There are multiple levels of image--ground truth connection. Look at our midterm report. We are after the atom number variance. Take 100 images of the same cloud, obtain the density distribution -- number of atoms in each pixel. Pick a region --superpixel--such that it contains reasonable number of atoms (Fig.2 left). Compare the atom number, NA, in this superpixel between the 100 images taken. Values of NA will not be the same if we believe in physics. There will be a mean and a variance of the mean. The ratio of the two determines the class, hot and cold. We plot the ratio for each superpixel as in Fig. 3.

Go back to high level: a raw image I2 in Fig.1. It's an image of a laser beam partially absorbed by atoms. The information (variance vs mean of atom number) is there but it's greatly obscured by other contributions (noises and the corresponding variances). Those contributions are 1. spatial variation of light in the laser beam (I1 has a pattern -- rings, it is not uniform). 2. Light has its own variance: take 100 images of the same light beam and all images would be different. 3. Technical noise of the camera: take 100 images with the camera blocked by a cap and all images would be different. Thus, to reveal the ground truth, we need to process the raw images to remove the above contributions 1-3.

For deep learning, we start with images as shown in the left panel in Fig.2. At this pre-processing stage, the contribution 1 from light beam patterns is removed. Images represent variable OD as in Eq.1 and Eq.2.

If we take the variance for those images, that would give the left-hand-side of Eq.2 and contain 4 contributions as in the right-hand-side of the equation. The last term is the information we are after. The first three terms got in there in the process of making the actual measurement in the lab and carry information about the laser beam and a camera we used in the measurement. The contributions from the "obscuring" first three terms and the last "information" term are about 50-50. The "obscuring" terms are the same for both cold and hot samples as they come from the measurement setup, while the fourth term is different as it depends on the sample class. We hope that CNN can "see" that difference.

There are 187 cold images and 145 hot. You might want to compute the mean (for each pixel) for each set and subtract it from every image of the corresponding set. That would produce images with zero mean and preserve the variance we are after. The two means, one for cold, one for hot are slightly different and we don't the machine to trivially classify based on hot pictures being a bit brighter than cold.