# Convolutional Neural Networks for Predicting Daily Orographic Precipitation Gradients for Precipitation Downscaling

CIROH Developers Conference

Hydrological Applications of Machine Learning Workshops

Savanna Wolvin

05.30.24

# Importance of Western CONUS Snowpack

- Seasonal snowpack acts as a natural water tower – storing wintertime precipitation
- 71% of total runoff in the mountainous west is from snowmelt

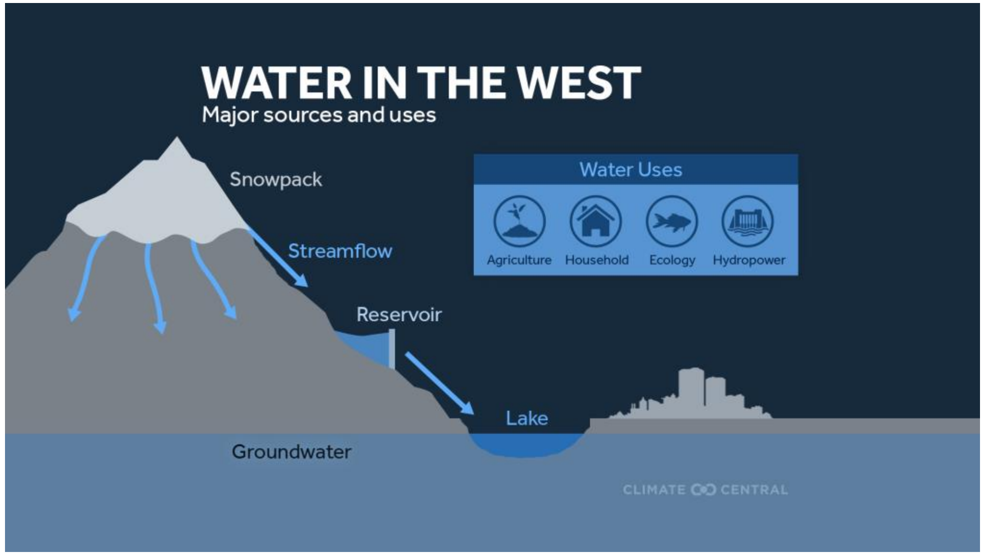



Fig. Water in the West (Courtesy of climatecentral.org)

**Multiple avalanches reported in Blaine County, Extreme Avalanche Danger Warning issued**

According to the Sawtooth Avalanche Center, there is a level 5 - Extreme Avalanche Danger in the Soldier and Wood River Mountains.



By KMVT News Staff

*Published: Mar. 10, 2023 at 12:53 PM MST*

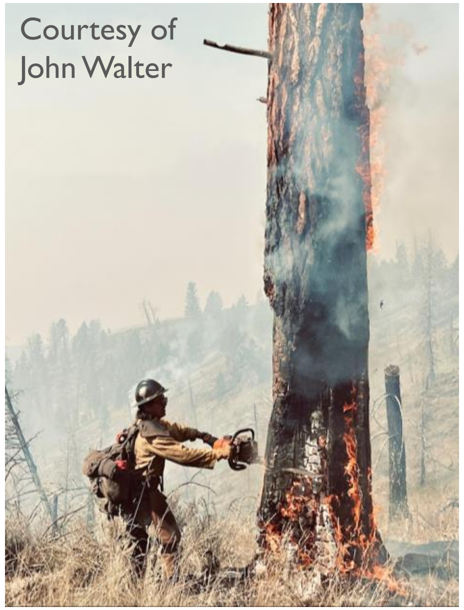
**Big Cottonwood Canyon closed in both directions due to multiple weather-related crashes**

by Kayla Winn, KUTV | Mon, December 12th 2022



*Multiple weather-related crashes forced both directions of Big Cottonwood Canyon closed Monday, Dec. 12, 2022. (Photo: UDOT Cottonwood Canyons)*

Courtesy of Grace Cawley



Courtesy of John Walter

(Li et al. 2017; Burakowski and Magnusson 2012; Black and Mote 2015; Westerling 2016)

# Current Orographic Downscaling Method

Precipitation-elevation Regressions on Independent Slopes Model (PRISM)

- Downscales using the climatological distribution of precipitation with elevation
- Downscales the GEFS from ~33 km to ~800 m horizontal resolution
- Does not account for event-to-event variability of precipitation

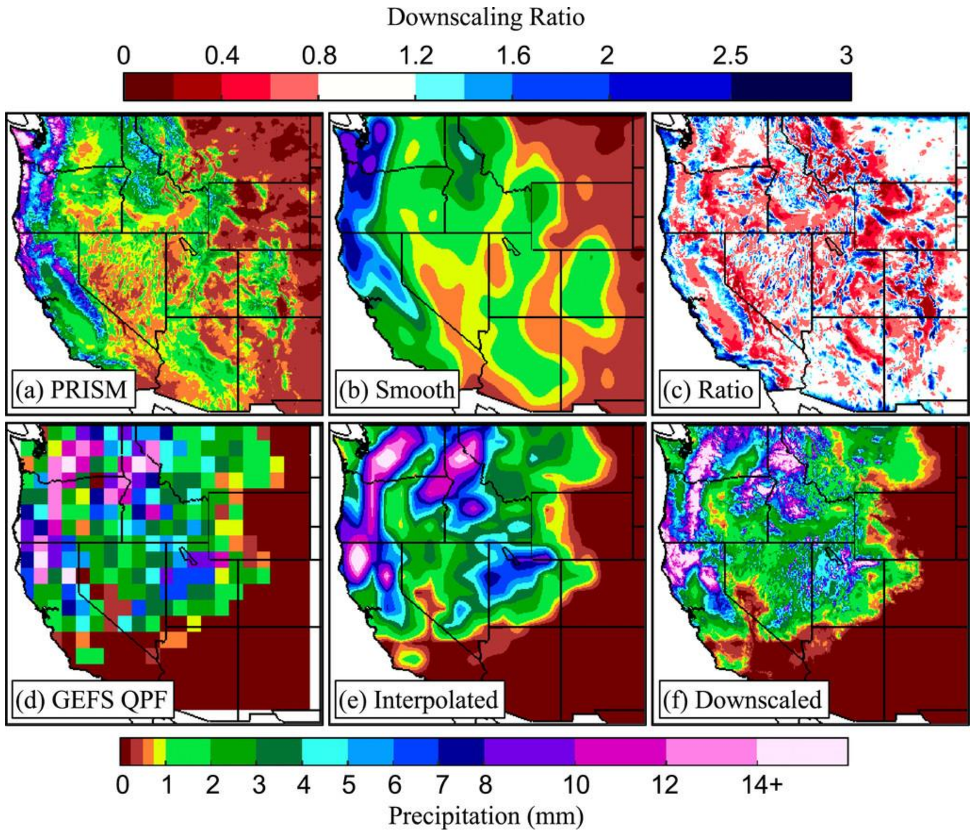



Fig. 24 Jan 2016 PRISM climatological precipitation (a), smoothed precipitation (b), downscaling ratio (c), 24 Jan 2016 GEFS forecast (d), interpolated GEFS forecast (e), downscaled precipitation (f).

(Lewis et al. 2017; Daly et al. 1994)

# Precipitation and Terrain → Faceting Algorithm → Orographic Precipitation Gradients (OPG)

Global Historical Climatology Network-Daily

- 1979-2018

Global 30 Arc-Second Elevation (GTOPO30)

- Drawn to approximately 4-km resolution

Terrain orientations are binned into 8 secondary intercardinal coordinates

- Regions deemed flat are removed
- Regions with the same orientation are called facets

