Using Semi-Supervised Learning in Cellular Automata for Edge Detection of Image Segmentation

Draiya A. Alaswad*[1], Yasser F. Hassan[2]

[*1]Department of Mathematics and computer science Faculty of science
Alexandria University, Alexandria, Egypt
[2]Department of Mathematics and computer science Faculty of science
Alexandria University, Alexandria, Egypt

Abstract:
Semi-Supervised Learning is an area of increasing importance in Machine Learning techniques that make use of both labeled and unlabeled data. The goal of using both labeled and unlabeled data is to build better learners instead of using each one alone. Semi-supervised learning investigates how to use the information of both labeled and unlabeled examples to perform better than supervised learning. In this paper we present a new method for edge detection of image segmentation using cellular automata with modification for game of life rules and K-means algorithm. We use the semi-supervised clustering method, which can jointly learn to fusion by making use of the unlabeled data. The learning aim consists in distinguishing between edge and no edge for each pixel in image. We have applied the semi-supervised method for finding edge detection in natural image and measured its performance using the Berkeley Segmentation Dataset and Benchmark dataset. The results and experiments showed the accuracy and efficiency of the proposed method.

Keywords*:* Semi-Supervised Learning, Semi-Supervised Clustering, Cellular Automata, K-means, Image Segmentation, Edge Detection.

1. Introduction

1.1 Semi-Supervised Learning (SSL)

Semi-Supervised Learning a learning paradigm in machine learning (ML), introduced by Valiant (1984). Semi-supervised learning has become an area of increased research attention over the past decades and seen as a viable alternative and trade-off between the more popular training approaches; unsupervised learning and supervised learning [1].

The main idea behind the semi-supervised learning is take advantage of huge unlabeled data, which is abundant and easily accessible, where the labeled data are few, difficult to obtain, and require special tools and experts [2]. Unlabeled data improve performance because they provide additional knowledge of data, such as density data in the same area in terms of clustering and dispersion [3].

Early, there is limited number of applications of semi-supervised learning, but currently there is lot of work going on in this field. Some applications are: using semi-supervised learning in computer vision [4], semi-supervised learning for detecting human trafficking [5], so also semi-supervised learning applies for natural language processing [6], and semi-supervised learning used in medical fields [7].

Recently, many algorithms for semi-supervised learning presented to images segmentation of which: a semi-supervised multi-view teaching algorithm based on Bayesian Learning [8], the algorithm uses two views and learns a separate classifier on each view. The first view contains the coordinates of the pixels and the second the RGB values of the points in the image, where it has added modifications on the standard co-training algorithm.

Nowadays, the semi-supervising learning used for biomedical image segmentation [9], the authors proposed a Forest Oriented Super Pixel (Voxel) method (FOSV), to be discriminant to the segmentation task. The FOSP (FOSV) method segment the initial estimation image into atomic

regions, where these atomic zones have been used to successfully train the semi-supervised learning algorithm. One of the most prevalent algorithms for semi-supervised learning is semi-supervised clustering,

1.2 Semi-Supervised Clustering

Semi-supervised clustering selects a subset of labeled data to facilitate unsupervised learning. This technique is one of the most efficient ML techniques that have found extensive use in image processing. Also semi-supervised clustering named constrained clustering or clustering with side information [10].

The main purposes for semi-supervised clustering are clustering many unlabeled data, and to guide clustering by few labeled data [11]. Clustering technique defined as the grouping of data points through some similarity or commonality between the points. This technique is one of the most important techniques to solve image segmentation problem.

Existing method of semi-supervised clustering for image segmentation as such: In paper [12] the semi supervised image segmentation using optimal hierarchical clustering algorithm introduced, the algorithm divides image to segment based on color image. In paper [13] presented an image segmentation of color images by semi-supervised clustering method based on modal analysis and mutational agglomeration algorithm in combination with the self-organization feature map (SOM) neural network. Also one of the most popular clustering algorithms is K-means.

1.3 K-means Algorithm

K-means a widely used clustering technique that seeks to reduce the average square distance between points in the same cluster. It was developed by J. MacQueen (1967), it is considered is a partitional clustering algorithm [14]. From the different technique one of the most efficient methods for image segmentation is the clustering method.

