Whale Classification via Siamese Neural Network

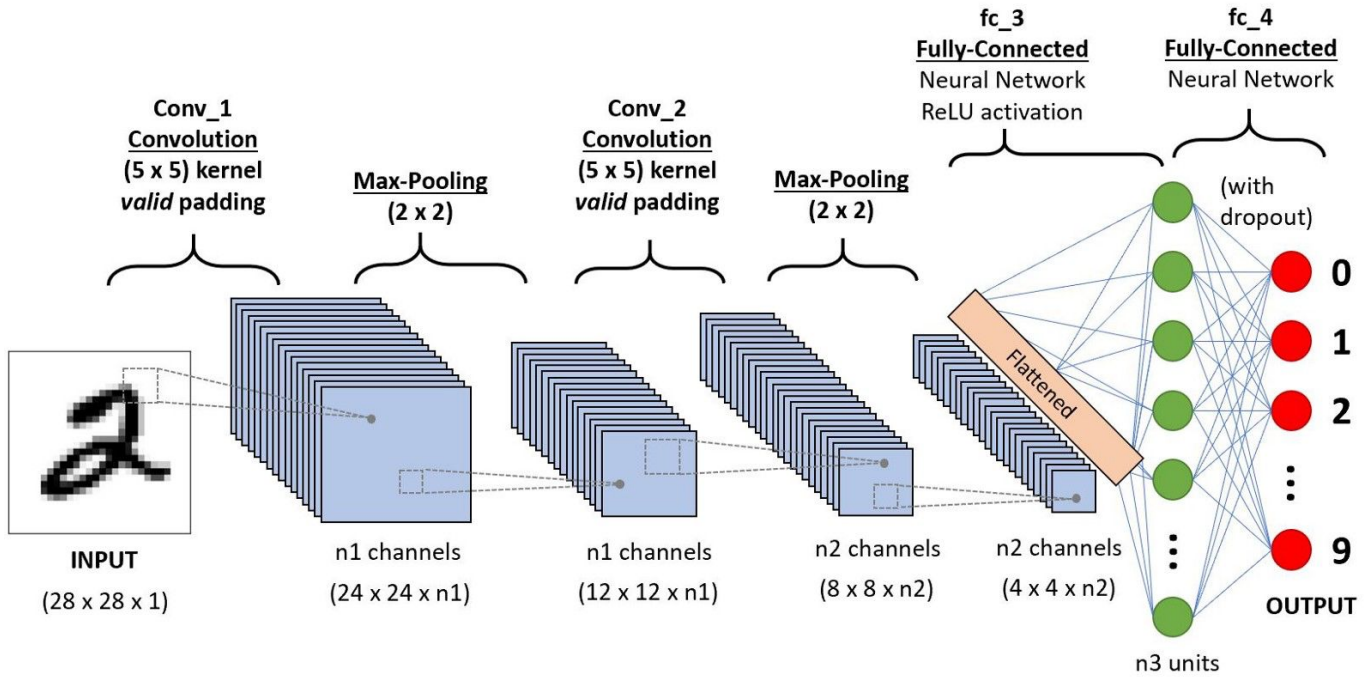
Matthew Wirtz

Overview

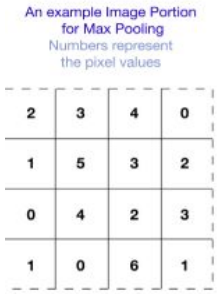
- Given a picture of a whale fluke, determine if it is a humpback whale
- This model achieves 0.78563
- Flow: preprocessing, Siamese neural network, training data construction, training procedure, and results

