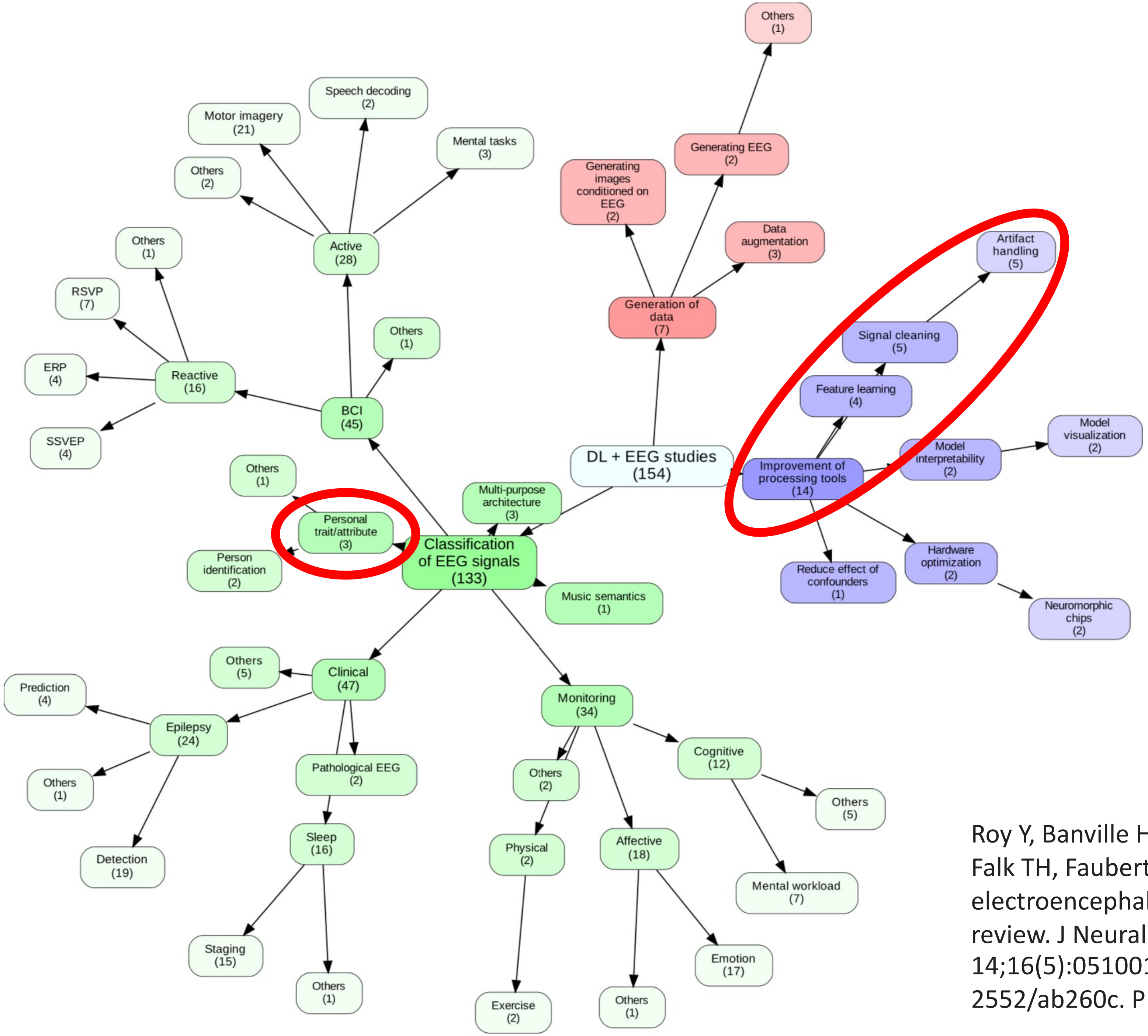
Deep Learning applied to EEG

Arnaud Delorme & Dung Truong



Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J. Deep learning-based electroencephalography analysis: a systematic review. J Neural Eng. 2019 Aug 14;16(5):051001. doi: 10.1088/1741-2552/ab260c. PMID: 31151119.

Data Data Data Data Data Data Data Data

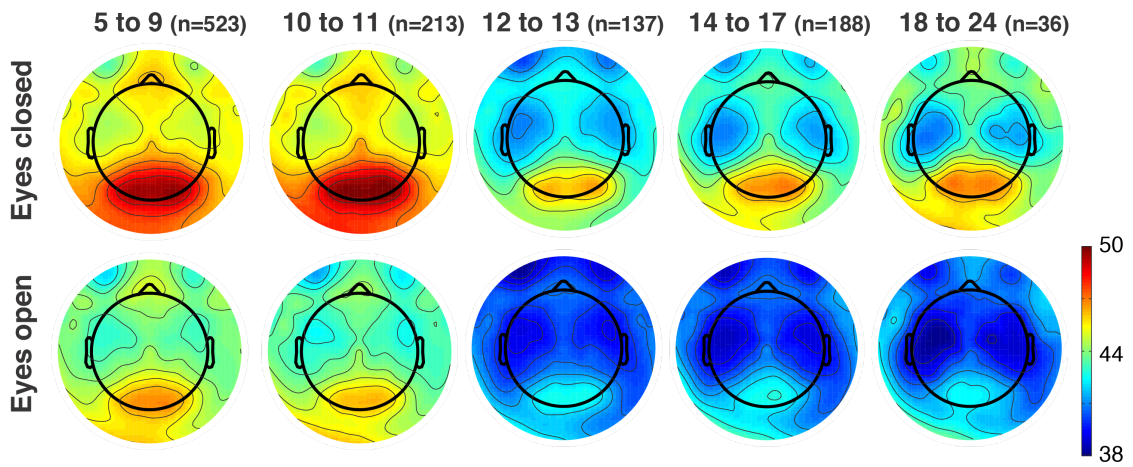
CHILD MIND® INSTITUTE

Michael Milham, PI

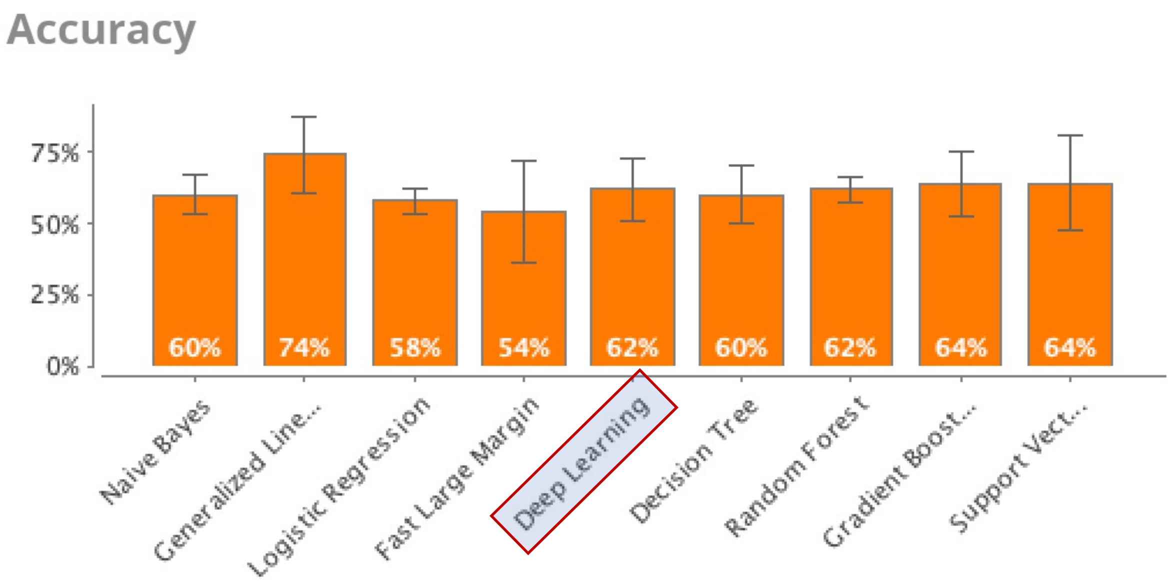
- 128 channels EEG
- ~3,000 EEG datasets (planed 10,000)
- Tasks involving emotions (The Gift movie)
- Rest (eyes open and eyes closed)

Eyes open vs closed

Different age groups eyes' open and eyes' closed from Child Mind Database (n=1,097)



Rapid Miner (3 hidden layers 100, 200, 100)



The Open EEGLAB portal. **Arnaud Delorme**, Amitava Majumdar, Subhashini Sivagnanam, Ramon Martinez-Cancino, Kenneth Yoshimoto, Scott Makeig. *9th International IEEE EMBS Conference on Neural Engineering,* March 20-23, 2019.

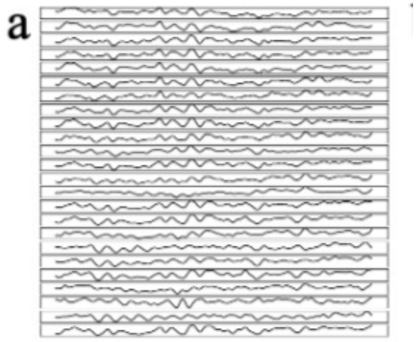
Machine learning/Deep learning

- 1574 participants (50% female) 128-channels
- 2-second extracted epochs eyes open and eyes closed
- 5 predictors: sex, handedness, eyes open/closed, age, segment count

