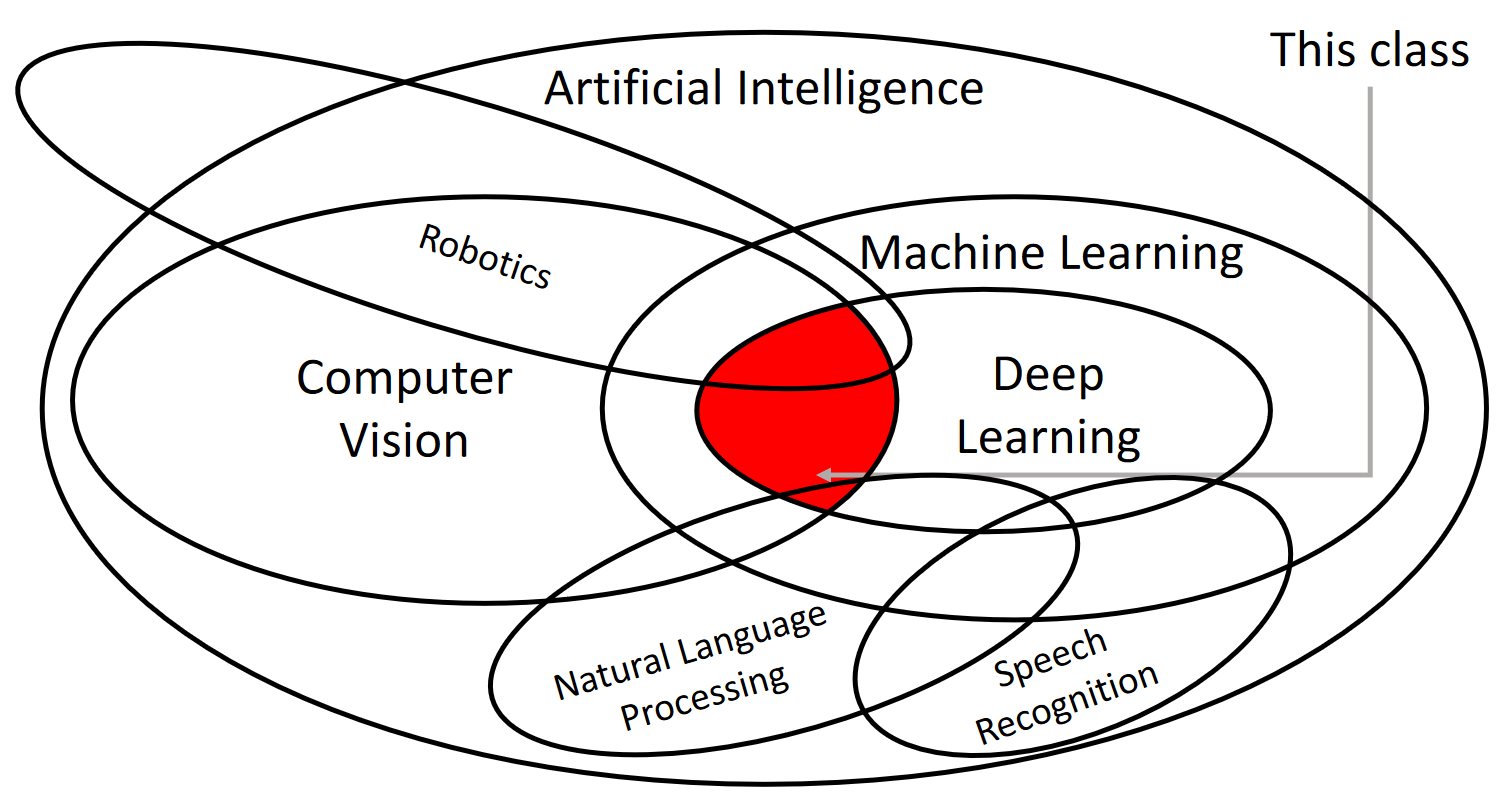
Deep Learning for Computer Vision

**Computer Vision** works towards building artificial systems that process, perceive, and reason about visual data.

**Learning** for artificial systems is the process of using data and experience to learn patterns.

**Deep Learning** is a learning process that uses a hierarchy of algorithms with many layers.



## Lessons and Advancements in Computer Vision History

* Edges are important to vision as evidenced by increased brain activity (Hubel and Wiesel, 1959)
* First use of edges as input features (Roberts, 1963)
* Study of stages of visual representation (Marr, 1970)
* Recognition of human movement based on position of joints (Brooks and Binford, 1979; Fischler and Elshlager, 1973)
* Recognition of objects using edges (Canny, 1986; Lowe, 1987)
* Recognition of objects by grouping common areas (Shi and Malik, 1997)
* Recognition of objects despite scale variances (Lowe, 1999)
* Face detection (Viola and Jones, 2001)
* Start of PASCAL Visual Object Challenge – Segmentation of specific classes of objects (2001)
* Start of IMAGENET Large Scale Visual Recognition Challenge – Image classification with 1,000 classes (2009)
* Development of AlexNet – Start of Deep Learning (Krizhevsky et al., 2012)

The creation of AlexNet was actual the result of a large number of advancements in the field of Machine Learning, which takes us back in history again.

* AlexNet uses the concept of a Perceptron (Rosenblatt, 1957), a hardware implementation of what is now known as a linear classifier. The Perceptron was able to recognize letters of the alphabet.
* The delay in the use of Perceptron was caused by proof that they could not be used as non-linear classifiers (Minsky and Papert, 1969).
* The issue was theoretically resolved by the Neocognitron (Fukushima, 1980), but the model could not be practically trained due to hardware limitations. AlexNet is very similar to this model.
* AlexNet was able to get past the hardware limitation issue due to the introduction of backpropagation (Rumelhart, Hinton, and Williams, 1986).
* The first use of backpropagation to a Neocognitron-like architecture was with Convolutional Neural Networks (LeCun et al., 1998). This architecture was able to recognize handwritten digits.
* The use of deeper networks started (Hinton and Salakhutdinov, 2006; Bengio et al., 2007; Lee et al., 2009; Glorot and Bengio, 2010).
* Explosion of deep learning research (2012 - )

The current explosion of deep learning research is mostly due to a combination of three factors, algorithms, data, and computation power. We have had the algorithms, or at least the starting of them, for several decades now. The availability of data only began when digital cameras became commonplace. Only recently has the computational resources required to train the algorithms on enough data for them to learn properly become cheap enough to be widely accessible.