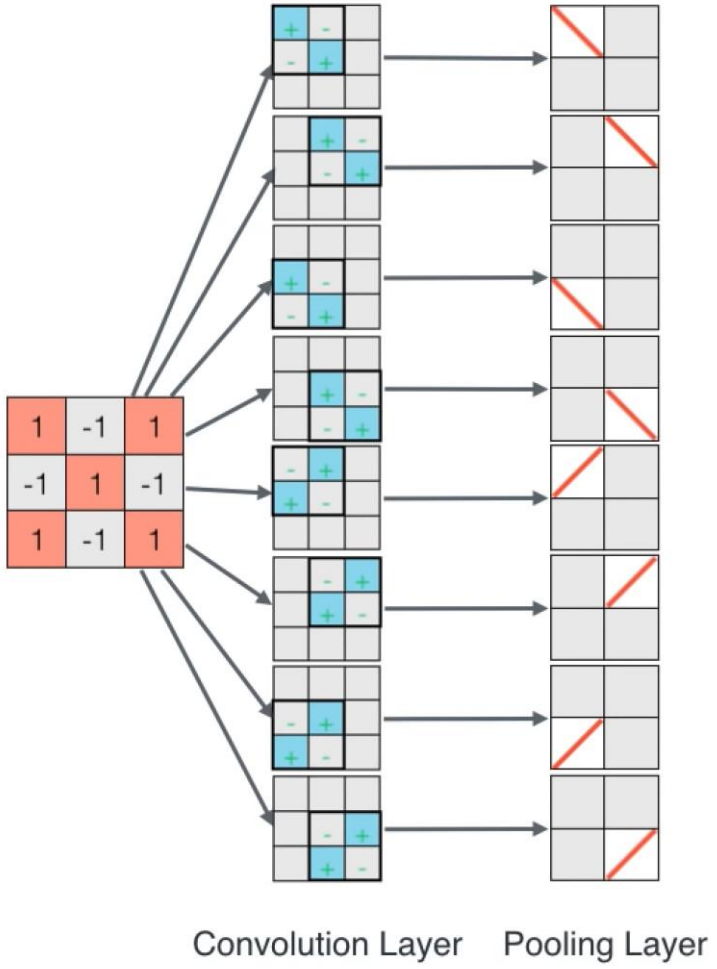
A little more complex world



Convolution Layer

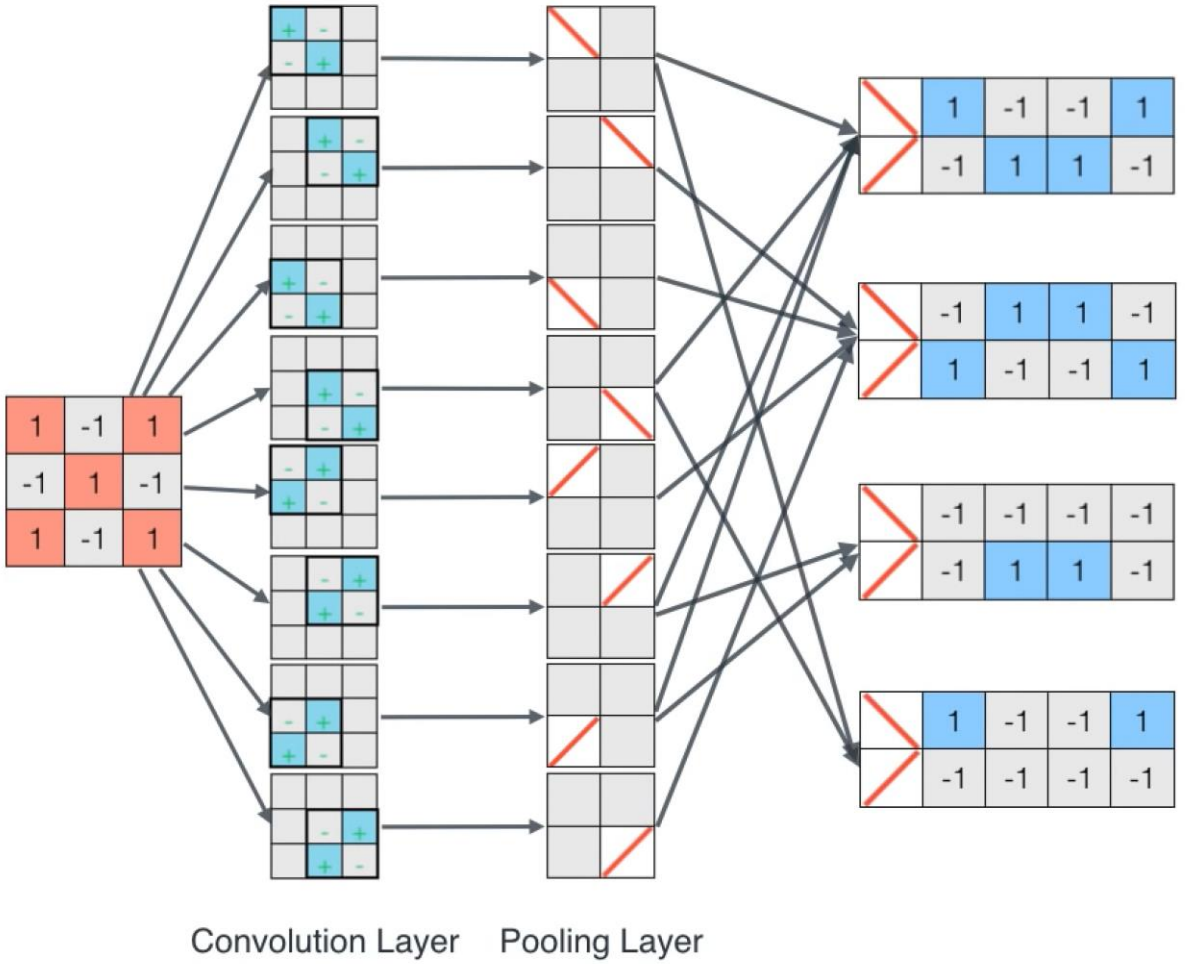
Convolution Neural Network (CNN)

A little more complex world



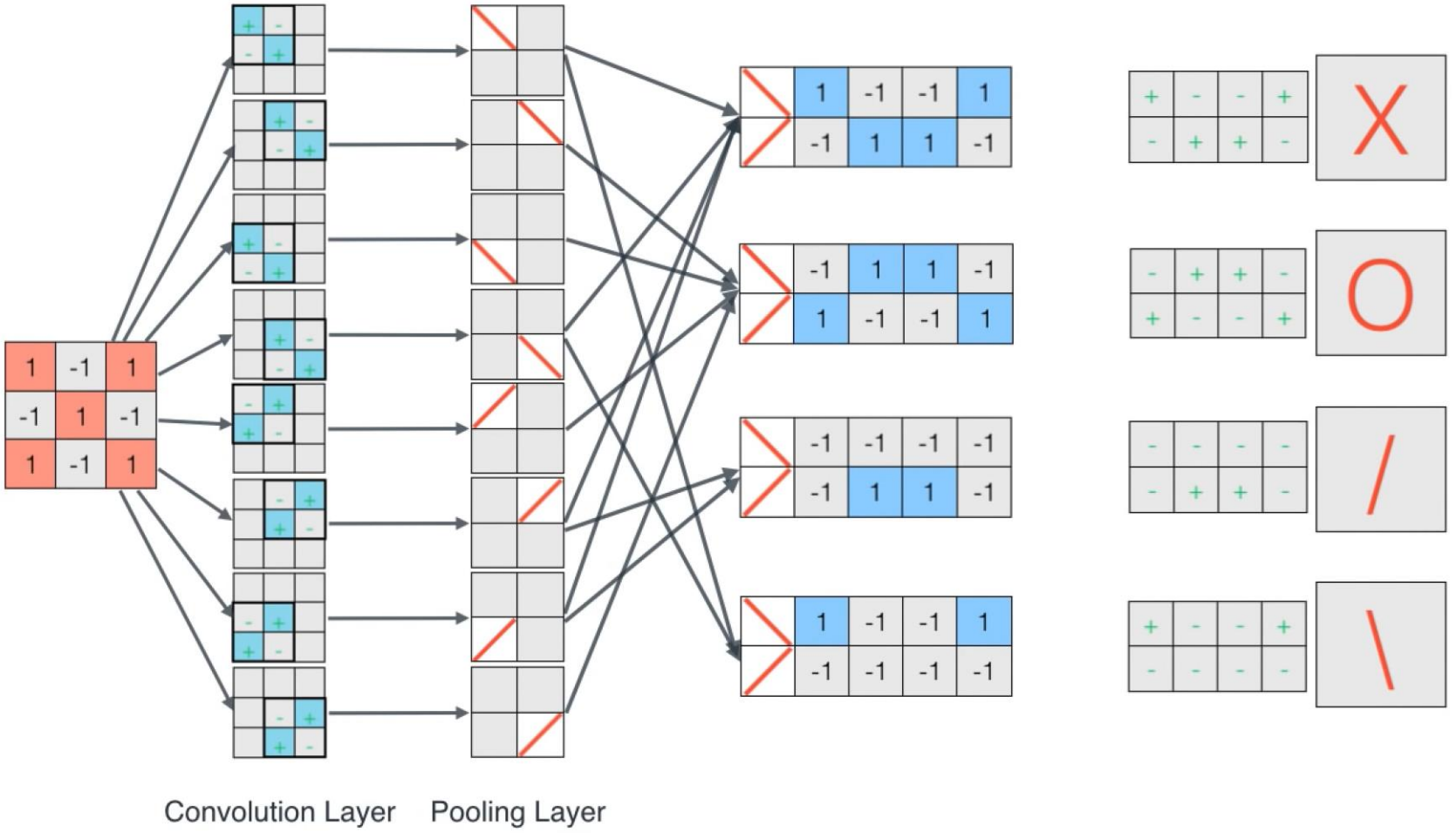
Convolution Neural Network (CNN)

