Multi Kernel Boosting Algorithm for Image Segmentation and Analysis

S. Sanjuna

Student, Electronics and Communication Engineering, Sri Krishna College of Engineering and Technology,

Coimbatore, Tamil Nadu

Abstract: The bootstrap learning algorithm for salient object detection in which both weak and strong models are exploited. First, a weak saliency map is constructed based on image priors to generate training samples for a strong model. Second, a strong classifier based on samples directly from an input image is learned to detect salient pixels. Results from multiscale saliency maps are integrated to further improve the detection performance. The proposes a novel framework to improve the saliency detection performance of an existing saliency model, which is used to generate the initial saliency map. First, a novel regional descriptor consisting of regional self-information, regional variance, and regional contrast on a number of features with local, global, and border context is proposed to describe the segmented regions at multiple scales. Then, regarding saliency computation as a regression problem, a multiple kernel boosting method based on support vector regression (MKB-SVR) is proposed to generate the complementary saliency map. Finally, an adaptive fusion method via learning a quality prediction model for saliency maps is proposed to effectively fuse the initial saliency map with the complementary saliency map and obtain the final saliency map with improvement on saliency detection performance. Experimental results on two public datasets with the state-of-the-art saliency models validate that the proposed method consistently improves the saliency detection performance of various saliency models.

Keywords: saliency detection, saliency map, saliency computation, multiple kernel boosting, adaptive fusion.

I. INTRODUCTION

Salient region detection is important in image understanding and analysis. Its goal is to detect salient regions in an image in terms of a saliency map, where the detected regions would draw humans' attention. Many previous studies have shown that salient region detection is useful, and it has been applied to many applications including segmentation, object recognition, image re-targeting, photo rearrangement, image quality assessment, image thumb nailing, and video compression. The development of salient region detection has often been inspired by the concepts of human visual perception. One important concept is how "distinct to a certain extent" the salient region is compared to the other parts of an image. As color is a very important visual cue to human, many salient region detection techniques are built upon distinctive color detection from an image.

Natural scenes always contain high dynamic range areas in comparison with the limited dynamic range capabilities of cameras or displays. The dynamic range is defined by the ratio between the maximum and minimum light intensities of the scene. An HDR image is commonly obtained by fusing multi-exposure images. The fused HDR image always exceeds the dynamic range of displays. So, some mapping is needed here to compress the intensity distribution of the HDR image. The compression is based on the feature of the human visual system (HVS) that it is less sensitive to the low-frequency components than to the high frequency components. The low-frequency components are compressed while the high-frequency components are retained. Through this reproduction process, here propose can hardly discern the difference between the artificial image and the real scene. Special considerations are also noted here to avoid artifacts (e.g., halo, the brighter or darker bands around edges). The difference between the luminance of a point and the average luminance of its surrounding points is visible for humans; This method terms the difference as the visible contrast. In the proposed approach use visible contrast and gradient of the whole scene to describe scene details. Given an SDR capture device, the scene luminance is recorded under a specific exposure level as the image luminance, which is determined by the response function of the film or the charge-coupled device.

Edge-preserving becomes an important property in filtering design to avoid halo artifacts. This technique decomposes an image into a piecewise smooth base layer and a detail layer. The base layer no longer only contains low frequency band, but it also has salient edges (high frequency). Multi-scale is used here to decompose progressively another detail layer from the last decomposed base layer. In other words, the high-frequency information is progressively decomposed from the original image. There is an important property in the decomposition, which is the residual base layer matches the large-scale shape of the original image signal. The tone mapped images using these edge-preserving filters give state-of-the-art quality, and they are visually appealing.

The scene luminance, which is perceived by us without involving any device between the scene and eyes, can greatly vary between 10-3and 105 cd/m2 from region to region. However, the dynamic range of a normal digital camera/monitor, termed as the standard dynamic range (SDR) device, is only about 102. As a consequence, the proposed system need professional high dynamic range (HDR) devices to fully capture and display scenes. However, these devices are expensive and are thus not accessible to most people. Compared with HDR devices, SDR devices are cheaper, and thus, it is important and attractive to capture and display HDR scenes by using SDR devices. When HDR capture devices are not available, existing systems typically first apply algorithms to synthesize an HDR image via a set of SDR images, which record scene details, and then reproduce an SDR image via the synthesized HDR image, so that it can display the scene on an SDR device. This two-stage approach usually requires complex manual interactions and damages some scene details in both stages.