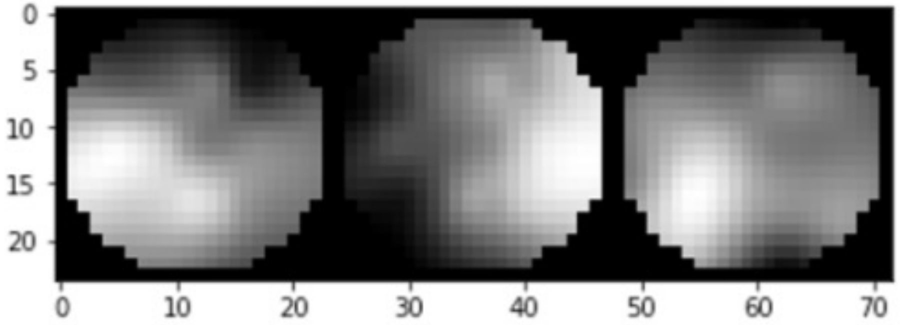
Categorical

Continuous

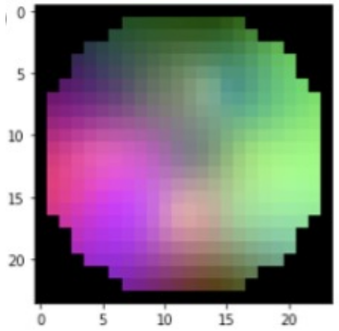
2-sec



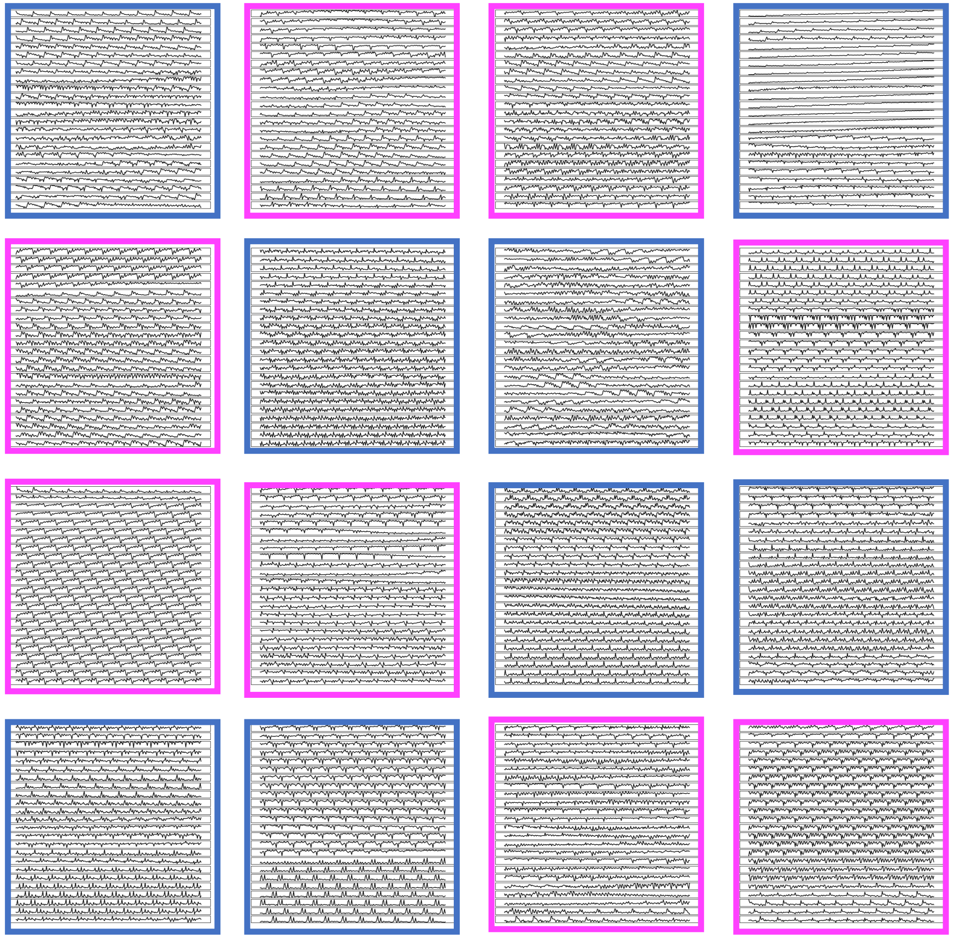
Achromatic image



Chromatic image

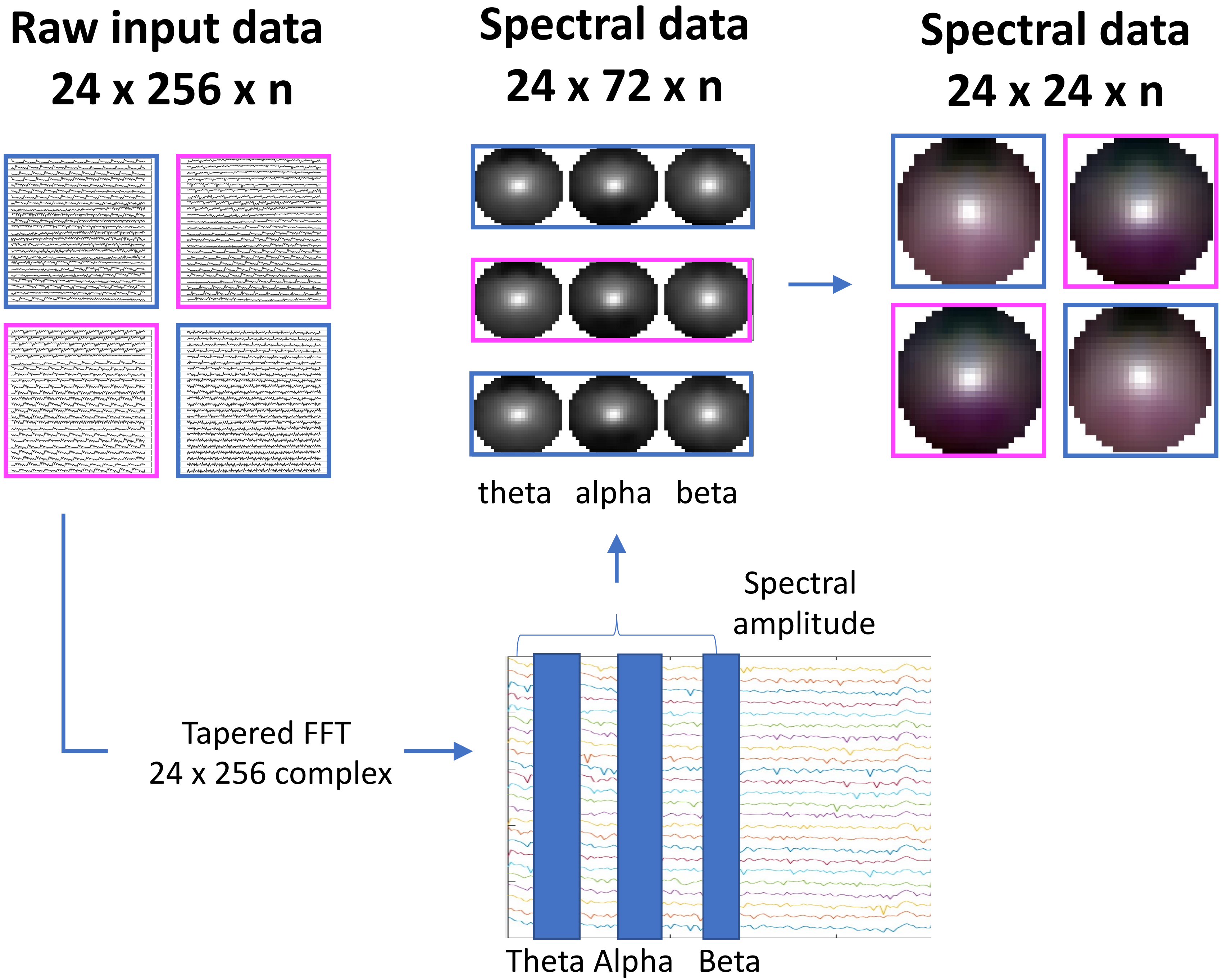


Input data 24 x 256 x n

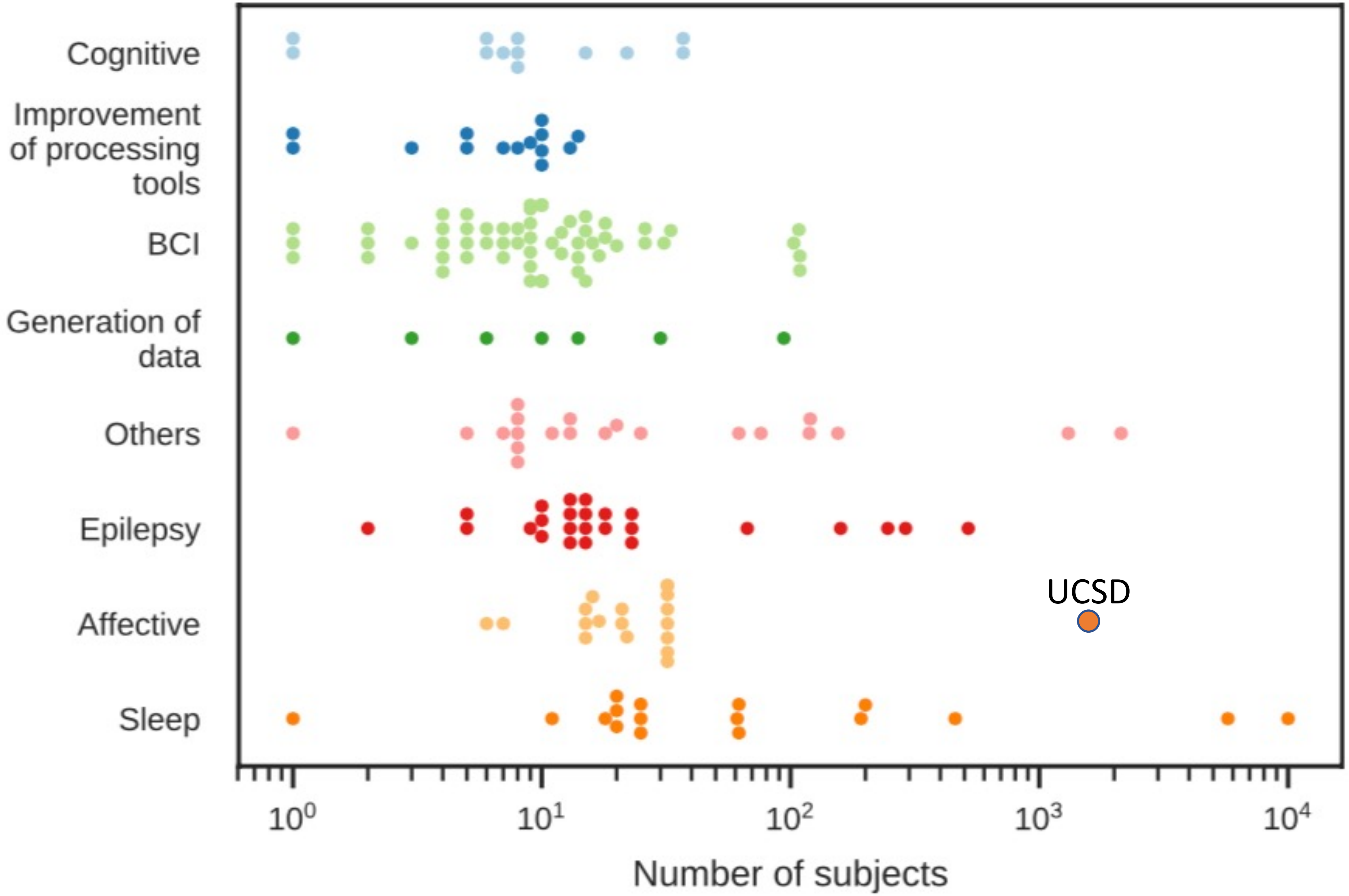


Training
944 subjects for training set (50% male)
472 subjects for validation set (50% male)

Testing
198 subjects for set 1 (50% male)
650 subjects for set 2 (100% male)



Raw input data
24 x 256 x n
Spectral data
24 x 72 x n
Spectral data
24 x 24 x n
theta
alpha
beta
Spectral
amplitude
Tapered FFT
24 x 256 complex
Theta
Alpha
Beta



Cognitive
Improvement of processing tools
BCI
Generation of data
Others
Epilepsy
Affective
Sleep
UCSD
10^0
10^1
10^2
10^3
10^4
Number of subjects

Sex classification

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

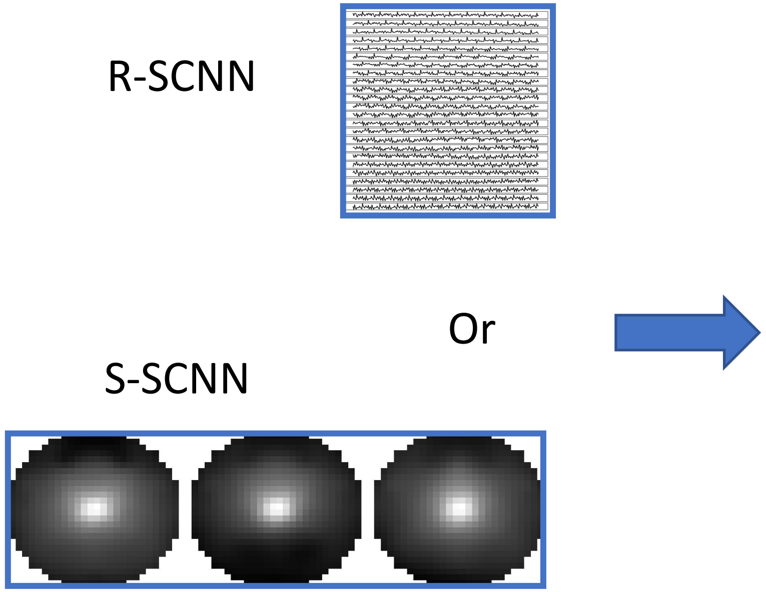
Predicting sex from brain rhythms with deep learning

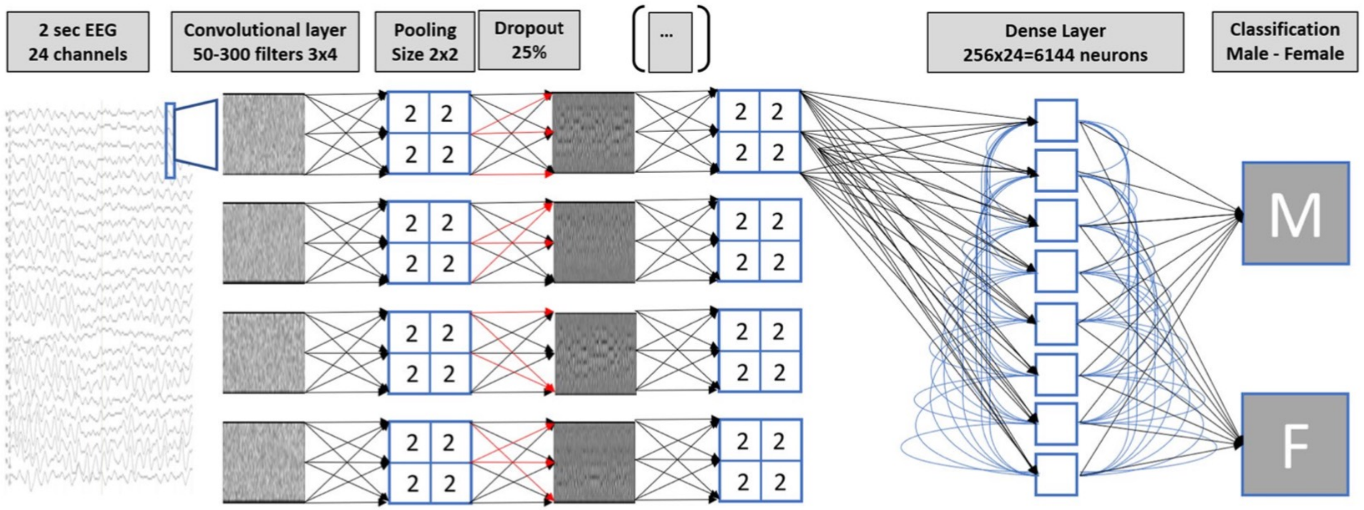
Michel J. A. M. van Putten[1], Sebastian Olbrich[2] & Martijn Arns[3]

Received: 12 October 2017
Accepted: 6 February 2018
Published online: 15 February 2018

We have excellent skills to extract sex from visual assessment of human faces, but assessing sex from human brain rhythms seems impossible. Using deep convolutional neural networks, with unique potential to find subtle differences in apparent similar patterns, we explore if brain rhythms from either sex contain sex specific information. Here we show, in a ground truth scenario, that a deep neural net can predict sex from scalp electroencephalograms with an accuracy of $>80\%$ ($p < 10^{-5}$), revealing that brain rhythms are sex specific. Further, we extracted sex-specific features from the deep net filter layers, showing that fast beta activity (20–25 Hz) and its spatial distribution is a main distinctive attribute. This demonstrates the ability of deep nets to detect features in spatiotemporal data unnoticed by visual assessment, and to assist in knowledge discovery. We anticipate that this approach may also be successfully applied to other specialties where spatiotemporal data is abundant, including neurology, cardiology and neuropsychology.