A little more complex world



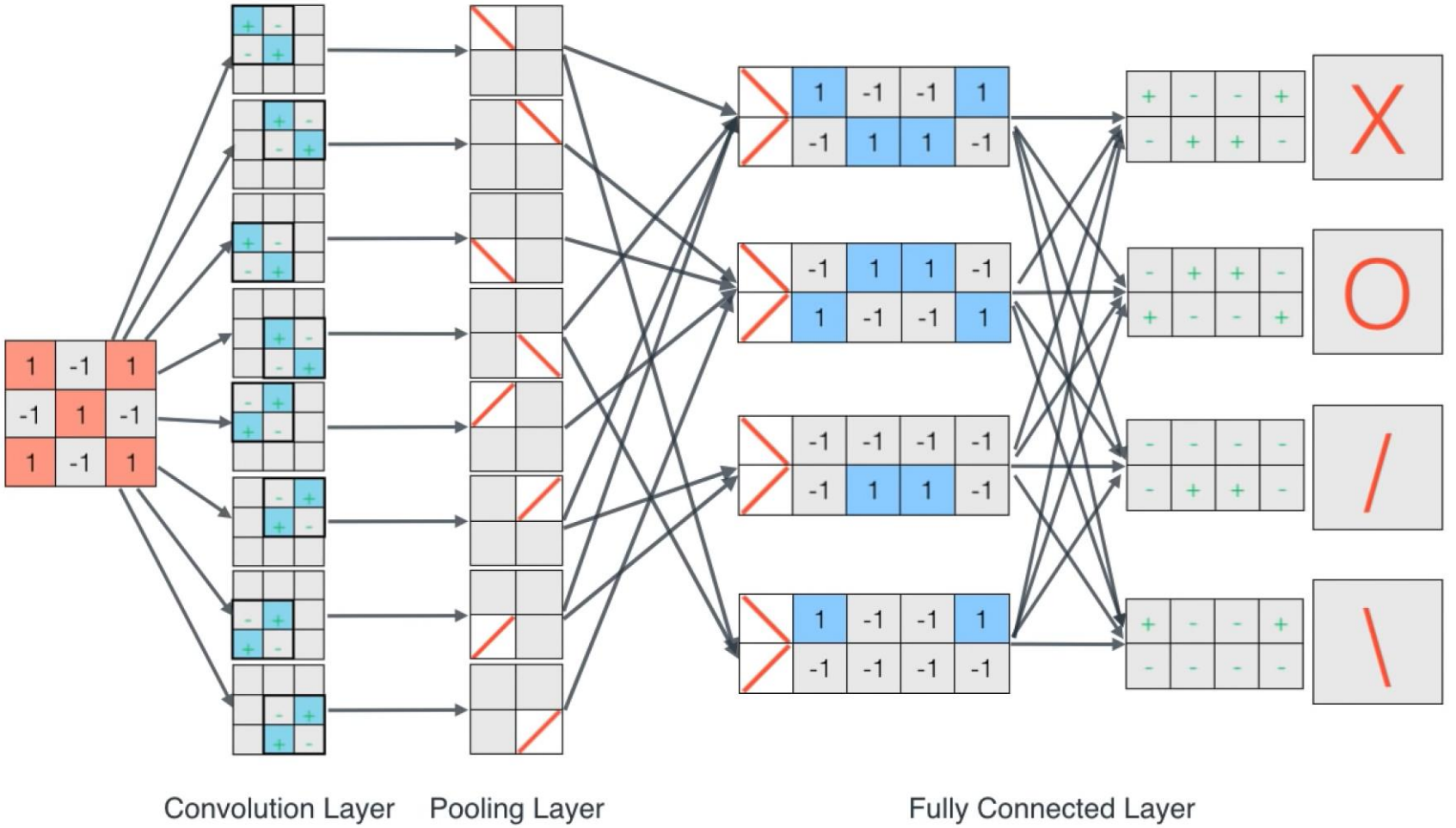
Convolution Neural Network (CNN)

A little more complex world



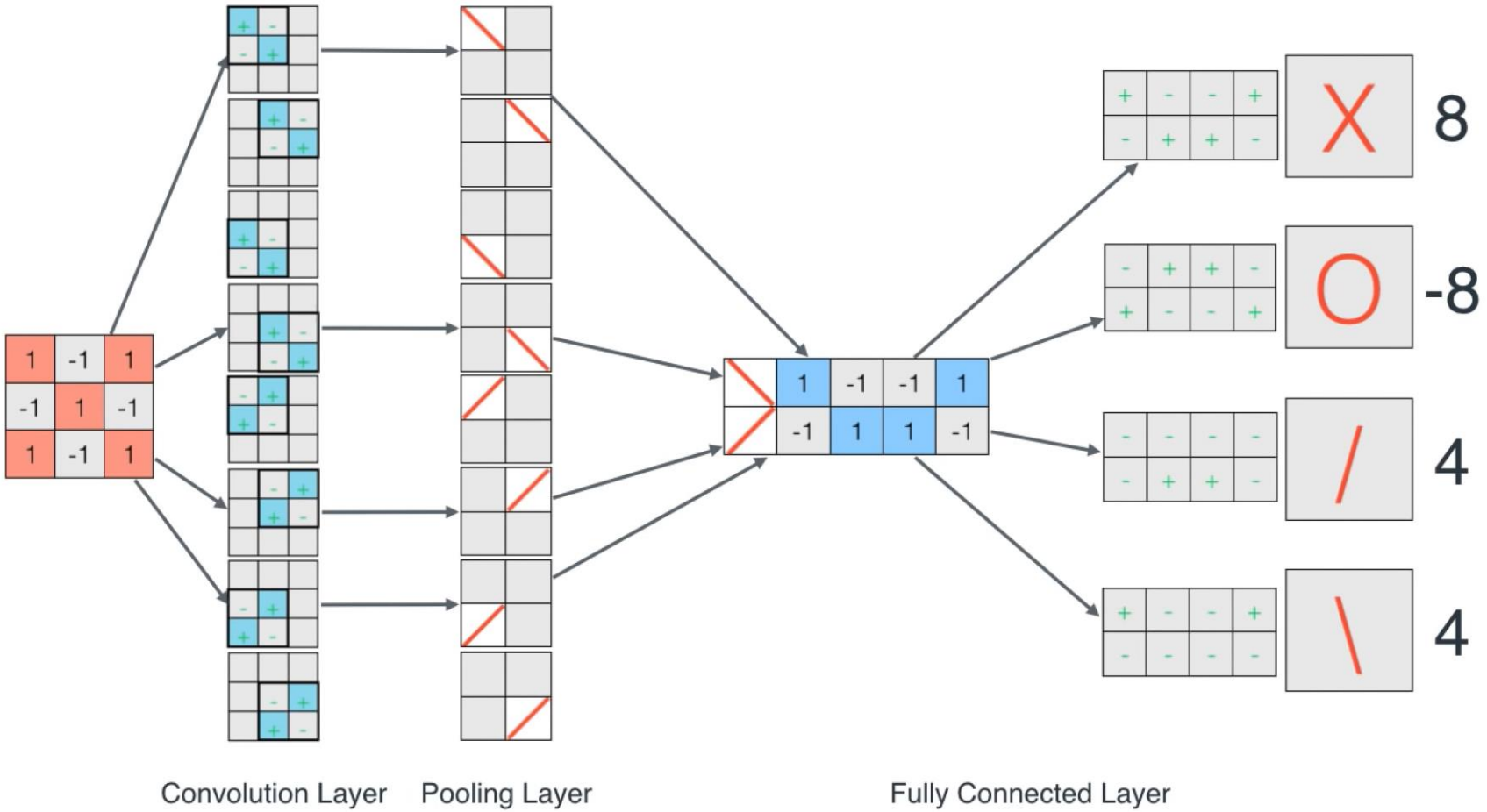
Convolution Neural Network (CNN)

A little more complex world



Convolution Neural Network (CNN)

A little more complex world

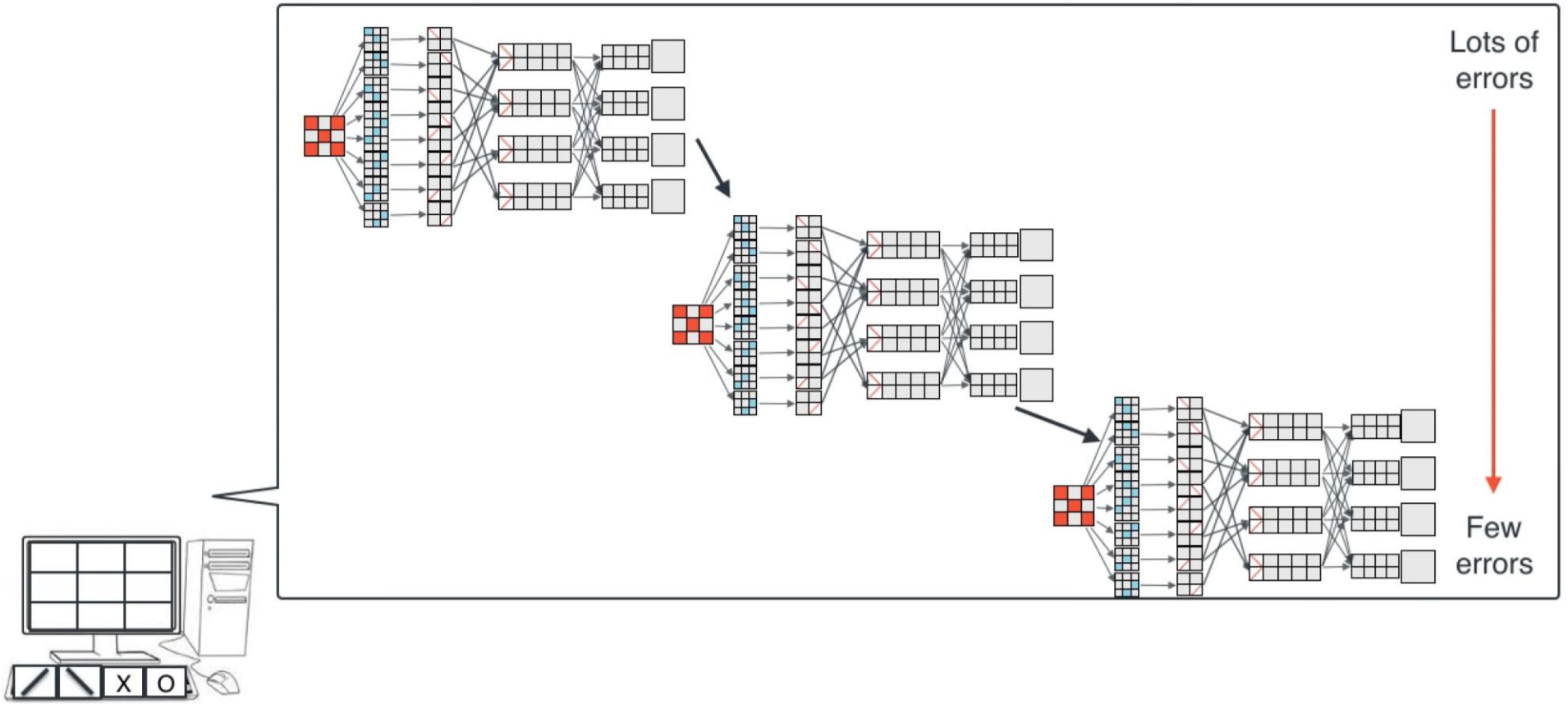


Convolution Neural Network (CNN)

A little more complex world

Convolution Neural Network (CNN)

A little more complex world



Convolution Neural Network (CNN)