Linear regression of the precipitation-elevation relationship

$$P_i = b_1 z_i + b_0$$

Fig. Precipitation observations stations and 4-km topography

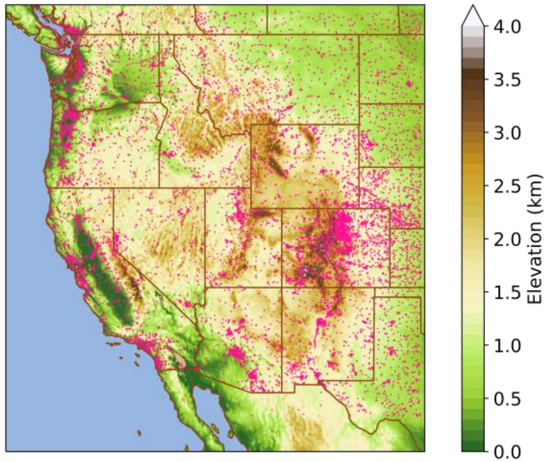



Fig. Northern Utah Facets

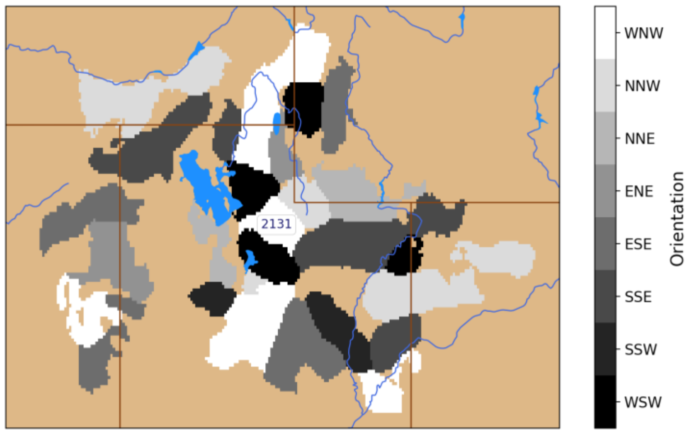
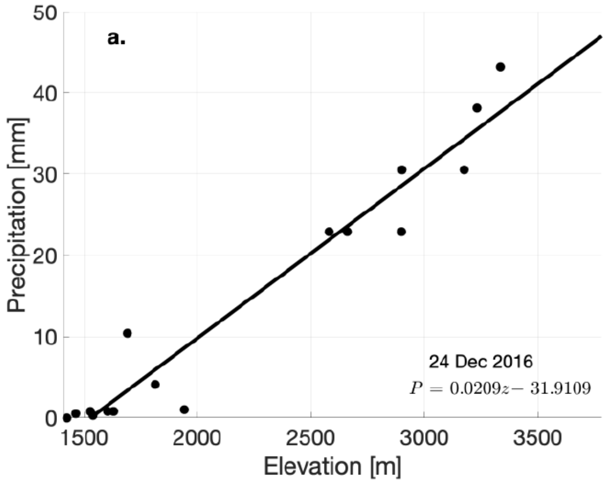



Fig. Precipitation-elevation regression




(Bohne et al. 2020)

# Predictor Dataset

## ECMWF ERA5 at 0.5° Grid-Spacing

| Variables | Levels ("Images") |
|---|---|
| Integrated Vapor Transport | Formulated from Specific Humidity, U-Winds, and V-Winds |
| U-Winds | 700 hPa |
| V-Winds | 700 hPa |
| Geopotential Height | 500 hPa |
| Accumulated Precipitation | Surface |
| Temperature | 700 hPa |

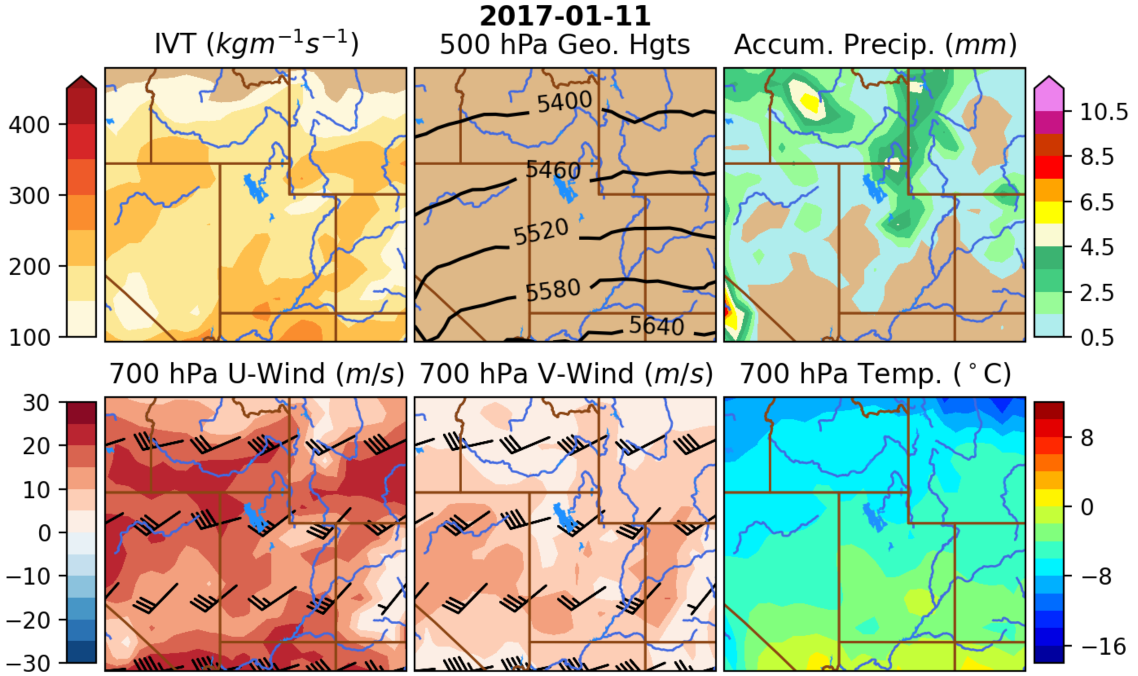
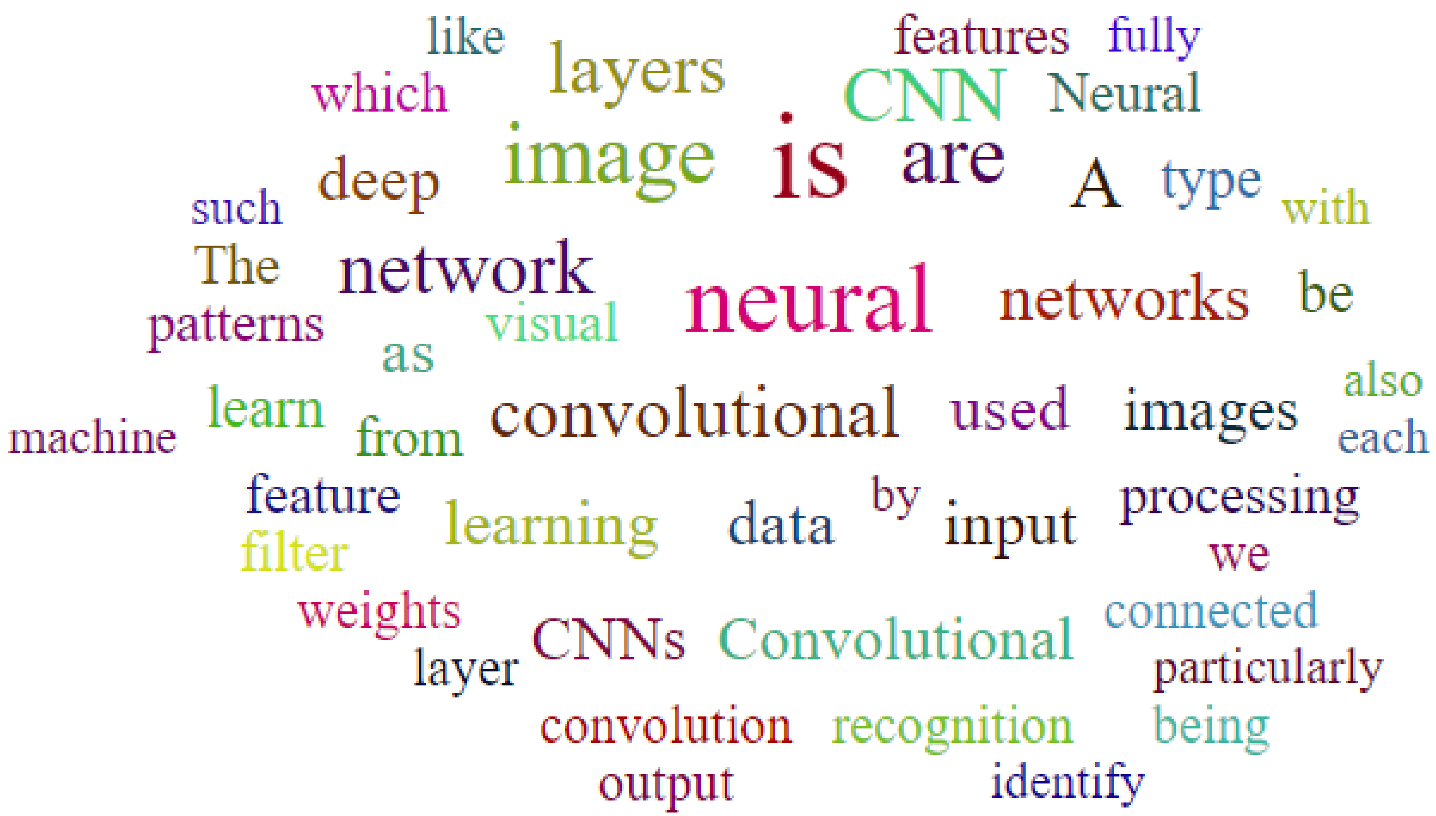
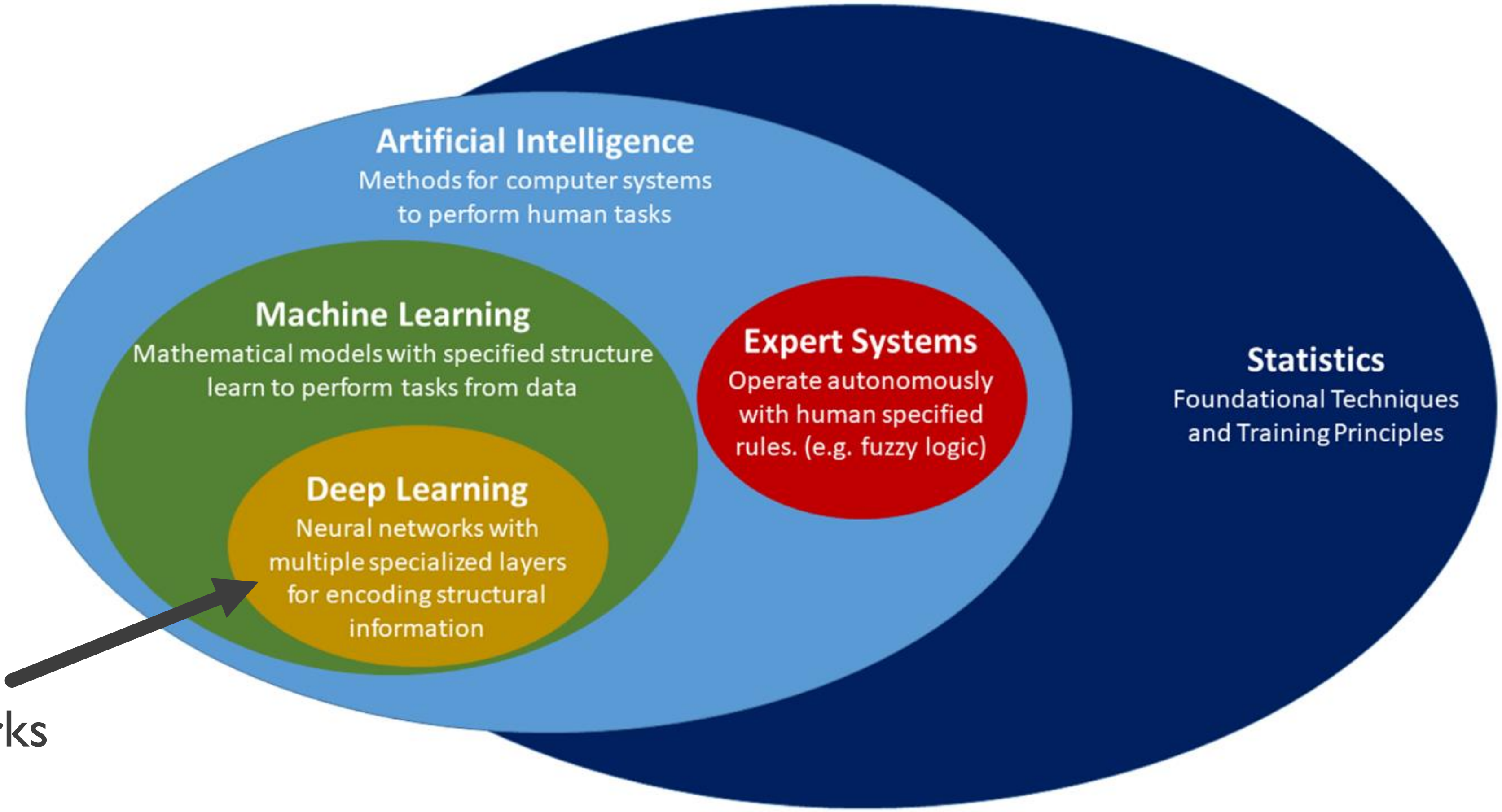