There are developmental manner to segment image by applying K-means clustering, such as: in paper [15] the K-means clustering algorithm deployed to segment an image into different colors, where the number of K = 7 based on colors in image, and the color image was divided based on number K.

In paper [16], the K-means algorithm was also applied for edge detection, it was applied to any type of gray scale image, and the similarity between the pixels and their eight neighbors was obtained by imposing three new features and providing them as attributes to K-means algorithm.

Also K-means algorithm is used to brain tumor detection [17], the MRI image of the affected area is detected and also reduces the time for analysis. The methodology helps to find perform the automatic segmentation of MRI brain image, where the image is segmented According to the number of pixel density in each segment. The rest of this paper is organized follows: Firstly section 2 describes the cellular automata technique. In section 3 describes the image segmentation task. In section 4 present methodology of proposed method. In Section 5 describes the experiments and results for proposed method. Finally section 6 presents conclusion.

2. Cellular Automata Technique

Cellular Automata (CA) have been successfully used in the area of image segmentation for the last couple of years. Cellular Automata were invented in the 1940s by the mathematicians John Von Neumann and Stanislaw Ulam. Cellular Automata deals with abstract mathematical models for systems comprising of substantial number of locally interacting simple but identical components known as "cell". The cell-like nature of the points which can be likened to points in a lattice or square boxes of a checker board forms the reason behind the name "cellular" and it is called automata due to the simple rules it follows [18].

Cellular Automata enable fast, parallel computation and have thus found application in image processing as well an image can be viewed as 2-dimensional cellular automata where each cell represents a pixel in the image and the intensity of the pixel is represented by the state of that cell. The states of the cells are updated synchronously at a discrete time step. So the time complexity to do any image processing task is the least [19]. Cellular Automata can be effectively applied in many image segmentation task, several algorithms have been introduced to this task using cellular automata model of which: in paper [20] approach for edge detection with noise filtering, where it applied game of life rules on digital image which represented as cellular automata, to

improve performance the parameters of the game of life rules is changed. Also cellular automata used for Digital Image Scrambling [21], where it continues investigations into its sensitivity to changes in the lattice which presented in previous paper. In paper [22], it discussed the most commonly used image segmentation techniques using cellular automata, where it presented many studies using cellular automata for image processing.

Many applications based on cellular automata, where using in medical image processing [23], a semi-automatic segmentation method based on cellular automata is proposed. The developed approach allows for the GTV segmentation and computes the lesion volume to be treated. The method was evaluated on 10 brain cancers, using both area-based and distance-based metrics.

There is way to update rules for multi–state cellular automata is investigated in the context of gray scale image segmentation [22]. Cellular Automata approach for efficient salt-n-pepper noise filtering in gray scale images is presented. Using a 2D Moore neighborhood, the classified "noisy" cells are corrected by averaging the non-noisy neighboring cells, algorithm was proposed to eliminate the salt and pepper noise from images using CA [24].

3. Image Segmentation Task

Image segmentation encompasses the entire process of splitting an image into multiple segments (sets of pixels). In other words; segmentation means division of an image into different objects or regions with the aim of simplifying the image representation. The process can make the image more meaningful and ease further ease analysis.

Image segmentation varies in types which include Translation, Rotation, Zooming, Noise reduction, and Edge detection. There are various CA-based segmentation algorithms, the most commonly used techniques are discussed in [25].

Many purposed algorithms and techniques have been developed for image segmentation. Since there is no general solution to the image segmentation problem, these techniques often have to be combined with domain knowledge in order to effectively solve an image segmentation problem for a problem domain. In paper [26] a comparative study of the basic block-based image segmentation techniques.

4. Methodology

In the segmentation problems we are dealing with two different kinds of data that can be used to train the image segmentation:

- Input data set provided to the image segmentation together with the class they belong to (supervised data).

- Input data set provided to the image segmentation without the class they belong to (unsupervised data).

We need a segmentation method that segments an image automatically without much interaction of the user and provide accurate results. This automated method makes use of cellular automata with K-means logarithm. The basic concept behind using cellular automata for automatic segmentation using is that an image is composed of pixels that can be considered as cells. The image represented as 2dimention cellular automata. In 2D cellular automata the cells are arranged in a 2dimensional lattice with connections among the neighboring cells, each cell has eight neighborhoods.

