# Deep Superpixel Cut for Unsupervised Image Segmentation

*Qinghong Lin[1,2], Weichan Zhong[1], Jianglin Lu[1,2]*

[1]*Shenzhen University, China*

[2]*Kazan Federal University, Russia*

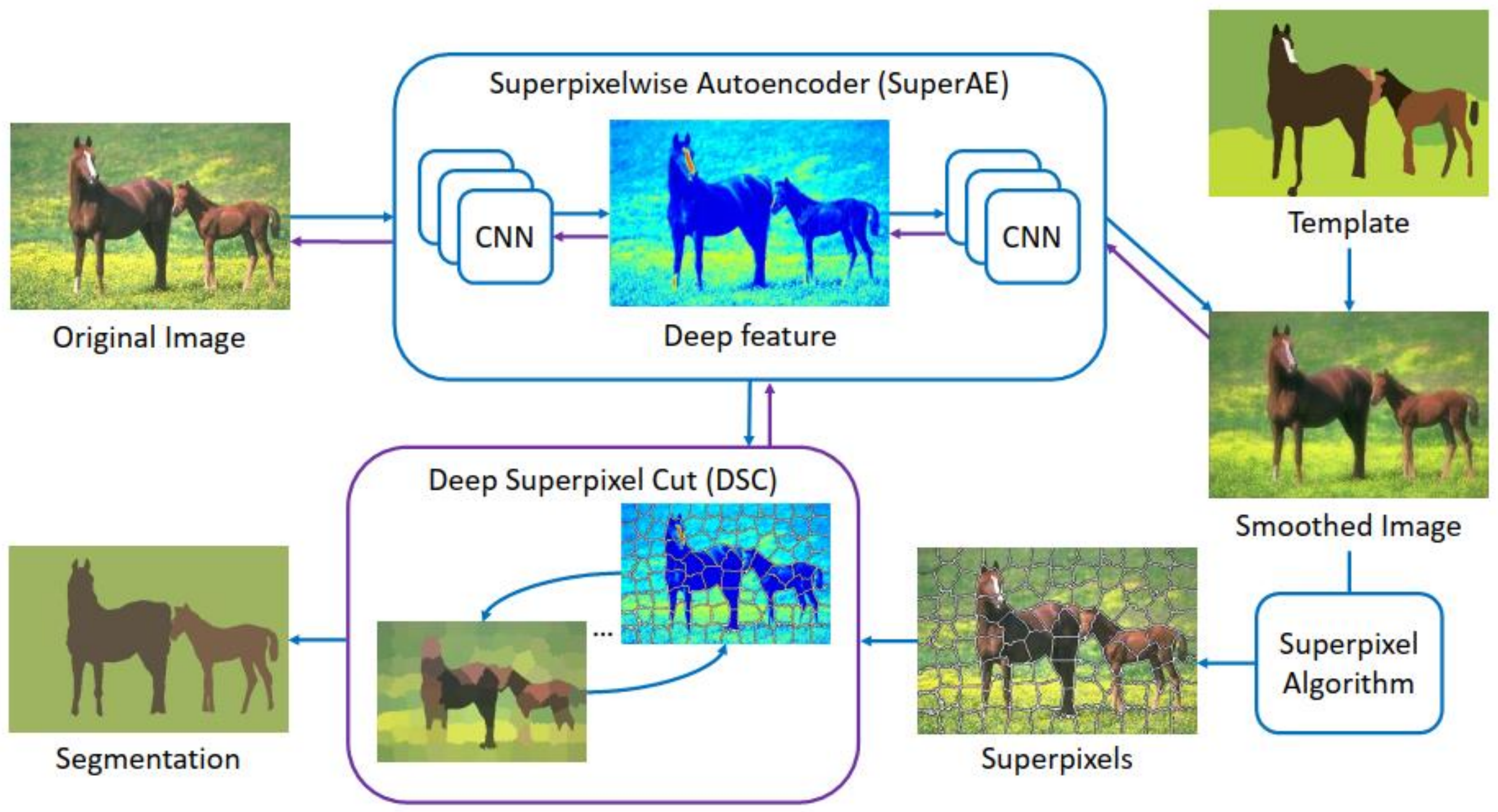
## OVERVIEW

We propose a **deep learning architecture for unsupervised image segmentation**, which contains two stages:

First stage, a **Superpixelwise Autoencoder (SuperAE)** is designed to learn the deep embedding and reconstruct a smoothed image, then the smoothed image is passed to generate superpixels.