Fig. CNN input domain and variables

(Hersbach et al. 2023a; Hersbach et al. 2023b)

# What is a Convolutional Neural Network






Convolutional Neural Networks

Fig. 1. Venn diagram of the relationship between AI, ML, DL, expert systems, and statistics.

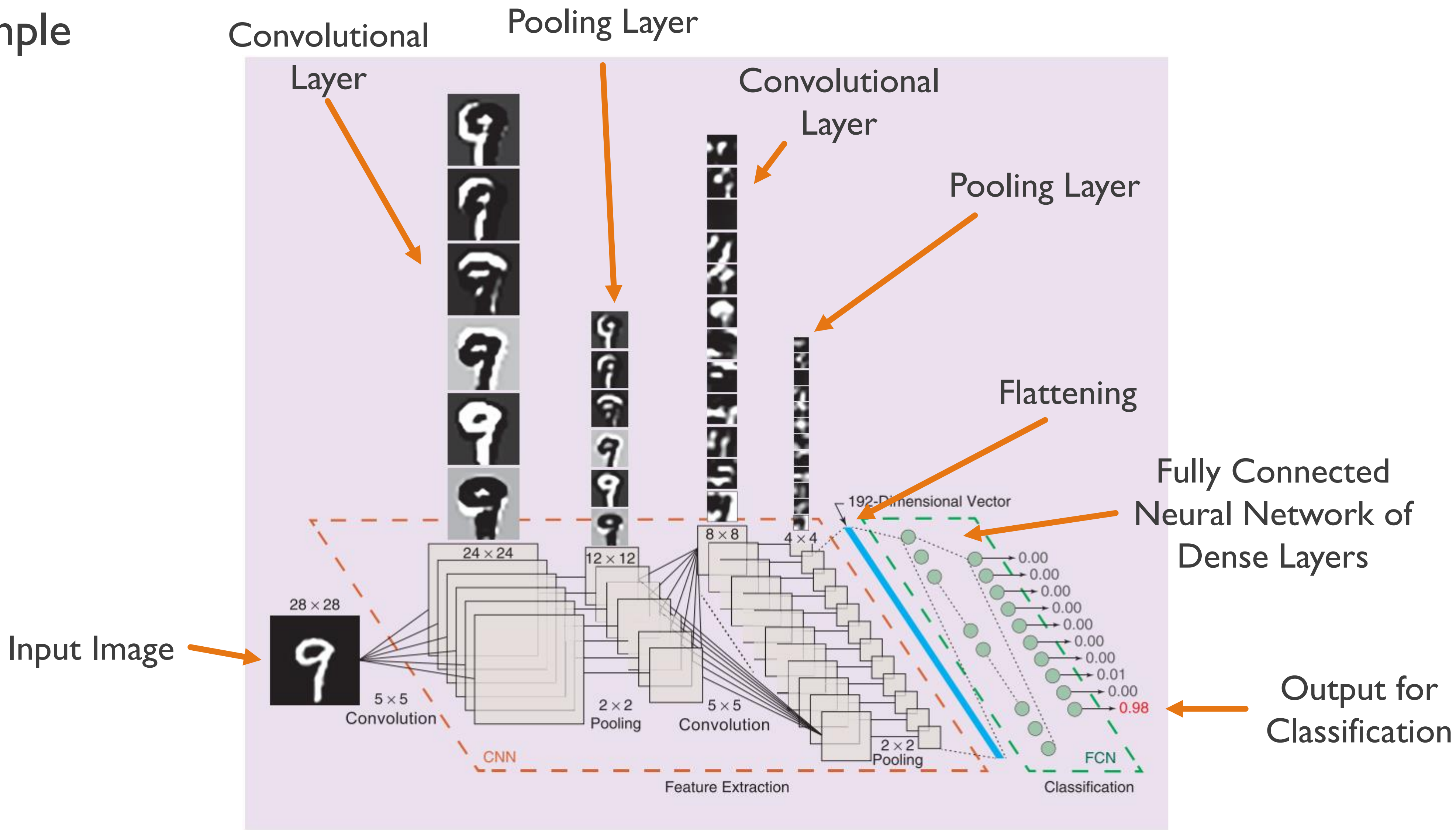
(S. E. Haupt et al. 2022)