II. RELATED WORK

Introduce a novel technique to automatically detect salient regions of an image via high dimensional cool transform. This main idea is to represent a saliency map of an image as a linear combination of high-dimensional color space where salient regions and backgrounds can be distinctively separated.

A. Local-contrast-based models
There are many previous methods that utilize low-level features such as color and texture for saliency detection. The saliency detection method based on color contrast called "center-surround difference." suggested a graph-based visual saliency (GBVS) model based on the Markovian approach. The viewed color and luminance in the frequency domain to compute an image's saliency. Combined global and local contrast saliency to improve the detection performance. The applied Kullback-Leibler divergence to the center-surround difference to combine different image features to compute saliency. The salient object segmentation with multi-scale super pixel-based saliency and closed boundary prior. The uniqueness and distribution of the CIELab color to find the salient region. The divided an image into a low-rank matrix and sparse noises to detect an object. A hierarchical model by computing contrast features of different scales of an image, and fused them into a single map using a graphical model.

Several methods estimated saliency in super pixel level instead of pixel-wise level to reduce the computational time. The salient object segmentation with multiscale super pixel-based saliency and a closed boundary prior. Their approach iteratively updates both the saliency map and the shape prior under an energy minimization framework. The decomposed an image into compact and perceptually homogeneous elements, and then considered the uniqueness and spatial distribution of these elements in the CIELab color to detect salient regions. The background measure that characterizes the spatial layout of image regions with a novel optimization framework.

These models tend to give a higher saliency at around edges and texture areas that have high contrasts, where humans tend to focus on in an image. However, these models tend to catch only parts of an object. Also, they tend to give non-uniform weight to the same salient object when different features presented in the same salient object.

B. Global-contrast-based models
The global color contrast with respect to the entire image to detect salient regions. These models can detect salient regions of an image uniformly with low computational complexity. A frequency-tuned approach to determine the center-surround contrast using the color and luminance in the frequency domain as features. The divided an image into two parts a low-rank matrix and sparse noise where the former explains the background regions and the latter indicates the salient regions. A Gaussian mixture model (GMM)-based abstract representation method that simultaneously evaluates the global contrast differences and spatial coherence to capture perceptually homogeneous elements and improve the salient region detection accuracy. The unique refocusing capability of light fields can robustly handle challenging saliency detection problems such as similar foreground and background in a single image. A pair of flash and no-flash images, inspired by the brightness of foreground objects for salient region detection.

These global-contrast-based models provide reliable results at low computational cost as they mainly consider a few specific colors that separate the foreground and the background of an image without using spatial relationships.

C. Statistical-learning-based models
Jointly estimates the segmentation of objects learned by a trained classifier called the auto-context model to enhance an appearance-based energy minimization framework for salient region detection. The ranked the similarity of image regions with foreground cues and background cues using graph-based manifold ranking based on affinity matrices and successfully conducted saliency detection. An unsupervised approach to learn patches that are highly likely to be parts of salient objects from un-labelled images and then sampled the object saliency map to find object locations and detect saliency regions. A saliency measure via dense and sparse representation errors of each image region using a set of background templates as the basis for reconstruction, and they constructed the saliency map by integrating multiscale reconstruction errors. A bottom-up saliency detection algorithm that considers the appearance divergence and spatial

distribution of salient objects and the background using the time property in an absorbing Markov chain. An optimal set of salient seeds obtained by learning a large margin formulation of the discriminant saliency principle.

These models are usually highly accurate and have a simple detection structure. However, they tend to require a lot of computational time. Therefore, super pixel-wise saliency detection is used to overcome the high computational complexity.