Dung Truong
MS Deep Learning

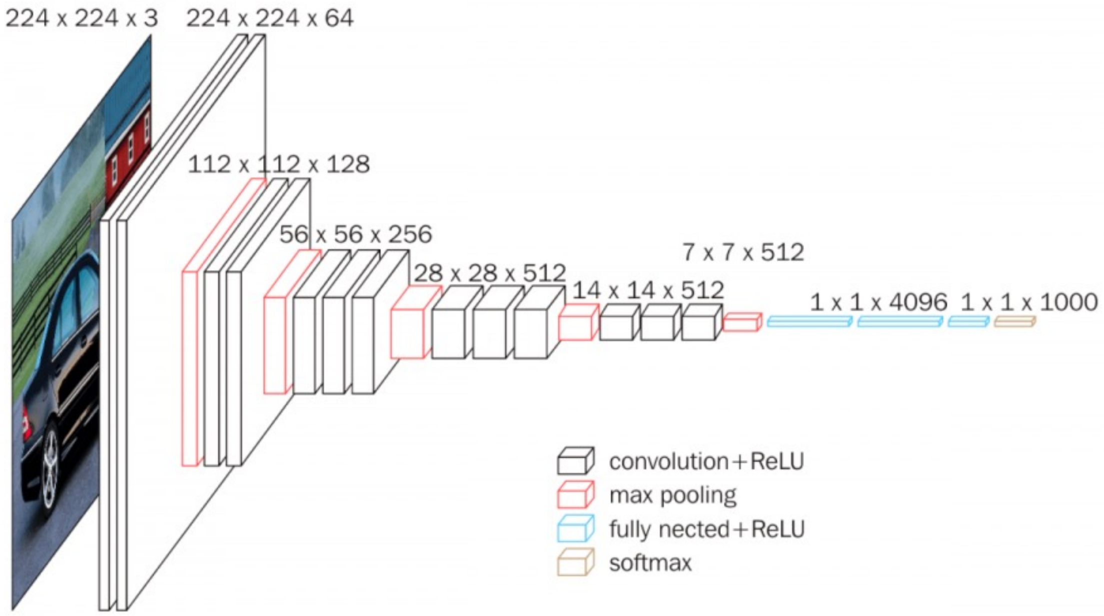




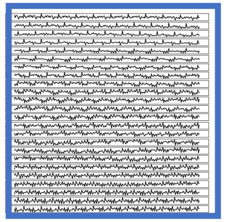
Layer	Filter size	# of filters/hidden units
Convolutional	3x3	100
MaxPooling Dropout (25%)		
Convolutional	3x3	100
MaxPooling Dropout (25%)		
Convolutional	2x3	300
MaxPooling Dropout (25%)		
Convolutional†	1x7	300
MaxPooling* Dropout (25%)		
Convolutional†	1x3	100
Convolutional†	1x3	100
Fully connected		6144
Fully connected		2
Softmax		

Modified VGG16

- Changed input size
- Keep the same scaling between layers
- Number of convolutions divided by 8 for each l
- Dropped layers 19 to 32
- Change number of output classes to 2
- Retrain the whole network

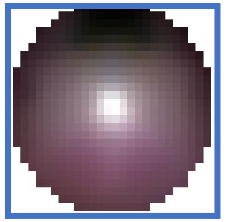


R-VGG



Or

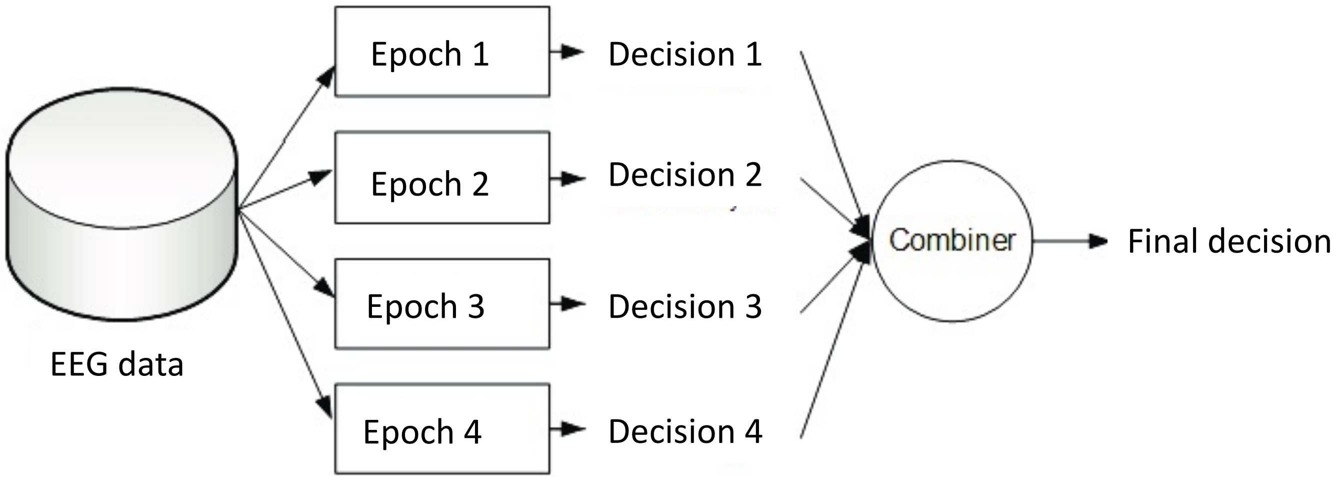
S-VGG

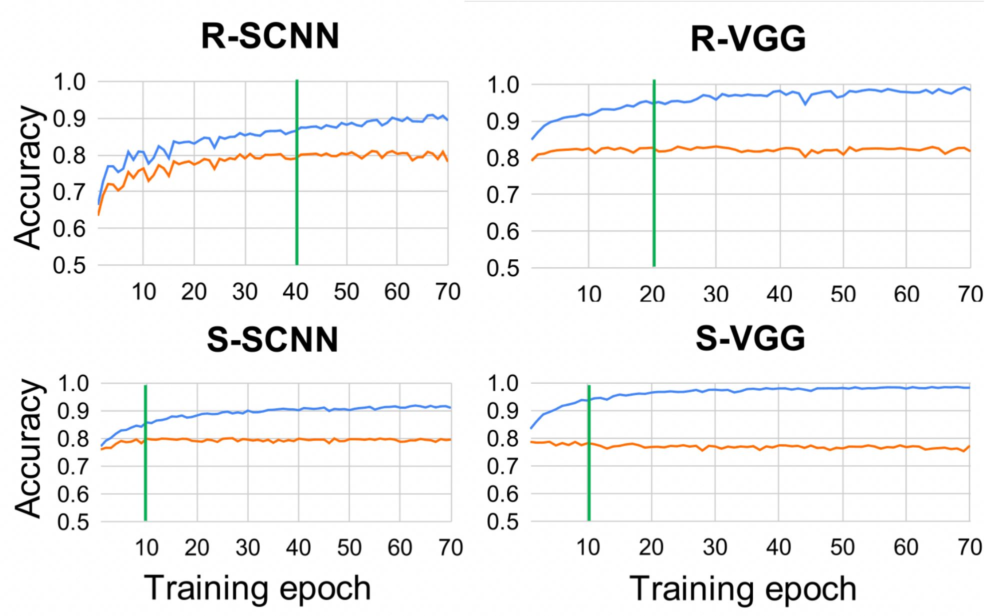


Layer	Filter size	# of filters/hidden units
Convolutional†	3x3	16
Convolutional	3x3	16
MaxPooling		
Convolutional	3x3	32
Convolutional	3x3	32
MaxPooling		
Convolutional	3x3	64
Convolutional	3x3	64
Convolutional	3x3	64
MaxPooling		
Fully connected		1024
Dropout (50%)		
Fully connected		1024
Dropout (50%)		
Fully connected		2
Softmax		

Majority voting

- Estimates are on 2-second epochs
- We have about 85 per subject
- Boost performance by about 8% (see next slide)

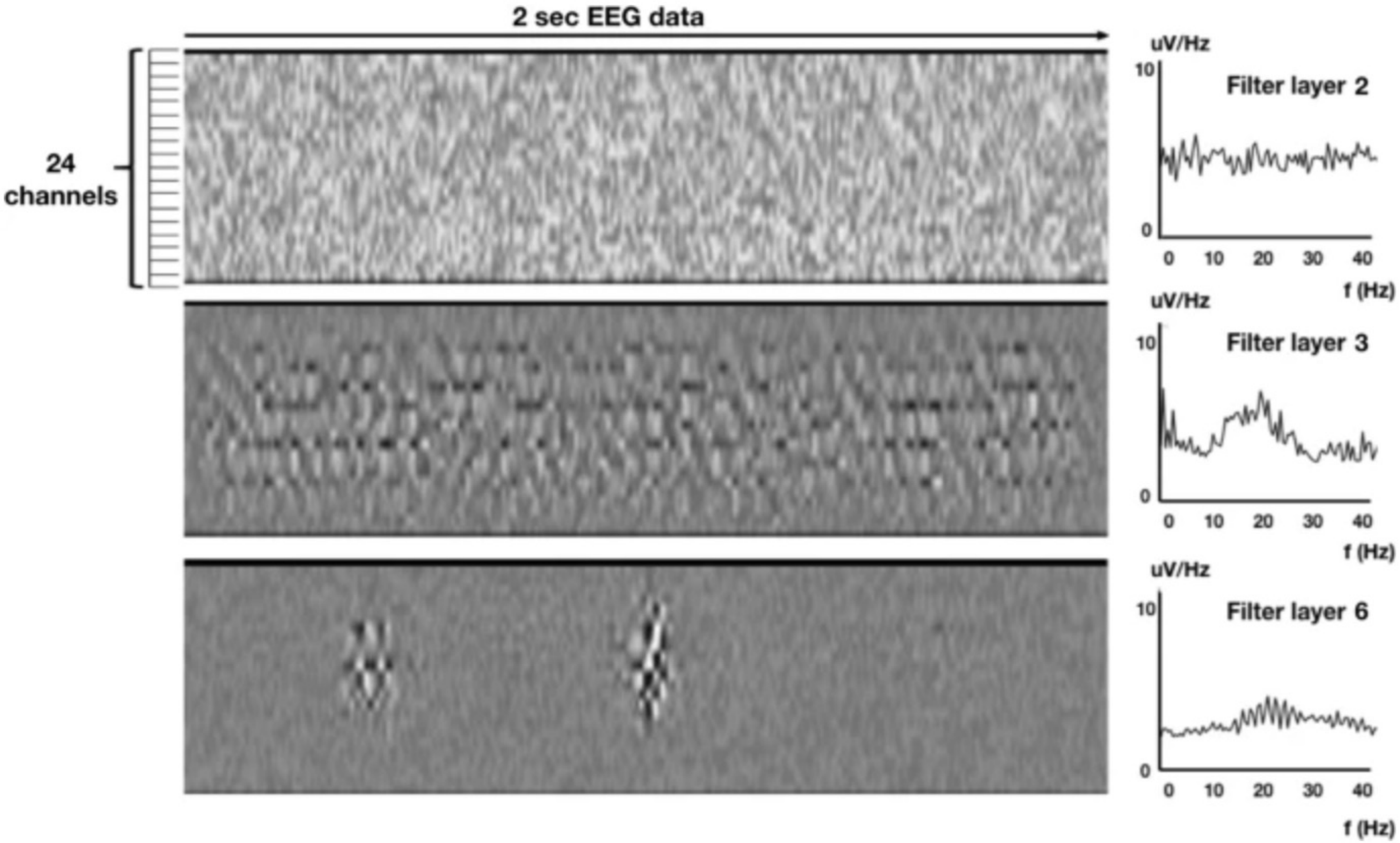




	2-sec	2-sec x n with vote
Model	**Per-sample**	**Per-subject**
R-SCNN	80.6 (79.7 to 81.5)	85.1 (84.3 to 85.9)
R-VGG	**83.1** (82.7 to 83.4)	**87.0** (86.6 to 87.4)
S-SCNN	79.0 (78.7 to 79.3)	83.2 (82.1 to 84.3)
S-VGG	77.1 (76.8 to 77.4)	81.3 (80.0 to 82.6)

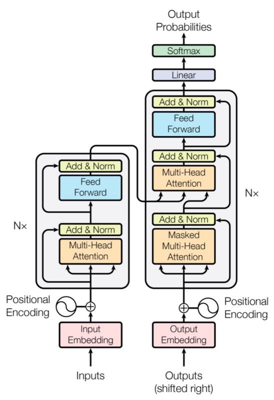
Table 3. Models classification accuracy. 95% confidence interval is indicated in parenthesis. Bolded values indicate best performance.

Dung Truong and Michael Milham and Scott Makeig and Arnaud Delorme (2021) Deep Convolutional Neural Network Applied to Electroencephalography: Raw Data vs Spectral Features, https://arxiv.org/abs/2105.04762



van Putten MJAM, Olbrich S, Arns M. Predicting sex from brain rhythms with deep learning. Sci Rep. 2018 Feb 15;8(1):3069. doi: 10.1038/s41598-018-21495-7. PMID: 29449649; PMCID: PMC5814426.