# Really, what is it?

## Convolutional Neural Networks are…

| | | |
|---|---|---|
| ▪ A supervised learning algorithm | → | Trains on labeled datasets to classify or predict |
| ▪ A deep learning neural network | → | A neural network with more than a few layers with the functionality to learn its own parameters/weights |
| ▪ For computer vision | → | Artificial intelligence algorithm which allows a computer to interpret images or visual data |

# Handwriting Classification Example

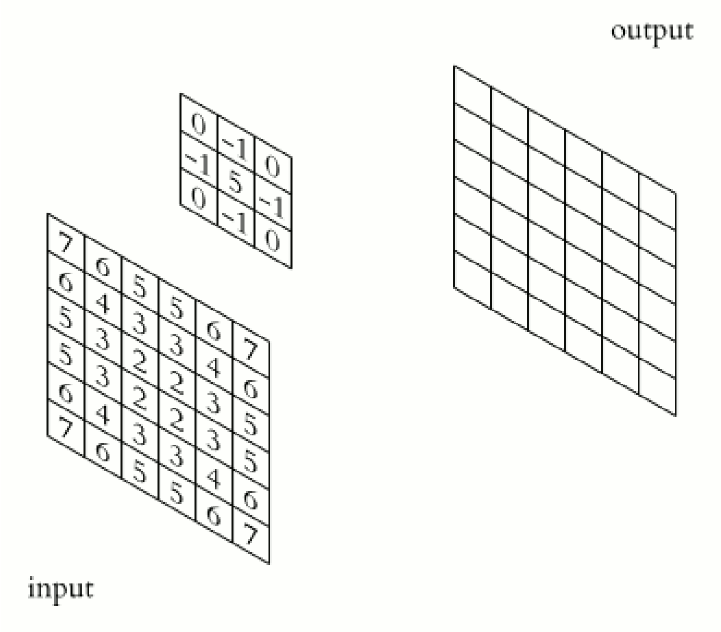



(Gonzalez 2018)

# Extracting Image Features

## Convolutional Layer

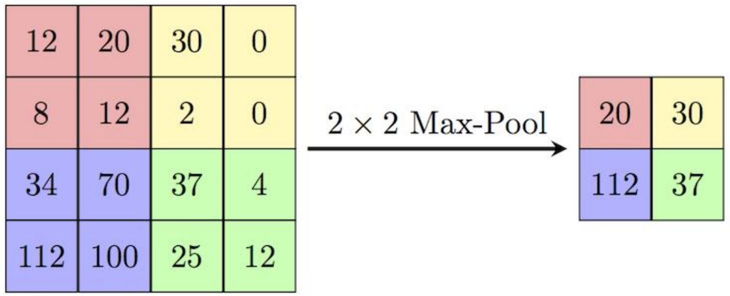
- Alters an image with a small learned filter (also known as a kernel), that highlights different patterns, edges, or values




https://en.m.wikipedia.org/wiki/File:2D_Convolution_Animation.gif

## Pooling Layers

- Reduces the resolution of an image by taking the mean or maximum values from a passing window
- Reduces processing power needed and forces the convolutional layers to evaluate consecutively larger areas

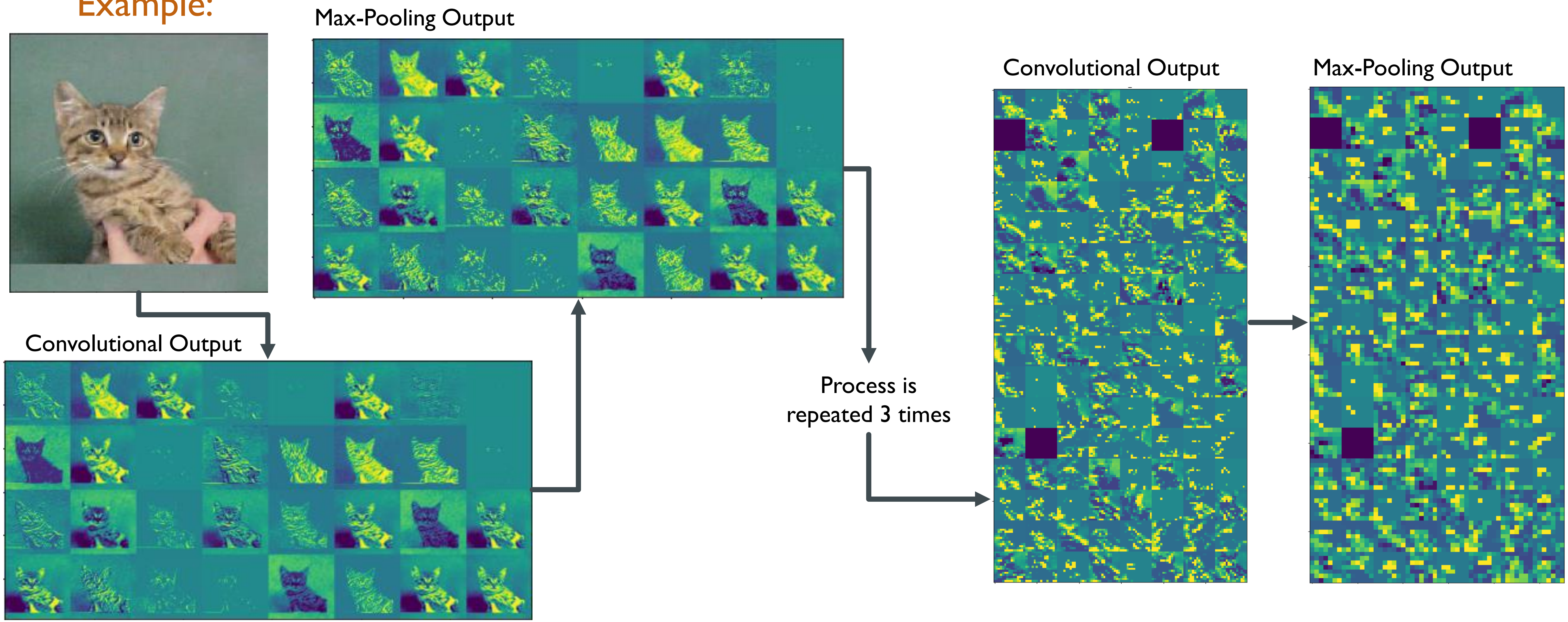



https://paperswithcode.com/method/max-pooling

(Chase et al. 2023; Gonzalez 2018; Chollet 2021)

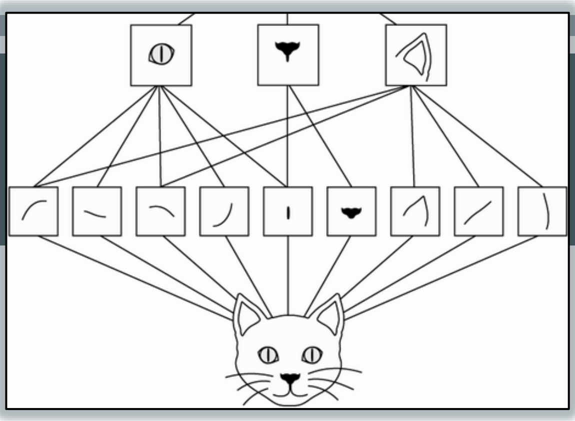
# Extracting Image Features

Example:



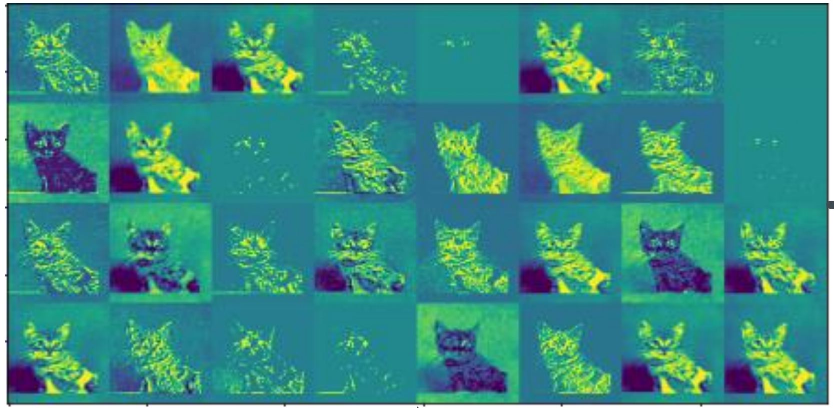

(https://www.geeksforgeeks.org/visualizing-representations-of-outputs-activations-of-each-cnn-layer/)