Advanced world

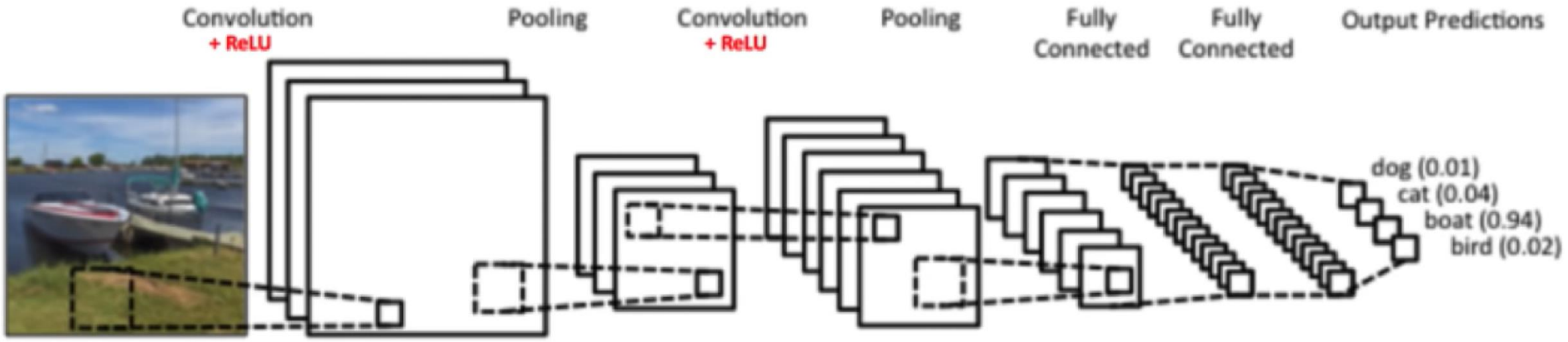
Análise facial

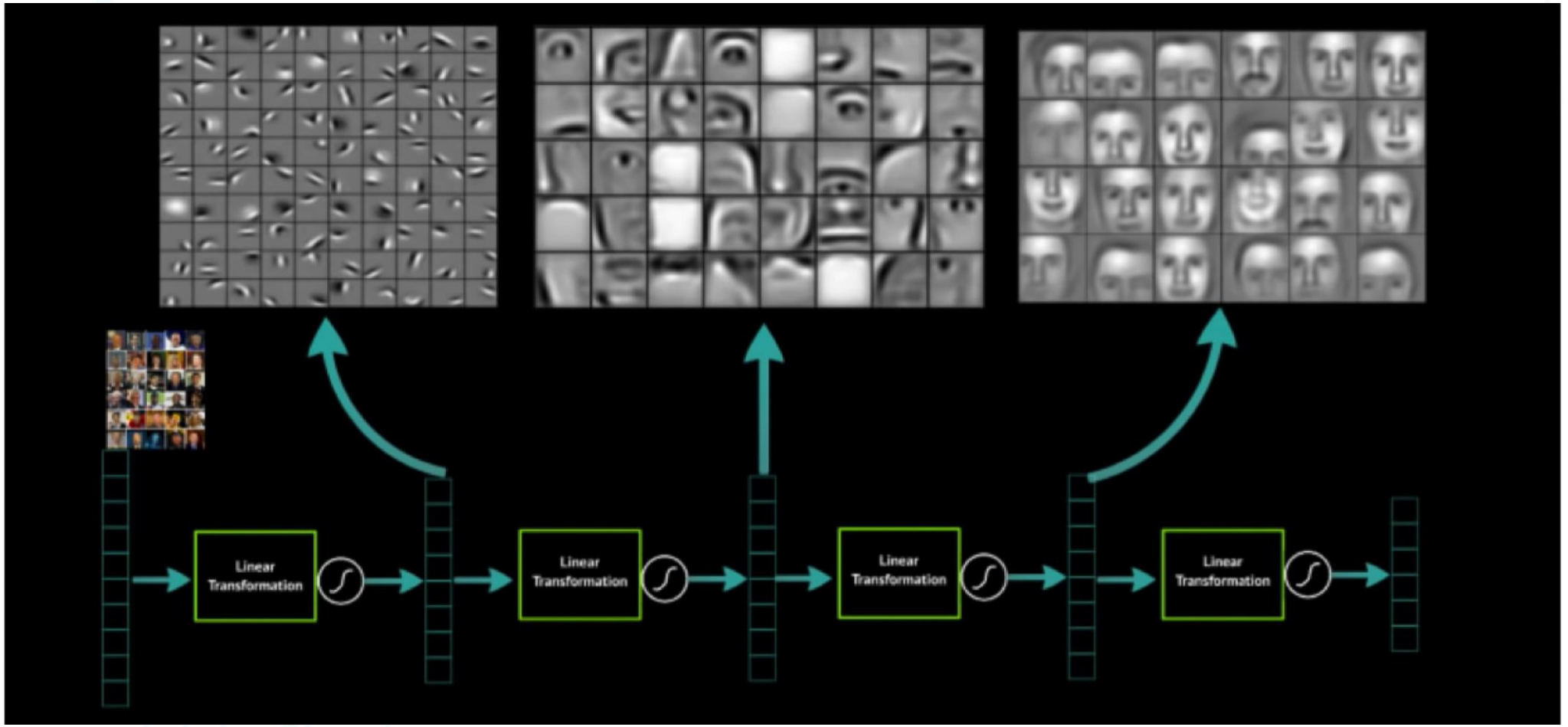
Determinação de caminhos

Detecção de objetos, cenas e atividades

Reconhecimento facial

Convolution Neural Network (CNN)



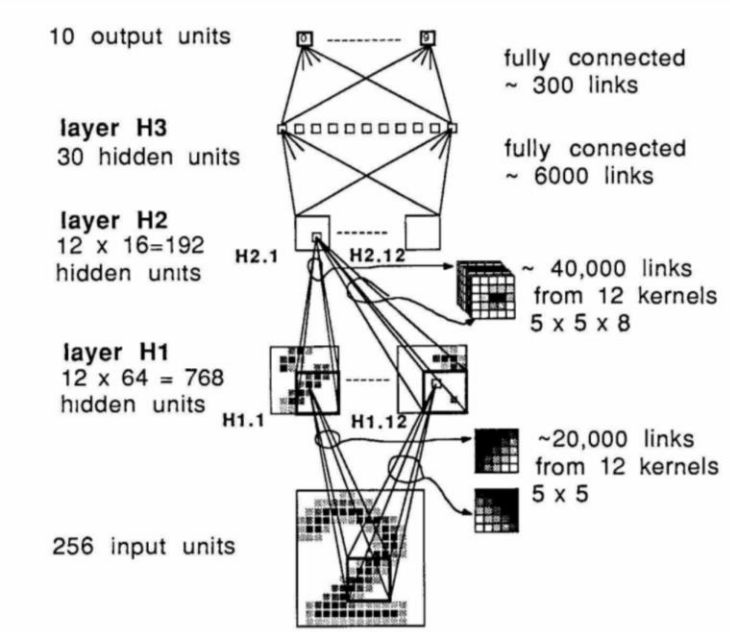

Linear Transformation
Linear Transformation
Linear Transformation
Linear Transformation

Convolution Neural Network (CNN)

Convolution Neural Network (CNN)

Year 1989 Kicked-Off Convolution Neural Nets

Ten-Digit Classifier using a Modest Neural Network with Three Hidden Layers



	Hidden Units	Connections	Params
Out – H3 (FC)	10 Visible	10 x (30W +1B)= 310	10 x (30W +1B)= 310
H3 – H2 (FC)	30	30 * (192 Weights + 1 Bias) = 5790	30 * (192 W + 1 B) = 5790
H2 – H1 (Conv)	12 X 4 x 4 = 192	192 x (5 x 5 x 8 + 1)= 38592	5 x 5 x 8 x 12 + 192 Biases = 2592
H1 – Input (Conv)	12 x 8 x 8 = 768	768 x (5 x 5 x 1 + 1) = 19968	5 x 5 x 1 x 12 + 768 Biases = 1068
Totals	16 x 16 In + 990 Hidden + 10 Out	64660 Connections	9760 Params

Backpropagation Applied to Handwritten Zip Code Recognition. LeCun, et. al.
http://yann.lecun.com/exdb/publis/pdf/lecun-89e.pdf

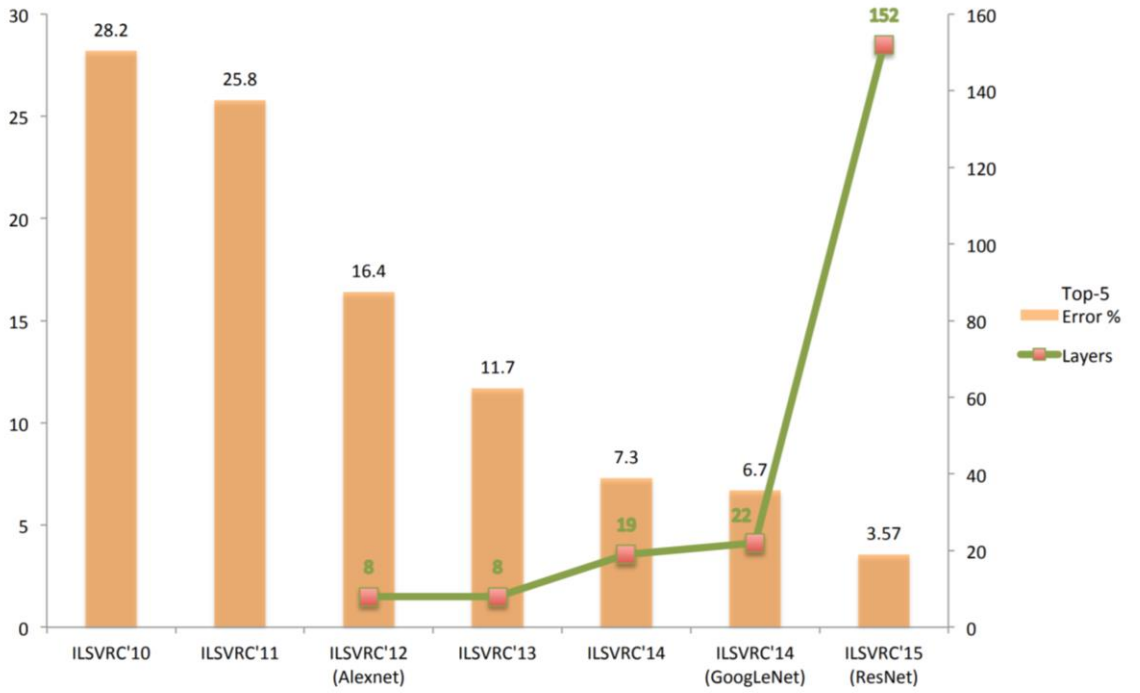
Each of the units in H2 combines local information coming from 8 of the 12 different feature maps in H1.