Image segmentation is the process of specify a label to every pixel in an image such that pixels with the same label share certain visual characteristics, this method divide every pixels in image into edge pixels and non-edge pixels. The problem this work attempts to solve is to employ a semi-supervised learning based mechanism for this process in order to correctly extract the desired edge.

This method is applied on image, where the image represented as a 2D cellular automata, each pixel of which is taken as a cell that connects with eight neighbors pixels, which may be labeled (0 or 1) or unlabeled (no value). The new method named "Semi CA".

4. 1 The Structure of Semi CA Technique

A Semi CA can be defined as follows:
Definition: A 2D is a structure (Z^d, $\varnothing$, X,N,F,L,) where:
- Z^d is a lattice of d-tuples of integer numbers.
- $\varnothing$ is set of states for cells.
- X is the set of states of K-means.
- N = {x_1, x_2,…,x_k} is the neighbourhood for each cell in cellular automata.

- F: $\varnothing^k \rightarrow \beta$ is the transition rule of CA, where β is the set of values which can takes.
- L: $X \times \varnothing \times \beta \rightarrow \varnothing$ is a learning algorithm for cellular automata.

4. 2 The Pseudo Code of Semi CA Technique

In this technique, instead of manually labeling the unlabeled pixels, we give approximate labels on the basis of the labeled pixels. This method is briefly described in steps as following in Algorithm 1:

1: **Input**: labeled image.
2: Divide the labeled image into labeled pixels and
   unlabeled pixels randomly.
3: Assume a 2-D MxN window of size 3 x 3
4: Apply CA transition rules
5: **For** each data x in unlabeled data **do**
6:      Choose randomly for K initial center points
7:      Cluster data using distance metric
8:      Calculate new center points of each cluster
9:      Re-Cluster all data using the new center points
10:     Repeat steps 8 and 9 until the center points have moved and until its convergence
11:     Add x to the labeled data set
12:     **End for**
13: **Output**: labeling unlabeled pixels.

4. 3 Transition Rules of Semi CA Technique

- **Transition rules of new method (Semi CA model) are modification for game of life rules as follows:**
1. If the central cell in state 0 and has three, four and five neighbors in state 1, then central cell becomes in state 1, in the next time step, otherwise becomes in state 0.
2. If the central cell in state 1 and has three, four and five neighbors in state 1, then central cell becomes in state 1, in the next time step, otherwise becomes in state 0.

- **The mathematical representation for transition rules of new method (Semi CA model) can be described as follows:**

If $S^t = 1$, then $S^{t+1} = \begin{cases} 1, & \text{Sum} = 3, 4 \text{ and } 5 \\ 0, & \text{otherwise} \end{cases}$

If $S^t = 0$, then $S^{t+1} = \begin{cases} 1, & \text{Sum} = 3, 4 \text{ and } 5 \\ 0, & \text{otherwise} \end{cases}$

Where, S^t is cell state at time t, S^{t+1} is the cell state at time t+1.

4. 4 The Construction of Semi CA model

This proposed method uses unlabeled pixels for training to enhance the performance of segmentation which uses only labeled pixels. Semi-supervised algorithms uses training data, which some of them are labeled, while the others are unlabeled of each data specifies its class, and semi-supervised algorithm should use such data to build segmentation. The following is the construction of Semi CA method for edge detection:

A. Image Acquisition
Image acquisition is the first method of image processing and the images have been taken from Berkeley University repository for image segmentation [27].

B. Image Pre-processing
The process of image pre-processing is aimed to improve the image data contained unwanted distortions or to improve processing time. This process implements several techniques on the image such as image resizing, noise reduction and image enhancement. In this work, we used various MATLAB code to resize image. Also we divide image to (labeled pixel, unlabeled pixel).