# Extracting Image Features

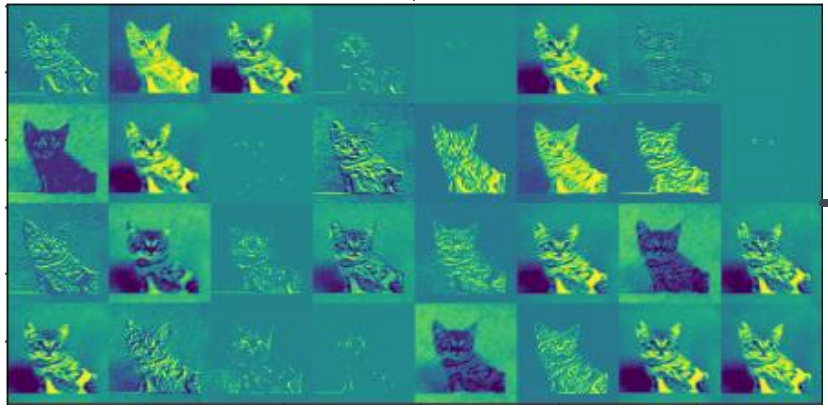


Cat image is our *Sample* or *Input*

Max-Pooling Output



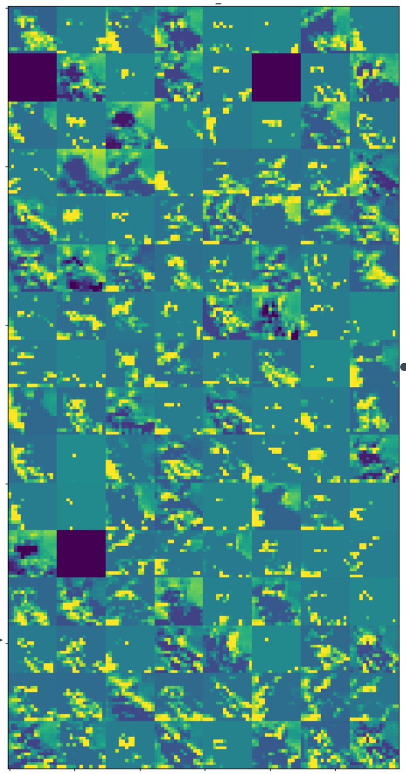
Convolutional Output



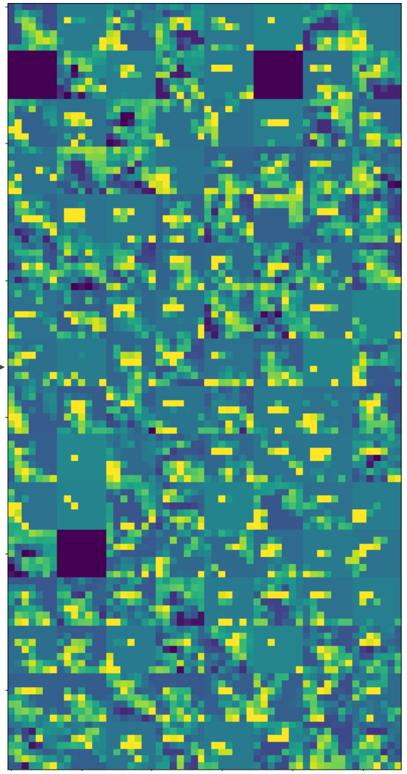
Each new image is called a *Feature Map*

Process is repeated 3 times

Convolutional Output

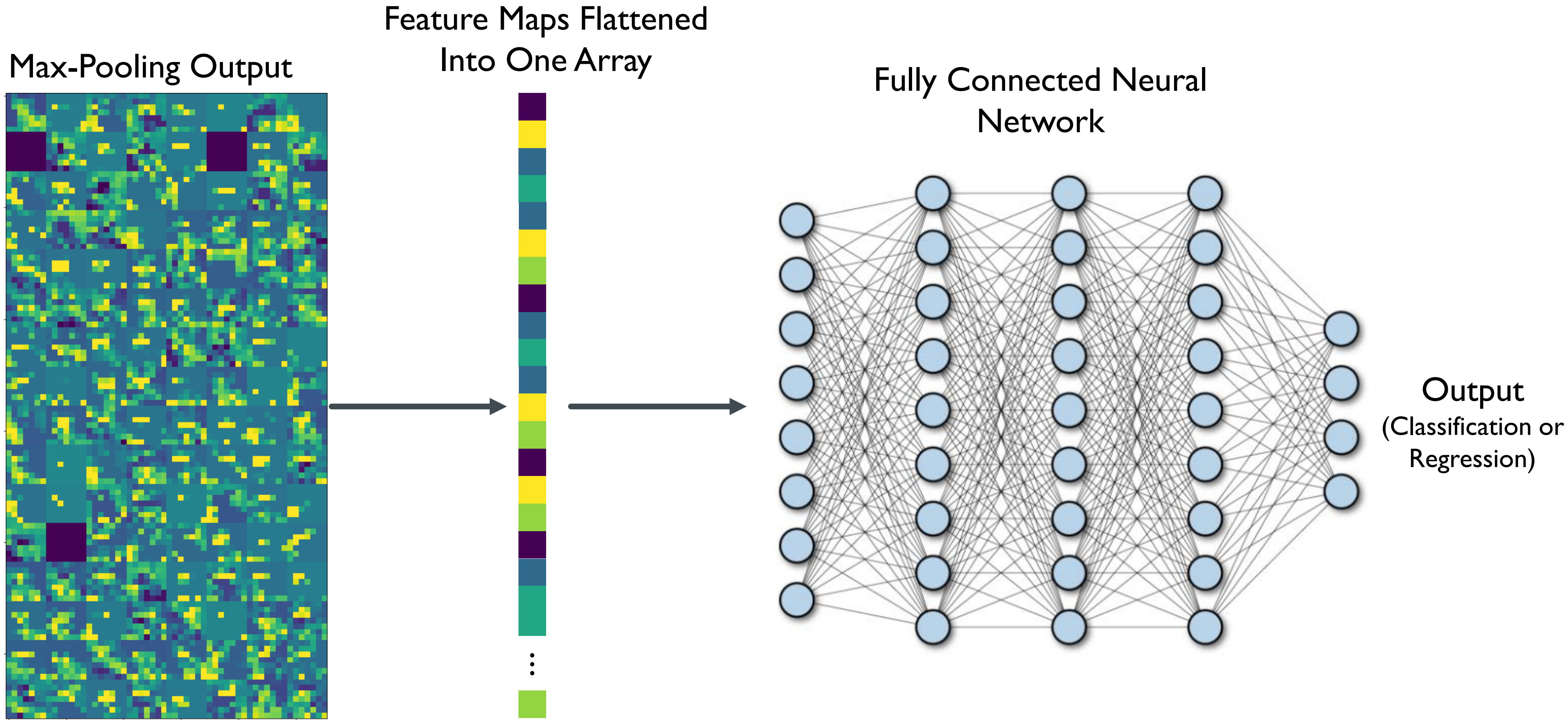


Max-Pooling Output



(https://www.geeksforgeeks.org/visualizing-representations-of-outputs-activations-of-each-cnn-layer/; Chollet 2021)

# Flattening

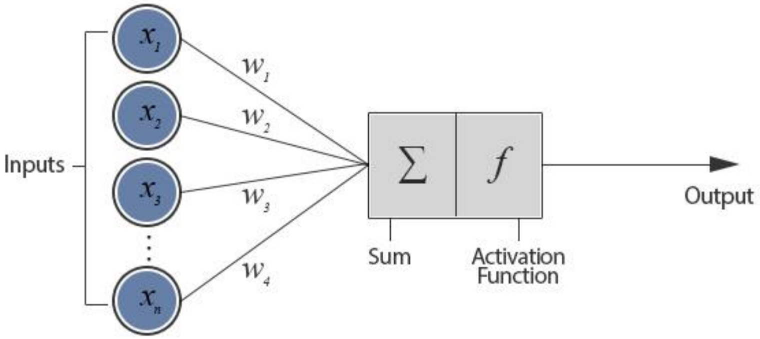



(https://www.geeksforgeeks.org/visualizing-representations-of-outputs-activations-of-each-cnn-layer/)