Second stage, we present a novel clustering algorithm called **Deep Superpixel Cut (DSC)**, which measures the deep similarity between superpixels and formulates image segmentation as a soft partitioning problem.

## PIPELINE




## QUANTITATIVE RESULT

| Method | | SC | | RI | | VI | |
|---|---|---|---|---|---|---|---|
| | | ODS | OIS | ODS | OIS | ODS | OIS |
| SLIC | IMG | 0.354 | 0.402 | 0.698 | 0.726 | 2.391 | 2.306 |
| | SuperAE | 0.376 | 0.415 | 0.710 | 0.745 | 2.360 | 2.286 |
| | DSC | 0.428 | 0.485 | 0.709 | 0.742 | 2.262 | 2.079 |
| MS | IMG | 0.515 | 0.535 | 0.771 | 0.805 | 2.450 | 2.429 |
| | SuperAE | 0.497 | 0.550 | 0.777 | 0.795 | 2.162 | 2.077 |
| | DSC | 0.504 | 0.543 | 0.739 | 0.762 | 2.014 | 1.849 |
| EGB | IMG | 0.458 | 0.467 | 0.770 | 0.775 | 2.334 | 2.325 |
| | SuperAE | 0.491 | 0.500 | 0.774 | 0.782 | 2.190 | 2.180 |
| | DSC | 0.506 | 0.548 | 0.737 | 0.779 | 1.989 | 1.853 |

Evaluate the effectiveness of SuperAE and DSC on different superpixels algorithm.

## SUPERPIXELWISE AUTOENCODER

The real world image usually contain noise, which would affected the segmentation and needs to be smoothed. We observe that pixels within a superpixel are more continuous in color space, so we regard this as constraint in reconstruction.

Our SuperAE formulation:

$$\mathcal{L}_r = \|x - x^R\|^2 + \lambda \sum_{i=1}^{N}\sum_{j=1}^{M} t_{ij}\|x_i^R - v_j\|^2$$

$$v_j = \frac{\sum_{i=1}^{N} t_{ij} x_i}{\sum_{i=1}^{N} t_{ij}} \quad \text{(Superpixel RGB vector)}$$

SuperAE aim to minimize the variance of pixel features within a superpixel, thus smoothing the regions

## DEEP SUPERPIXEL CUT

Regard the segmentation as a pixelwise K-class classification problem by generate pseudo-labels.

$$\mathcal{L}_1 = -\sum_{n=1}^{N}\sum_{k=1}^{K} y_n^k \log(p_n^k) \quad \text{(Cross-entropy Loss)}$$

Regard the segmentation as a superpixelwise graph-partition problem, based on their deep similarity.