C. Image Segmentation
Image segmentation is the method for conversion of image into several segments and rendering of an image into something for easier analysis. The image segmentation process is used for locating the objects and bounding line of that image. In image segmentation, also we will use K-means clustering method for partitioning of images into clusters in which at least one part of cluster contain image with major area of unlabeled pixels. The K-means clustering algorithm is applied to classify the objects into K number of classes according to set of unlabeled pixels. The segmentation is done by minimizing sum of square of distances between data objects and the corresponding cluster. This algorithm exploits the

relationships among semi supervised learning and cellular automata for images segmentation.
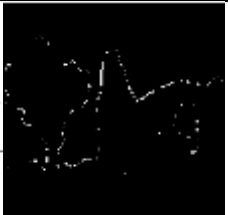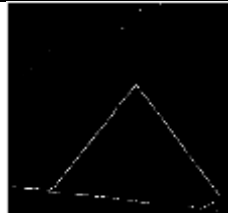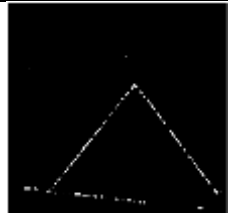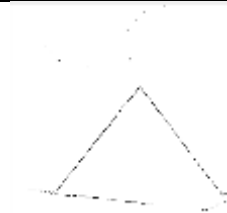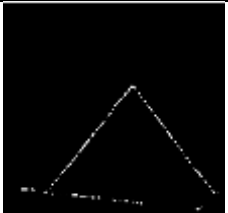
5. Experiments and Results

The performance of the proposed method is given in term of accuracy, by comparison new method with regular cellular automata and unsupervised learning algorithm (K-means). We compare the performance in the Semi CA image setting, i.e. the transition rule of K-mean in cellular automata technique are trained on both labeled and unl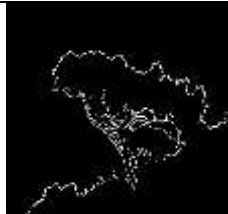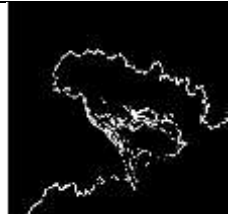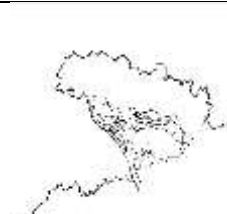abeled data, and the performance is evaluated on the unlabeled data. We employ the accuracy on classifying the unlabeled data as a performance indicator of different methods. The optimal values can be obtained by using the case of 10% labeled data pixels.

We used K-means clustering method to divide

cluster contain image with major area of unlabeled pixels. The K-means clustering algorithm is applied to labeling unlabeled pixels into K number of classes according to unlabeled of pixels, and it aimed to organize unlabeled pixel into clusters.

All the experiments are conducted using MATLAB R2016a on a 2.53 GHz CPU with 4GB of memory running in Windows 7.The performance of the image segmentation accuracy heavily depends on the size of the labeled training data. Precisely labeled pixels with explicit correspondence is however, difficult to create or obtain in large quantities, whereas, unlabeled pixels are easier to collect. Hence, we proposed to use the semi-supervised learning which attempts to exploit the unlabeled pixels to improve the annotation performance.

Dataset					
1					
2					
3					
	(a) 0rginal Image	(b) Semi-Supervised Image	(c) Edge Detection using Cellular Automata	(d) Edge Detection using K-means	(e) Edge Detection using Semi CA

an image into clusters in which at least one part of

Figure 1: Result of applying the proposed Semi CA versus cellular automata and K-means on the image dataset using 10% labeled pixels and 90% unlabeled pixels.

We have shown that cellular automata and K-means clustering algorithms are quite useful for

edge detection. Performance of the labeling process can be improved more, such as the every

dataset consists of 481x321 pixels, to improve processing time, we resize image 300 ×300. At each step only a small amount of labeled pixels are used. Multiple tests were held depending on the number of labeled pixel (10%, 20%, 30% pixels) and the (90%, 80%, 70% pixels) are used as unlabeled pixels. We assume K (here 2) numbers of pixels in an image then deploy K-means clustering over it. Each unlabeled pixel from the input image was assigned to one of the clusters whose centroid was nearest.

In dataset 1 the process elapsed time for every configuration is 1.92, 2.13, and 2.17 seconds respectively. In dataset 2 the process elapsed time for every configuration is 1.73, 1.35, and 2 seconds respectively, and dataset 3 the process elapsed time for every configuration is 1.90, 1.96, and 2.91 seconds respectively.

The results indicate that the proposed approach offers significant performance improvements in labeling unlabeled pixel.

In Figure1, comparing the results of edge detection using cellular automata and K-means with new method using Semi CA method for edge detection, this result where test image for training on the semi-supervised dataset with 10%labeled pixels and 90% unlabeled pixels.