# Fully Connected Neural Network

## Perceptrons

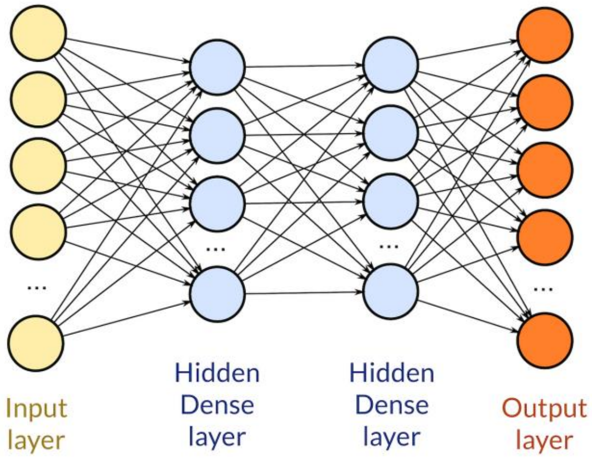
- Input data is multiplied by learned weights, aggregated, then passed through an activation function




https://sites.cc.gatech.edu/~san37/post/dlhc-fnn/

## Dense Layers

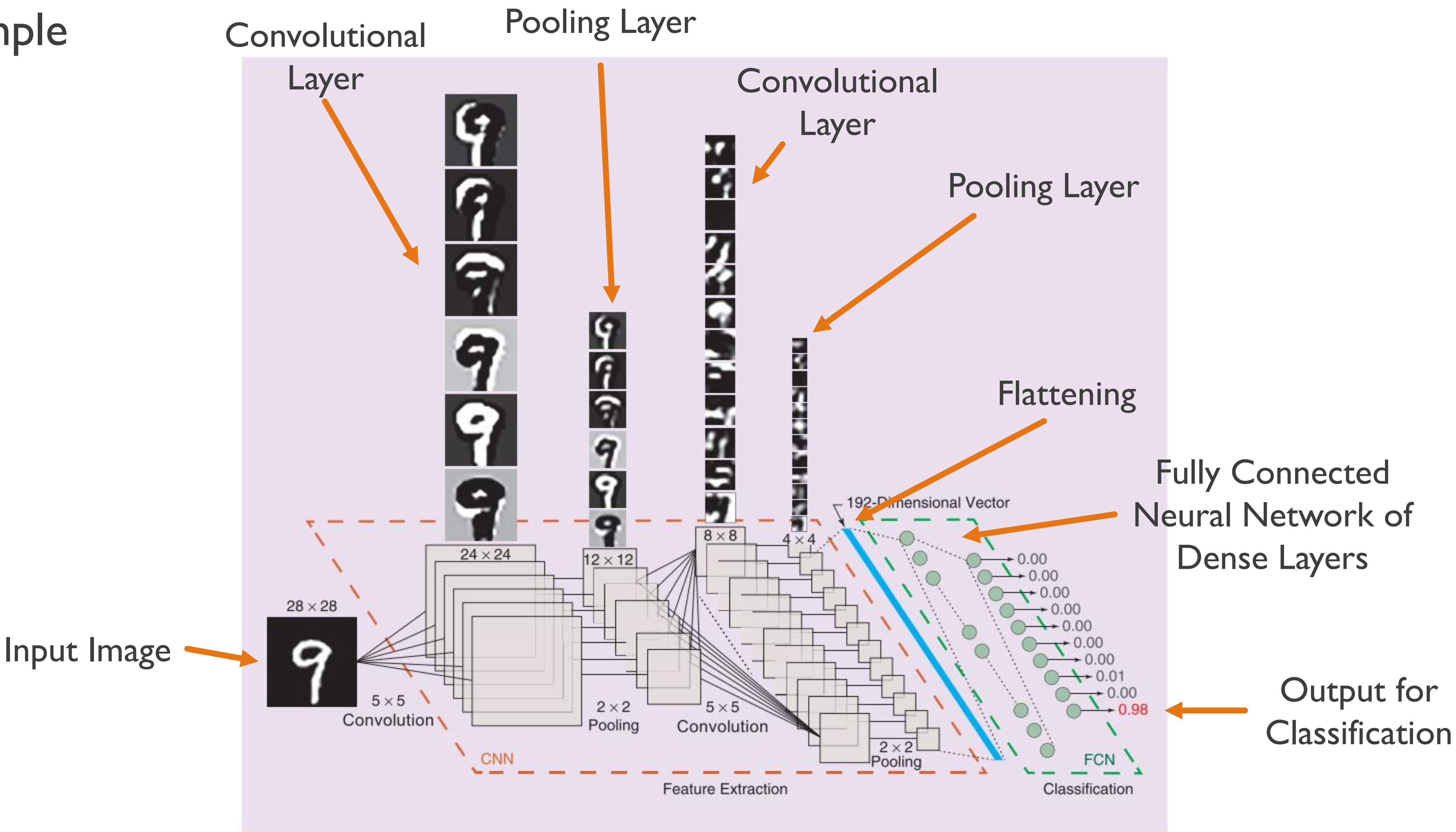
- Layers of perceptrons which transfer information from layer to layer, to the output value(s)




https://medium.datadriveninvestor.com/custom-layers-in-keras-de5f793217aa

(Chase et al. 2023; Gonzalez 2018)

# Handwriting Classification Example

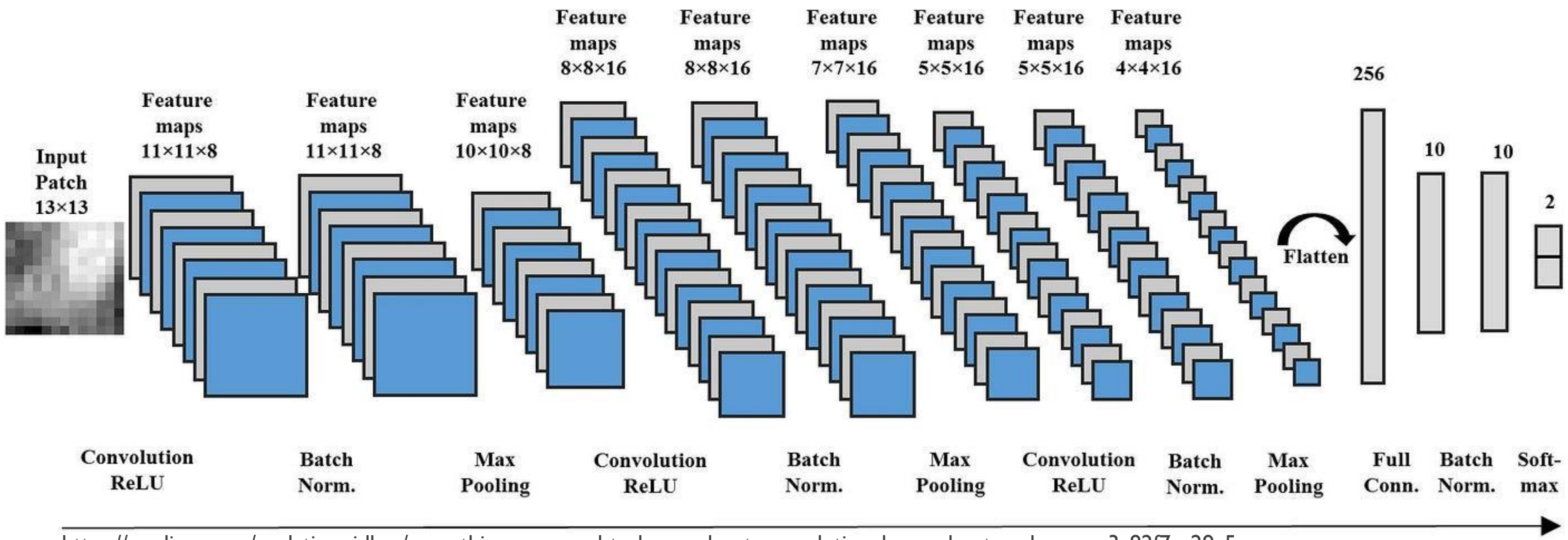



(Gonzalez 2018)

# Batch Normalization

During training, predictor data is sent in small subsets, called a batch. But these batches are not always representative of the entire dataset

- A batch normalization layer is added to normalize the feature maps between layers
- Batch normalization has been found to help smoothen and quicken training