III.PROPOSED SYSTEM

The proposed weak and strong saliency maps have complementary properties. The weak map is likely to detect fine details and to capture local structural information due to the contrast-based measure. In contrast, the strong map works well by focusing on global shapes for most images except the case when the test background samples have similarity with the positive training set or large differences compared to the negative training set, or vice versa for the test foreground sample. Then, regarding saliency computation as a regression problem, a multiple kernel boosting method based on support vector regression (MKB-SVR) is proposed to generate the complementary saliency map. Finally, an adaptive fusion method via learning a quality prediction model for saliency maps is proposed to effectively fuse the initial saliency map with the complementary saliency map and obtain the final saliency map with improvement on saliency detection performance. Experimental results on two public datasets with the state-of-the-art saliency models validate that the proposed method consistently improves the saliency detection performance of various saliency models.

The key contributions of MKB-SVR work are as follows:
- The propose a novel regional descriptor with the introduction of more effective features, the incorporation of local, global, and border context for computing regional contrast and the concatenation of regional self-information, regional variance, and regional contrast on the basis of each feature.
- The proposed regional descriptor is considerably different from those used in the previous works nd can better describe the region for saliency computation.
- Different from the BL method which also uses the multiple kernel boosting, it regards saliency computation as a regression problem and propose a novel multiple kernel boosting method based on support vector regression (MKB-SVR), which is consistent with multiple kernel learning for taking full advantage of multiple features via the definition of regression error and the update of sample weights.
- The propose an adaptive fusion method via learning a quality prediction model to obtain the fusion weights for initial saliency map and complementary saliency map and finally enable the proposed method to consistently improve the saliency detection performance of various saliency models.

A. Image Features

The image feature has been used extensively in vision tasks as the basic units to capture the local structural information. In this paper, we compute a fixed number of super pixels from an input image using the Simple Linear Iterative Clustering (SLIC) method. Three descriptors including the RGB, CIELab and Local Binary Pattern (LBP) features are used to describe each super pixel. The rationale to use two different color representations is based on empirical results were better detection performance is achieved when both are used, which can be found in the supplementary document. We consider the LBP features in a 3×3 neighborhood of each pixel. Next, each pixel is assigned to a value between 0 and 58 in the uniform pattern. To construct an LBP histogram for each super pixel, i.e., a vector of 59 dimensions ($\{hi\}$, i = 1, 2, ...59, where hi is the value of the i-th bin in an LBP histogram).

B. Super pixel segmentation

The features from super pixels are effective and efficient for salient object detection. For an input image I, the first perform over-segmentation to form super pixels $X = \{X_1, X_2 .... X_N\}$. The proposed system use the SLIC super pixel because of its low computational cost and high performance, and set the number of super pixels to N = 500. To build feature vectors for saliency detection, it combines multiple information that are commonly used in saliency detection. Here first concatenate the super pixels' x- and y-locations into feature vector. The location feature is used because humans tend to focus more on objects that are located around the center of an image. Then, it concatenates the color features, as this is one of the most important cues in the human visual system and certain colors tend to draw more attention than others. here compute the average pixel color and represent the color features using different color space representations. Next, concatenate histogram features as this is one of the most effective measurements for the saliency feature, as demonstrated. The histogram features of the ith super pixel D_{Hi} is measured using the chi-square distance between other super pixels' histograms. It is defined as

$$D_{Hi} = \sum_{j=1}^{N} \sum_{k=1}^{b} \frac{(h_{ik} - h_{jk})^2}{(h_{ik} + h_{jk})} \quad (1)$$

where b is the number of histogram bins. In this work, used eight bins for each histogram. The have also used the global contrast and local contrast as color features. The global contrast of the ith super pixel D_{Gi} is given by

$$D_{Gi} = \sum_{j=1}^{N} d(c_i, c_j) \quad (2)$$

where $d(c_i, c_j)$ denotes the Euclidean distance between the ith and the jth superpixels' color values, c_i, and c_j, respectively. It uses the RGB, CIELab, hue, and saturation of eight color channels to calculate the color contrast feature so that it has eight dimensions. The local contrast of the color features D_{Li} is defined as