Convolution Neural Network (CNN)

Advanced world

Year 2012 Marked The Inflection Point

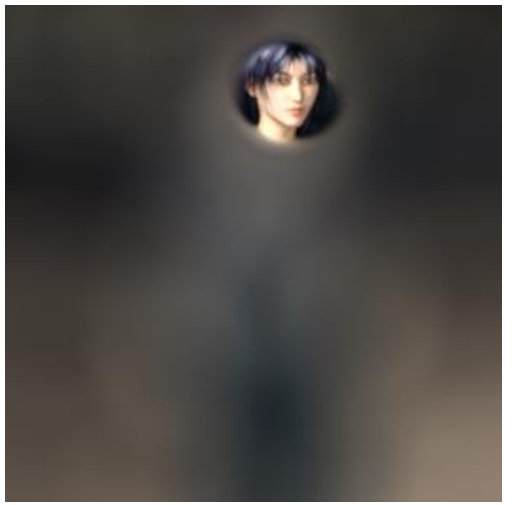
Reintroducing CNNs Led to Big Drop in Error for Image Classification.
Since Then, Deeper Networks Continued to Reduce Error

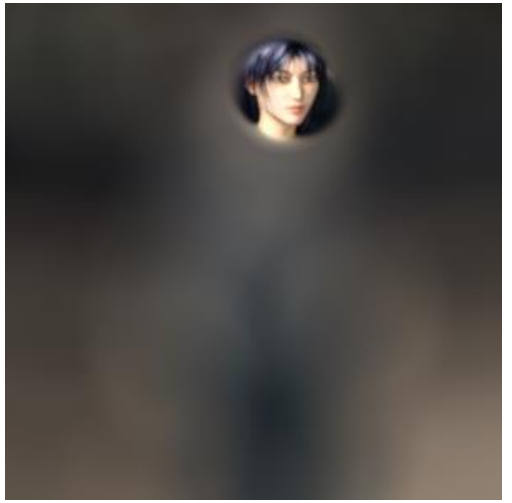


www.shatterline.com

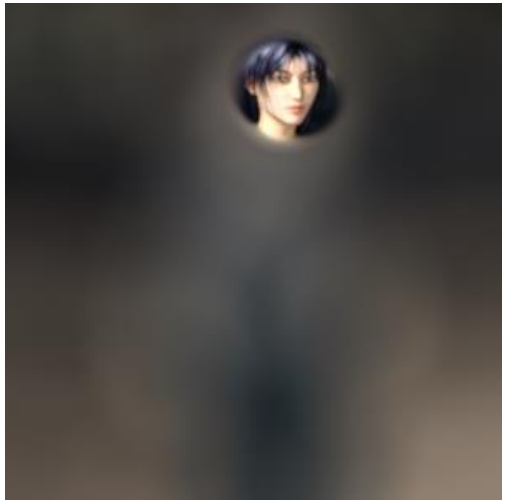
+

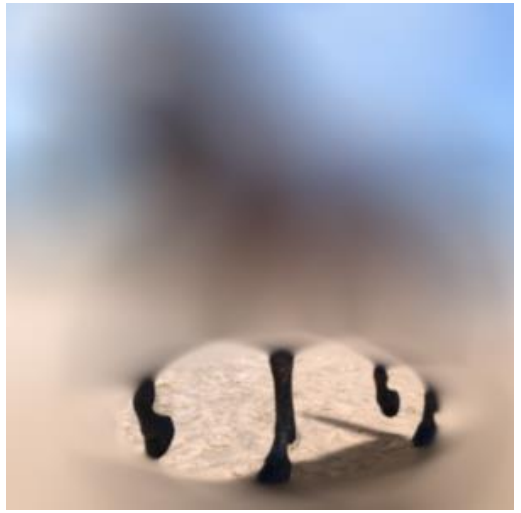
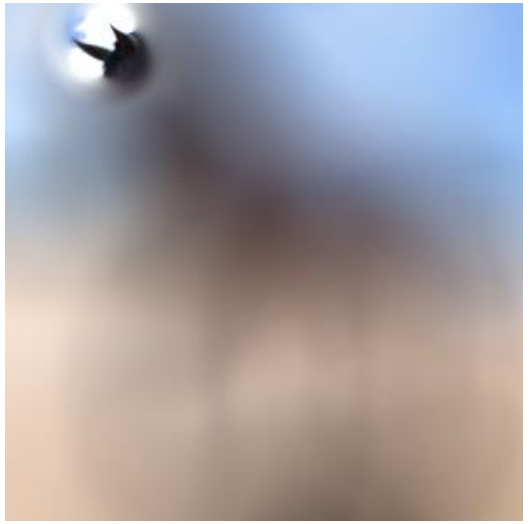
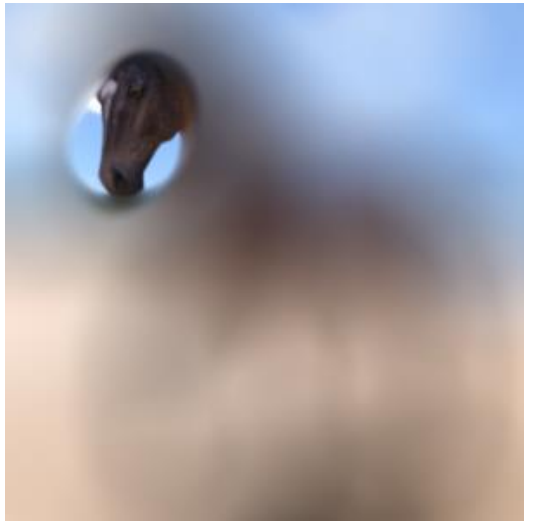

+

+


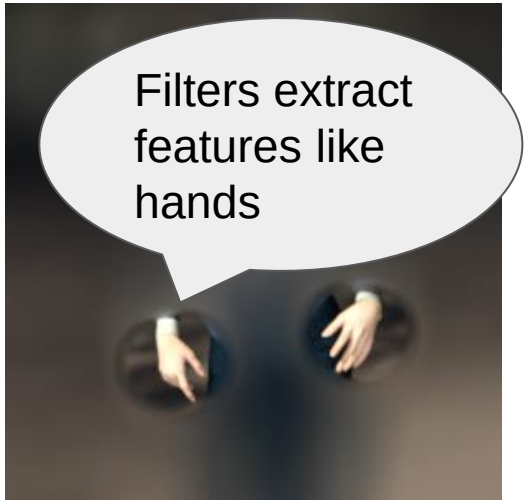
 +  +  = HUMAN


+

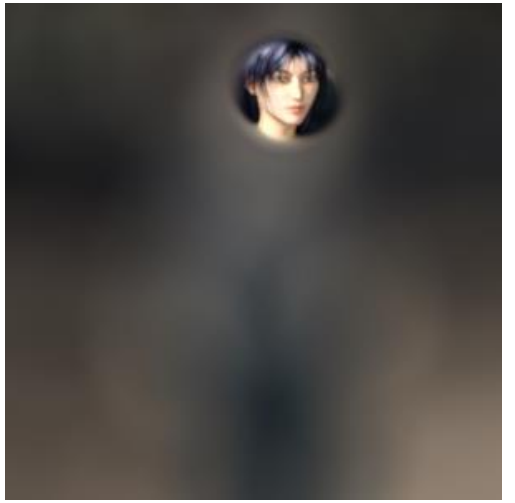
+

=
HUMAN


+

+

=
HORSE

+

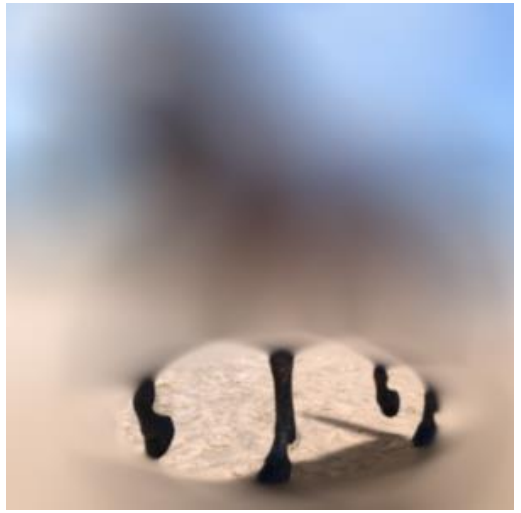


+

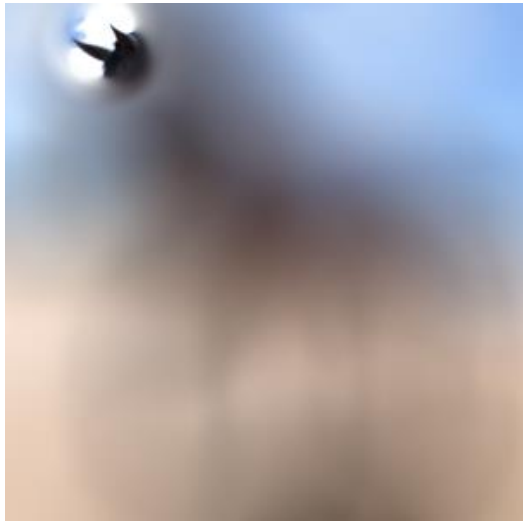


=

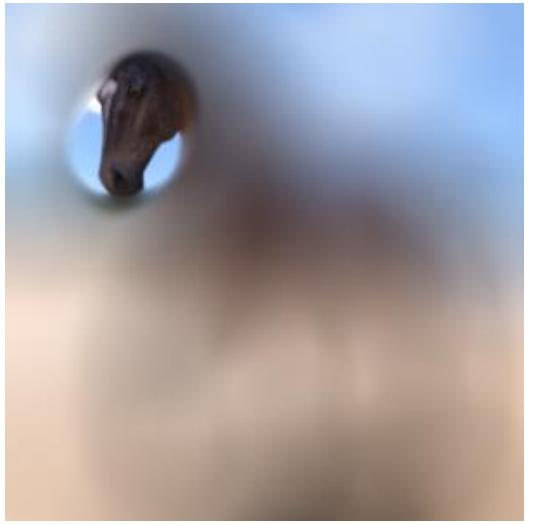
HUMAN



+



+

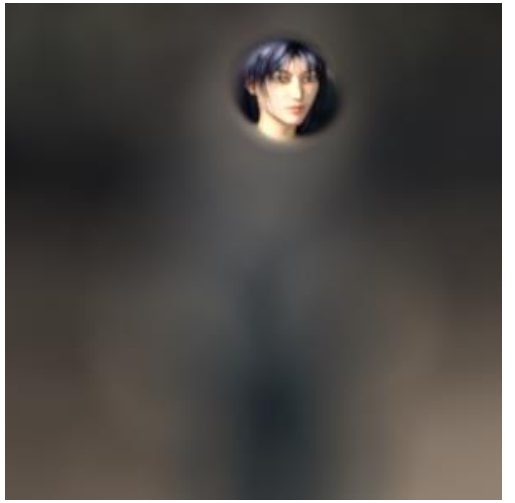


=

HORSE

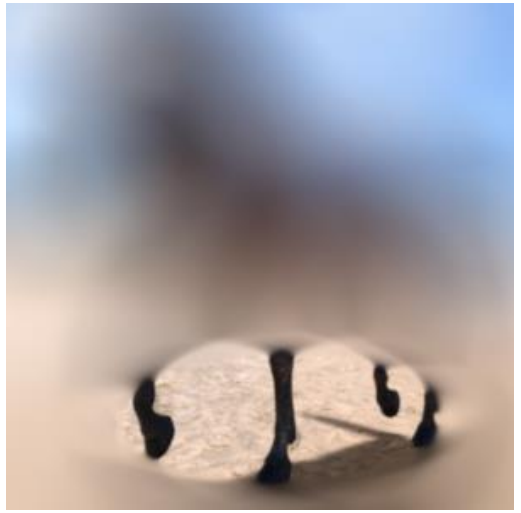
+

+



=

HUMAN

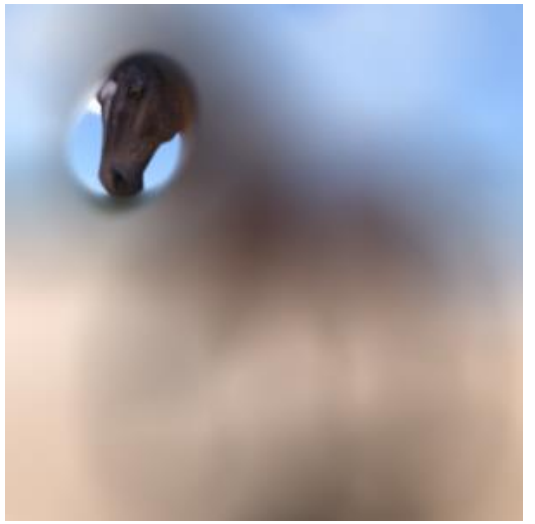


+

=

HORSE

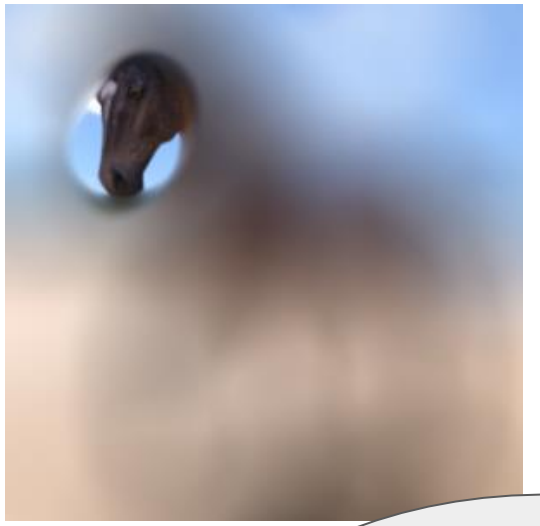

+

+

=
HORSE

Filters can then be combined with labels to make a prediction of the image contents...