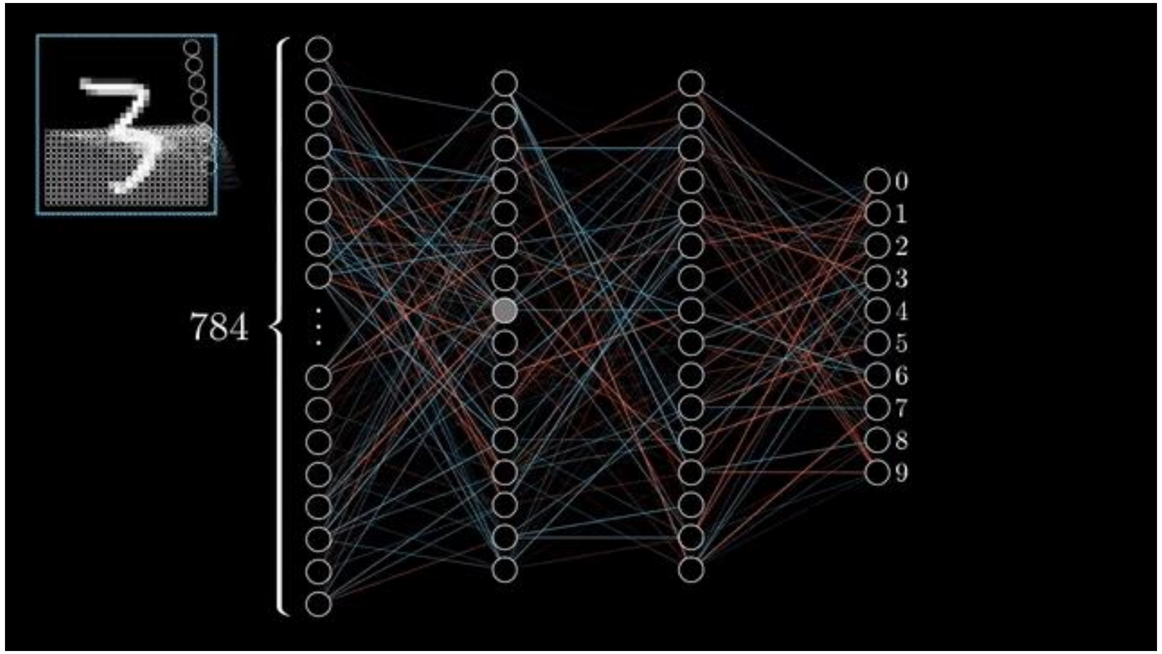



https://medium.com/analytics-vidhya/everything-you-need-to-know-about-convolutional-neural-networks-cnns-3a82f7aa29c5

(Chollet 2021; Ioffe and Szegedy 2015)

# How does the Model Learn? Formulating Error

*Prediction error* or *loss* – The measurement of the distance between your model's prediction and the target




Minimizing the loss function is how the algorithm *"learns"*

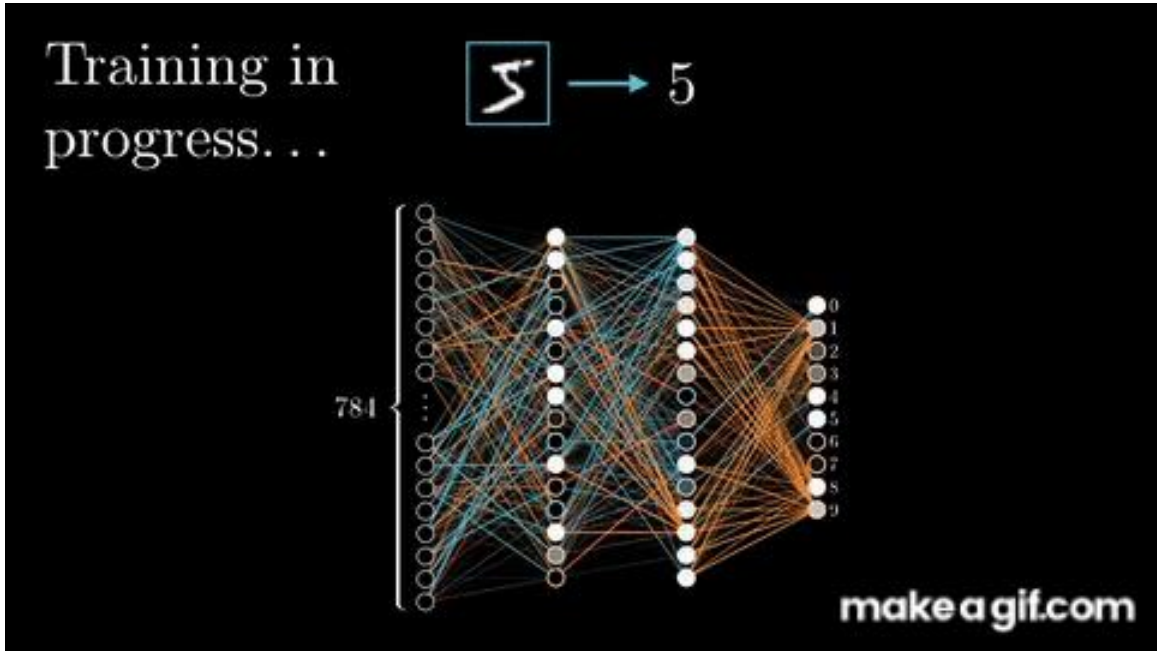
Common Loss Metrics:

| **Binary** | **Categorical** | **Numerical** |
|---|---|---|
| Binary Cross Entropy | Categorical Cross Entropy | Mean Squared Error<br><br>Mean Absolute Error |

(Chollet 2021)

# How does the Model Learn? Backpropagation

*Backpropagation* – updates the weights within the model to minimize the loss using the gradient

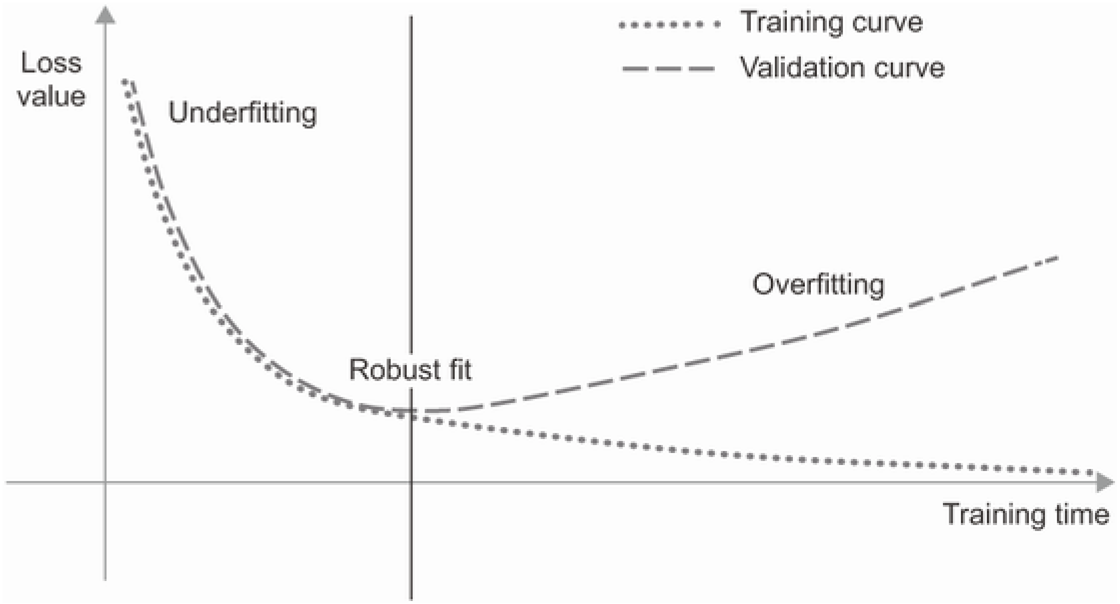


# Overfitting to the Training Subset

*Optimization* – Adjusting the model for the best performance

*Generalization* – How well the trained model performs on data it has never seen

*Overfitting* – Learning representations only specific to the training data and do not generalize to the validation data