$$D_{Li} = \sum_{j=1}^{N} \omega_{i,j}^{p} d(c_i, c_j) \quad (3)$$

$$\omega_{i,j}^{p} = \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_P^2} \|P_i - P_j\|_2^2\right) \quad (4)$$

where pi $\in [0, 1] \times [0, 1]$ denotes the normalized position of the ith super pixel and Z_i is the normalization term. The weight function in Eq. (4) is widely used in many applications including spectral clustering. The adopt this function to give more weight to neighboring super pixels. In this experiment, here set $\sigma_P^2 = 0.25$. In addition to the global and local contrast, it further evaluates the element distribution by measuring the compactness of colors in terms of their spatial color variance.

C. Weak Saliency Model

The center-bias prior has been shown to be effective in salient object detection. Based on this assumption, we develop a method to construct a weak saliency model by exploiting the contrast between each region and the regions along the image border. However, existing contrast-based methods usually generate noisy results since low-level visual cues are limited. In this paper, we exploit the center-bias and dark channel priors to better estimate saliency maps. The dark channel prior is proposed for the image haze removal task. The main observation is that, for regions that do not cover the sky (e.g., ground or buildings), there exist some pixels with low intensity values in one of the RGB color channels. Thus, the minimum pixel intensity in any such region is low. The dark channel of image patches is mainly generated by colored or dark objects and shadows, which usually appear in the salient regions. The sky region of an image usually belongs to the background, which is just consistent with the dark channel property for the sky region. Therefore, it exploits the dark channel property to estimate saliency of pixels. In addition, for situations where the input image has dark background or bright foreground, it uses an adaptive weight computed based on the average value on the edge of dark channel map. In the proposed method, it defines the dark channel prior of an image on the pixel level. For a pixel p, the dark channel prior $S_d(p)$ is computed by

$$S_d(p) = 1 = \min_{q \in patch\ (p)} \left(\min_{ch \in \{r,g,b\}}(I^{ch}(q))\right) \quad (5)$$

where patch(p) is the 5×5 image patch centered at p and Ich(q) is the color value of pixel q on the corresponding color channel ch. Note that all the color values are normalized into [0, 1]. We achieve pixel-level accuracy instead of the patch-level counterpart.

An input image is segmented into N super pixels, $\{c_i\}$, $i = 1, \dots N$. The regions along the image border are represented as $\{n_j\}$, $j = 1, \dots N_B$, where N_B is the number of regions along the image border. We compute the dark channel prior for each region ci using $\frac{1}{N_{ci}} \sum_{p \in c_i} s_d(p)$, where n_{c_i} is the number of pixels within the region c_i. The coarse saliency value for the region ci is constructed by

$$f_o(c_i) = g(c_i) \times s_d(c_i) \times \sum_{k \in \{F_1, F_2, F_3\}} \left(\frac{1}{N_B} \sum_{j=1}^{N_B} d_k(c_i, n_j)\right) \quad (6)$$

where $d_k(c_i, n_j)$ is the Euclidean distance between region ci and nj in the feature space that κ represents, i.e., the RGB (F_1), CIELab (F_2) and LBP (F_3) texture features respectively. Note that all the distance values in each feature space are normalized into [0, 1]. In addition, $g(c_i)$ is computed based on the center prior using the normalized spatial distance between the center of the super pixel ci and the image center. Thus, the saliency value of the region closer to the image center is assigned a higher weight. We generate a pixel-wise saliency map $\mathcal{M}_0$, where the saliency value of each super pixel is assigned to the contained pixels.

Given an input image, we construct an undirected graph G = (V, E, T), where E is a set of undirected edges that connect the nodes V (pixels) while T is the set of the weights of nodes connected to the background and foreground terminals.