+

+

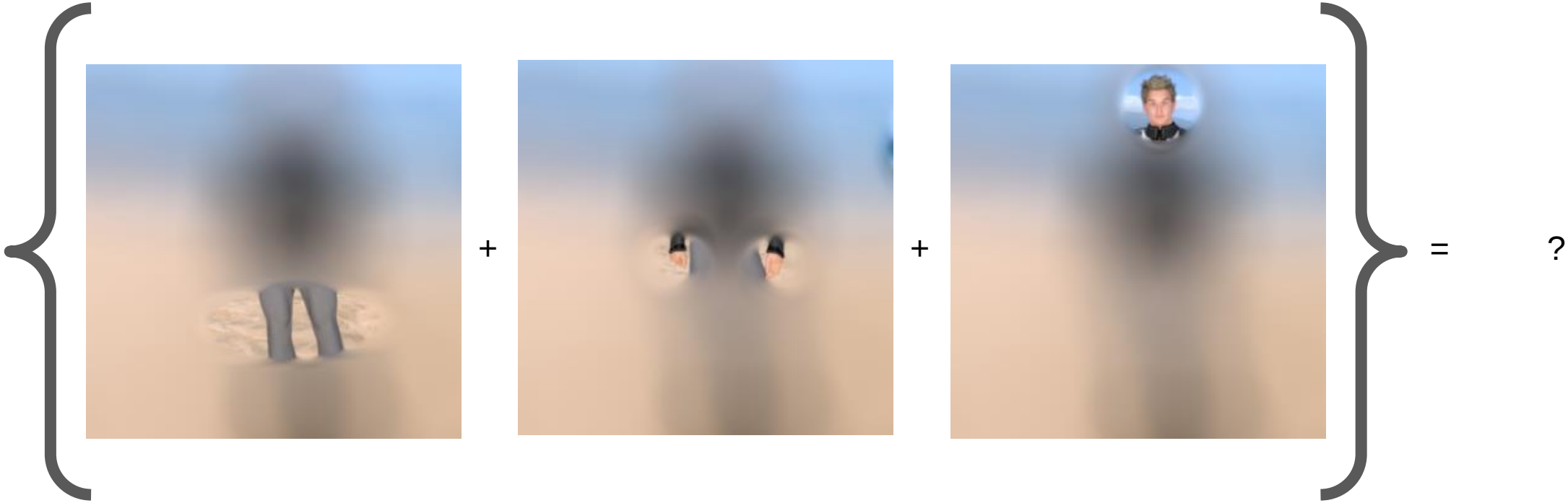


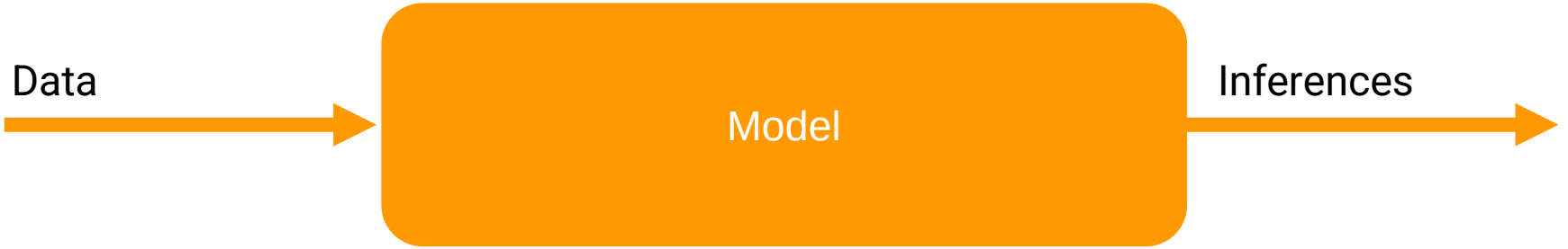
=

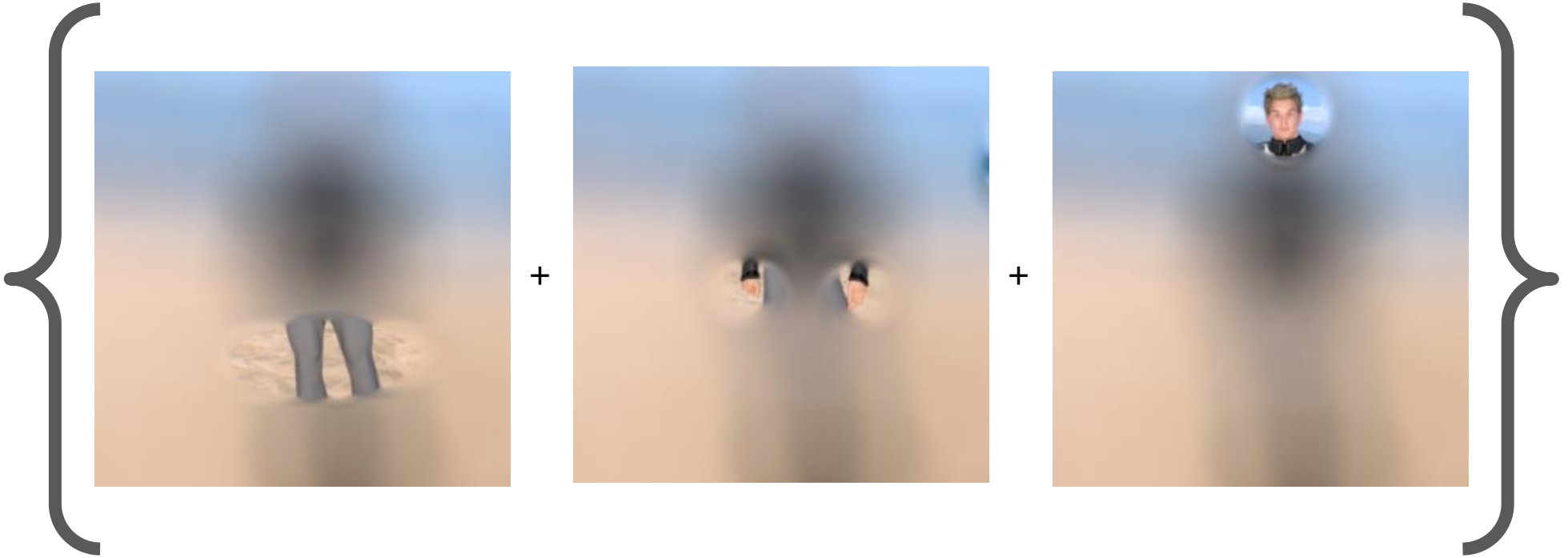
HORSE

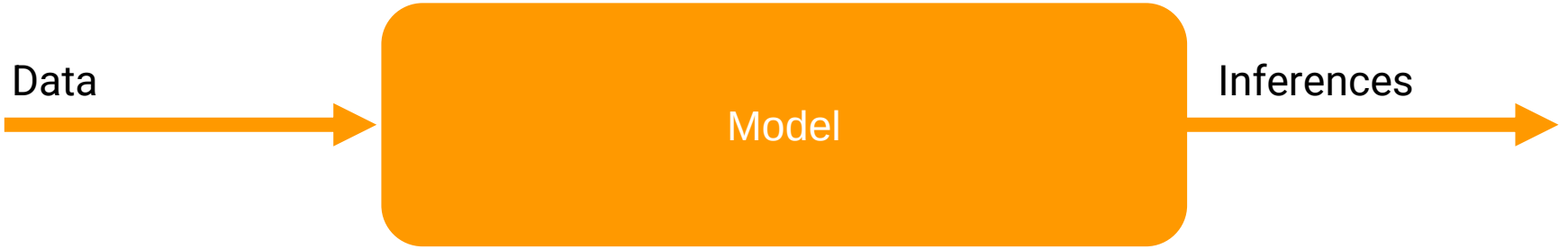
Filters can then be combined with labels to make a prediction of the image contents...

The filters that match the label are learned over time!


+
+
=
?