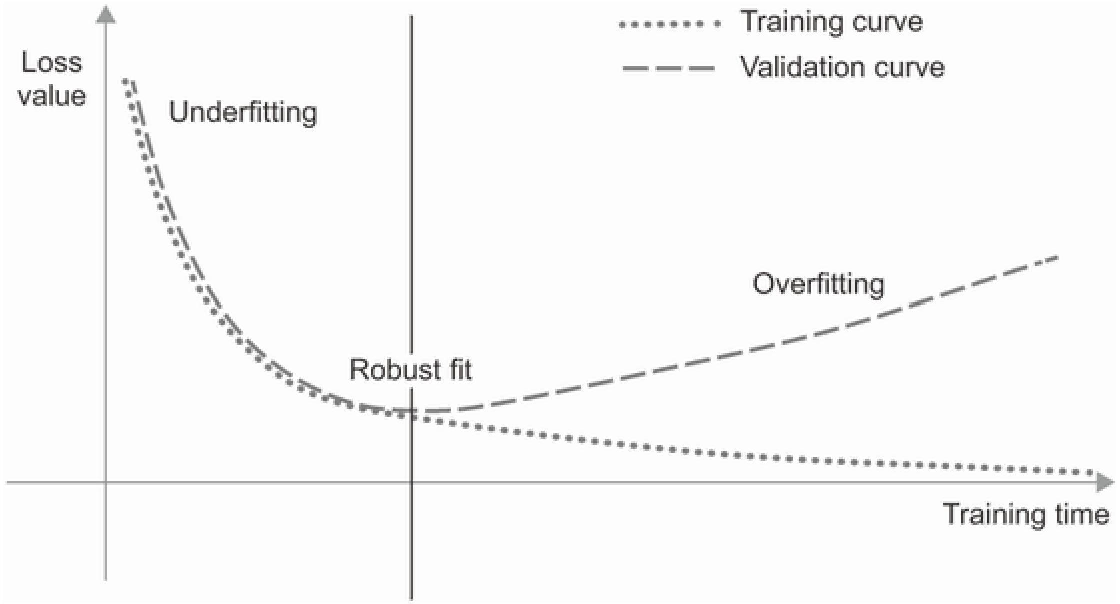

(Chollet 2021)

# Overfitting to the Training Subset

*Optimization* – Adjusting the model for the best performance

*Generalization* – How well the trained model performs on data it has never seen

*Overfitting* – Learning representations only specific to the training data and do not generalize to the validation data



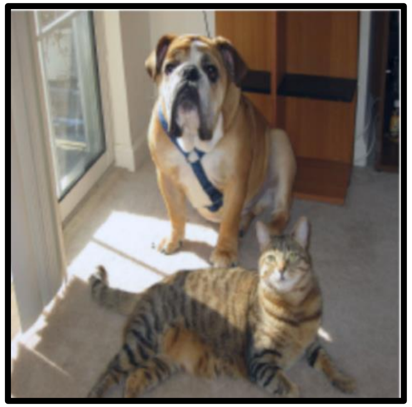

Challenges:

- Noisy training data
- Ambiguous features
- Rare features
- Spurious correlations (correlation is not causation)

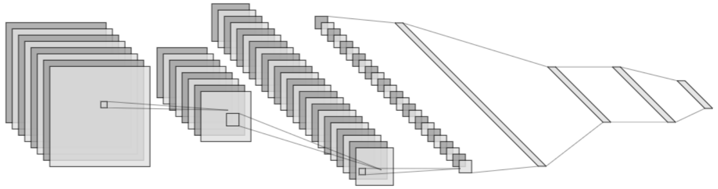
(Chollet 2021)

# Grad-CAM (Gradient-weighted Class Activation Maps)

What is in this image?



Animal Detector CNN



96% Dog
92% Cat

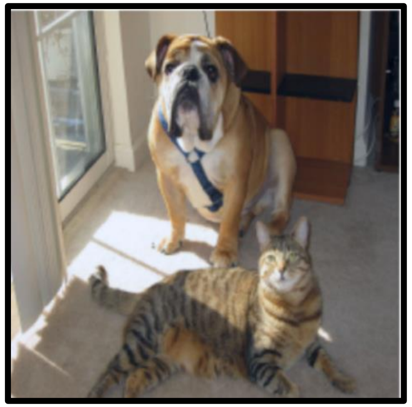
Where in the image is the CNN looking to identify the Dog and Cat?

- An Explainable AI technique to create ‘visual explanations’ for the decisions within a CNN
- Formulate heatmap from the weighted summation of gradients from an output node to all the feature maps of the last convolutional layer
- In short, Grad-CAMs highlight the grid points **where** the CNN is ‘looking’ during prediction

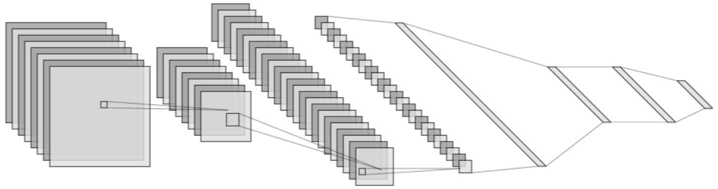
(Selvaraju et al., 2017; http://alexlenail.me/NN-SVG/LeNet.html)

# Grad-CAM (Gradient-weighted Class Activation Maps)

What is in this image?



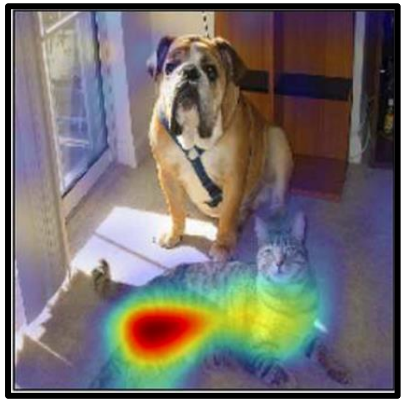
Animal Detector CNN



96% Dog
92% Cat

Where in the image is the CNN looking to identify the Dog and Cat?

Grad-CAM:
Weighted summation of the gradients from an output node to all feature maps



Grad-CAM for 'Cat'

Grad-CAM for 'Dog'

Strong Gradients = High Importance

(Selvaraju et al., 2017; http://alexlenail.me/NN-SVG/LeNet.html)

# References


Black, A. W., and T. L. Mote, 2015: Characteristics of winter-precipitation-related transportation fatalities in the United States. *Wea. Climate Soc.*, 7, 133–145.

Bohne, L., C. Strong, and W. J. Steenburgh, 2020: Climatology of orographic precipitation gradients in the contiguous western United States. *J. Hydrometeor.*, 21, 1723–1740.

Burakowski, E., and M. Magnusson, 2012: *Climate Impacts on Winter Tourism Economy in the United States*. 36 pp.

Chase, R. J., D. R. Harrison, G. M. Lackmann, and A. McGovern, 2023: A Machine Learning Tutorial for Operational Meteorology. Part II: Neural Networks and Deep Learning. *Wea. Forecasting*, **38**, 1271–1293, https://doi.org/10.1175/WAF-D-22-0187.1.

Chollet, F., and Coauthors, 2015: Keras. https://keras.io.

Chollet, F., 2021: Deep Learning with Python. Manning Publications Co., 478 pp.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor. Climatol.*, **33**, 140–158.

Gonzalez, R. C., 2020: Deep Convolutional Neural Networks [Lecture Notes]. *IEEE Signal Processing Magazine*, **35**(6), 79-87, https://doi.org/10.1109/MSP.2018.2842646.

Haupt, S. E., and Coauthors, 2022: The History and Practice of AI in the Environmental Sciences. *Bull. Amer. Meteor. Soc.*, **103**, E1351–E1370, https://doi.org/10.1175/BAMS-D-20-0234.1.

Hersbach, H., and Coauthors, 2023a: ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed 18 January 2023, doi: 10.24381/cds.bd0915c6.476

Hersbach, H., and Coauthors, 2023b: ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed 28 January 2023, doi: 10.24381/cds.adbb2d47.

Ioffe, S., and C. Szegedy, 2015: Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proc. 32nd Int. Conf. on Machine Learning*, https://doi.org/10.48550/arXiv.1502.03167.

Lewis, W. R., W. J. Steenburgh, T. I. Alcott, and J. J. Rutz, 2017: GEFS precipitation forecasts and the implications of statistical downscaling over the western United States. *Wea. Forecasting*, **32**, 1007–1028.

Li, D., M. L. Wrzesien, M. Durand, J. Adam, and D. P. Lettenmaier, 2017: How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, **44**, 6163–6172.

Selvaraju, R. R., M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, 2020: Grad-cam: Visual explanations from deep networks via gradient-based localization. *Int. J. of Comput. Vis.*, **128**, 336-359, https://doi.org/10.1007/s11263-019-01228-7.

Westerling, A. L., 2016: Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Phil. Trans. R. Soc. B*, **371**, http://doi.org/10.1098/rstb.2015.0178.