An Unsupervised Method for Sketch Recognition Using Jigsaw Framework

Emre Karaman and T. Metin Sezgin
Koç University, Istanbul, Turkey
{emrekaraman14, mtsezgin}@ku.edu.tr

Abstract

The Jigsaw Model is a patch-based model which learns shape, size and appearance of patches automatically from the repeated structures without supervision. The model extracts irregularly sized and shaped patches from a latent image called jigsaw. In this project we propose how to apply the jigsaw model into sketch recognition domain to classify the repeated strokes in a sketch scene without supervision. The results show that our newly model is able to learn and classify hand-crafted repeated structures in a sketch scene.

Method

The Jigsaw Model

Jigsaw Model [1] learns the repeated structures by running an EM algorithm iteratively. The model assigns each pixel of a training image to a jigsaw pixel by an offset map which enforces adjacent pixels in the training image to be adjacent in the jigsaw. Finally, the training image is reconstructed from the jigsaw image by using the offset map.

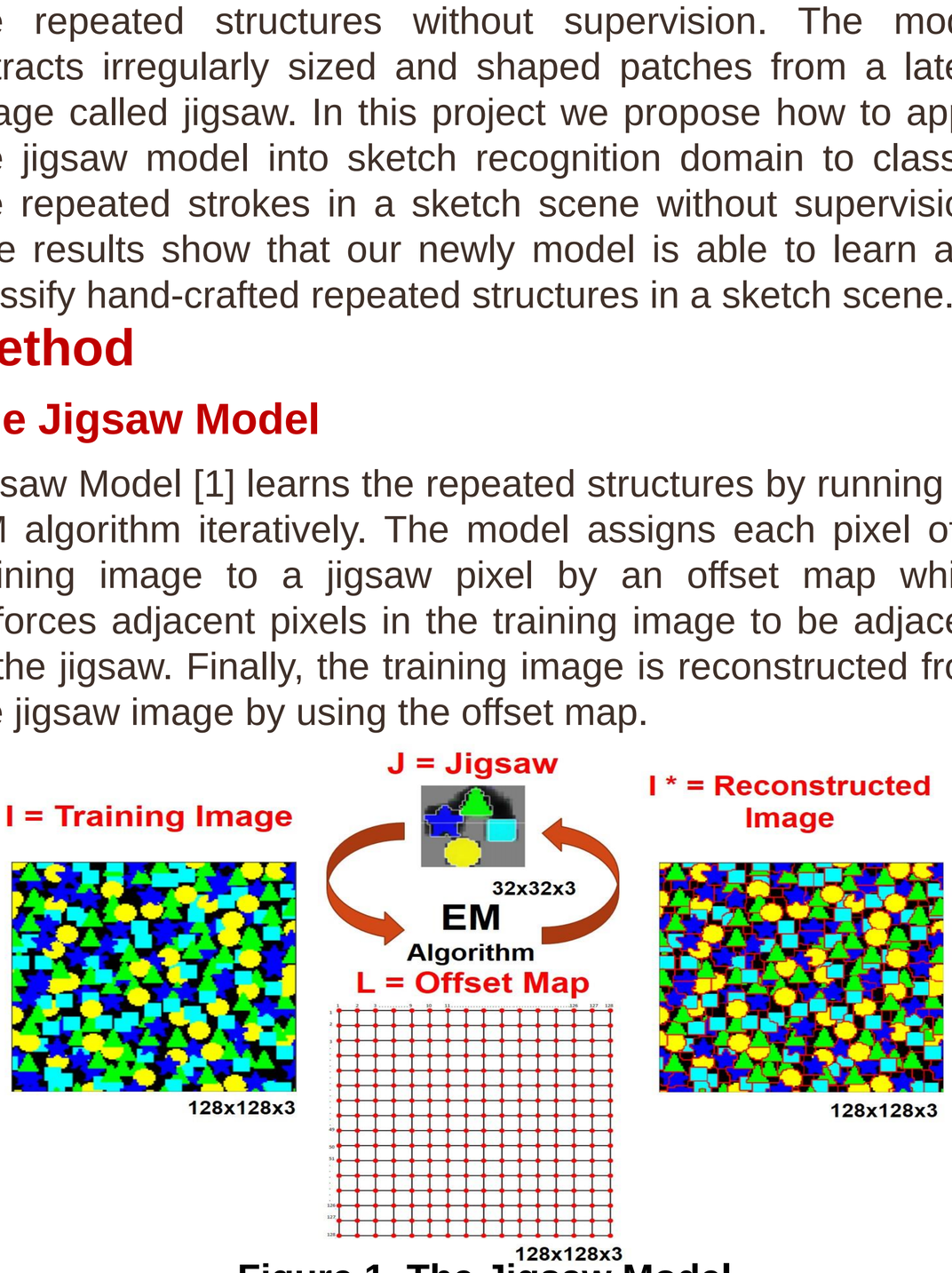

Figure 1. The Jigsaw Model

The Jigsaw Model for Sketch Recognition

In the jigsaw model [1] each pixel value is represented by RGB intensity values. In our model pixels correspond to strokes and the strokes are represented as Image Deformation Model (IDM) features [2] in a 1x720 feature vector.

The jigsaw model is founded on a 4 connected grid. However strokes in a sketch scene do not lie on a clear-cut grid. Thus we define a method to create connection between similar strokes. Additionally, we reposition strokes to make the relative positions of similar stroke pairs same.

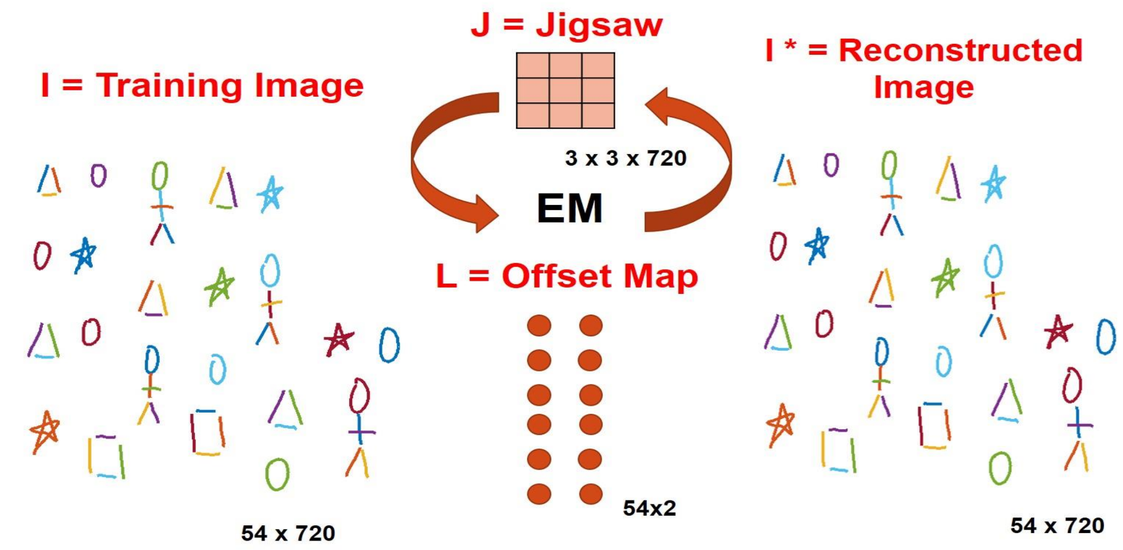

Figure 2. The Unsupervised Model for Sketch Recognition

Results

In the jigsaw model the hyperparameters especially interaction potential between adjacent pixels plays an important role to reach a consistent jigsaw image. The test results with different interaction potential values are shown in Figure 3.

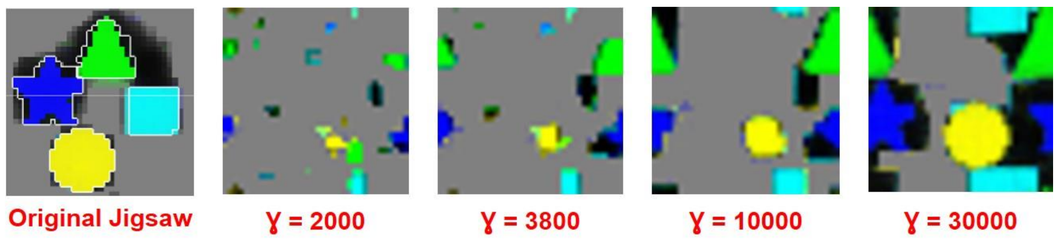

Figure 3. Jigsaw images with different parameters.