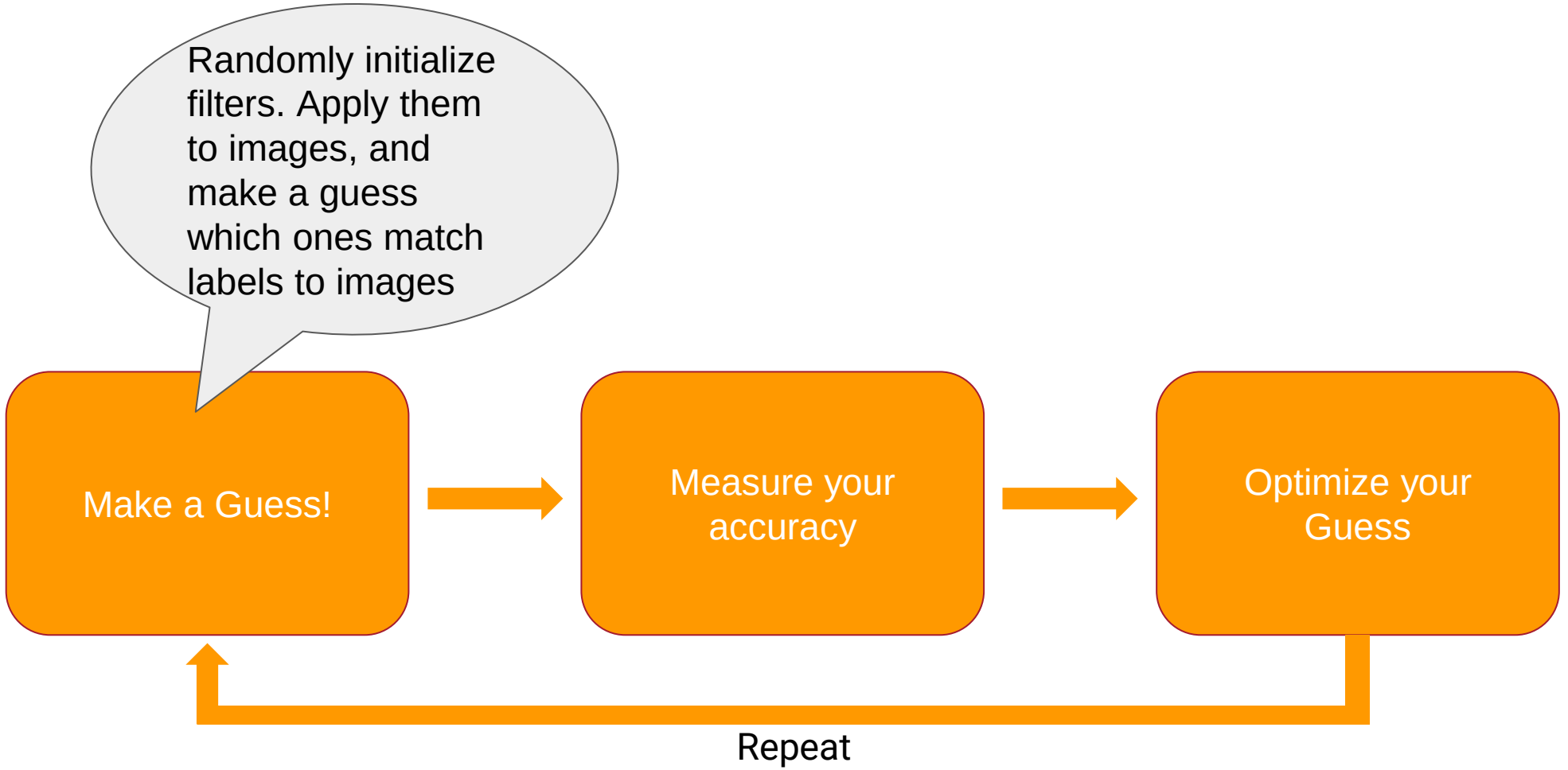

Data
Model
Inferences


+
+
=
HUMAN

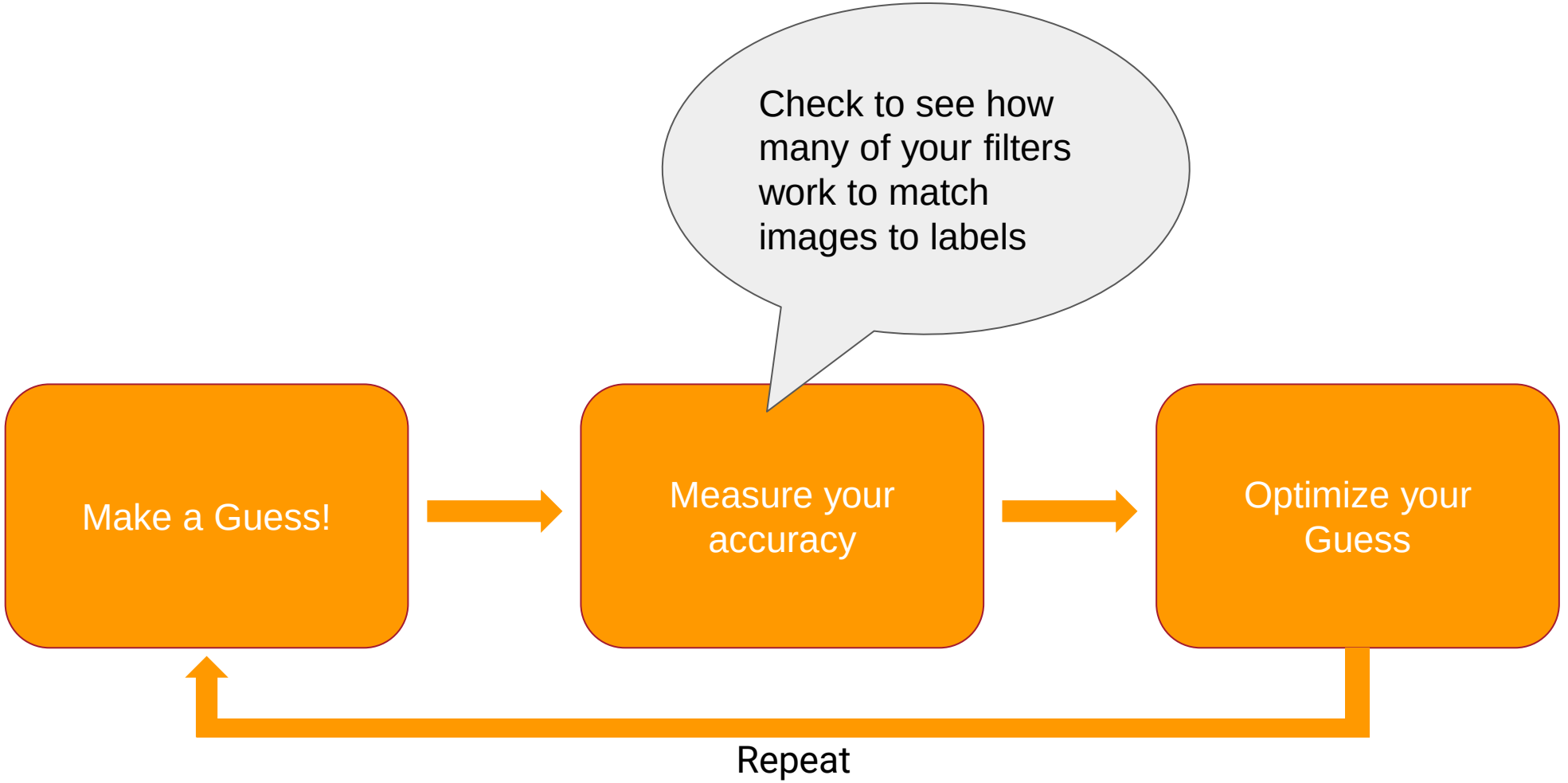

Data
Model
Inferences

The Machine Learning Paradigm

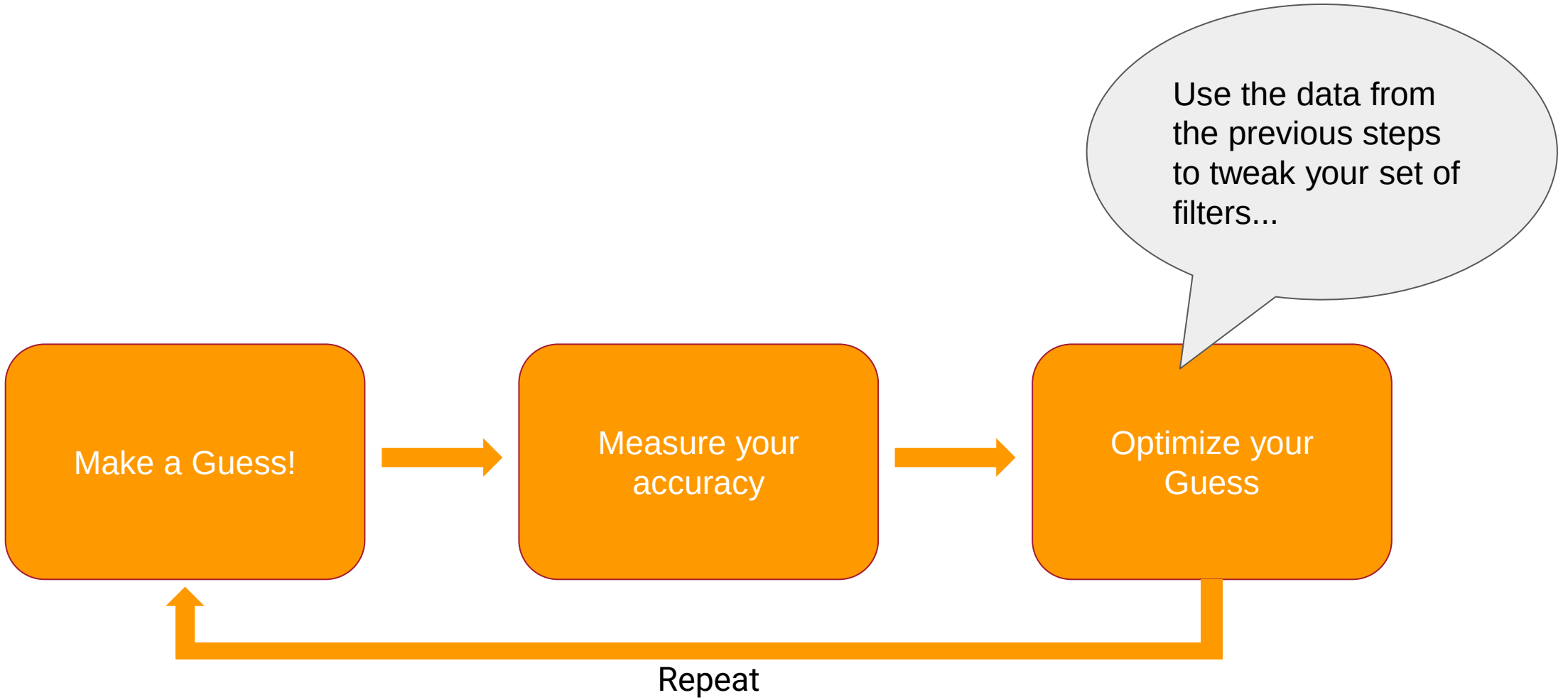
The Machine Learning Paradigm



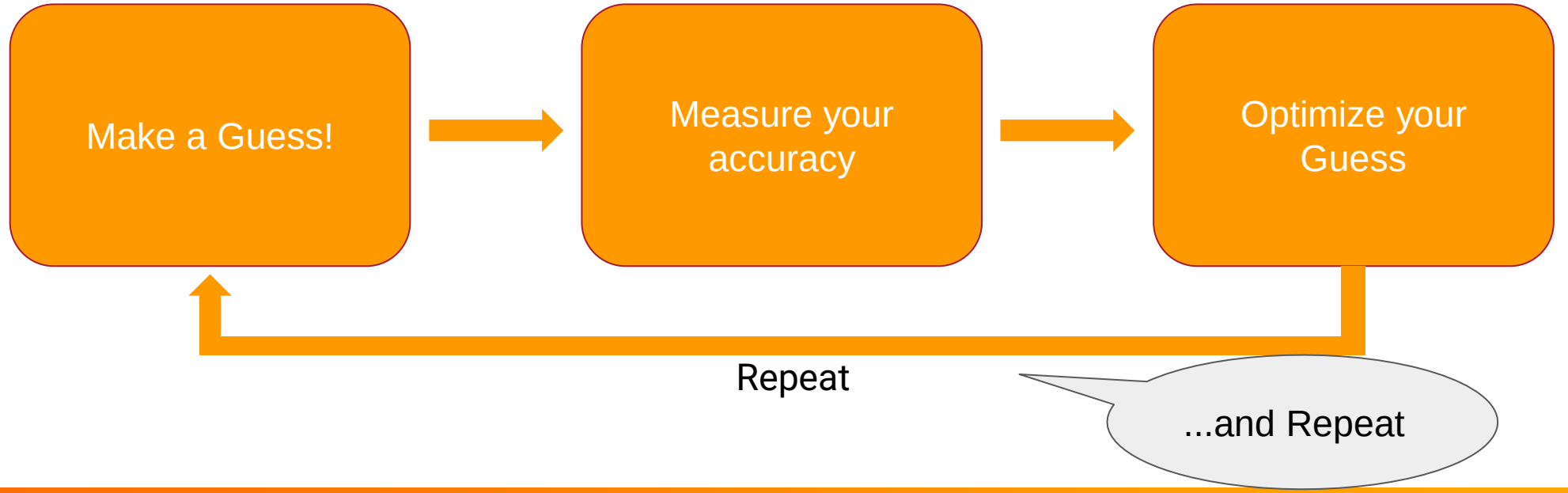
The Machine Learning Paradigm

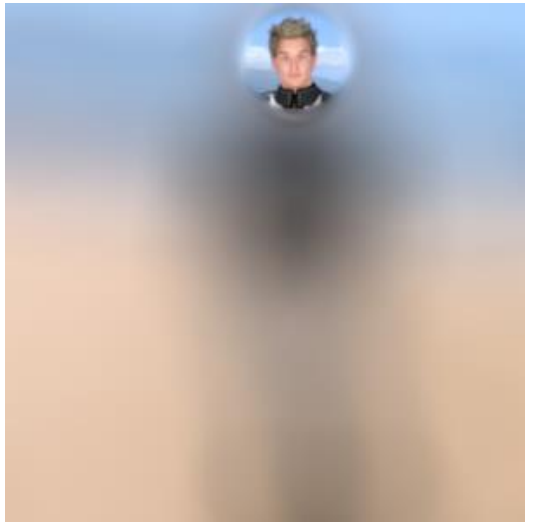


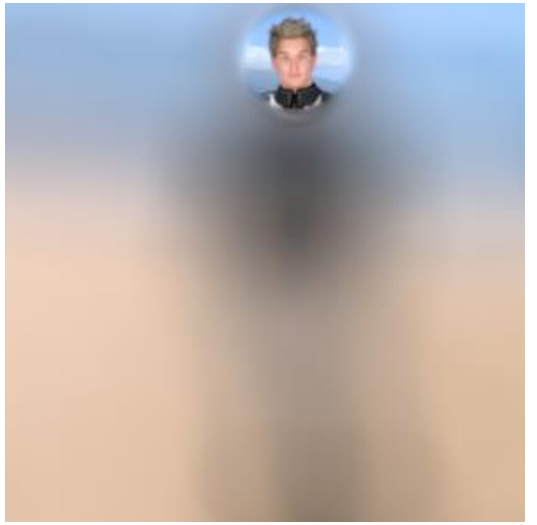
The Machine Learning Paradigm

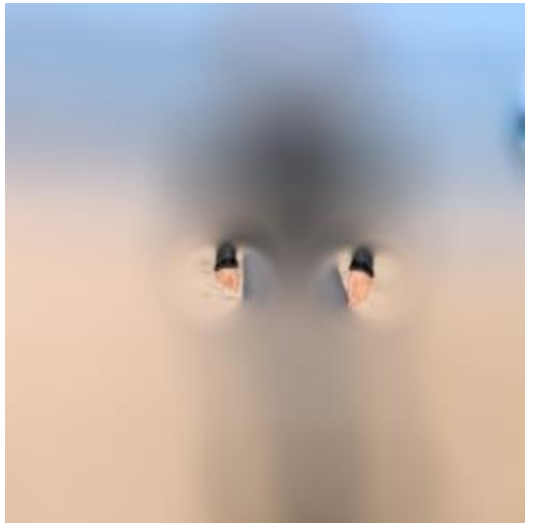
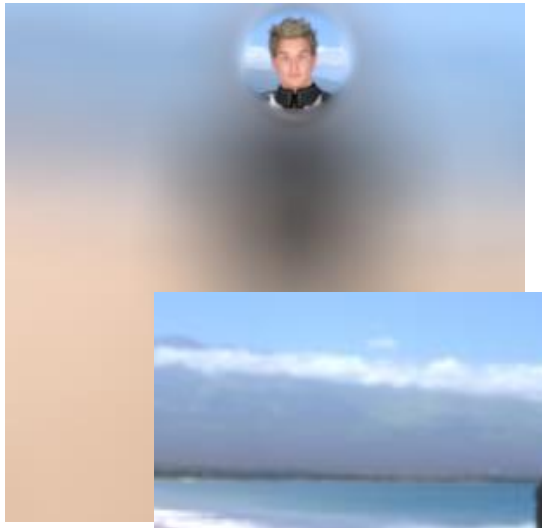


The Machine Learning Paradigm



 +  +  = ?

 +  +  = HUMAN


+

+

=
HUMAN


```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(16, (3, 3), activation='relu',
                           input_shape=(300, 300, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(512, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])
```

```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(16, (3, 3), activation='relu',
                           input_shape=(300, 200, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(512, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])
```

