**Deep learning for cellular image analysis**

Deep learning: a set of machine-learning methods—specifically, neural networks—that are capable of learning representations from data with increasing levels of abstraction. To illustrate the workfow for training a deep learning model in a supervised manner, here we consider the case of training a linear classifer to recognize grayscale images of cats and dogs. Each image is an array of size (Nx, Ny, 1), where Nx and Ny are the number of pixels in the x and y dimensions, respectively, and 1 is the number of channels in the image. For this exercise, we collapse the image into a vector of size (NxNy, 1). The classification task is to construct a function that takes this vector as input and predicts a label (0 for cats, 1 for dogs). Mathematically, this is written as below

Diagram

Description automatically generated

A linear classifier performs this task by producing class scores that are a linear function of each pixel value. where **y0 and y1** are the class scores, **W** is a matrix of class weights, and **x** is the image vector. The class with the highest score is the predicted class.

We have 4 main biological applications for deep learning namely image classification, image segmentation, object tracking, augmented microscopy. **Image classification** deals with assigning exact labels to images such as differentiation of images of dogs and cats. Image classifiers have also been used to identify changes in cell state: in a recent study, scientists used a fluorescent marker of differentiation to establish a ground truth and then trained a classifier to identify differentiated cells directly from brightfield images. Deep learning has also been used to classify spatial patterns in fluorescence images and to determine protein localization in large datasets from yeast and humans. These studies highlight the fact that deep learning is an accessible tool that can help biologists understand their imaging data. **Image segmentation** is the task of partitioning an image into several parts to identify meaningful objects or features. One specific biological example is the need to identify single cells in microscope images. **Object tracking** is the task of following objects through a series of time-lapse images. One example of a biological application of this is the tracking of single cells in live-cell imaging movies. During a typical movie, cells can move from one side of the imaging chamber to the other. **Augmented microscopy** is the extraction of latent information from biological images, such as the identification of the locations of cellular nuclei in bright-field images.

The barriers to spreading deep learning throughout biology labs are both cultural and technical. The mathematics renders some of the inner workings of deep learning algorithms opaque; the unique requirements of deep learning necessitate a different way of thinking about writing software. Specifically, the need for annotated data. Although the application of deep learning to biological image analysis is still in its early days, there has already been remarkable progress in adapting deep learning to biological discovery.