Image processing

So now that we have an image, what can we do with it? You have probably played with Photoshop, or some other image manipulation program such as GIMP, and you know that there are thousands of operations, filters, changes, and tricks you can play on images. What are a few of them? We can make an image brighter or darker by adjusting the brightness.
We can increase the contrast between the white parts of the image and the dark parts. We can make an image blurry, usually by applying a Gaussian blur filter. We can also make an image sharper (somewhat) by using a filter such as unsharp mask. You may have also tried an **edge detector** filter, such as the Canny filter, to isolate the edges of an image, where color or value changes. We will be using all of these techniques to help the computer identify images.

What on earth is **unsharp mask**? This image enhancing process was invented in the 1930s in Germany, and was originally used with film. The original image is blurred slightly, and then inverted (black becomes white and vice-versa). This positive image is used as a mask to block where the negative image is combined. The net result is to amply the high frequency (in other words, fine detail) in the image. The effect in Photoshop or GIMP uses a threshold in the difference of the blurred and normal images to simulate the masking process.

What we are trying a achieve with all this manipulation for the computer is to not have the computer be sensitive to the size of the image – which is called **scale invariance**, the angle at which the photograph was taken – **angle invariance**, and the lighting available – illumination invariance. This is all very desirable in a computer vision system – we would not want an AI system that only recognized our toys from exactly the same angle and distance as the original image. Remember that we are going to train our vision system to recognize toys based on a training set of images we take in advance, and the robot will have to recognize objects based on what it learned from the training set.

What we are going to do is use features about the image that mostly don't change based on size, angle, distance, or lighting. What sorts of features might that be?

If we look at a common household object, such as a chair, and inspect if from several angles, what about the chair does not change? The easy answer is the edges and corners. The chair has the same number of corners all of the time, and we can see a consistent number of them from most angles. It also has a consistent number of edges.

Admittedly, that is a bit of an oversimplification of the approach. We will be training our ANN to example a whole host of image features that may or may not be unique to this object, and let it decide which ones work and which do not. We will accomplish this by means of a generic approach to image manipulation called **convolution**.