We tested our model with a hand-crafted training image which contains repeated structures. Our model creates a connection between similar stroke pairs and adjust their origin to have the same displacement value. Thus, we can learn repeated structures and classify them into the same group.

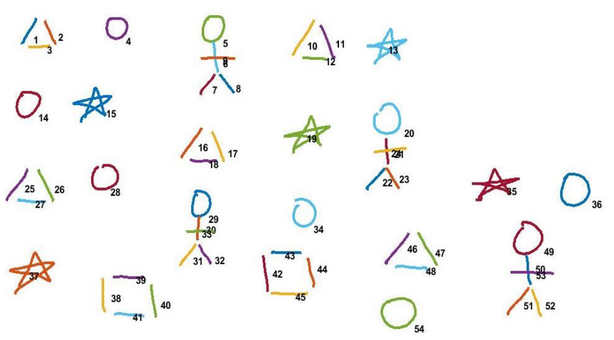

Figure 4. Training image

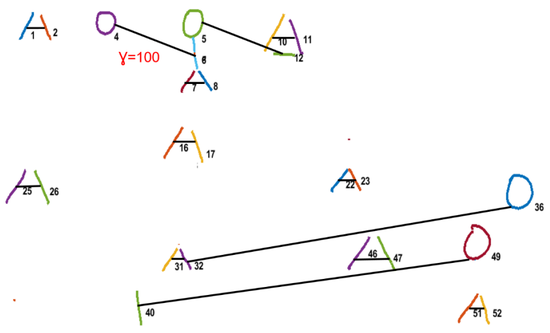

Figure 5. Connected graph

Upon constructing the connected graph we run EM algorithm until convergence. The jigsaw of training sketch scene and its reconstruction from the jigsaw is shown below.

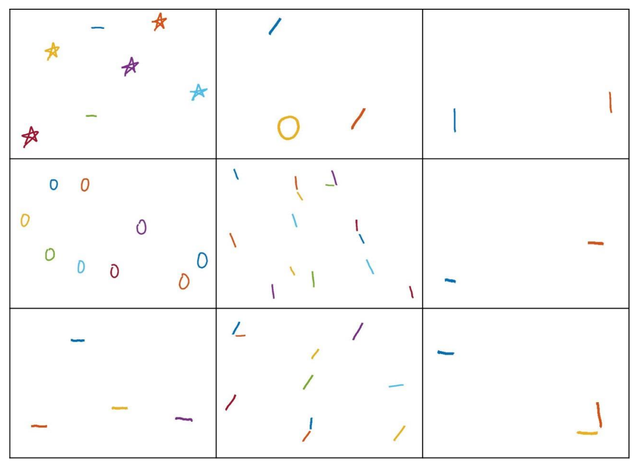

Figure 6. Sketch scene jigsaw.

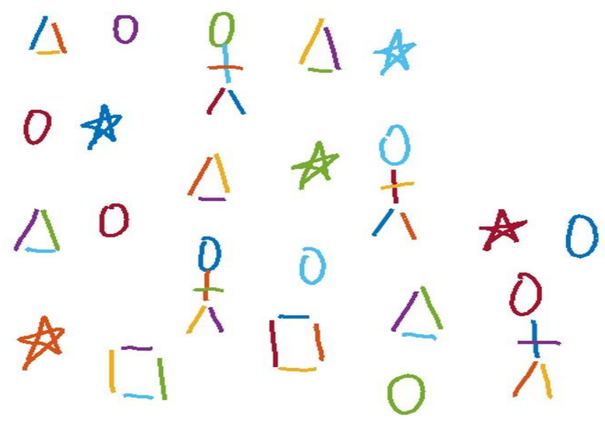

Figure 7. Reconstructed image scene.

References

[1] Kannan, A., Winn, J., Rother, C., 2006. “Clustering appearance and shape by learning jigsaws”, Advances in Neural Information Processing Systems, 657–664.

[2] Ouyang, T.Y., Davis, R., 2009. “A visual approach to sketched symbol recognition”, International Joint Conferences on Artificial Intelligence.