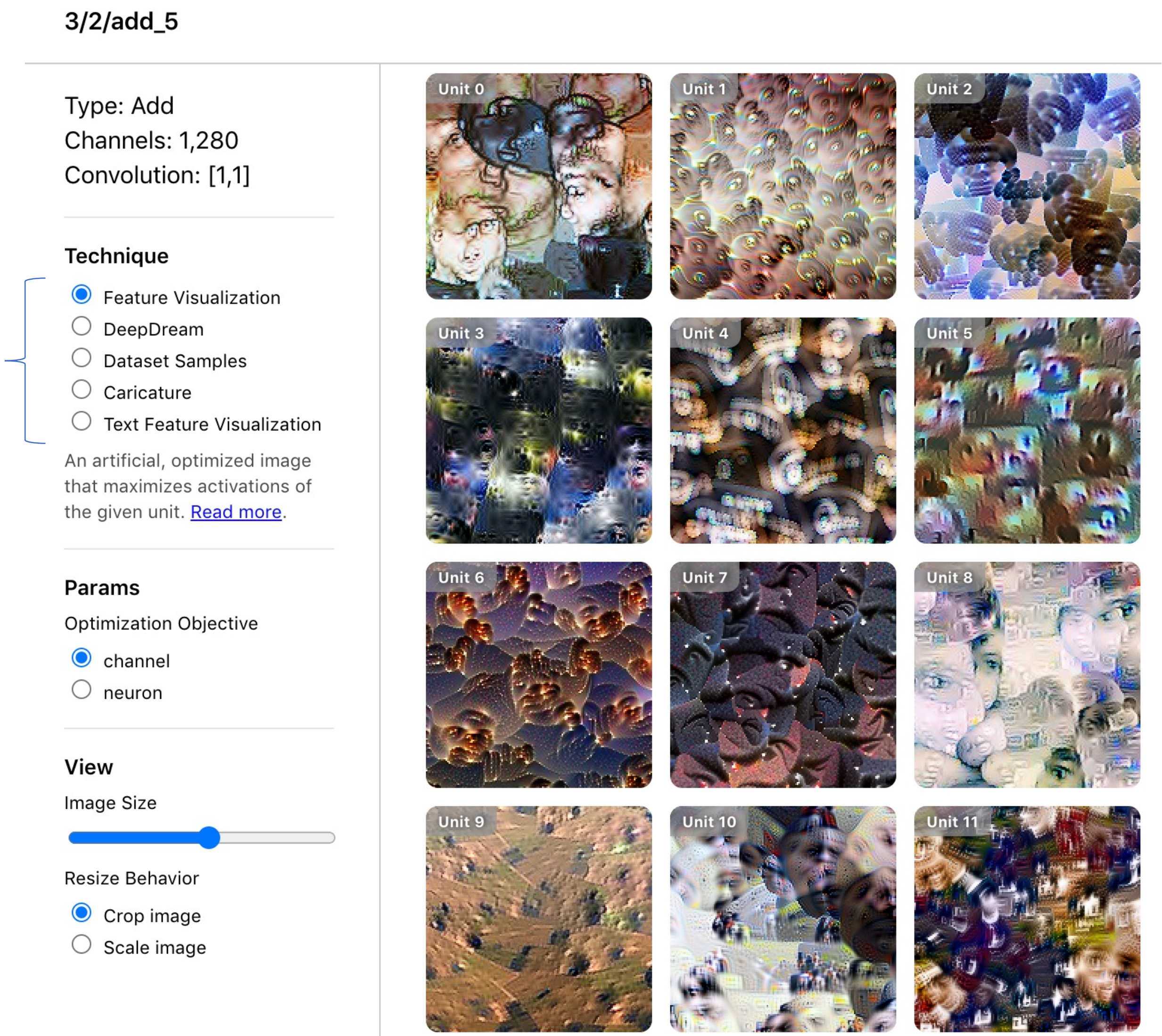
Future developments



- New interpretations/features for EEG
- New architectures
- Benchmark data and scripts (Matlab and Python)
- EEGLAB data exporter

Interpretation

Reconstruction

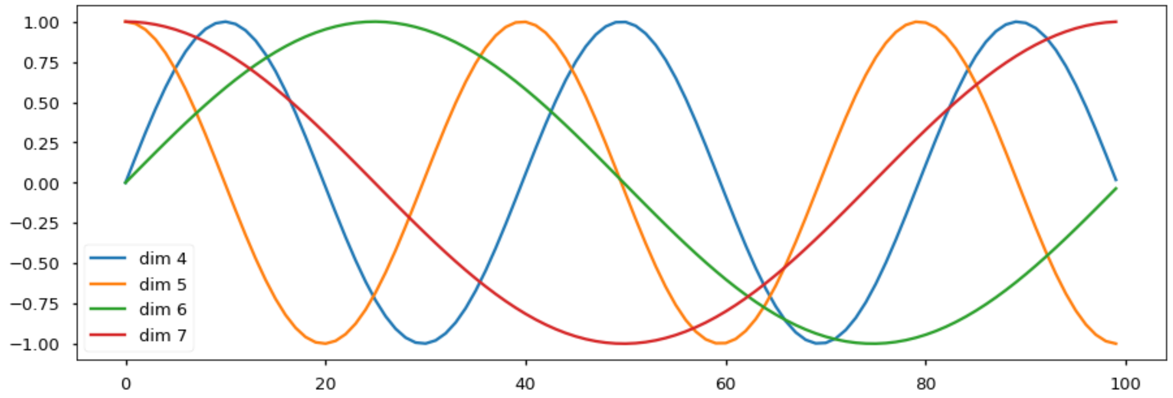


Perspective

- Attention networks

Vaswani et al. (2017) All you need is attention. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Positional encoding



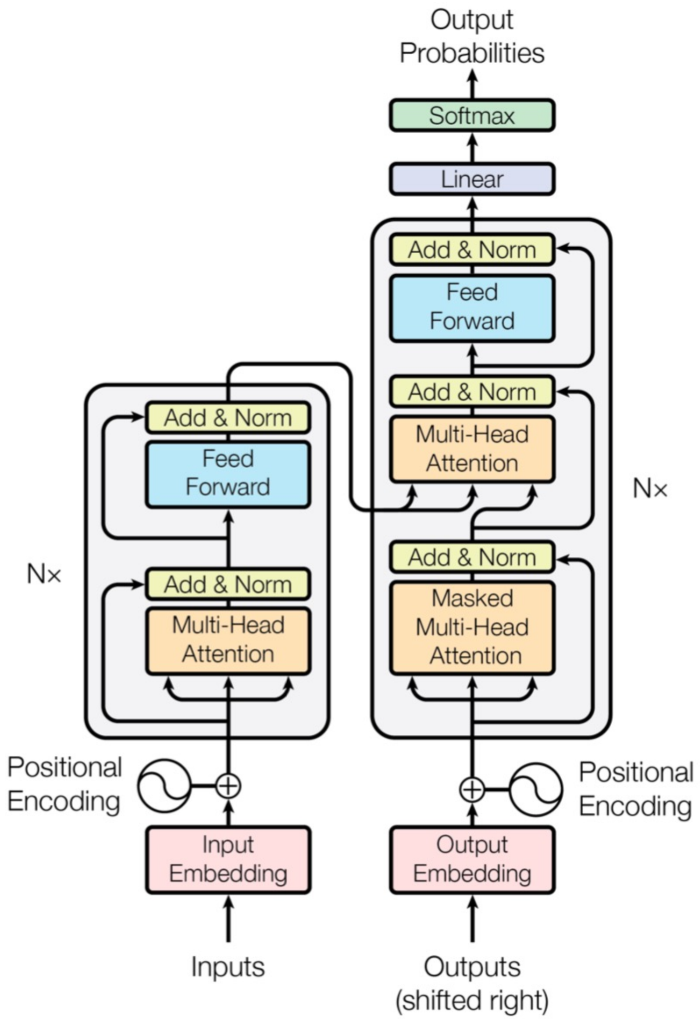
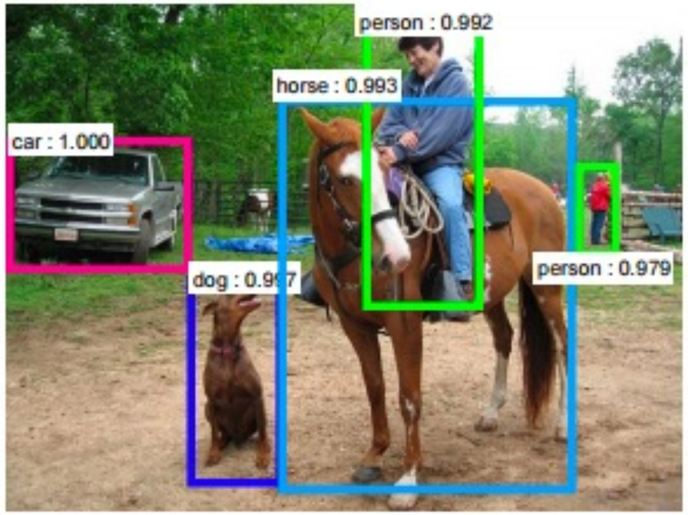


Figure 1: The Transformer - model architecture.

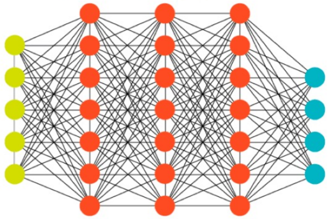


- EEG
- Other bio sensors
- Motion
- Eye tracker

- Automated object labeling
- Automated sound labeling

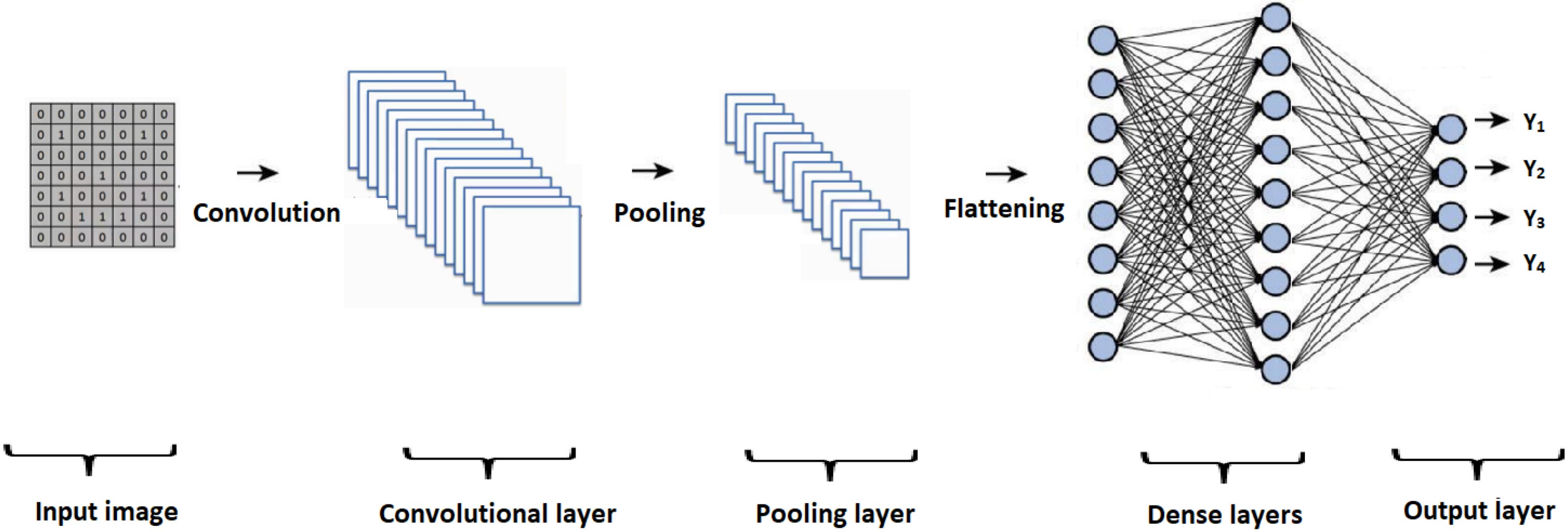
Event labels

Deep learning



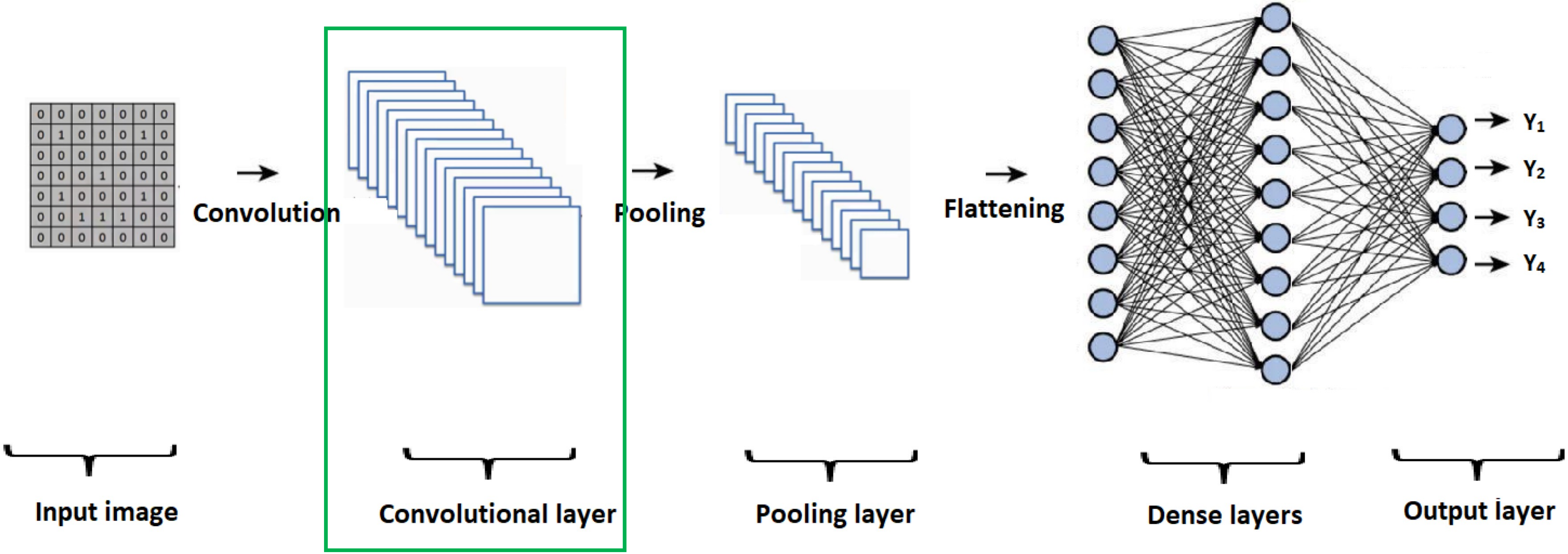
Scriptures
written
Holy

Convolutional Neural Network Architecture



https://www.mdpi.com/2076-3417/10/4/1245/htm

Convolutional Neural Network Architecture



https://www.mdpi.com/2076-3417/10/4/1245/htm

Convolutional layer

- Convolution operation

Input * Kernel/Filter = Feature/Activation map

$$
\begin{bmatrix} 7 & 2 & 3 & 3 & 8 \\ 4 & 5 & 3 & 8 & 4 \\ 3 & 3 & 2 & 8 & 4 \\ 2 & 8 & 7 & 2 & 7 \\ 5 & 4 & 4 & 5 & 4 \end{bmatrix}
*
\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}
=
\begin{bmatrix} 6 & & \\ & & \\ & & \end{bmatrix}
$$