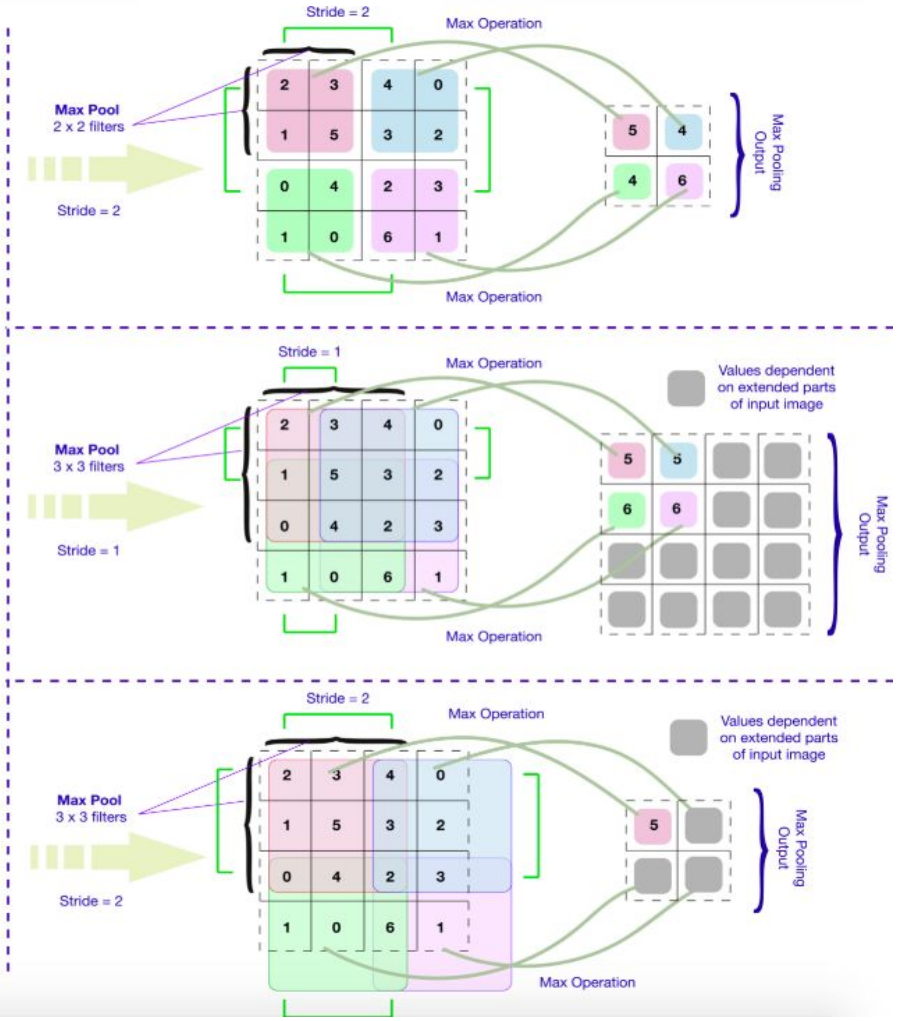
Terminology

- Bounding Box Model
- Batch Normalization
- Convolutional Neural Network (CNN)
- Max Pooling
- Overfitting
- Reshaping
- Siamese Neural Network



-1	-1	-1
-1	8	-1
-1	-1	-1





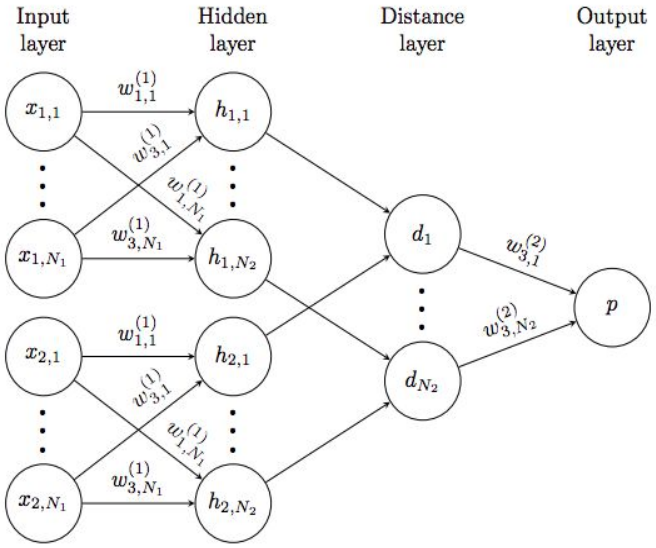
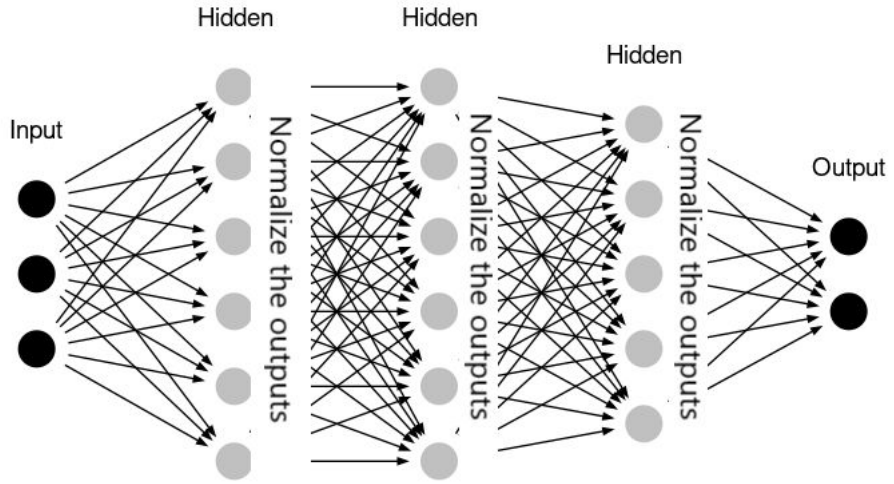


Figure 3.1: A simple 2 hidden layer siamese network for binary classification with logistic prediction p. The structure of the network is replicated across the top and bottom sections to form twin networks, with shared weight matrices at each layer.



References

- https://www.kaggle.com/martinpiotte/whale-recognition-model-with-score-0-78563/notebook
- https://towardsdatascience.com/a-beginners-guide-to-convolutional-neural-networks-cnns-14649dbddce8
- https://towardsdatascience.com/batch-normalization-in-neural-networks-1ac91516821c
- https://www.quora.com/What-is-max-pooling-in-convolutional-neural-networks
- https://ai.stackexchange.com/questions/6875/what-is-the-purpose-of-reshaping-it-into-the-shape-the-network-expects-and-scal
- https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/
- https://ai.stackexchange.com/questions/10177/whats-the-role-of-bounding-boxes-in-object-detection/10186
- http://www.cs.utoronto.ca/~gkoch/files/msc-thesis.pdf
- https://medium.com/@ilango100/batch-normalization-speed-up-neural-network-training-245e39a62f85
- https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53