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

The weight of each node (pixel) p connected to the foreground terminal is assigned with the saliency value in the pixelwise map $\mathcal{M}_0$. Thus, for each pixel p, the set T consists of two components, defined as $\{T^f(p)\}$ and $\{T^b(p)\}$, and is computed by

$$T^f(p) = \mathcal{M}_0(p), T^b(p) = 1 - \mathcal{M}_0(p) \quad (7)$$

where $T^f(p)$ is the weight of pixel p connected to the foreground while $T^b(p)$ is the weight to the background? The minimum cost cut generates a foreground mask $\mathcal{M}_1$ using the Max-Flow method to measure the probability of each pixel being foreground. binary map which may contain noise in both foreground and background. Thus, we consider both the binary map $\mathcal{M}_1$ and the map $\mathcal{M}_0$ to construct the continuous and smoothed weak saliency map $\mathcal{M}_w$ by

$$\mathcal{M}_w = \frac{\mathcal{M}_0 + \mathcal{M}_1}{2} \quad (8)$$

The training set for the strong classifier is selected from the weak saliency map. The compute the average saliency value for each super pixel and set two thresholds to generate the training set containing both positive and negative samples. The super pixels with saliency values larger than the high threshold are labeled as the positive samples with +1 while those with saliency values smaller than the low threshold as the negative samples labeled with −1.

D. Multiple Kernel Boosting with Strong Saliency Model
Support Vector Machine (SVM) is to determine the appropriate kernel for the given dataset. This problem is more complicated when the dataset contains thousands of diverse images with different properties. While numerous saliency detection methods based on various features have been proposed, it is still not clear how these features can be well integrated. To cope with these problems, we present a method similar to the Multiple Kernel Boosting (MKB) method to include multiple kernels of different features. The SVMs with different kernels as weak classifiers and then learn a strong classifier using the boosting method. Note that we restrict the learning process to each input image to avoid the heavy computational load of extracting features and learning kernels for a large amount of training data.

The MKB algorithm is a boosted Multiple Kernel Learning (MKL) method, which combines several SVMs of different kernels. For each image, we have the training samples $\{r_i, l_i\}_{i=1}^H$ from the weak saliency map $\mathcal{M}_w$ where r_i is the i-th sample, l_i represents the binary label of the sample and H indicates the number of the samples. The linear combination of kernels $\{k_m\}_{m=1}^M$ is defined by

$$K(r, r_i) = \sum_{m=1}^M \beta_m k_m(r, r_i), \sum_{m=1}^M \beta_m = 1, \beta_m \in \mathbb{R}_+ \quad (9)$$

where β_m is the kernel weight and M denotes the number of the weak classifiers, and $M = N_f \times N_k$. Here, N_f is the number of the features and N_k indicates the number of the kernels. For different feature sets, the decision function is defined as a convex combination,

$$Y(r) = \sum_{m=1}^M \beta_m \sum_{i=1}^H \alpha_i l_i k_m(r, r_i) + \bar{b} \quad (10)$$

where α_i is the Lagrange multiplier while $\bar{b}$ is the bias in the standard SVM algorithm The parameters $\{\alpha_i\}$, $\{\beta_m\}$ and $\bar{b}$ can be learned from a joint optimization process? The conventional function for the MKL method. The proposed boosting algorithm instead of the simple combination of multi-kernel SVMs in the MKL method.

$$Y(r) = \sum_{m=1}^M \beta_m \left(\alpha^T k_m(r) + \bar{b}_m \right) \quad (11)$$

By setting the decision function of a multi-kernel SVM the parameters can be learned straightforwardly.

$$y(r) = \sum_{j=1}^J \beta_j z_j(r) \quad (12)$$

In order to compute the parameters β_j, it use the Adaboost method and the parameter J denotes the number of iterations of the boosting process. We consider each SVM as a weak classifier and the final strong classifier Y (r) is the weighted combination of all the weak classifiers.
Starting with uniform weights, $\omega_1(i) = \frac{1}{H}, i = 1,2 \dots H$, for the SVM classifiers, we obtain a set of decision functions $\{z_m(r)\}$, m=1,2….M At the j-th iteration, we compute the classification error for each of the weak classifiers,

$$\epsilon_m = \frac{\sum_{i=1}^{H} \omega(i)|z_m(r_i)|(sgn(-l_i z_m(r_i)) + 1)/2}{\sum_{i=1}^{H} \omega(i)|z_m(r_i)|} \quad (13)$$

where sgn(x) is the sign function, which equals to 1 when x > 0 and −1 otherwise. In addition, update the weight using the following equation,

$$\omega_{j+1}(i) = \frac{\omega_j(i)e^{-\beta_j l_i z_j(r_i)}}{2\sqrt{\epsilon_i(\epsilon_j - 1)}} \quad (14)$$