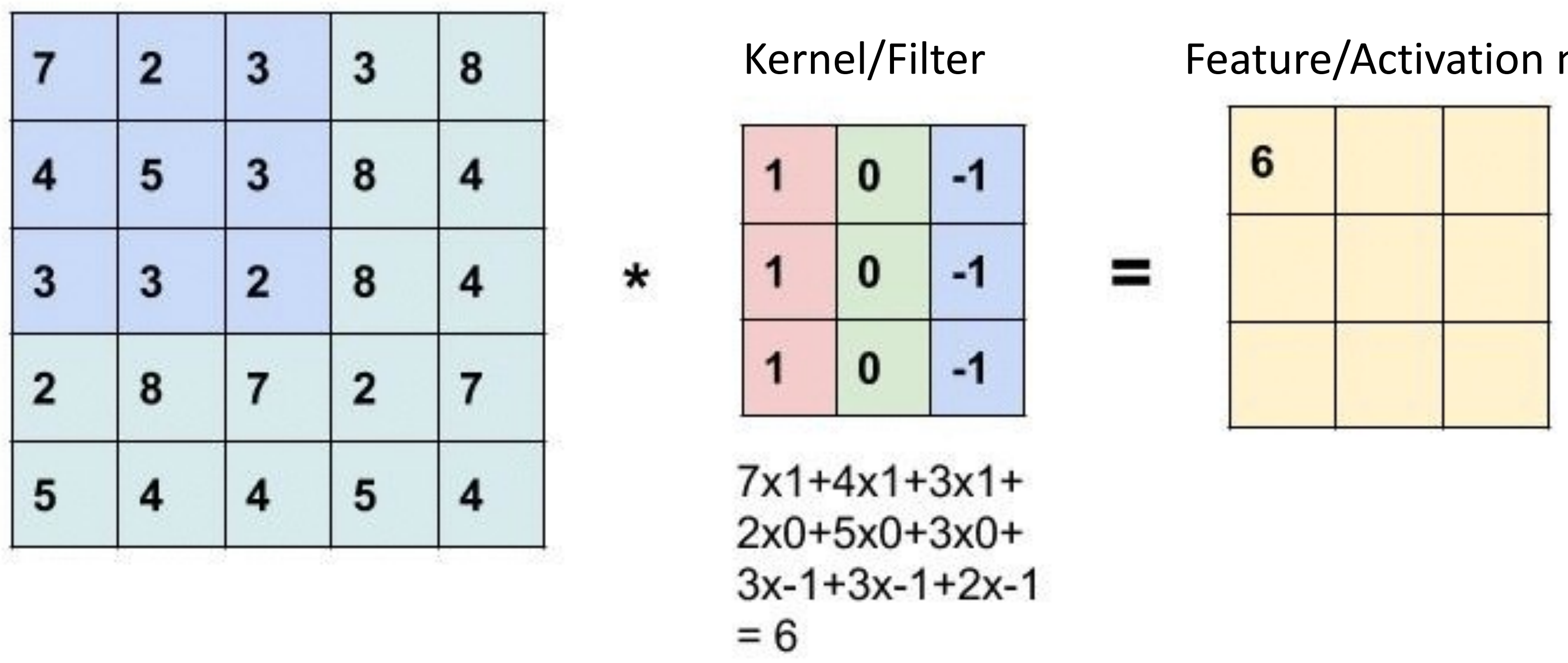
7x1+4x1+3x1+
2x0+5x0+3x0+
3x-1+3x-1+2x-1
= 6