Conv2D stands for 2D Convolution -- another word for a filter

```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(16, (3, 3), activation='relu',
                           input_shape=(300, 300, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(512, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])
```

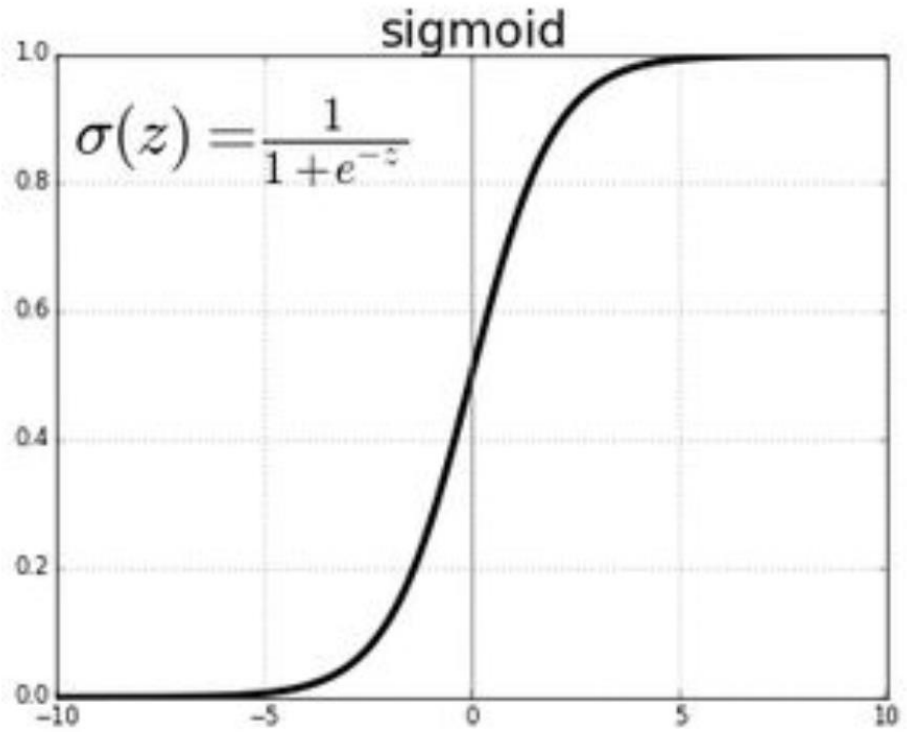
MaxPooling is a way of compressing the image while enhancing features

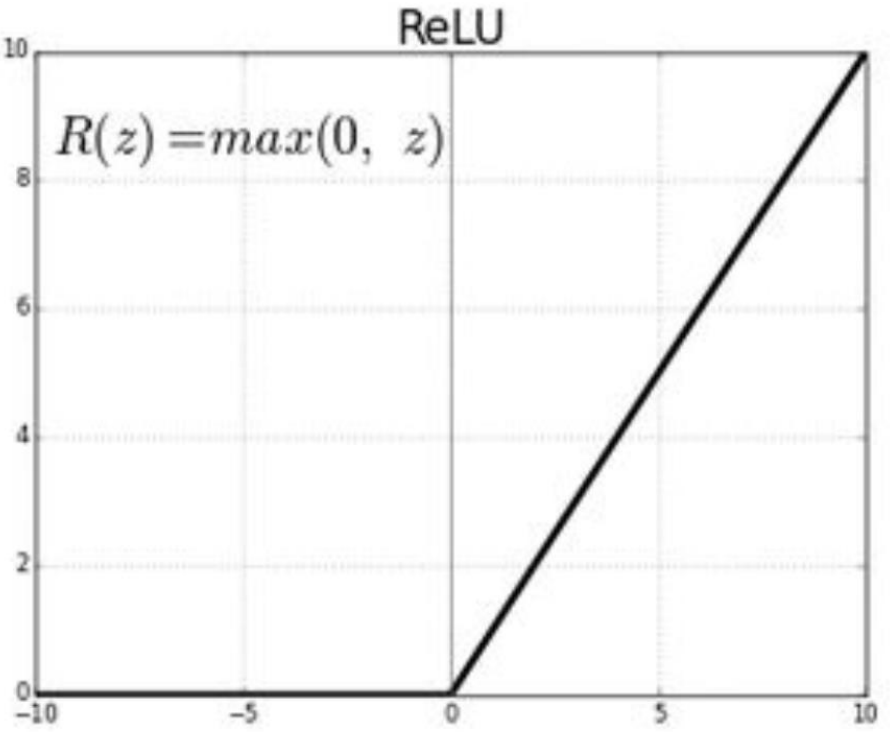
```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(16, (3, 3), activation='relu',
                           input_shape=(300, 300, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(512, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])
```

Dense is a neural network that matches the filters to the labels

```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(16, (3, 3), activation='relu',
                           input_shape=(300, 300, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(512, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])
```

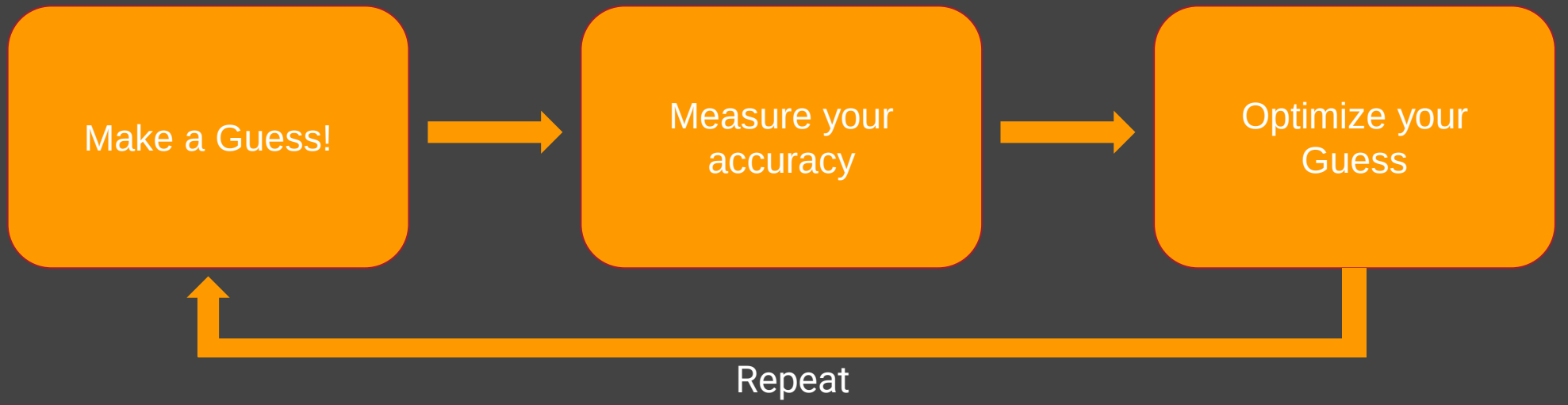
The final Dense represents the labels - Horse = 0, Human = 1