After J iterations, all the β_j and $z_j(r_i)$ are computed and we have a boosted classifier as the saliency model learned directly from an input image. It applies this strong saliency model to the test samples (based on all the super pixels of an input image), and a pixel-wise saliency map is thus generated.

E. Adaptive Fusion

The proposed weak and strong saliency maps have complementary properties. The weak map is likely to detect fine details and to capture local structural information due to the contrast-based measure. In contrast, the strong map works well by focusing on global shapes for most images except the case when the test background samples have similarity with the positive training set or large differences compared to the negative training set, or vice versa for the test foreground sample. The fusion these two maps by a weighted

$$\mathcal{M} = \sigma\mathcal{M}_s + (1 - \sigma)\mathcal{M}_w \quad (15)$$

where σ is a balance factor for the combination, and σ =0.7 to weigh the strong map more than the weak map, and M is the final saliency map via bootstrap learning.

IV. EXPERIMENTAL RESULTS

The experimental results of saliency detection methods including the proposed algorithms on six benchmark datasets. The ASD dataset, selected from a bigger image database, contains 100 images, and is labeled with pixel-wise ground truth. The THUS dataset consists of 100 images where all images are labeled with pixel-wise ground truth. The SOD dataset is composed of 300 images from the Berkeley segmentation dataset where each one is labeled with salient object boundaries, based on which the pixel level ground truth is built. Some of the images in the SOD dataset include more than one salient object. The SED2 dataset contains 100 images which are labeled with pixel-wise ground truth annotations. It is challenging due to the fact that every image has salient objects. The Pascal-S dataset contains 850 images which are also labeled with pixel-wise ground truth.
1) Precision-Recall Evaluation: The precision is also called the positive predictive value, and it is defined as the ratio of the number of ground-truth pixels retrieved as a salient region to the total number of pixels retrieved as the salient region. The recall rate is also called the sensitivity, and it is defined as the ratio of the number of salient regions retrieved to the total number of ground-truth regions.
2) F-Measure Rate Evaluation: The second evaluation index is the F-measure rate. The F-measure combines the precision and the recall rate for a comprehensive evaluation.

The Precision and Recall (P-R) curve to evaluate all the methods. In addition, we measure the quality of the saliency maps using the F-Measure by adaptively setting a segmentation threshold for binary segmentation. The adaptive threshold is twice the average value of the whole saliency map. Each image is segmented with super pixels and masked out if the mean saliency values are lower than the adaptive threshold. The average precision and recall values are computed based on the generated binary masks and the ground truth while the F-Measure is computed by

$$F_\eta = \frac{(1 + \eta^2) \times \text{Precision} \times \text{Recall}}{\eta^2 \times \text{Precision} \times \text{Recall}} \quad (16)$$

and η^2 is set to 0.3 to weigh precision more than recall.
The proposed weak saliency model may generate less accurate results than several state-of-the-art methods. The performance of the proposed bootstrap learning method hinges on the quality of the weak saliency model.