https://medium.datadriveninvestor.com/convolutional-neural-networks-3b241a5da51e

Convolutional layer

- Convolution operation

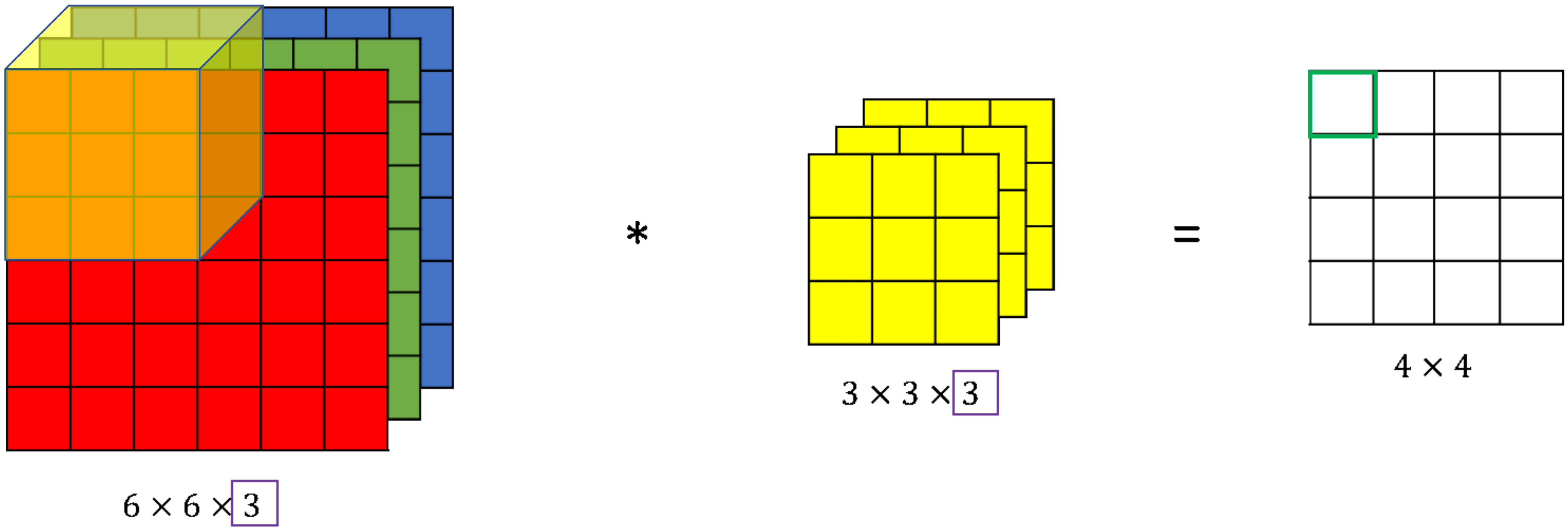
Input



https://medium.datadriveninvestor.com/convolutional-neural-networks-3b241a5da51e

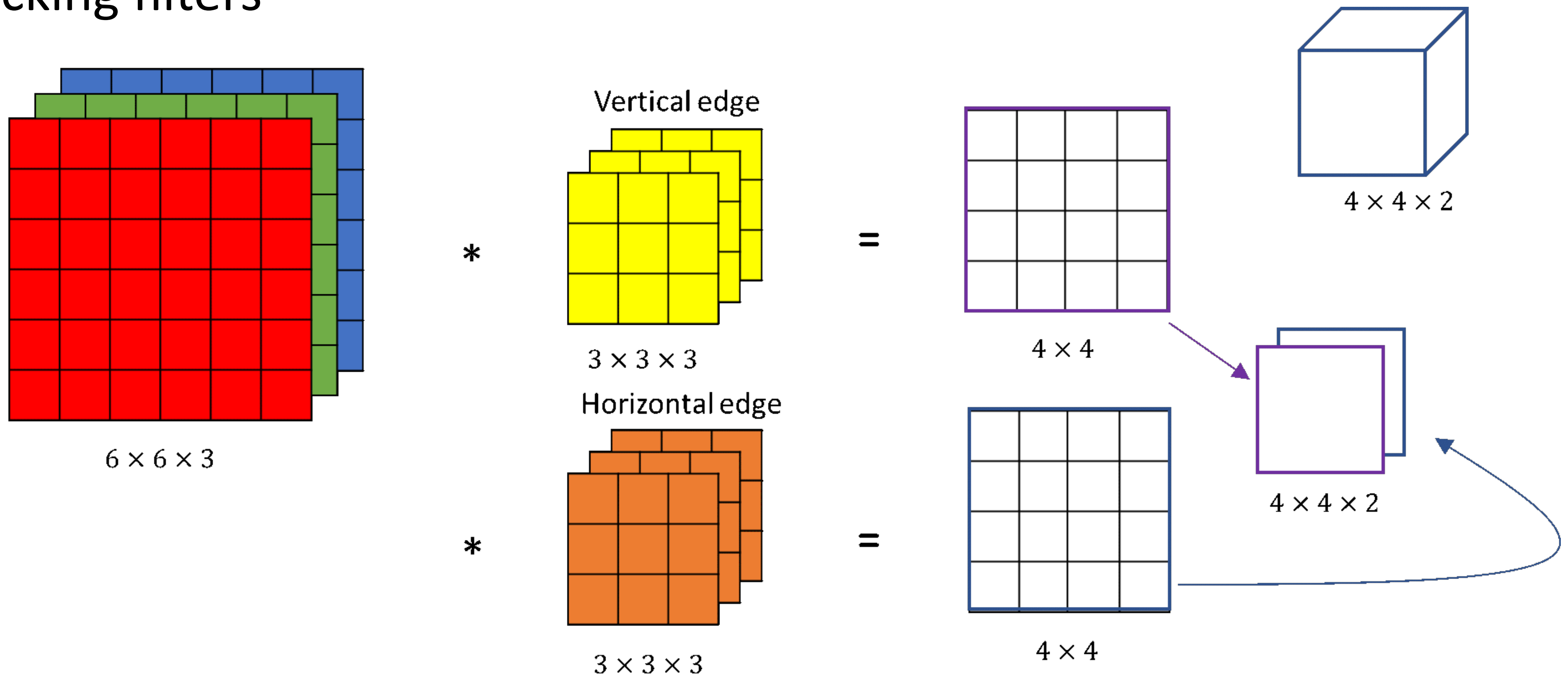
Convolutional layer

- Convolution over volume

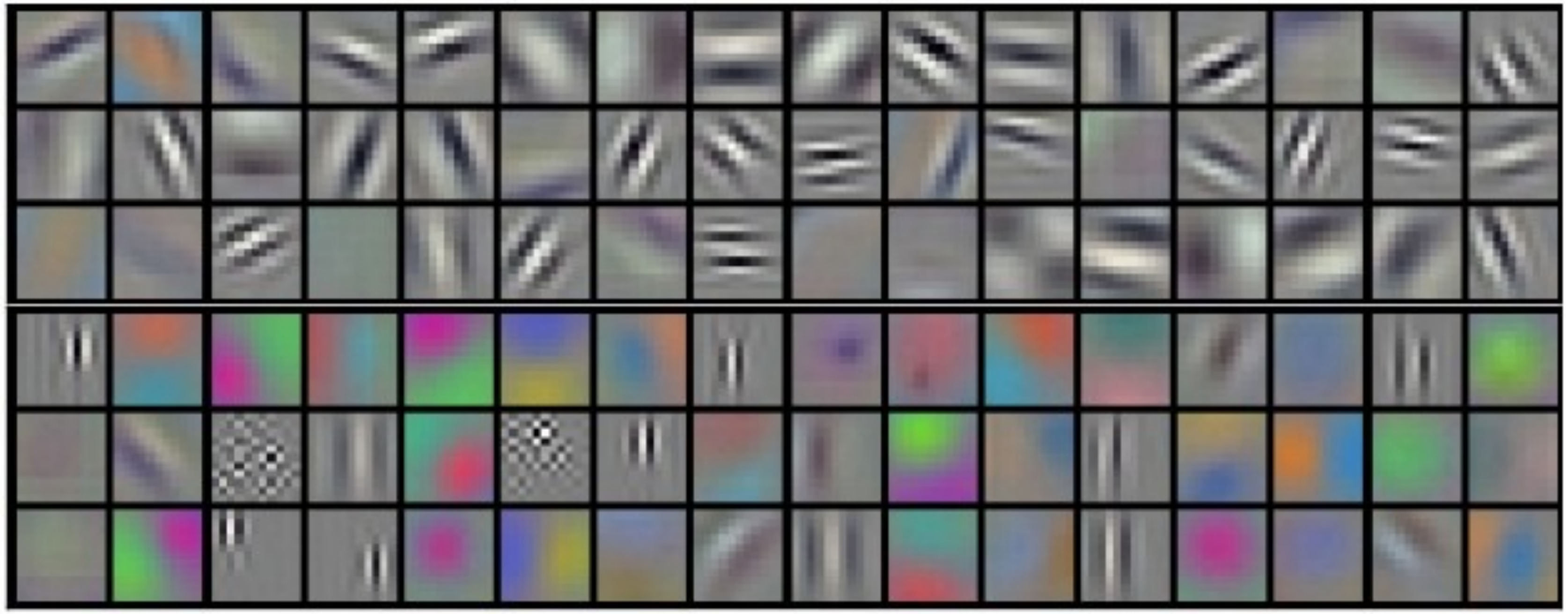


Convolutional layer

- Stacking filters

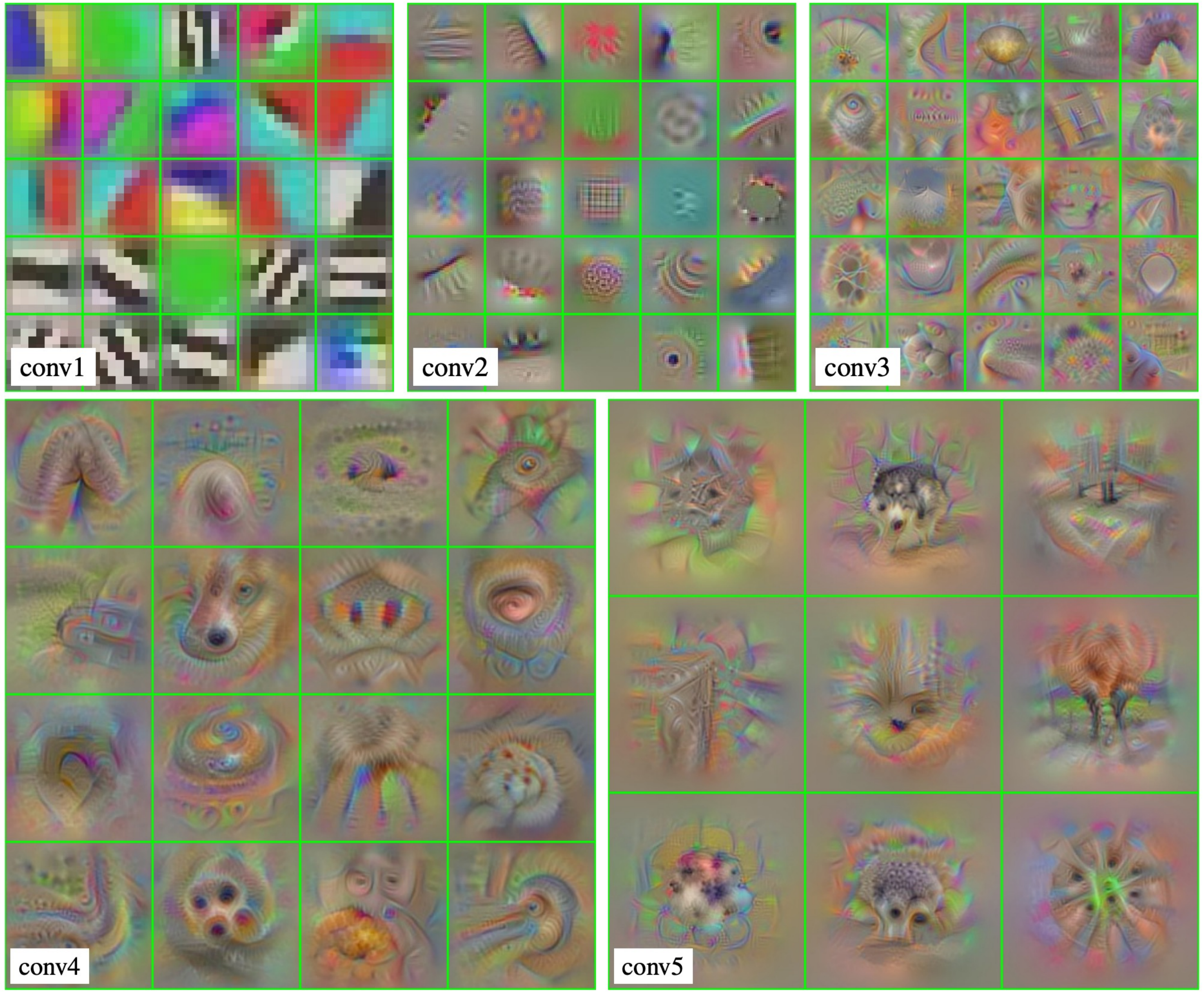


Convolutional layer



Krizhevsky et al., 2012

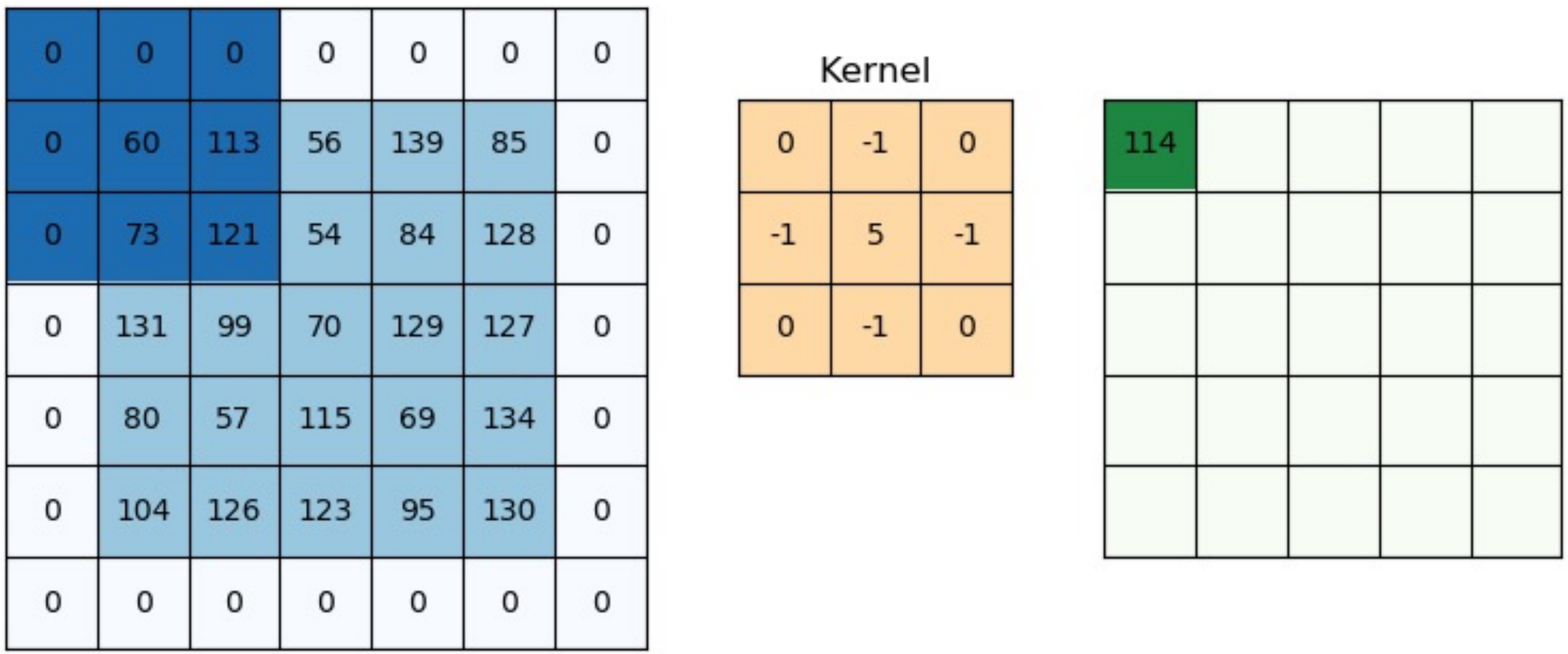
Convolutional layer



Mahendran & Vedaldi, 2016

Convolutional layer

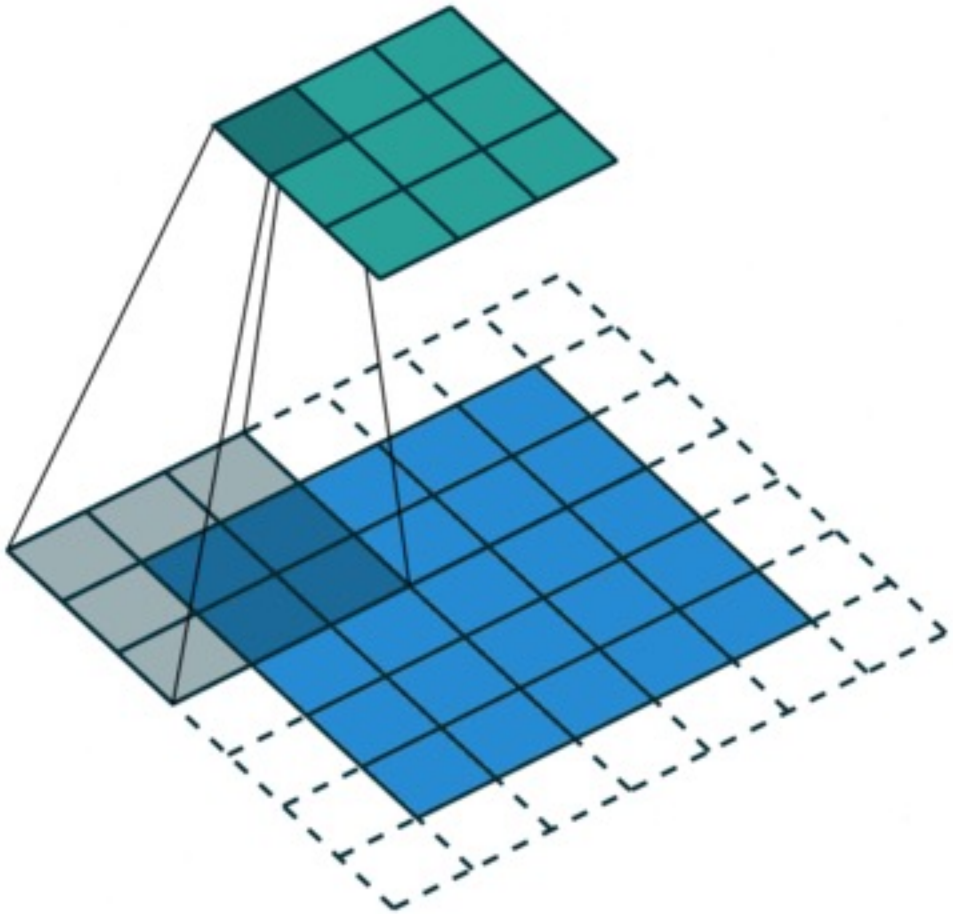
- Padding



https://www.pyimagesearch.com/2018/12/31/keras-conv2d-and-convolutional-layers/

Convolutional layer

- Stride

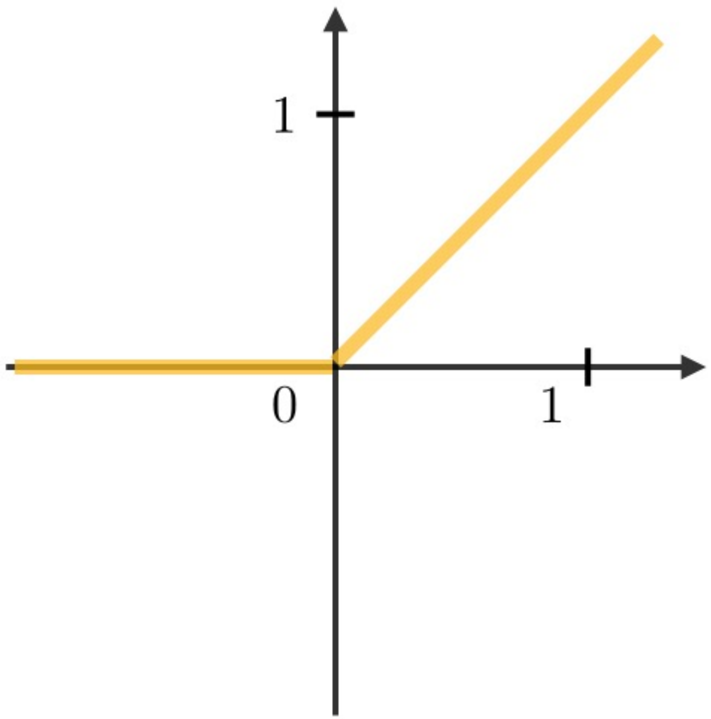


Example: Padding = 1, Stride = 2

Activation functions

- Rectified Linear Unit (ReLU)