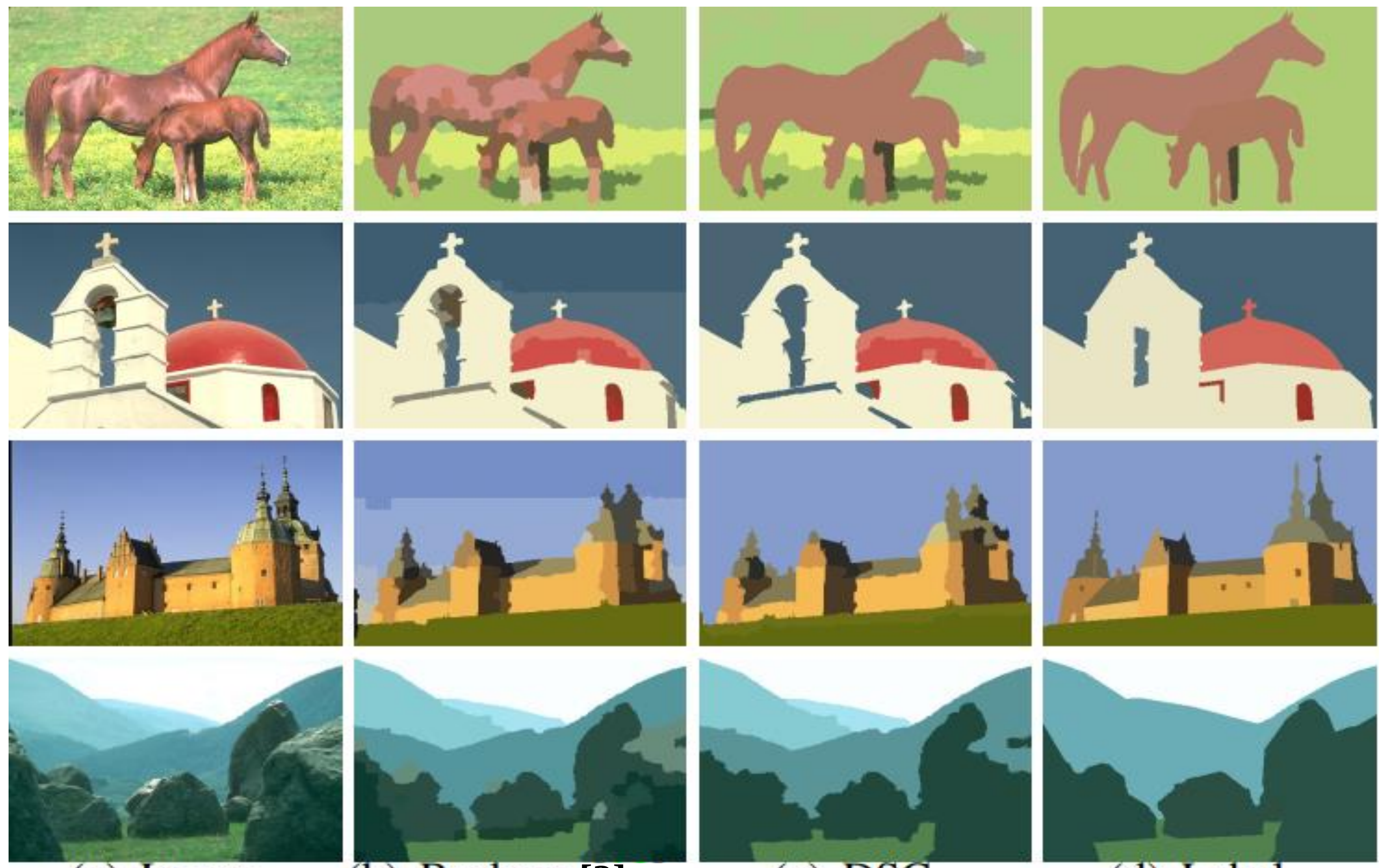
$$\begin{aligned}\mathcal{L}_2 &= \sum_{k=1}^{K}\sum_{a,b\in V} \mathcal{P}(a \in A_k, b \notin A_k) W(a,b) \\ &= \sum_{k=1}^{K}\sum_{a,b\in V} q_k^a w_{ab}(1 - q_k^b) \\ &= \sum_{k=1}^{K} Q_k^T \mathbf{W}(\mathbf{1} - Q_k)\end{aligned} \quad \text{(Differential)}$$

Intuition: If two superpixel a and b are more similar in deep features, then the probability they belong to different partitions should be lower.

### Loss function

$$\begin{aligned}\mathcal{L}_s &= \alpha\mathcal{L}_1 + \beta\mathcal{L}_2 \\ &= \alpha\sum_{n=1}^{N}\sum_{k=1}^{K} \hat{y}_n^k \log(p_n^k) + \beta\sum_{k=1}^{K} Q_k^T \mathbf{W}(\mathbf{1} - Q_k)\end{aligned}$$

## QUALITATIVE RESULT



(a) Image (b) Backpro[2] (c) DSC (d) Label

## REFERENCES

1. Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation[J]. *IEEE transactions on pattern analysis and machine intelligence, 2010*, 33(5): 898-916.
2. A. Kanezaki, "Unsupervised Image Segmentation by Backpropagation," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, 2018, pp. 1543-1547, doi: 10.1109/ICASSP.2018.8462533.