Figure 1 show the output of standard criteria for evaluating the salient region detection algorithm: precision-recall rate. The precision is also called the positive predictive value, and it is defined as the ratio of the number of ground-truth pixels retrieved as a salient region to the total number of pixels retrieved as the salient region. The recall rate is also called the sensitivity, and it is defined as the ratio of the number of salient regions retrieved to the total number of ground-truth regions. it uses two different approaches to examine the precision-recall rate.
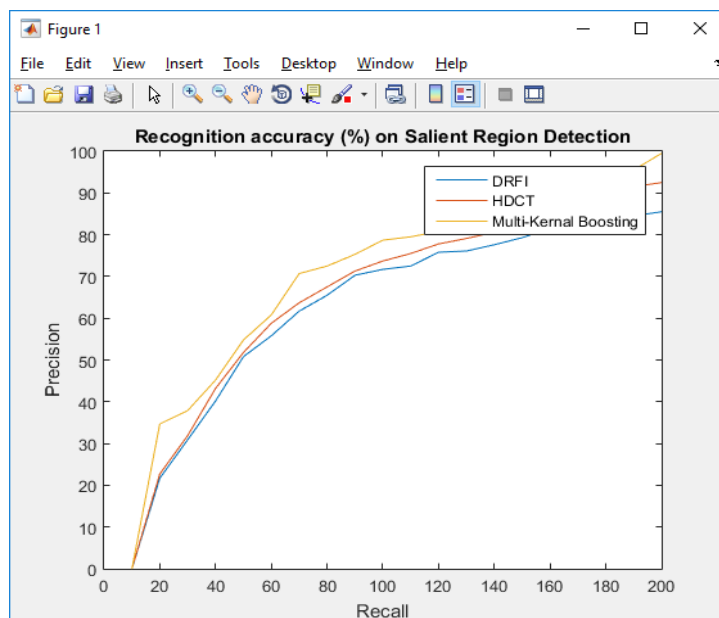
Fig. 1. Comparison of existing and proposed algorithm

V. CONCLUSION

The propose a bootstrap learning model for salient object detection in which both weak and strong saliency models are constructed and integrated. The proposed multi learning process is restricted within multiple scales of the input image and is unsupervised since the training examples for the strong model are determined by a weak saliency map based on contrast and image priors. The strong saliency model is constructed based on the MKB algorithm which combines all the weak classifiers. The novel regional descriptor better describes a region for saliency estimation, the MKB-SVR algorithm effectively integrates the most discriminative features and the corresponding kernels to generate the complementary saliency map, and the adaptive fusion via learning a quality prediction model generates the final saliency map with the higher quality. In the future, we plan to investigate efficient algorithms that incorporate spatial relationships in saliency computation while preserving fine details in the resulting saliency maps. Also, it is desirable to develop saliency detection algorithms to handle cluttered and textured background, which can introduce artifacts to our global histogram based approach (although we did not encounter such images in the database). Finally, it may be beneficial to incorporate high level factors like human faces, symmetry to saliency maps. We believe the proposed saliency maps can be used for efficient object detection, reliable image classification, leading to improved image retrieval.

REFERENCES

[1]   SJiwhan Kim, Dongyoon Han, Yu-Wing Tai and Junmo Kim, "Salient Region Detection via High-Dimensional Color Transform and Local Spatial Support" IEEE transactions on image processing, vol. 25, no. 1, january 2016

[2]   J. Kim, D. Han, Y.-W. Tai, and J. Kim, "Salient region detection via high-dimensional color transform," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2014, pp. 883–890.

[3]   R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1597–1604.

[4]   S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 2376–2383.

[5]   W. Hou, X. Gao, D. Tao, and X. Li, "Visual saliency detection using information divergence," Pattern Recognit., vol. 46, no. 10, pp. 2658–2669, Oct. 2013.

[6]   H. Jiang, J. Wang, Z. Yuan, T. Liu, and N. Zheng, "Automatic salient object segmentation based on context and shape prior," in Proc. Brit. Mach. Vis. Conf. (BMVC), 2011, pp. 110.1–110.12.

[7]   A. Borji and L. Itti, "Exploiting local and global patch rarities for saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 478–485.

[8]   Z. Liu, R. Shi, L. Shen, Y. Xue, K. N. Ngan, and Z. Zhang, "Unsupervised salient object segmentation based on kernel density estimation and two-phase graph cut," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1275–1289, Aug. 2012.

[9]   H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li, "Salient object detection: A discriminative regional feature integration approach," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2013, pp. 2083–2090.

[10]   C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2013, pp. 3166–3173.

[11]   X. Li, H. Lu, L. Zhang, X. Ruan, and M.-H. Yang, "Saliency detection via dense and sparse reconstruction," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2013, pp. 2976–2983.