$$g(z) = \max(0, z)$$



ReLU Layer

Filter 1 Feature Map

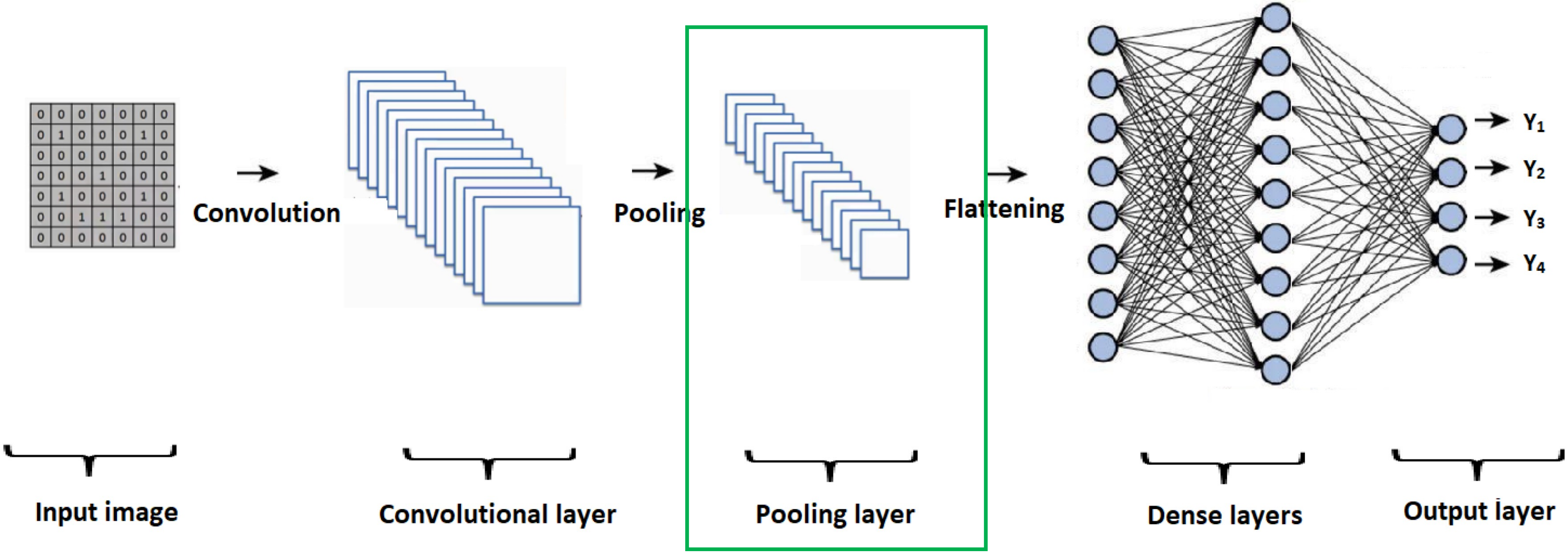
9	3	5	-8
-6	2	-3	1
1	3	4	1
3	-4	5	1

→

9	3	5	0
0	2	0	1
1	3	4	1
3	0	5	1

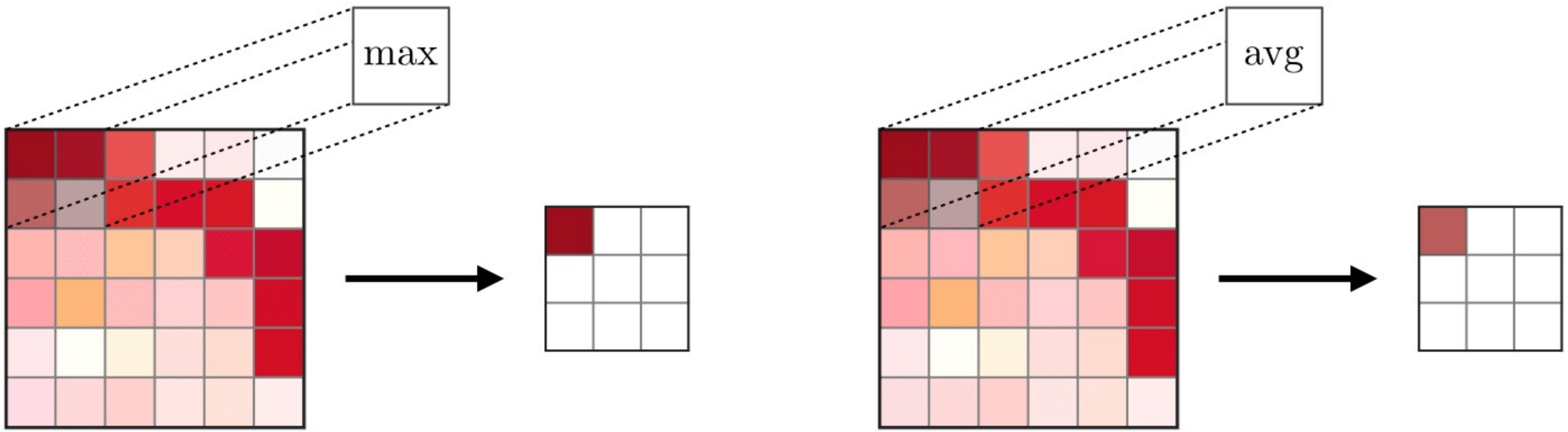
https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks#layer

Convolutional Neural Network Architecture



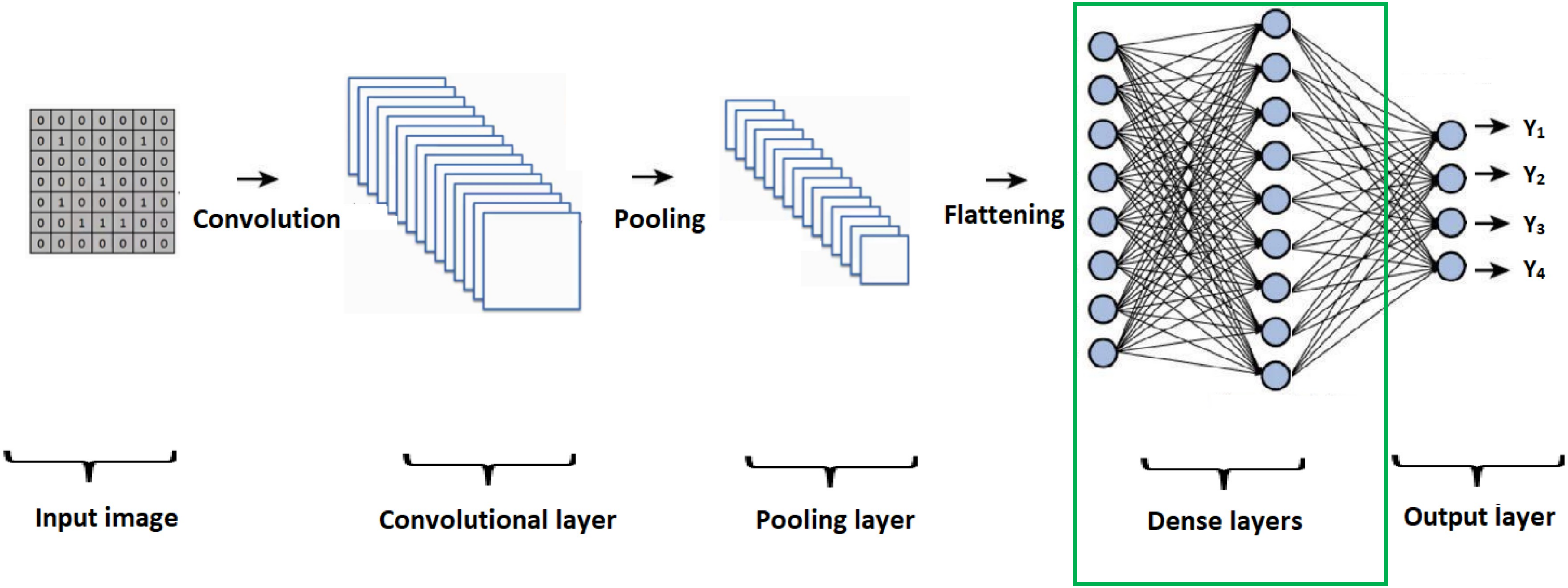
https://www.mdpi.com/2076-3417/10/4/1245/htm

Pooling layer



https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks#layer

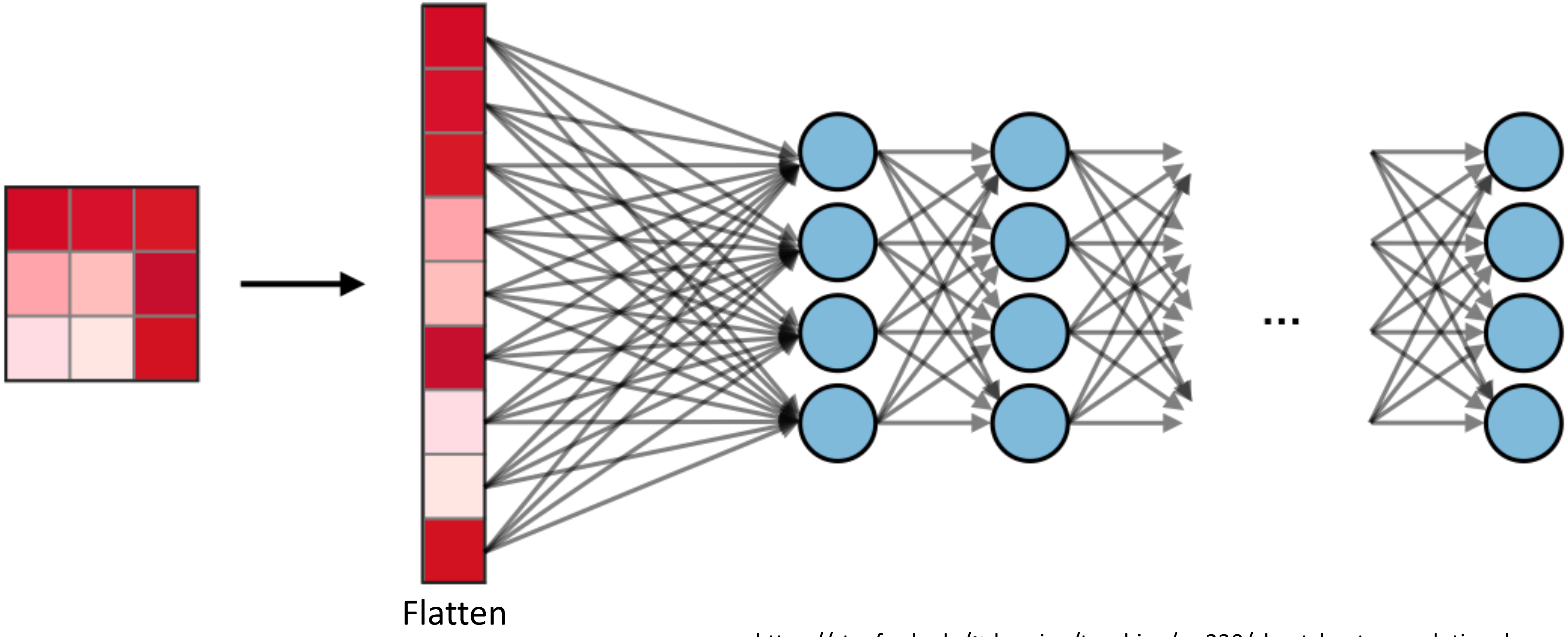
Convolutional Neural Network Architecture



https://www.mdpi.com/2076-3417/10/4/1245/htm

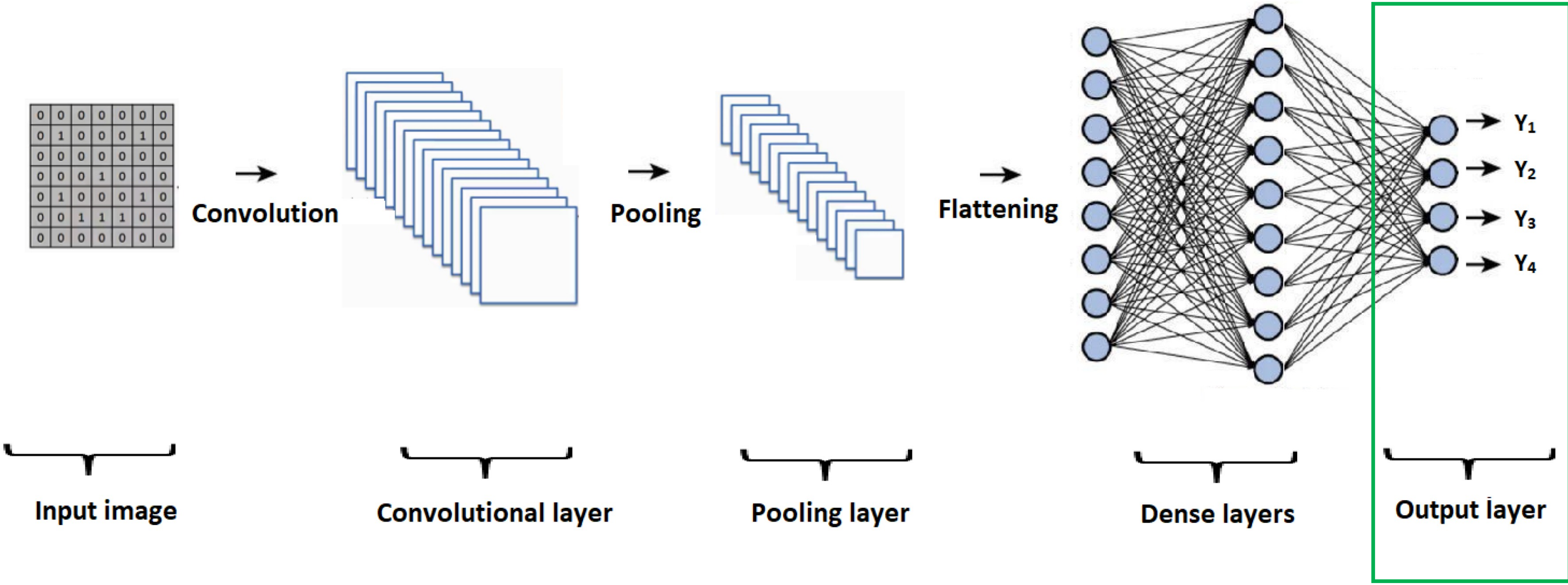
Fully connected layer

- Also known as dense layer
- Each input is connected to all hidden units



https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks#layer

Convolutional Neural Network Architecture



https://www.mdpi.com/2076-3417/10/4/1245/htm

Activation functions

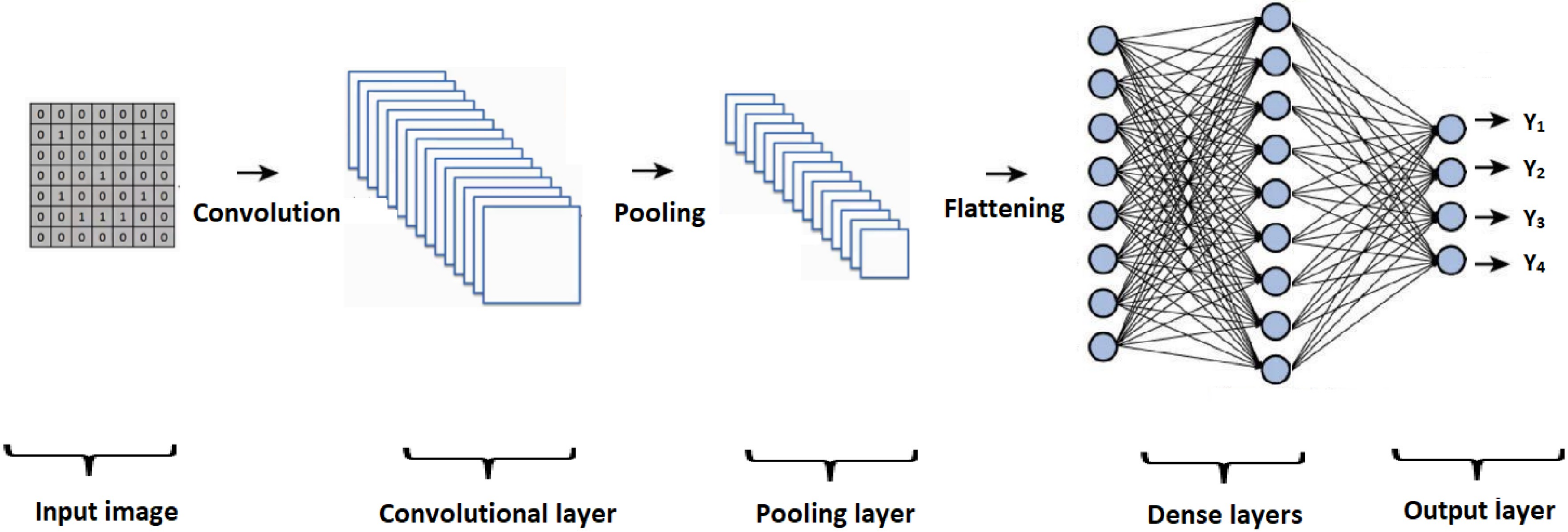
- Softmax

$$\boxed{p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}} \quad \text{where} \quad \boxed{p_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}}$$

