This paper provides an approach to learn the receptive fields with additional parameters. The goal is to increase selected features for fast learning.

The raw image is extracted into local image patches. And those patches are encodes to corresponding activation values. A pooled feature contains the spatial information and a code index. Pooling reduces the dimensionality of features and also keeps the spatial information for each feature. In the paper, receptive fields consists of rectangular regions because it worked empirically well. But I think the paper should give more explanation on how to select the shape of spatial regions, because, as what they claim, the receptive regions are dataset driven. The receptive field learning selects the pooled features, by multiple-class linear classification. To retrain the model, by adding one more feature on the previous model for each iteration, and we can find the one that gives the largest boosting. And as the retaining rates drops under certain threshold we can get enough selected features. Of course more features are preferred to improve the accuracy, but it will sacrifice the efficiency.

The method shows it beats the random selection and pyramid pooling. The interesting part is that random selection has fairly good accuracy, especially when the feature number is large. So the feature selection will be significant when feature number is comparatively small. It does help to avoid the overfitting and give higher performance.