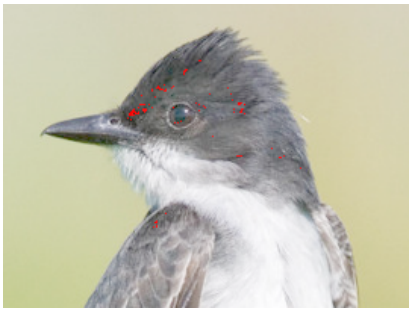
Boolean Point Features

by Sven Nilsen, 2018

In this paper I present a concept for continuously learning features from high dimensional data sets.



Boolean point features used to detect a bird eye. Notice that there are just a few points inside the eye.

Assume you have a dataset of images where each image is labeled with an image segment of some object to be recognized. A boolean point feature has the following properties:

- A boolean point feature is a function of the local neighborhood around a pixel
- At least one image segment in the dataset contains the boolean point feature
- The ideal fitness is appearing once for each label in the dataset within an image segment, but with few redundant pixels within and with extremal false positives outside image segments
- Ideal fitness implies that the boolean point feature is unique and identifies the label
- Attention can be directed by taking the union of boolean point features
- Structure can be learned by looking at internal relations between boolean point features
- Boolean point features are independent of each other and can be trained in parallel

New boolean point features are learned by applying the same learning algorithm over and over to a data set. Each time, the algorithm gets stuck in a local minima when trying to optimize the fitness, but that is the whole point: The algorithm tries to find a function which results in fewest possible points within some labeled region of the data, but appearing at least once. Whatever the local minima is found under training, there is something to gain from it. Outside the region it is possible to detect a false positive when there are too many of the boolean point feature clustered together.

For example, a tiger might be easily recognizable when moving in the open, but difficult to recognize when hiding in the grass. In the open, the stripes identifies the tiger, but in the grass, this often leads to a false positive.

If you have two boolean point features A and B for an object, then their union tells where to look:

A ∨ B This is a set that includes the points of both features A and B

This means that an agent can concentrate on a few pixels in an image: Those that are likely to identify the object on closer inspection. An object is more likely to be detected if there are many associated boolean point features near each other but few of each kind.