- Cross-entropy loss

$$L = -\frac{1}{m}\sum_{i=1}^{m} y_i \cdot \log(\hat{y}_i)$$

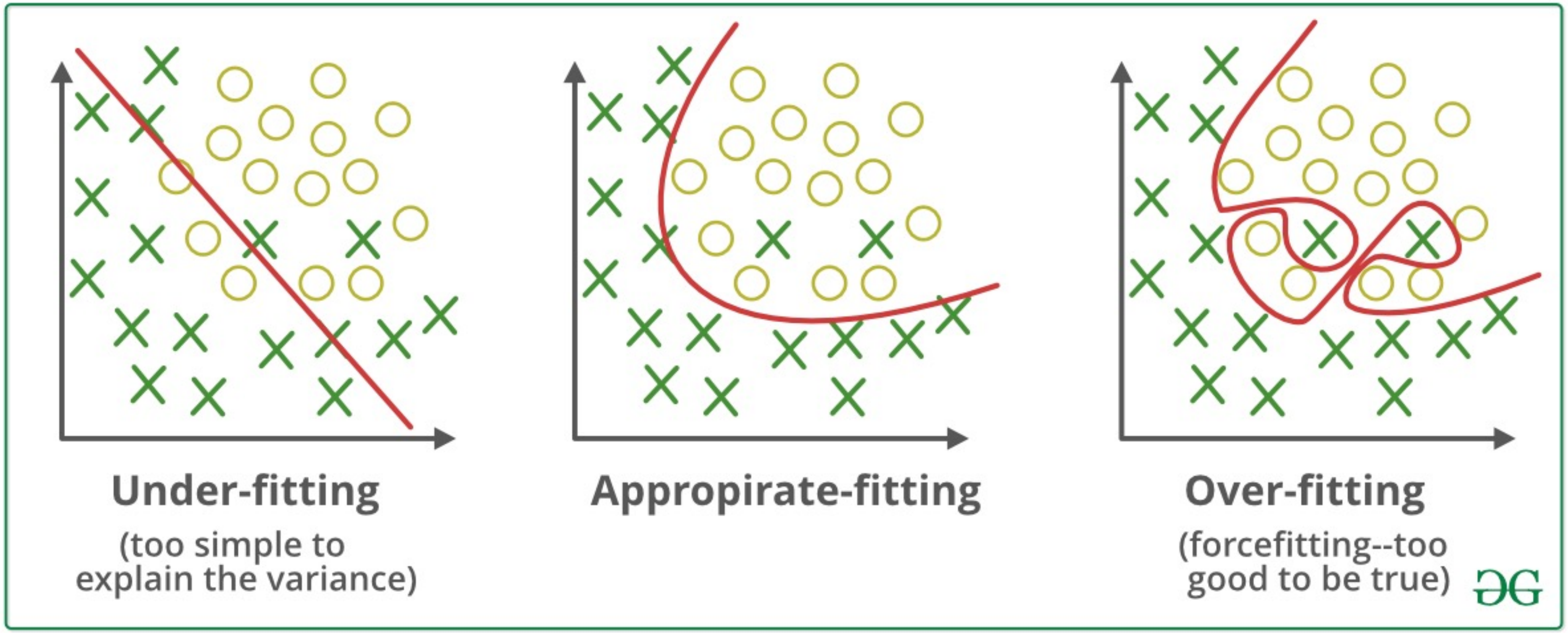
y is a one-hot vector ([0,0,1,....,0])



https://www.mdpi.com/2076-3417/10/4/1245/htm

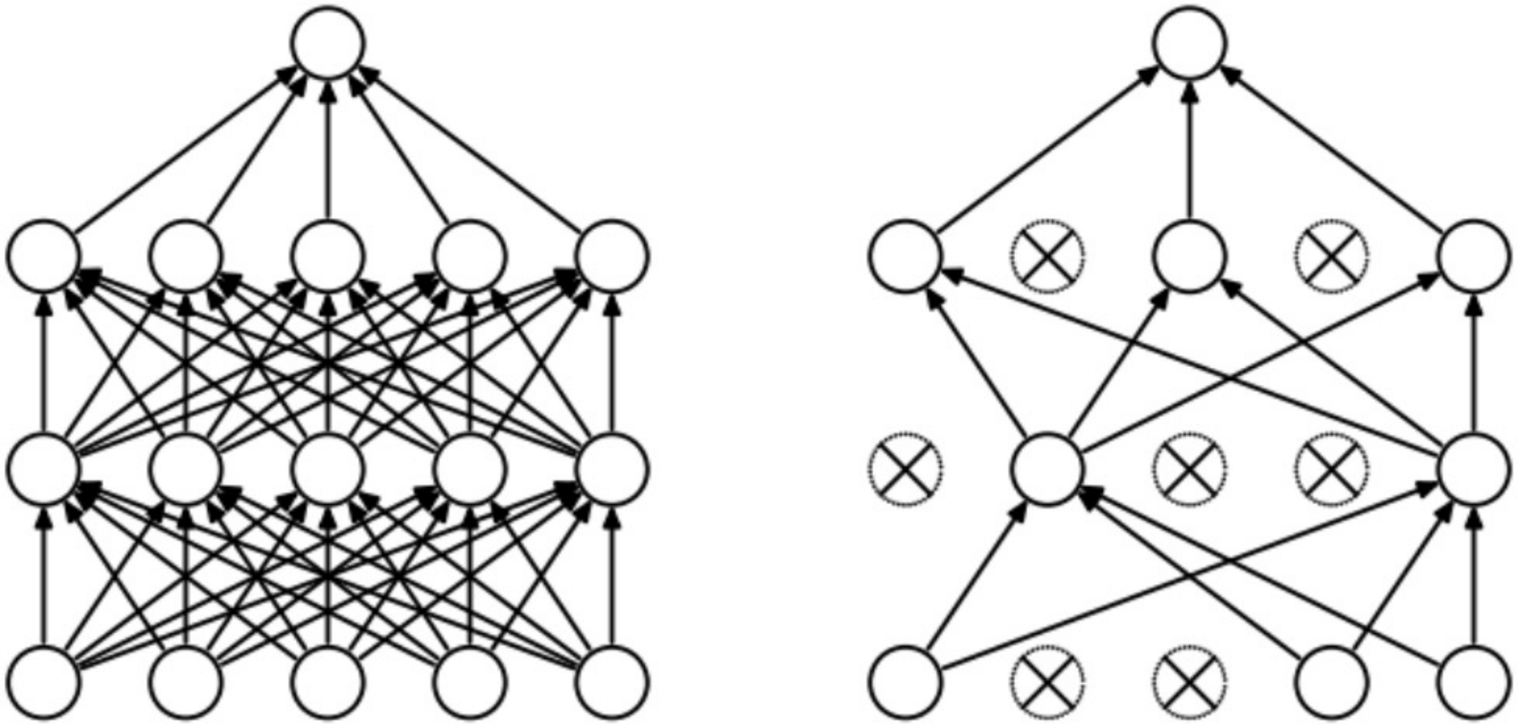
Dropout

- Overfitting



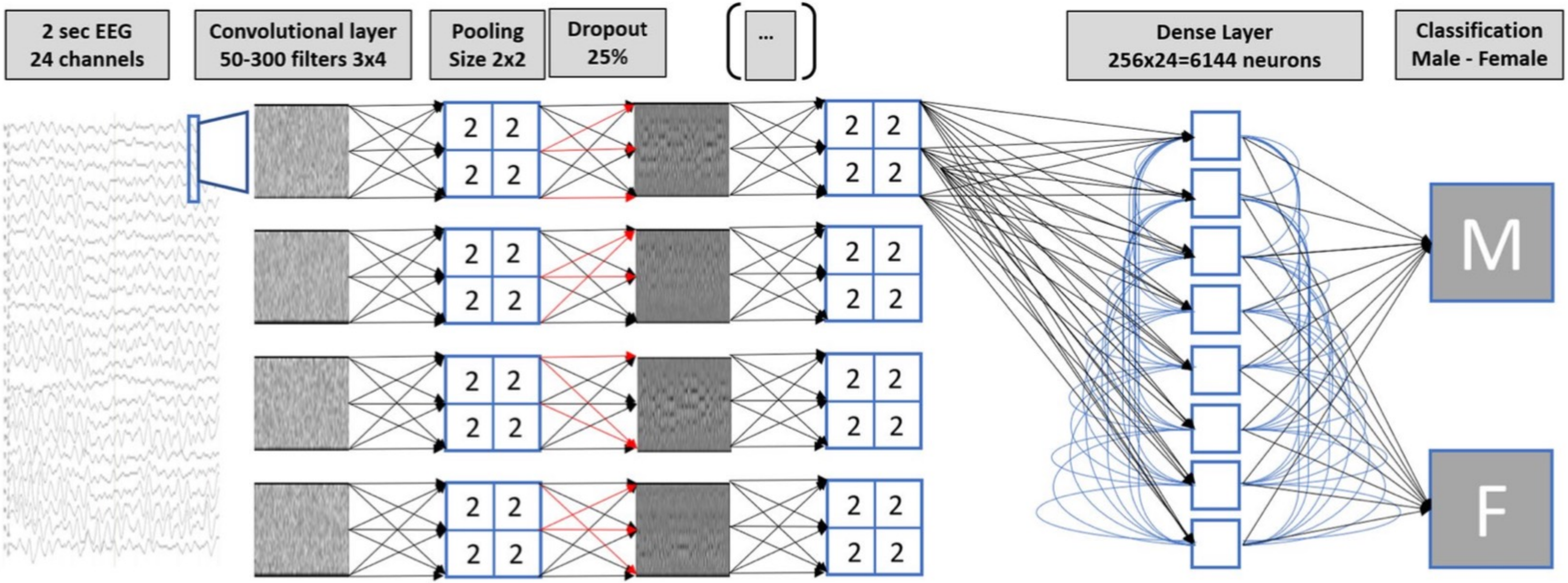
https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/

Dropout



(a) Standard Neural Net (b) After applying dropout.

https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5


2 sec EEG
24 channels
Convolutional layer
50-300 filters 3x4
Pooling
Size 2x2
Dropout
25%
...
Dense Layer
256x24=6144 neurons
Classification
Male - Female
2 2
2 2
M
F

The model

```
-------------------------------------------------------------------------
      Layer (type)          Input Shape         Param #     Tr. Param #
=========================================================================
          Conv2d-1      [1, 1, 24, 256]           1,000           1,000
            ReLU-2    [1, 100, 22, 254]               0               0
       MaxPool2d-3    [1, 100, 22, 254]               0               0
         Dropout-4    [1, 100, 11, 127]               0               0
          Conv2d-5    [1, 100, 11, 127]          90,100          90,100
            ReLU-6     [1, 100, 9, 125]               0               0
       MaxPool2d-7     [1, 100, 9, 125]               0               0
         Dropout-8      [1, 100, 4, 62]               0               0
          Conv2d-9      [1, 100, 4, 62]         180,300         180,300
           ReLU-10      [1, 300, 3, 60]               0               0
      MaxPool2d-11      [1, 300, 3, 60]               0               0
        Dropout-12      [1, 300, 1, 30]               0               0
         Conv2d-13      [1, 300, 1, 30]         630,300         630,300
           ReLU-14      [1, 300, 1, 24]               0               0
      MaxPool2d-15      [1, 300, 1, 24]               0               0
        Dropout-16      [1, 300, 1, 23]               0               0
         Conv2d-17      [1, 300, 1, 23]          90,100          90,100
         Conv2d-18      [1, 100, 1, 21]          30,100          30,100
        Flatten-19      [1, 100, 1, 19]               0               0
         Linear-20            [1, 1900]      11,679,744      11,679,744
         Linear-21            [1, 6144]          12,290          12,290
=========================================================================
Total params: 12,713,934
Trainable params: 12,713,934
Non-trainable params: 0
-------------------------------------------------------------------------
```