sigmoid
$\sigma(z) = \frac{1}{1+e^{-z}}$
1.0
0.8
0.6
0.4
0.2
0.0
−10
−5
0
5
10

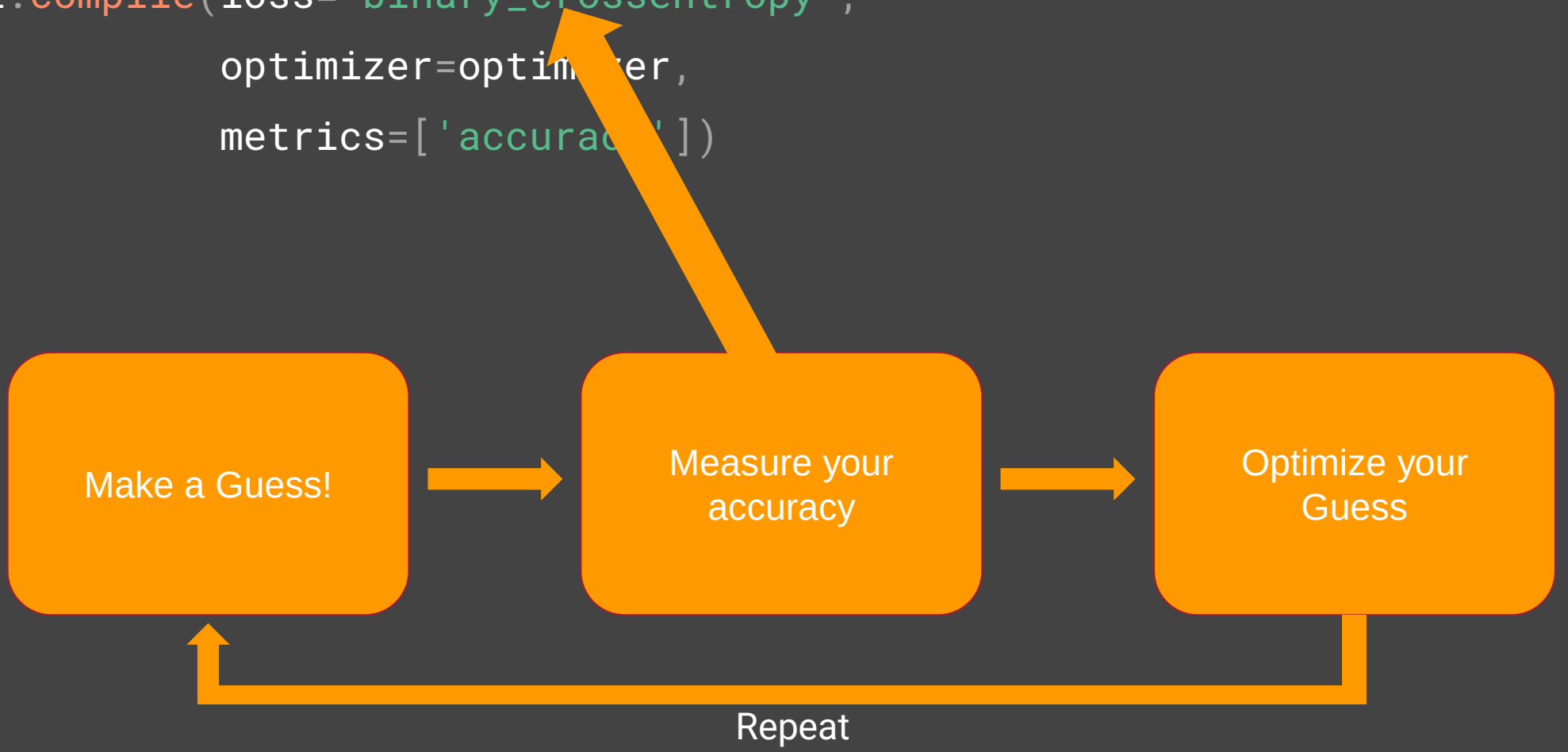

ReLU
$R(z) = max(0,\ z)$
10
8
6
4
2
0
−10
−5
0
5
10

```
optimizer = tf.keras.optimizers.RMSprop(lr=0.001)
model.compile(loss='binary_crossentropy',
              optimizer=optimizer,
              metrics=['accuracy'])
```

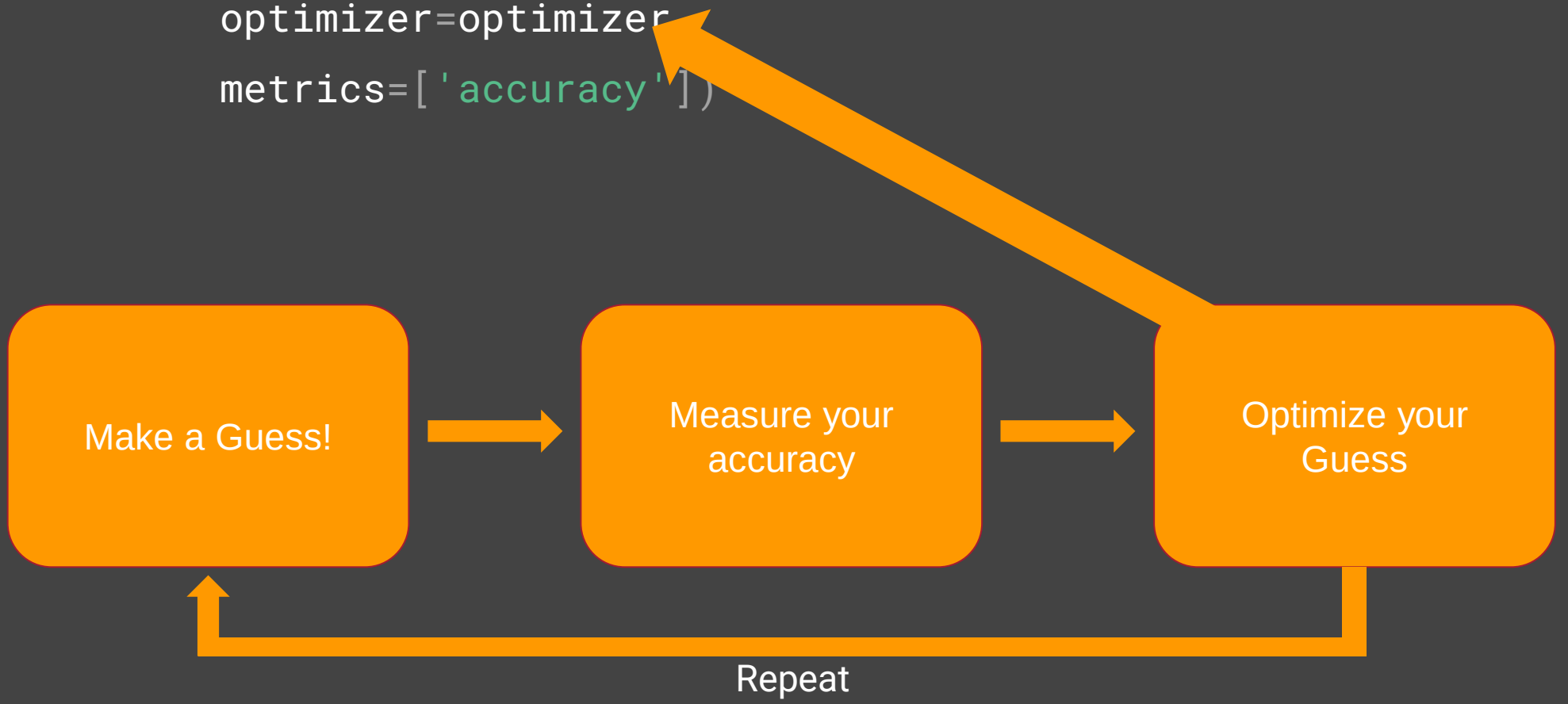
```
optimizer = tf.keras.optimizers.RMSprop(lr=0.001)
model.compile(loss='binary_crossentropy',
              optimizer=optimizer,
              metrics=['accuracy'])
```



```
optimizer = tf.keras.optimizers.RMSprop(lr=0.001)
model.compile(loss='binary_crossentropy',
              optimizer=optimizer,
              metrics=['accuracy'])
```

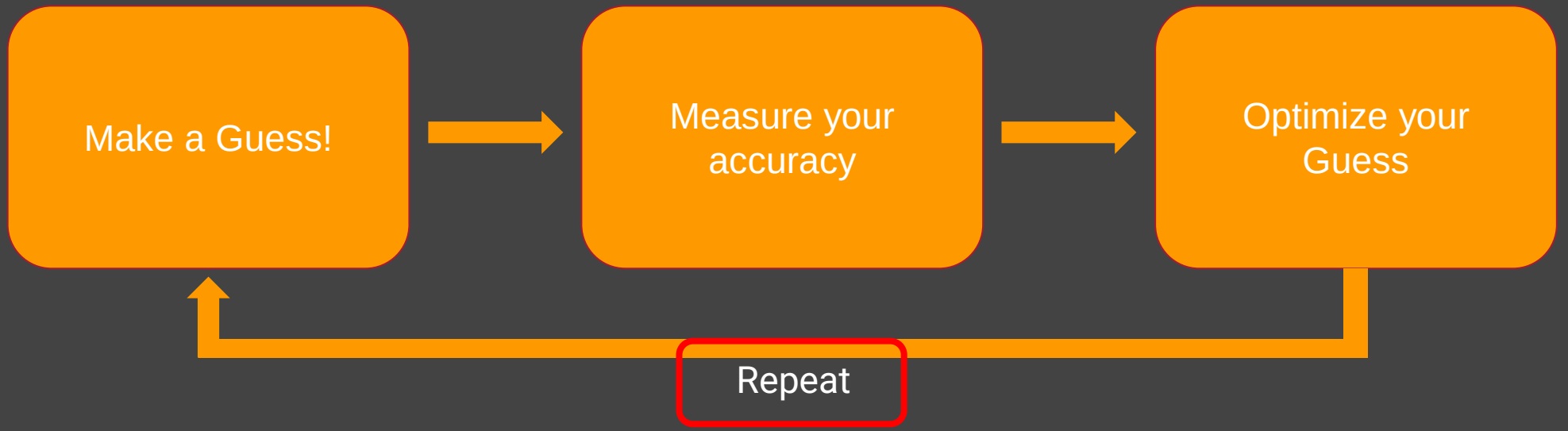


```
optimizer = tf.keras.optimizers.RMSprop(lr=0.001)
model.compile(loss='binary_crossentropy',
              optimizer=optimizer,
              metrics=['accuracy'])
```



```
model.fit(train_generator, epochs=12, ...)
```

```
model.fit(train_generator, epochs=12, ...)
```



VINICARIDÁ
Human-Centered Data & AI

Thank you!

@vinicius caridá

@vfcarida

vfcarida@gmail.com

https://linktr.ee/vfcarida