Image Parsing with a Three-State Series Neural Network Classifier

Mojtaba Seyedhosseini, Antonio R. C. Paiva and Tolga Tasdizen, Scientific Computing and Imaging Institute

THE UNIVERSITY OF UTAH

SCI www.sci.utah.edu

Introduction

- Image parsing is the problem of assigning an object label to each pixel in the image.
- Image parsing can be posed as a supervised learning problem where a classifier is learnt from training data.
- A series classifier is composed of stages which use features computed from the output of the previous classifier (context) as well as image features.
- We propose a three-state neural network in a series classifier architecture to improve the performance of the classifier.

Three-state neuron

- The idea is that having a third state allows for better uncertainty propagation.
- A three-state neuron is obtained by augmenting a two-state neuron with an additional bias term in the activation function:

$$g(x) = \frac{1}{2}\tanh(W^T x + b_1) + \frac{1}{2}\tanh(W^T x + b_2)$$

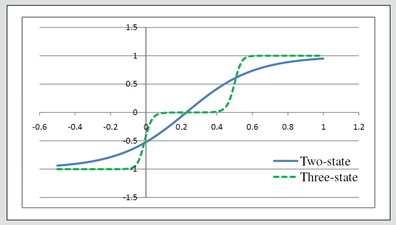

Fig 1: Three-state vs two-state

- The biases are learned using back-propagation. The third state appears as the separation between the biases increase. In training this happens in regions where the classes are not well separated.

Three-state series-ANN

- Using a single three-state ANN is pointless because the two-state ANN and the three-state ANN generate the same thresholded output.
- However, employing three-state neurons in the series-ANN structure improves the performance because it propagates the uncertainty to next stages.
- Then, later stages can use context information and the input image to decide about ambiguous pixels.

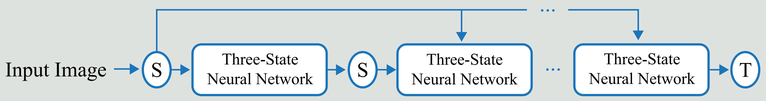

Fig 2: Series neural network

Results

- Texture segmentation

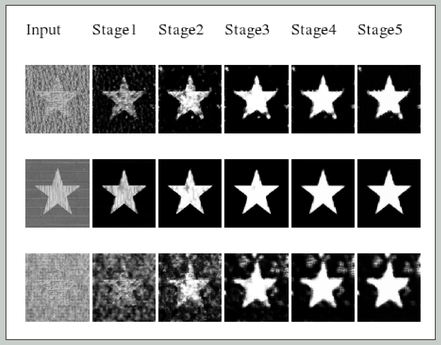

Fig 3: Test results for texture segmentation

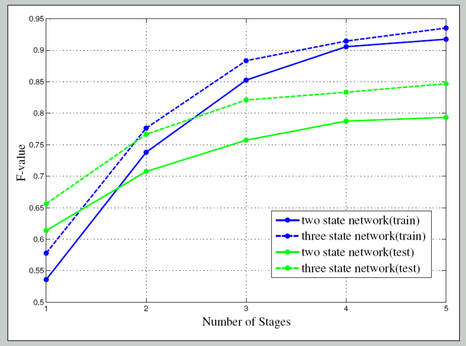

Fig 4: F-value* at different stages of the classifier

- Rectangle completion

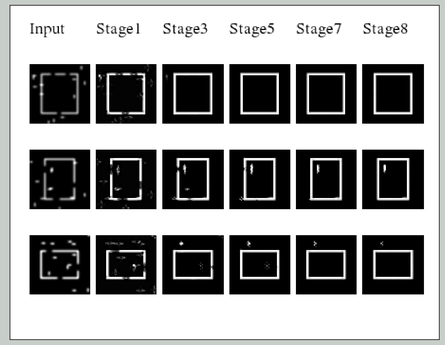

Fig 5: Test results for rectangle completion

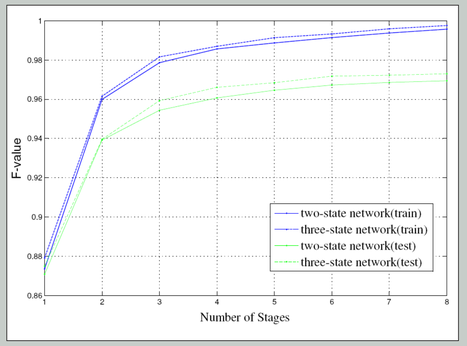

Fig 6: F-value* at different stages of the classifier

- Horse segmentation

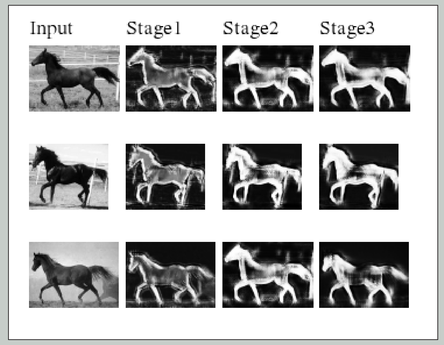

Fig 7: Test results for horse segmentation

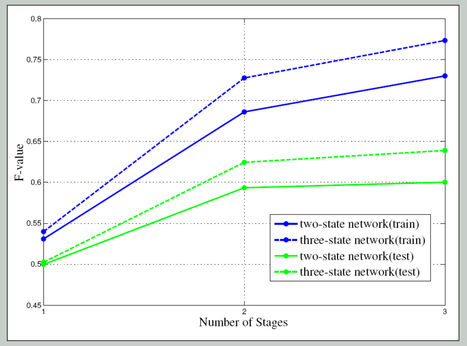

Fig 8: F-value* at different stages of the classifier

$$* F - value = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Conclusion

- Leaving the decision about some samples for later stages leads to better decision making according to context information.
- Our method outperforms the conventional two-state series-ANN.
- The third state reduces the mean square error and allows uncertainty propagation to next stages.
- Three-state neural networks can be used in general computer vision tasks.