The performance of the clustering teaching based on K-means algorithm is compared to its cellular automata and K-means algorithm, depending on the number of labeled pixels.

We test the proposed method with regular cellular automata and K-means algorithm. Tables 1, 2 and 3 illustrate the accuracy of the new method with regular cellular automata and K-means algorithm. We notice that, the Semi CA give high accuracy than cellular automata and K-means algorithm on the case of 10% labeled pixels. Specifically, the proposed algorithm achieves the best performance. From this we can see that the proposed segmentation mechanism is effective and efficient.

Table 1: Test image for training on the semi-supervised dataset 1 using 10%, 20%, 30%labeled pixels and 90%, 80%,70% unlabeled pixels.

Dataset 1	(Labeled, Unlabeled)%		
	(10 , 90)	(20 , 80)	(30 , 70)
K-means Algorithm	47%	45%	46%
Cellular Automata	70 %	72%	71%
Semi CA	**74%**	**72%**	**73%**

Table 2: Test image for training on the semi-supervised dataset 2 using10%, 20%, 30%labeled pixels and 90%, 80%, 70% unlabeled pixels.

Dataset 2	(Labeled, Unlabeled)%		
	(10 , 90)	(20 , 80)	(30 , 70)
K-means Algorithm	36%	32%	35%
Cellular Automata	82 %	80%	81%
Semi CA	**84%**	**83%**	**83%**

Table 3: Test image for training on the semi-supervised dataset 3 using 10%, 20%, 30%labeled pixels and 90%, 80%,70% unlabeled pixels .

Dataset 3	(Labeled, Unlabeled)%		
	(10 , 90)	(20 , 80)	(30 , 70)
K-means Algorithm	62%	64%	60%
Cellular Automata	85%	86%	87%
Semi CA	**90%**	**88%**	**89%**

From the results of K-means, labeling of each pixel in the image is done also segmented image.

6. Conclusion

In this work, we have introduced new semi-supervised learning method (Semi CA) for edge detection of image segmentation problem by using K-means clustering algorithm and cellular automata with modification of Conway's game of life rules. We also develop a new method for reducing the amount of labeled data in image. A comparison between the performance of the CA and the proposed method is introduced. The experimental results show that the proposed method overcomes regular cellular automata and K-means algorithm. The new method has good

performance as results and experiments have shown.

A semi-supervised learning algorithm is a new method for edge detection of image segmentation using both labeled dataset and unlabeled dataset. We applied semi-supervised learning in cellular automata, where we combined between K-means algorithm and cellular automata. The results show that the method based on semi-supervised learning with cellular automata has a desirable efficiency in edge detection.

In the future, we plan to apply the idea of new semi-supervised learning framework on apply medical image. The methodology of the semi-supervised learning may be developed by adding expectation maximization algorithms (EM), neural network algorithms, and genetic algorithm (GA), to get the highest accuracy for image segmentation.

References

[1]   Z. Zhang; F.Xing; X.Shi; L.Yang; "SemiContour: A Semi-Supervised Learning Approach for Contour Detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 251–259, 2016.

[2] H.Gan; Z.Luo; M.Meng; Y.Ma; Q.She, "A Risk Degree-Based Safe Semi-supervised Learning Algorithm", Internationl Journal of Machine Learning and Cybernetics, 7, 1, 85-94, 2016.

[3] X.Zhu, "Semi-Supervised Learning Literature Survey", Computer Sciences TR 1530, 1-60, 2008.

[4] D.C.Park, "Multiple Feature-Based Classifier and Its Application to Image Classification", On IEEE Computer Society, Sydney, 65–71, 2010.

[5]   H. Alvari; P. Shakarian; J. E. K. Snyder, "Semi-Supervised Learning for Detecting Human Trafficking", Springer Nature remains neutral with regard to jurisdictional claims in publishedmaps and institutional affiliations,6, 1, 1-14, 2017.

[6]   S.Bergsma, "Large-Scale Semi-Supervised Learning for Natural Language Processing", University of Alberta, 2010.

[7]   R.Meier1; S.Bauer; J.Slotboom; R.Wiest; M. Reyes. "Patient-Specific Semi-Supervised Learning for Postoperative Brain Tumor Segmentation", International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2014.

[8]   G.A. Lazarova, "Semi-Supervised Image Segmentation ", In: International Conference on Artificial Intelligence: Methodology, Systems, and Applications. Springer, Cham, 59-68, 2014 .

[9]   L. Gu; Y. Zheng; R. Bise; I. Sato; N. Imanishi; S. Aiso, "Semi-supervised Learning for Biomedical Image Segmentation via Forest Oriented Super Pixels (Voxels)", Springer International Publishing, 702-710, 2017.

[10] M.Bilenko; S.Basu; R.J.Mooney, "Integrating Constraints and Metric Learning in Semi-Supervised Clustering", In: The twenty-first international conference on Machine learning. ACM, 2004.

[11]   S.Basu, "Semi-Supervised Clustering: Learning With Limited User Feedback", Computer Science Department, University of Texas at Austin, 2003.

[12]   L.Sankari,   C.Chandrasekar, "Semi-Supervised Image Segmentation Using Optimal Hierarchical Clustering By Selecting Interested Region As Prior Information", Journal of Global Research in Computer Science, 2, 11,1-5, 2011.

[13] A.Halder; S.Dalmiya; T.Sadhu, " Color Image Segmentation Using Semi-Supervised Self-Organization Feature Map", Springer International Publishing Switzerland, 591-598, 2014.

[14] Zhao; Weizhong; H.Ma; Q. He, "Parallel K-means Clustering Based on Map Reduce", In: IEEE International Conference on Cloud Computing. Springer, Berlin, Heidelberg, 2009.

[15] Z.Reihani, H.A.Ghazijahani, A.Joudi, "Color Image Segmentation with K-means Clustering and Edge Detection Methods", In:

Conference on Electrical Engineering, September 2012.

[16] W. M. Khalaf; K. Al-Majdi; N. H. Hamed , " A Novel Edge Detection Method Using K-means Clustering", Journal of Engineering and Sustainable Development, 20, 06, 207-215, 2016.

[17] V.Surudhi; K.Sanjana; R.Saravanan; G.Santhosh; S.Kirubha, "Brain Tumor Detection Using K-Means Clustering Algorithm", International Journal of Computer Science and Engineering, 10-13, 2017.

[18] A.A.Abin; M.Fotouhi; S.Kasaei, "A New Dynamic Cellular Learning Automata-Based Skin Detector", Springer Multimedia systems, 15, 309-323, 2009.

[19] J.George; E.George, "Cellular Automata Cryptography Using Reconfigurable Computing Source", In: The International Conference on Developments in Applied Artificial Intelligence , 104–111, 2003.

[20] M.A.Peer; F.Qadir; K.A.Khan, " Investigations of Cellular Automata Game of Life Rules for Noise Filtering and Edge Detection", International Journal Information Engineering and Electronic Business, 2, 22-28, 2012.

[21] F.Qadir; M.A. Peer; K. A. Khan, "Digital Image Scrambling Based on Two Dimensional Cellular Automata", International Journal Computer Network and Information Security, 2, 36-41, 2013.

[22] D.R.Nayak; P.K.Patra; A.Mahapatra, "A Survey on Two Dimensional Cellular Automata and Its Application in Image Processing", 1-10, 2014.

[23] L.Rundo; C.Militello; G.Russo; P.Pisciotta; L.M.Valastro; M.G.Sabini; S.Vitabile; M.C.Gilardi; G.Mauri, "Neuro-Radiosurgery Treatments: MRI Brain Tumor Seeded Image Segmentation Based on a Cellular Automata Model", In: International Conference on Cellular Automata. Springer International Publishing, 323:333, 2016.

[24] D.Tourtounis; N.Mitianoudis; G.C.Sirakoulis, "Salt-n-Pepper noise Filtering using Cellular Automata", Journal of Cellular Automata", 13, 81-101, 2018.

[25] P.Shotrya; S.Bhardwaj, "Image Segmentation Using Cellular Automata: A Technical Survey", International Journal Of Engineering And Computer Science, 2, 4, 1268-1272, 2013.

[26] N.M.Zaitouna; M. J. Aqelb, "Survey on Image Segmentation Techniques", In: International Conference on Communication, Management and Information Technology (ICCMIT), 797-806, 2015.

[27] https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds.