



Saliency detection via double nuclear norm maximization and ensemble manifold regularization[☆]



Bing Liu^{a,b,c}, Yong Zhou^{a,b,*}, Peng Liu^{d,*}, Wei Sun^{e,*}, Shuang Li^f, Xinqiu Fang^g

^a School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, 221116, Jiangsu province, China

^b Mine Digitization Engineering Research Center of Ministry of Education, China

^c Institute of Electrics, Chinese Academy of Sciences, Beijing, 100190, China

^d Internet of Things Perception Mine Research Center, China University of Mining and Technology, Xuzhou 221008, Jiangsu, China

^e School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, 221116, Jiangsu province, China

^f School of Management, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China

^g School of Mines, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China

ARTICLE INFO

Article history:

Received 10 November 2018

Received in revised form 13 July 2019

Accepted 13 July 2019

Available online 17 July 2019

Keywords:

Salient object detection

Low-rank matrix recovery

Ensemble manifold regularization

Double nuclear norm penalty

Alternating direction method of multipliers

(ADMM)

ABSTRACT

In recent years, salient object detection via robust principal component analysis (RPCA) has received a significant amount of attention. Existing methods generally replace the rank function by the nuclear norm to obtain low-rank and sparse matrices, ignores the heavy-tailed distributions of singular values and over-penalizes large singular values of low rank matrices. In addition, although the manifold regularization is introduced into RPCA to obtain satisfactory low-rank representation of an original image, the graph hyperparameters selection lacks the ability to approximate the optimal solution for intrinsic manifold estimation. To solve these issues, we propose a novel low-rank matrix recovery model for salient object detection, which integrates double nuclear norm maximization with ensemble manifold regularization and can be formulated as a tractable optimization problem. By virtue of the alternating direction method (ADM), we develop an efficient algorithm to optimize the proposed model, which not only effectively fits the heavy-tailed distribution of singular values of low-rank matrices, but automatically learns the optimal linear combination of a set of predefined graph Laplacians. Experimental results on five challenging datasets show that our model achieves better performance than state-of-the-art unsupervised methods.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Salient object detection aims to accurately segment the objects of interest from a single scene, or perform co-saliency detection from multiple images [1,2]. In recent years, it has become a hot topic in computer vision due to its high potential for use in multimedia applications, such as image segmentation [3,4], content-based image retrieval [5], image compression [6], image cropping [7], etc. According to whether labeled information is used or not, salient object detection algorithms can be roughly divided into supervised and unsupervised methods. Although supervised methods, especially those based on deep learning [8–10], can generally achieve satisfactory performance, they heavily

rely on a large amount of labeled training images. By contrast, unsupervised methods have the advantage of the greater flexibility based on high-level prior knowledge or low-level saliency [11–15]. In this paper, we mainly discuss unsupervised methods.

Recently, a lot of salient object detection algorithms have been developed by integrating bottom-up low-level saliency cues (color, gradient, boundaries, etc.) with top-down high-level prior knowledge [16–30]. Achanta et al. [11] constructed a frequency-tuned model that computes the saliency map based on color contrast prior. The methods, proposed in [31,32], effectively combined the center prior with low-level features to localize saliency objects. Besides the center and color priors, the background prior has also been incorporated into saliency detection models. For instance, Wei et al. [26] measured saliency by virtue of the geodesic distance to the boundary regions. Shen et al. [28] introduced low-level features and top-down priors into a unified model based on robust principal component analysis (RPCA). Lang et al. [29] integrated multiple priors into region-level saliency generation. Tang et al. [31] made use of the color, location and boundary connectivity to infer the likelihood that each image region belongs to the background. Huo et al. [13] built a specialized linear feedback

[☆] No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.knosys.2019.07.021>.

* Corresponding authors.

E-mail addresses: yzhou@cumt.edu.cn (Y. Zhou), liupeng@cumt.edu.cn (P. Liu), zc@cumt.edu.cn (W. Sun).

control system model, in which a saliency map is generated by inputting multiple saliency cues and image features. Liu et al. [2] proposed a nonparametric saliency detection model based on kernel density estimation (KDE), where likelihood measures of pixels and saliency measures are used to compute the pixel-wise saliency map. Goferman et al. [12] developed a context-aware salient region extraction method based on local and global contrast. Then, the multi-scale contrast, center-surround histogram and color spatial distribution were introduced into the conditional random field model to improve the quality of the saliency map [33]. In addition, the global contrast of the features extracted from the transform domain, such as the wavelet and Fourier transform [34,35], has also been applied in saliency detection.

Among the salient object detection approaches mentioned above, the low-rank matrix recovery based methods have attracted considerable attention owing to its efficiency and robustness. Motivated by the recent efforts in both robust principal component analysis (RPCA) and manifold learning, this paper proposes a novel salient object detection model based on the Schatten-1/2 quasi-norm and ensemble manifold regularization. It fuses multiple graph Laplacians to select the optimal intrinsic manifold for the proposed model. Furthermore, to enhance the performance of salient object detection, the Schatten-1/2 norm, instead of the nuclear norm, is introduced to fit the heavy-tailed distribution of singular values of low-rank matrices. The low-rank and structured sparse matrix decomposition model for salient object detection (SMD) is closely relevant to our proposed method [36]. However, it has two disadvantages: (1) it still relaxes the rank function by the nuclear norm, which has negative effects on the performance of salient object detection; (2) it lacks the ability to approximate the optimal solution for intrinsic manifold estimation. The proposed method can avoid overpenalizing large singular values of low-rank matrices via the Schatten-1/2 quasi-norm. Thus, the first disadvantage of SMD is addressed. On the other hand, our method incorporates ensemble manifold regularization to automatically learn the optimal intrinsic manifold. Therefore, the second disadvantage is addressed.

The main contributions of this work are listed below:

(1) By introducing the hyper-Laplacian prior and ensemble manifold regularization, we propose a novel low-rank recovery model for salient object detection, which is formulated as the tractable and scalable optimization problem.

(2) We derive an efficient optimization algorithm by virtue of the alternating direction method (ADM). The proposed algorithm only requires SVDs on two much smaller factor matrices and learns both the low-rank matrix and the optimal intrinsic manifold jointly.

(3) We apply the proposed model to salient object detection, and the experimental results validate that our algorithm outperforms the state-of-the-art methods on the benchmark datasets.

The remainder of this paper is organized as follows. We briefly review some related works in Section 2. The framework of salient object detection and the ADM-based algorithm for low-rank matrix recovery are proposed in Section 3. Section 4 shows the experimental results, including a thorough comparison with recently proposed unsupervised salient object detection algorithms. Finally, Section 5 concludes this paper.

2. Related work

Salient object detection based on unsupervised low-rank matrix recovery has been an active research topic in the recent years [37–47]. Compared with supervised approaches, they do not require a large amount of training beforehand and have advantages of flexibility and scalability. In this paper, we mainly discuss the salient object detection methods based on low-rank matrix recovery.

2.1. Salient object detection via convex nuclear norm minimization

In the low-rank matrix recovery theory, given the feature matrix $\mathbf{F} \in \mathbb{R}^{m \times n}$ ($m \geq n$) of an input image, it can be decomposed into two parts: a low-rank matrix $\mathbf{L} \in \mathbb{R}^{m \times n}$ (corresponding to background regions) and a sparse matrix $\mathbf{S} \in \mathbb{R}^{m \times n}$ (corresponding to foreground object regions). The low-rank matrix recovery model for salient object detection can be formulated as a non-convex optimization problem [28]:

$$\min_{\mathbf{L}, \mathbf{S}} \text{rank}(\mathbf{L}) + \lambda \|\mathbf{S}\|_1 \quad \text{s.t. } \mathbf{F} = \mathbf{L} + \mathbf{S}, \quad (1)$$

where $\text{rank}(\cdot)$ denotes the matrix rank function, $\|\cdot\|_1$ is the ℓ_1 -norm which promotes sparsity and $\lambda > 0$ is the parameter tradeoff. Since solving (1) is NP-hard, $\text{rank}(\mathbf{L})$ is usually replaced by the nuclear norm. Thus, the problem (1) is reformulated as follows [29]:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{L}\|_* + \lambda \|\mathbf{S}\|_1 \quad \text{s.t. } \mathbf{F} = \mathbf{L} + \mathbf{S}, \quad (2)$$

where $\|\cdot\|_*$ represents the nuclear norm of a matrix, defined as the sum of the singular values of a matrix, which is a convex relaxation of the matrix rank function $\text{rank}(\cdot)$. Finally, a saliency score is evaluated for each patch of the input image, and the patch corresponding to the large score has a high probability to be salient.

Subsequently, the unified low-rank matrix recovery model (ULR) was proposed in [28], which utilized a linear transformation on the original feature space and high-level priors to improve the performance of salient object detect. In addition, the segmentation priors were introduced in [30] to guide the low-rank matrix recovery. Different from traditional LR models, low-rank representation (LRR) was originally developed to reconstruct a low rank matrix by self-representation and the $\ell_{2,1}$ -norm [38]. Lang et al. [29] extended low-rank representation (LRR) to the multi-task learning model, which incorporates multiple features into the generation of saliency maps. However, these models do not consider the inter-correlation between elements in the sparse matrix \mathbf{S} . To address this issue, a structured matrix decomposition model was proposed to boost the quality of saliency maps [32]. But, this model still substituted the nuclear norm minimization for the matrix rank function, which inevitably performs iterative singular value decomposition (SVD) and results in a biased solution.

2.2. Non-convex models for RPCA

Nie et al. [39] reformulated RPCA as the joint ℓ_p -norm and Schatten- q norm minimization problem, which can be solved by an efficient augmented Lagrange multiplier method. Lai et al. [40] and Lu et al. [37] replaced the rank functions by the non-convex Schatten quasi-norms and developed iteratively reweighted least squares methods to solve optimization problems. However, these algorithms have to perform SVDs iteratively, which is time-consuming. Although some partial SVD methods were proposed to reduce time cost, it was validated that they cannot achieve satisfactory performance in many real applications [41,42].

To address this issue, Shen et al. [43] and Jiang et al. [44] factorized the low-rank matrix \mathbf{L} of RPCA into two smaller factor matrices, i.e., $\mathbf{L} = \mathbf{U}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m \times d}$, $\mathbf{V} \in \mathbb{R}^{n \times d}$, and usually $d \ll \min(m, n)$. In [45,46], Problem (2) was transformed into a smaller matrix nuclear norm minimization problem by further imposing the column-orthonormal constraints on \mathbf{U} . In [47], the matrix tri-factorization model was designed to solve RPCA, which factorizes the low-rank matrix \mathbf{L} into three smaller factor matrices, i.e., $\mathbf{L} = \mathbf{U}\mathbf{W}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m \times d}$, $\mathbf{W} \in \mathbb{R}^{d \times d}$ and $\mathbf{V} \in \mathbb{R}^{n \times d}$.

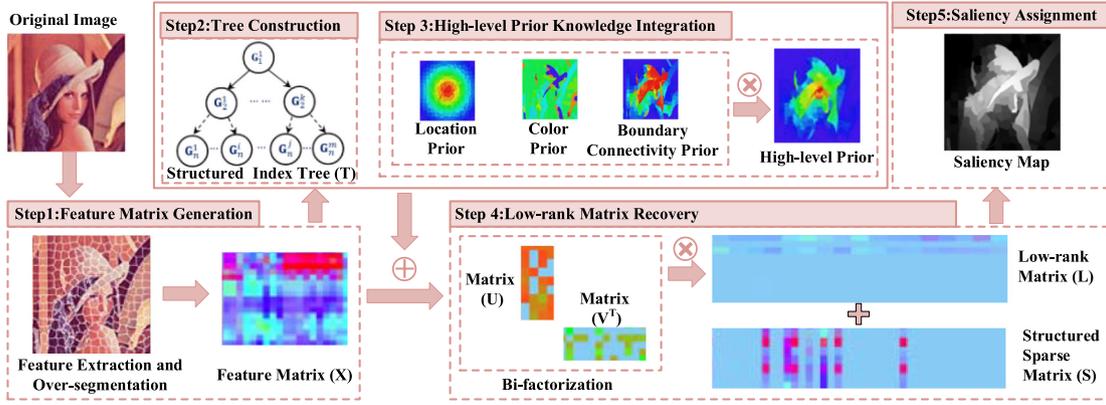


Fig. 1. The proposed salient object detection framework.

3. Saliency detection via low-rank matrix recovery and top-down visual features

In this section, we integrate the low-rank matrix recovery model and top-down visual features to construct the framework for salient object detection. Our method not only makes a closer approximation by fitting the heavy-tailed distribution of singular values of low-rank matrices, but more effectively segments the object regions from background by ensemble manifold regularization. In addition, our algorithm adopts the top-down technique to exploit high-level prior knowledge. Fig. 1 shows the proposed salient object detection framework.

As can be seen from Fig. 1, the proposed framework includes five steps: feature matrix generation, index tree construction, high-level prior knowledge integration, low-rank matrix recovery and saliency map generation. Details are given in Fig. 1.

3.1. Feature matrix generation and index tree construction

For fair comparison, following [28,32], we transform an original image into a feature matrix. Firstly, a 53-dimension feature vector is generated by extracting the low-level features based on RGB color, Gabor filter and steerable pyramids [32]. Secondly, the simple linear iterative clustering (SLIC) algorithm [48] is carried out to over-segment the input image into N nonoverlapping superpixels $\mathbf{P} = \{P_1, P_2, \dots, P_N\}$. Thirdly, each patch P_i is represented as a D -dimension feature vector \mathbf{x}_i by the low-level feature representation. Finally, the feature matrix $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{D \times N}$ is composed of extracted N feature vectors, and its affinity matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ is defined as [32]

$$W_{i,j} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), & \text{if } (P_i, P_j) \in \Omega, \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where Ω represents the set of adjacent superpixel pairs.

The non-structural sparsity-inducing norms used in RPCA, such as the ℓ_1 -norm, ℓ_2 -norm and $\ell_{2,1}$ -norm, induce the sparsity of the columns of the sparse matrix \mathbf{S} independently, which ignore the spatial contiguity among the superpixels [49]. To solve this problem, we construct an index tree \mathbf{T} to induce the structural sparsity of the matrix \mathbf{S} . Specifically, based on the affinity matrix \mathbf{W} , a graph-based image segmentation algorithm [50] is used to partition superpixels into different groups. By adjusting an affinity threshold \mathcal{T} , we can obtain hierarchically fine-to-coarse segmentation results of the input image, and assign the segment of each granularity to the node at the corresponding layer in the index tree.

As shown in Fig. 2, an index tree \mathbf{T} with depth 3 over indices $\{1, 2, \dots, 7\}$ consists of hierarchical nodes, where each node contains a set of the indices of the superpixels, and G_i^j denotes

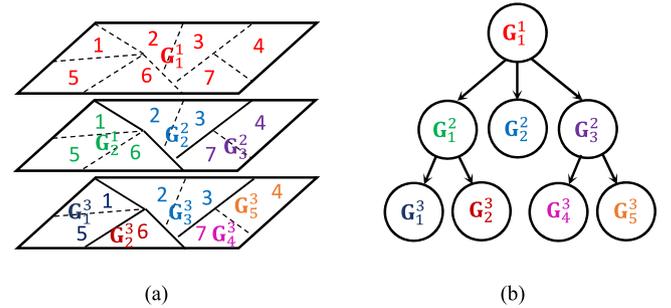


Fig. 2. The construction of an index tree based on superpixels. (a) The hierarchical segmentation of an input image. (b) An index tree corresponding to (a).

the j th node at the i th level. According to [32], the index tree-structured sparsity-inducing norm can more precisely represent hierarchical structure of image patches, which contributes to the completeness of saliency maps. In our model, we introduce a weighted variant defined as follows [32]:

$$\mathcal{T}(\mathbf{S}) = \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_i^j}\|_{\infty}, \quad (4)$$

where n_i is the number of nodes at the i th level and $w_j^i \geq 0$ represents the weight corresponding to the node G_i^j , $\mathbf{S}_{G_i^j} \in \mathbb{R}^{D \times |G_i^j|}$ ($|\cdot|$ denotes set cardinality) is the sub-matrix of \mathbf{S} corresponding to the node G_i^j , $\|\mathbf{S}_{G_i^j}\|_{\infty}$ represents that the saliency value of G_i^j is calculated by the ℓ_{∞} -norm, i.e., the maximum saliency value of the superpixels in G_i^j . Thus, the pixels within the same group have similar saliency values. Consequently, Eq. (4) tends to induce structurally consistent saliency maps. In addition, w_j^i can be integrated with high-level prior knowledge to further improve the quality of saliency maps.

3.2. High-level prior knowledge integration

To effectively improve the quality of saliency maps, we incorporate multiple high-level priors (location, color and boundary connectivity priors) into our models. To begin with, we generate a location prior map by calculating the distances of the pixels from the image center, since salient objects are generally located near the image center [6]. Second, we directly use the same color prior utilized in [32]. Third, the boundary connectivity prior is also incorporated into our framework, and the corresponding prior map is constructed by computing the probabilities of different

image regions connected to the borders of the image [51]. Finally, we produce the high-level prior map by multiplying three high-level priors together (see Fig. 1). Specifically, the high-level prior of each patch $P_i \in \mathbf{P}$, denoted as h_i , is normalized to $[0, 1]$, which represents the probability of P_i belonging to a salient object. In practice, h_i can be used to assign a weight value to the corresponding patch. Thus, it can be naturally integrated into the tree-structured sparsity-inducing norm. Following [32], we define w_j^i as

$$w_j^i = 1 - \max(\{h_m : m \in G_j^i\}). \quad (5)$$

If one node of the index tree is assigned a small weight w_j^i , the corresponding saliency value would be inclined to be large. Furthermore, the superpixels from the same node tend to have identical saliency values via the tree-structured sparsity-inducing norm. Thus, salient objects can be highlighted more accurately by means of high-level priors.

3.3. Low-rank matrix recovery model based on Schatten-1/2 quasi-norm and ensemble manifold regularization

Combined with the index tree, high-level priors and ensemble manifold regularization, a novel low-rank matrix recovery model is proposed to effectively decompose the input feature matrix into the low-rank matrix \mathbf{L} and the structured-sparse matrix \mathbf{S} . Compared with existing methods, the proposed model can fit the heavy-tailed distribution of singular values much better, which contributes to improve the performance of saliency detection.

3.3.1. Ensemble manifold regularization

The manifold regularization technique is common used to enhance the performance of the salient object detection [52,53]. Specifically, the manifold regularization term $\mathcal{R}(\mathbf{S})$ to measure the smoothness of vectors in the sparse matrix \mathbf{S} is defined as [54]

$$\mathcal{R}(\mathbf{S}) = \frac{1}{2} \sum_{i,j=1}^N \|\mathbf{s}_i - \mathbf{s}_j\|_2^2 w_{i,j} = \text{Tr}(\mathbf{S} \mathbf{L}_g \mathbf{S}^T), \quad (6)$$

where \mathbf{s}_i is the i th column of \mathbf{S} , $\mathbf{L}_g \in \mathbb{R}^{N \times N}$ represents the graph Laplacian, i.e., $\mathbf{L}_g = \mathbf{D} - \mathbf{W}$ (\mathbf{D} denotes degree matrix) and $\text{Tr}(\cdot)$ is the trace of a matrix.

According to Ref. [32], the Laplacian regularization term can effectively smooth the columns in \mathbf{S} , which in essence enlarges the distance between subspaces induced by \mathbf{L} and \mathbf{S} . However, the hyperparameters for creating the data adjacency graph are empirically specified. Our framework for salient object detection aims to automatically learn the optimal graph Laplacian. Therefore, it is essential to introduce ensemble manifold regularization, which assumes that the intrinsic manifold lies in the convex hull of the predefined manifold candidates. Specifically, given a set of pre-given graph Laplacians $\{\mathbf{L}_g^1, \mathbf{L}_g^2, \dots, \mathbf{L}_g^K\}$, the ensemble manifold is formulated as

$$\mathbf{L}_g^E = \sum_{i=1}^K \tau_i \mathbf{L}_g^i, \text{ s.t. } \sum_{i=1}^K \tau_i = 1, \tau_i \geq 0, \text{ for } i = 1, \dots, K, \quad (7)$$

By replacing Eq. (6) by Eq. (7), the ensemble manifold regularization term is formulated as

$$\mathcal{R}^E(\mathbf{S}) = \beta \text{Tr} \left(\mathbf{S} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{S}^T \right) + \gamma \|\boldsymbol{\tau}\|^2, \quad (8)$$

where the $\|\boldsymbol{\tau}\|^2$ term prevents the parameter $\boldsymbol{\tau}$ from overfitting to one manifold, β and γ are tradeoff parameters.

It is worth noting that ensemble manifold regularization is a kind of ensemble learning methods. Consequently, diversified

ensemble learning schemes can be introduced into ensemble manifold regularization in the supervised or semi-supervised cases [55]. In this paper, we mainly discuss the unsupervised low-rank matrix recovery model based on ensemble manifold regularization, and the corresponding models with other cases are left to future work.

3.3.2. The proposed low-rank matrix recovery model

Since the singular values of nonlocal matrices in natural images usually exhibit a heavy-tailed distribution [36], existing methods, based on the nuclear norm (i.e., the Schatten-1 norm), overpenalize large singular values of low-rank matrices, which results in a biased solution. In addition, the nuclear norm minimization generally involves iterative singular value decomposition (SVD), which limits the scalability for large scale feature matrices. Consequently, it is essential to approximate the rank function by using the Schatten- q quasi-norm i.e., imposing the ℓ_q -norm ($0 < q < 1$) on singular values. However, the use of such hyper-Laplacian distributions makes the optimization problems non-convex, non-smooth and non-Lipschitz. Fortunately, in the special case, such as $q = 1/2$, the Schatten- q quasi-norm minimization can be transformed into the equivalent, tractable and scalable form [36].

For completeness, we introduce the double nuclear norm defined as follows:

Definition 1 ([34]). For any matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ of rank at most $r \leq d$, we decompose it into two factor matrices $\mathbf{U} \in \mathbb{R}^{m \times d}$ and $\mathbf{V} \in \mathbb{R}^{n \times d}$ such that $\mathbf{X} = \mathbf{UV}^T$. Then the double nuclear norm penalty of \mathbf{X} is defined as

$$\|\mathbf{X}\|_{\text{D-N}} = \min_{\mathbf{U}, \mathbf{V}: \mathbf{X} = \mathbf{UV}^T} \frac{1}{4} (\|\mathbf{U}\|_* + \|\mathbf{V}\|_*)^2. \quad (9)$$

Analogous to the well-known Schatten- q quasi-norm [39,40], the double nuclear norm is a kind of quasi-norms, as stated in the following theorems.

Theorem 1 ([34]). The double nuclear norm penalty $\|\cdot\|_{\text{D-N}}$ is

$$\|\mathbf{X}\|_{\text{D-N}} = \|\mathbf{X}\|_{S_{1/2}}, \quad (10)$$

where $\|\cdot\|_{S_{1/2}}$ denotes the Schatten-1/2 quasi-norm.

Since the singular values of nonlocal matrices in natural images and scenes usually exhibit a heavy-tailed distribution, it is feasible to replace the nuclear norm of \mathbf{L} by the Schatten-1/2 norm, which makes a closer approximation to the rank function. By introducing the structured-sparsity and ensemble manifold regularization, the Schatten- q quasi-norm based low-rank matrix recovery model for salient object detection is formulated as follows:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{L}\|_{S_{1/2}}^{1/2} + \alpha \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_{\infty} + \beta \text{Tr} \left(\mathbf{S} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{S}^T \right) + \gamma \|\boldsymbol{\tau}\|^2 \quad (11)$$

s.t. $\mathbf{L} + \mathbf{S} = \mathbf{X}$, $\sum_{i=1}^K \tau_i = 1$, $\tau_i \geq 0$, for $i = 1, \dots, K$, where the feature matrix $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{D \times N}$, $\|\mathbf{L}\|_{S_{1/2}}^{1/2}$ denotes the Schatten-1/2 quasi-norm of \mathbf{L} , which makes the solution closer to the original solution and avoids over-penalizing large singular values. $\sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_{\infty}$ is the structured sparsity regularization term, which reflects the spatial relations of elements in \mathbf{S} . The ensemble manifold term, $\beta \text{Tr} \left(\mathbf{S} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{S}^T \right) + \gamma \|\boldsymbol{\tau}\|^2$, is used to effectively separate the sparse matrix \mathbf{S} from the low-rank matrix \mathbf{L} , and α , β and γ are positive tradeoff parameters.

According to Eqs. (9) and (10), the optimization problem (11) can be reformulated as the following tractable and scalable optimization problem:

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{S}, \mathbf{L}} \frac{1}{2} (\|\mathbf{U}\|_* + \|\mathbf{V}\|_*) + \alpha \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_\infty + \beta \text{Tr} \left(\mathbf{S} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{S}^T \right) + \gamma \|\boldsymbol{\tau}\|^2 \quad (12)$$

s.t. $\mathbf{L} + \mathbf{S} = \mathbf{X}$, $\mathbf{UV}^T = \mathbf{L}$, $\sum_{i=1}^K \tau_i = 1$, $\tau_i \geq 0$, for $i = 1, \dots, K$, where $\mathbf{U} \in \mathbb{R}^{D \times r}$ and $\mathbf{V} \in \mathbb{R}^{N \times r}$, $r \ll \min(m, n)$ is an upper bound on the rank of the matrix \mathbf{L} .

3.3.3. Alternating optimization algorithm

The optimization problem (12) can be efficiently solved by means of the ADM method [56]. By introducing the auxiliary variables, the optimization problem (12) can be transformed into the following equivalent optimization problem:

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{S}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{E}, \boldsymbol{\tau}} \frac{1}{2} (\|\hat{\mathbf{U}}\|_* + \|\hat{\mathbf{V}}\|_*) + \alpha \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_\infty + \beta \text{Tr} \left(\mathbf{E} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{E}^T \right) + \gamma \|\boldsymbol{\tau}\|^2 \quad (13)$$

s.t. $\hat{\mathbf{U}} = \mathbf{U}$, $\hat{\mathbf{V}} = \mathbf{V}$, $\mathbf{UV}^T + \mathbf{S} = \mathbf{X}$, $\mathbf{S} = \mathbf{E}$, $\sum_{i=1}^K \tau_i = 1$, $\tau_i \geq 0$, for $i = 1, \dots, K$,

which is non-convex and can be efficiently solved by the alternating direction method. Specifically, $(\mathbf{U}, \mathbf{V}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{S}, \mathbf{E})$ and $\boldsymbol{\tau}$ are alternately optimized in each iteration. Because the proposed method integrates double nuclear norm with ensemble manifold regularization, we name our model DNN-EMR, which iteratively performs matrix factorization and learns the optimal composite manifold. The optimization procedure of DNN-EMR is described in Algorithm 1.

(1) On Optimizing $(\mathbf{U}, \mathbf{V}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{S}, \mathbf{E})$

By fixing $\boldsymbol{\tau}$, the optimization problem (13) is transformed into

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{S}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{E}, \boldsymbol{\tau}} \frac{1}{2} (\|\hat{\mathbf{U}}\|_* + \|\hat{\mathbf{V}}\|_*) + \alpha \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_\infty + \beta \text{Tr} \left(\mathbf{E} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{E}^T \right) \quad (14)$$

s.t. $\hat{\mathbf{U}} = \mathbf{U}$, $\hat{\mathbf{V}} = \mathbf{V}$, $\mathbf{UV}^T + \mathbf{S} = \mathbf{X}$, $\mathbf{S} = \mathbf{E}$.

The corresponding augmented Lagrangian function to the problem (14) is given by

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \mathbf{V}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{S}, \mathbf{E}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3, \mathbf{Y}_4, \mu) &= \frac{1}{2} (\|\hat{\mathbf{U}}\|_* + \|\hat{\mathbf{V}}\|_*) \\ &+ \alpha \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_\infty + \beta \text{Tr} \left(\mathbf{E} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{E}^T \right) \\ &+ \text{Tr}(\mathbf{Y}_1^T (\hat{\mathbf{U}} - \mathbf{U})) + \text{Tr}(\mathbf{Y}_2^T (\hat{\mathbf{V}} - \mathbf{V})) \\ &+ \text{Tr}(\mathbf{Y}_3^T (\mathbf{UV}^T + \mathbf{S} - \mathbf{X})) + \text{Tr}(\mathbf{Y}_4^T (\mathbf{S} - \mathbf{E})) \frac{\mu}{2} (\|\hat{\mathbf{U}} - \mathbf{U}\|_F^2 \\ &+ \|\hat{\mathbf{V}} - \mathbf{V}\|_F^2 + \|\mathbf{UV}^T + \mathbf{S} - \mathbf{X}\|_F^2 + \|\mathbf{S} - \mathbf{E}\|_F^2), \end{aligned} \quad (15)$$

where $\mathbf{Y}_1 \in \mathbb{R}^{D \times r}$, $\mathbf{Y}_2 \in \mathbb{R}^{N \times r}$ and $\mathbf{Y}_3, \mathbf{Y}_4 \in \mathbb{R}^{D \times N}$ are Lagrange multipliers, and $\mu > 0$ is the penalty parameter. The general ADM method minimizes \mathcal{L} with respect to $\mathbf{U}, \mathbf{V}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{S}, \mathbf{E}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3$ and \mathbf{Y}_4 alternately in each iteration.

(2) Updating \mathbf{U}^{k+1} and \mathbf{V}^{k+1} . By keeping other variables fixed, \mathbf{U}^{k+1} and \mathbf{V}^{k+1} can be updated by the following optimization

subproblems (16) and (17), respectively.

$$\min_{\mathbf{U}} \|\hat{\mathbf{U}}^k - \mathbf{U} + \mathbf{Y}_1^k / \mu^k\|_F^2 + \|\mathbf{U} (\mathbf{V}^k)^T + \mathbf{S}^k - \mathbf{X} + \mathbf{Y}_3^k / \mu^k\|_F^2, \quad (16)$$

$$\min_{\mathbf{V}} \|\hat{\mathbf{V}}^k - \mathbf{V} + \mathbf{Y}_2^k / \mu^k\|_F^2 + \|\mathbf{U}^{k+1} \mathbf{V}^T + \mathbf{S}^k - \mathbf{X} + \mathbf{Y}_3^k / \mu^k\|_F^2. \quad (17)$$

The optimal solutions of the problems (16) and (17) are derived by solving two least squares problems as follows:

$$\mathbf{U}^{k+1} = (\hat{\mathbf{U}}^k + \mathbf{Y}_1^k / \mu^k + \mathbf{Q}^k \mathbf{V}^k) (\mathbf{I}_r + (\mathbf{V}^k)^T \mathbf{V}^k)^{-1}, \quad (18)$$

$$\mathbf{V}^{k+1} = (\hat{\mathbf{V}}^k + \mathbf{Y}_2^k / \mu^k + (\mathbf{Q}^k)^T \mathbf{U}^{k+1}) (\mathbf{I}_r + (\mathbf{U}^{k+1})^T \mathbf{U}^{k+1})^{-1}, \quad (19)$$

where $\mathbf{Q}^k = \mathbf{X} - \mathbf{S}^k - \mathbf{Y}_3^k / \mu^k$ and \mathbf{I}_r denotes an identity matrix of size $r \times r$.

(3) Updating $\hat{\mathbf{U}}^{k+1}$ and $\hat{\mathbf{V}}^{k+1}$. To update the variables $\hat{\mathbf{U}}^{k+1}$ and $\hat{\mathbf{V}}^{k+1}$, we fix the other variables and solve the following optimization subproblem:

$$\min_{\hat{\mathbf{U}}} \frac{1}{2} \|\hat{\mathbf{U}}\|_* + \frac{\mu^k}{2} \|\hat{\mathbf{U}} - \mathbf{U}^{k+1} + \mathbf{Y}_1^k / \mu^k\|_F^2, \quad (20)$$

$$\min_{\hat{\mathbf{V}}} \frac{1}{2} \|\hat{\mathbf{V}}\|_* + \frac{\mu^k}{2} \|\hat{\mathbf{V}} - \mathbf{V}^{k+1} + \mathbf{Y}_2^k / \mu^k\|_F^2. \quad (21)$$

According to Ref. [56], the least squares problems (20) and (21) based on nuclear norm regularization have closed-form solutions by the SVT operator as follows:

$$\hat{\mathbf{U}}^{k+1} = \text{SVT}_{1/(2\mu^k)} (\mathbf{U}^{k+1} - \mathbf{Y}_1^k / \mu^k), \quad (22)$$

$$\hat{\mathbf{V}}^{k+1} = \text{SVT}_{1/(2\mu^k)} (\mathbf{V}^{k+1} - \mathbf{Y}_2^k / \mu^k). \quad (23)$$

(4) Updating \mathbf{E}^{k+1} and \mathbf{S}^{k+1} . Fixing all other variables, \mathbf{E}^{k+1} can be updated by solving the following subproblem:

$$\begin{aligned} \mathbf{E}^{k+1} &= \text{argmin}_{\mathbf{E}} \beta \text{Tr} \left(\mathbf{E} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{E}^T \right) + \text{Tr} \left((\mathbf{Y}_4^k)^T (\mathbf{S}^k - \mathbf{E}) \right) \\ &+ \frac{\mu^k}{2} \|\mathbf{S}^k - \mathbf{E}\|_F^2. \end{aligned} \quad (24)$$

By taking derivative of the objective function in Eq. (24), the solution can be calculated as follows:

$$\mathbf{E}^{k+1} = (\mu^k \mathbf{S}^k + \mathbf{Y}_4^k) \left(2\beta \sum_{i=1}^K \tau_i \mathbf{L}_g^i + \mu^k \mathbf{I} \right)^{-1}. \quad (25)$$

After updating \mathbf{E}^{k+1} , we update \mathbf{S}^{k+1} with other variables fixed and obtain the following tree-structured sparsity optimization subproblem:

$$\mathbf{S}^{k+1} = \text{arg min}_{\mathbf{S}} \lambda \sum_{i=1}^d \sum_{j=1}^{n_i} w_j^i \|\mathbf{S}_{G_j^i}\|_\infty + \frac{1}{2} \|\mathbf{S} - \mathbf{K}_1\|_F^2, \quad (26)$$

where $\mathbf{K}_1 = (\mathbf{X} - \mathbf{U}^{k+1} (\mathbf{V}^{k+1})^T + \mathbf{E}^{k+1} - (\mathbf{Y}_3^k + \mathbf{Y}_4^k) / \mu^k) / 2$ and $\lambda = \alpha / (2\mu^k)$. The problem (26) can be solved by the hierarchical proximal operator [57], as outlined in Algorithm 2.

(5) On optimizing $\boldsymbol{\tau}$. To update the variable $\boldsymbol{\tau}$, we fix the other variables and solve the following optimization subproblem:

$$\min_{\boldsymbol{\tau}} \text{Tr} \left(\mathbf{E} \sum_{i=1}^K \tau_i \mathbf{L}_g^i \mathbf{E}^T \right) + \gamma / \beta \|\boldsymbol{\tau}\|^2 \quad (27)$$

s.t. $\sum_{i=1}^K \tau_i = 1$, $\tau_i \geq 0$, for $i = 1, \dots, K$.

The optimization subproblem (27) is a constrained quadric programming problem, which can be solved by the quadric optimization solver in matlab toolbox.

Algorithm 1. ADMM for Solving Problem (13)

Input: Feature matrix \mathbf{X} , parameters α, β and γ , index tree $\mathbf{T} = \{G_j^i\}$, tree node weight w_j^i , the rank upper bound r, K candidate graph Laplacians.

1: Initialize $\mathbf{U}^0 = \mathbf{0}, \mathbf{V}^0 = \text{eye}(N, r), \mathbf{S}^0 = \mathbf{0}, \mathbf{E}^0 = \mathbf{0}, \mathbf{Y}_1^0 = \mathbf{0}, \mathbf{Y}_2^0 = \mathbf{0}, \mathbf{Y}_3^0 = \mathbf{0}, \mathbf{Y}_4^0 = \mathbf{0}, \rho = 1.1, \mu^0 = 0.1, \mu_{max} = 10^{10}, \varepsilon = 10^{-5}, \tau_i = 1/K$ and $k = 0$.

2: **While** not converged **do**

3: **While** not converged **do**

4: $\mathbf{U}^{k+1} = \text{argmin}_{\mathbf{U}} \mathcal{L}(\mathbf{U}, \mathbf{V}^k, \mathbf{S}^k, \mathbf{E}^k, \hat{\mathbf{U}}^k, \hat{\mathbf{V}}^k, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

5: $\mathbf{V}^{k+1} = \text{argmin}_{\mathbf{V}} \mathcal{L}(\mathbf{U}^{k+1}, \mathbf{V}, \mathbf{S}^k, \mathbf{E}^k, \hat{\mathbf{U}}^k, \hat{\mathbf{V}}^k, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

6: $\hat{\mathbf{U}}^{k+1} = \text{argmin}_{\hat{\mathbf{U}}} \mathcal{L}(\mathbf{U}^{k+1}, \mathbf{V}^{k+1}, \mathbf{S}^k, \mathbf{E}^k, \hat{\mathbf{U}}, \hat{\mathbf{V}}^k, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

7: $\hat{\mathbf{V}}^{k+1} = \text{argmin}_{\hat{\mathbf{V}}} \mathcal{L}(\mathbf{U}^{k+1}, \mathbf{V}^{k+1}, \mathbf{S}^k, \mathbf{E}^k, \hat{\mathbf{U}}^{k+1}, \hat{\mathbf{V}}, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

8: $\mathbf{E}^{k+1} = \text{argmin}_{\mathbf{E}} \mathcal{L}(\mathbf{U}^{k+1}, \mathbf{V}^{k+1}, \mathbf{S}^k, \mathbf{E}, \hat{\mathbf{U}}^{k+1}, \hat{\mathbf{V}}^{k+1}, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

9: $\mathbf{S}^{k+1} = \text{argmin}_{\mathbf{S}} \mathcal{L}(\mathbf{U}^{k+1}, \mathbf{V}^{k+1}, \mathbf{S}, \mathbf{E}^{k+1}, \hat{\mathbf{U}}^{k+1}, \hat{\mathbf{V}}^{k+1}, \mathbf{Y}_1^k, \mathbf{Y}_2^k, \mathbf{Y}_3^k, \mathbf{Y}_4^k, \mu^k)$

10: **End While** //Inner loop

11: Update τ^{k+1} according to (27)

12: Update the multipliers by

13: $\mathbf{Y}_1^{k+1} = \mathbf{Y}_1^k + \mu^k (\hat{\mathbf{U}}^{k+1} - \mathbf{U}^{k+1})$

14: $\mathbf{Y}_2^{k+1} = \mathbf{Y}_2^k + \mu^k (\hat{\mathbf{V}}^{k+1} - \mathbf{V}^{k+1})$

15: $\mathbf{Y}_3^{k+1} = \mathbf{Y}_3^k + \mu^k (\mathbf{U}^{k+1} (\mathbf{V}^{k+1})^T + \mathbf{S}^{k+1} - \mathbf{X})$

16: $\mathbf{Y}_4^{k+1} = \mathbf{Y}_4^k + \mu^k (\mathbf{S}^{k+1} - \mathbf{E}^{k+1})$

17: Update μ^{k+1} by $\mu^{k+1} = \min(\rho \mu^k, \mu_{max})$

18: $k \leftarrow k + 1$

19: **End While** //Outer loop

20: **Return** $\mathbf{L}^k = \mathbf{U}^k (\mathbf{V}^k)^T$ and \mathbf{S}^k .

Algorithm 2. Solving the Tree-Structured Sparsity [57]

Input: The index tree \mathbf{T} with depth d and node $G_j^i (i = 1, 2, \dots, d; j = 1, 2, \dots, n_i)$, weight $w_j^i \geq 0$ (default as 1), the matrix \mathbf{K}_1 , parameters α , and set $\lambda = \alpha / (2\mu^k)$.

1: Initialize $\mathbf{S} = \mathbf{K}_1$

2: **For** $i = d$ to 1 **do**

3: **For** $j = 1$ to n_i **do**

4:
$$\mathbf{S}_{G_j^i}^{k+1} = \begin{cases} \frac{\|\mathbf{S}_{G_j^i}\|_1 - \lambda w_j^i}{\|\mathbf{S}_{G_j^i}\|_1} \mathbf{S}_{G_j^i}, & \text{if } \|\mathbf{S}_{G_j^i}\|_1 > \lambda w_j^i \\ 0, & \text{otherwise} \end{cases}$$

5: **End For**

6: **End For**

7: **Return** \mathbf{S}^{k+1}

The estimated upper bound r of the rank of \mathbf{L} is an important parameter in Algorithm 1. According to Lemma 1 and Lemma 2 in [34], the globally optimal solution of the problem (13) depends on a feasible value of r . Thus, the estimate of a good value \hat{r} is crucial to generate an accurate low-rank matrix \mathbf{L} . Fortunately, several researchers [43,47] have proposed some effective methods to perform rank estimation. In practical applications, we only set a relatively large estimate r so that $r \geq \hat{r}$.

3.3.4. Complexity analysis

Given an $D \times N$ matrix with $D \leq N$, existing Schatten quasi-norm and matrix nuclear norm minimization models, such as SMD [32] and LpSq [39], take $\mathcal{O}(ND^2)$ to perform SVD in each

iteration. In contrast, the time complexity of performing SVD in Algorithms 1 is only $\mathcal{O}(Nr^2 + Dr^2)$, where $r \ll D, N$. To update \mathbf{U}, \mathbf{V} and \mathbf{L} , the time cost of the matrix multiplications takes $\mathcal{O}(DNr + r^3)$ in each iteration. In addition, Solving the optimization subproblem (27) takes $\mathcal{O}(K^{1/2} \log K \log(K/\varepsilon))$. In our experiments, K is set as 12. Thus, the overall complexity of our Algorithm is $\mathcal{O}(tDNr)$, where t is the number of iterations. Consequently, the computational complexity of either DNN-SMD or FNN-SMD is the same as that of ROSL [45], factEN [58], RegL1 [59] and Unifying [60], although our methods are based on the Schatten quasi-norm.

3.4. Saliency map generation

Similar to [29,61], we use a function $\text{Sal}(\cdot)$ of each superpixel in \mathbf{P} to estimate saliency values:

$$\text{Sal}(P_i) = \|\mathbf{s}_i\|_1, \quad (28)$$

where \mathbf{s}_i is the i th column of the matrix \mathbf{S} and $\|\mathbf{s}_i\|_1$ represents the saliency value of the i th superpixel P_i . Thus, $\text{Sal}(P_i)$ denotes the possibility that P_i belongs to a salient object. After all saliency values are estimated, all superpixels are merged to generate the final saliency map of the original image.

4. Experiments

To validate the effectiveness and efficiency of our method for salient object detection, we carry out our algorithm and compare with recent 25 state-of-the-art methods. First, we perform different algorithms on five benchmark datasets under various scenarios, and evaluate their performance in terms of six metrics. Second, we compare visual results of different methods by generating saliency maps in different scenarios. Finally, we analyze our methods with fixed rank and compare with the best among 25 competitive methods.

4.1. Experimental setup

Specifically, MSRA10K [62] and DUTOMRON [63] are used to validate different algorithms for cases with a single salient object, iCoSeg [64] and SOD [65] for images with multiple salient objects, and ECSSD [66] for cases with complex scenes. The detailed descriptions of these benchmark datasets are listed in Table 1.

We compare our algorithms with ULR [28], LRR [29], SLR [30] and SMD [32], which are based on the low rank matrix recovery theory. In addition, we perform 21 recent proposed methods on five benchmark datasets, including CA [12], RC [14], SEG [15], SR [16], SS [17], LC [18], HCT [19], SVO [20], CB [21], TD [22], HS [23], GC [25], GS [26], RBD [51], DSR [61], MR [63], FT [67], SF [68], MC [69], DRFI [70] and PCA [71]. All algorithms are performed on five popular benchmark datasets in different scenarios. Specifically, we perform all algorithms on MSRA10K [62] and DUTOMRON [63] in the case of a single salient object, iCoSeg [64] and SOD [65] with multiple salient objects, and ECSSD [66] with complex scenes.

The parameters of the proposed algorithm are set as follows. To generate the feature matrix, the same features are extracted as [32] and the number of superpixels N is set as 200 in our experiments for fair comparison. We further analyze the performance with different numbers of superpixels. In index tree construction, the affinity thresholds are set as $\mathcal{T} = [100, 400, 2000]$ and the tree depth is set to 5 by performing the initial over-segmentation. To extract the superpixels, the parameter σ is empirically set as the standard derivation of the superpixels. In the problem (12), we search the optimal (α, β, γ) over the range of $\{10^{-3}, 10^{-2}, \dots, 10^3\}$. For ensemble manifold regularization, we use the hot kernel, dot-product and 0–1 weighting schemes and set the number of nearest neighbors $p = \{5, 10, 12, 15\}$ respectively, which lead to the total 12 predefined graph Laplacians.

To precisely decompose the matrix \mathbf{L} , we estimate the rank of the low rank matrix by utilizing the rank-revealing feature of QR factorization as proposed in [43]. In addition, we discuss the performance of our methods with fixed rank. For comparing competitive algorithms in different scenarios, we use the source or binary codes provided by the authors with default parameters. All the experiments have been performed in MATLAB R2017 running in a 3.0 GHz Intel Core i5-8500 with 16-GB RAM.

4.2. Evaluation metrics

To quantitatively evaluate different algorithms, we report the experimental results on six metrics, including the precision–recall (PR) curve, the F-measure curve, mean absolute error, area under the ROC (AUC), the overlapping ratio (OR) and the weighted F-measure (WF) score.

Specifically, as in [32], PR curves of different methods are generated by binarizing saliency maps as the threshold ranges from 0 to 255. F-measure is defined as [68]:

$$F_\beta = \frac{(\beta^2 + 1)P \cdot R}{\beta^2 P + R}, \quad (29)$$

where P and R represent precision and recall, respectively, and β^2 is set to 0.3 for fair comparison. The F-measure curves are computed by setting different saliency thresholds.

In addition, we use the MAE metric to evaluate the dissimilarity between the binary saliency map and the ground truth, defined as [68]

$$\text{MAE} = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |S(x, y) - G(x, y)|, \quad (30)$$

where W and H denote the width and height of the saliency map S , respectively, and G is the ground truth binary saliency map.

Following [72], the OR metric is utilized to measure the overlapping ratio between the salient object mask S_M and the ground truth G , defined as

$$\text{OR} = \frac{|S_M \cap G|}{|S_M \cup G|}. \quad (31)$$

Finally, we use the weighted F-measure (WF) metric [73], a weighted version of F-measure, to test different algorithms. WF overcomes the interpolation, dependency and equal importance flaws of traditional measures, since it adopts weighted precision and recall to measure exactness and completeness, respectively.

4.3. Comparison with state-of-the-arts

To evaluate the performance of our algorithms, we first compare with 25 recently proposed algorithms on the five benchmark datasets. The results of the top 12 methods are reported in Tables 2 to 6, where partial data are cited from [32] and the best 3 results are highlighted in bold. In addition, the PR and F-measure curves of all algorithms on five datasets are illustrated in Fig. 3.

As can be seen from Tables 2 to 6, our method has better performance than other algorithms in most cases. DRFI has superior performance than SMD and DNN-EMR on DUT-OMRON. But, DRFI is a supervised method and relies on a large amount of labeled images, while SMD and DNN-EMR are unsupervised models independent of extern datasets. Thus, they are more flexible than DRFI in real applications. In different scenarios, our methods achieve better or comparable performance compared with SMD. We will compare the performance of our methods with SMD and DRFI in detail.

(1) Results on Single-Object Images

First, we analyze the performance of different algorithms on the MSRA10K and DUT-OMRON datasets. Fig. 3A and B shows the PR and F-measure curves on two datasets, respectively. In addition, Tables 2 and 3 list the MAE, WF, AUC and OR scores.

As shown in Table 2, although DRFI achieves the best AUC score, SMD, DNN-EMR and FNN-SMD obtain the best or second best performance according to WF, OR and MAE. From Fig. 3A, we can see that the PR curves of SMD and DNN-EMR are almost overlapped and DNN-EMR is superior to SMD. They significantly outperform other algorithms in the PR curves. From the

Table 1
Descriptions of benchmark datasets.

Name	Size	Descriptions
MSRA10K	10,000	Single object, simple background, high contrast
DUT-OMRON	5,168	Single object, relatively complex background, more challenging
iCoSeg	643	Multiple objects, various number of objects with different sizes
SOD	300	Multiple objects, various size and location of objects, complex background
ECSSD	1,000	Structurally complex natural images, various object categories

Table 2
Results on MSRA10K in terms of MAE, WF, AUC and OR.

Metric	DNN-EMR	SMD	DRFI	RBD	DSR	MC	MR	HS	PCA	ULR	SLR	HCT
MAE	0.105	0.105	0.114	0.108	0.121	0.145	0.125	0.149	0.185	0.224	0.141	0.143
WF	0.701	0.699	0.666	0.685	0.656	0.576	0.642	0.604	0.473	0.425	0.601	0.582
AUC	0.851	0.846	0.857	0.834	0.825	0.843	0.824	0.833	0.839	0.831	0.840	0.847
OR	0.742	0.741	0.723	0.716	0.654	0.694	0.693	0.656	0.576	0.524	0.692	0.674

Table 3
Results on DUT-OMRON in terms of MAE, WF, AUC and OR.

Metric	DNN-EMR	SMD	DRFI	RBD	DSR	MC	MR	HS	PCA	ULR	SLR	HCT
MAE	0.175	0.166	0.138	0.144	0.139	0.186	0.187	0.227	0.207	0.260	0.161	0.164
WF	0.416	0.423	0.424	0.427	0.419	0.347	0.381	0.350	0.287	0.254	0.392	0.353
AUC	0.817	0.812	0.839	0.814	0.803	0.820	0.779	0.801	0.827	0.805	0.822	0.815
OR	0.436	0.444	0.444	0.432	0.408	0.425	0.420	0.397	0.341	0.318	0.429	0.393

Table 4
Results on iCoSeg in terms of MAE, WF, AUC and OR.

Metric	DNN-EMR	SMD	DRFI	RBD	DSR	MC	MR	HS	PCA	ULR	SLR	HCT
MAE	0.136	0.138	0.139	0.138	0.153	0.179	0.162	0.176	0.201	0.222	0.179	0.179
WF	0.621	0.610	0.592	0.599	0.548	0.461	0.554	0.536	0.407	0.379	0.473	0.464
AUC	0.832	0.822	0.839	0.827	0.801	0.807	0.795	0.812	0.798	0.814	0.805	0.833
OR	0.606	0.598	0.582	0.588	0.514	0.543	0.573	0.537	0.427	0.443	0.505	0.519

Table 5
Results on SOD in terms of MAE, WF, AUC and OR.

Metric	DNN-EMR	SMD	DRFI	RBD	DSR	MC	MR	HS	PCA	ULR	SLR	HCT
MAE	0.234	0.233	0.217	0.229	0.260	0.260	0.261	0.283	0.274	0.308	0.248	0.243
WF	0.465	0.456	0.456	0.428	0.390	0.390	0.406	0.410	0.343	0.322	0.395	0.385
AUC	0.744	0.739	0.742	0.706	0.722	0.746	0.709	0.731	0.730	0.713	0.712	0.720
OR	0.425	0.421	0.447	0.406	0.392	0.392	0.373	0.325	0.340	0.290	0.400	0.377

Table 6
Results on ECSSD in terms of MAE, WF, AUC and OR.

Metric	DNN-EMR	SMD	DRFI	RBD	DSR	MC	MR	HS	PCA	ULR	SLR	HCT
MAE	0.175	0.173	0.217	0.225	0.227	0.251	0.235	0.269	0.291	0.312	0.252	0.249
WF	0.544	0.539	0.517	0.490	0.489	0.441	0.480	0.449	0.358	0.351	0.442	0.430
AUC	0.819	0.812	0.780	0.752	0.754	0.779	0.761	0.766	0.759	0.755	0.764	0.755
OR	0.563	0.561	0.527	0.494	0.480	0.495	0.491	0.432	0.371	0.347	0.474	0.457

F-measure curves (Fig. 3A), we can observe that DNN-EMR has superior performance over a large range has better performance than SMD and DRFI. In particular, DNN-EMR performs best than other algorithms as the thresholds increase.

From Table 3, we can see that the performance of all the methods degrades obviously on DUT-OMRON owing to its relatively complex background. According to the OR metric, SMD and DNN-EMR achieve the best and second best results, respectively. DRFI achieves all the best scores based on supervised information and multi-level saliency maps. The PR curves of different algorithms on DUT-OMRON are shown in Fig. 3B, we can see that SMD and DNN-EMR achieve favorable performance. The PR curve of DNN-EMR is better than that of SMD at the high recall rates. Furthermore, as can be seen from the F-measure curves in Fig. 3B, DNN-EMR performs better than SMD and DRFI at high thresholds.

(2) Results on Multiple-Object Images

We further discuss the performance of different algorithms on multiple-object images. Fig. 3C and D display the PR and F-measure curves on iCoSeg and SOD, respectively.

As can be seen from Table 4, although DRFI achieves the best AUC score on iCoSeg, DNN-EMR achieve the best MAE, WF and OR scores, respectively. In addition, Fig. 3C shows that DRFI has lower precision than DNN-EMR. Moreover, the PR curve of DNN-EMR is obviously superior to that of SMD when the recall rate is large than 0.5. In the F-measure curves, shown in Fig. 3C, DNN-EMR performs best among all methods. It should be noted that the F-measure values of our methods are better and more stable than those of other methods over a wide range.

From Table 5, we can see that DNN-EMR achieves the best WF score and the second best AUC and OR scores on SOD. The PR

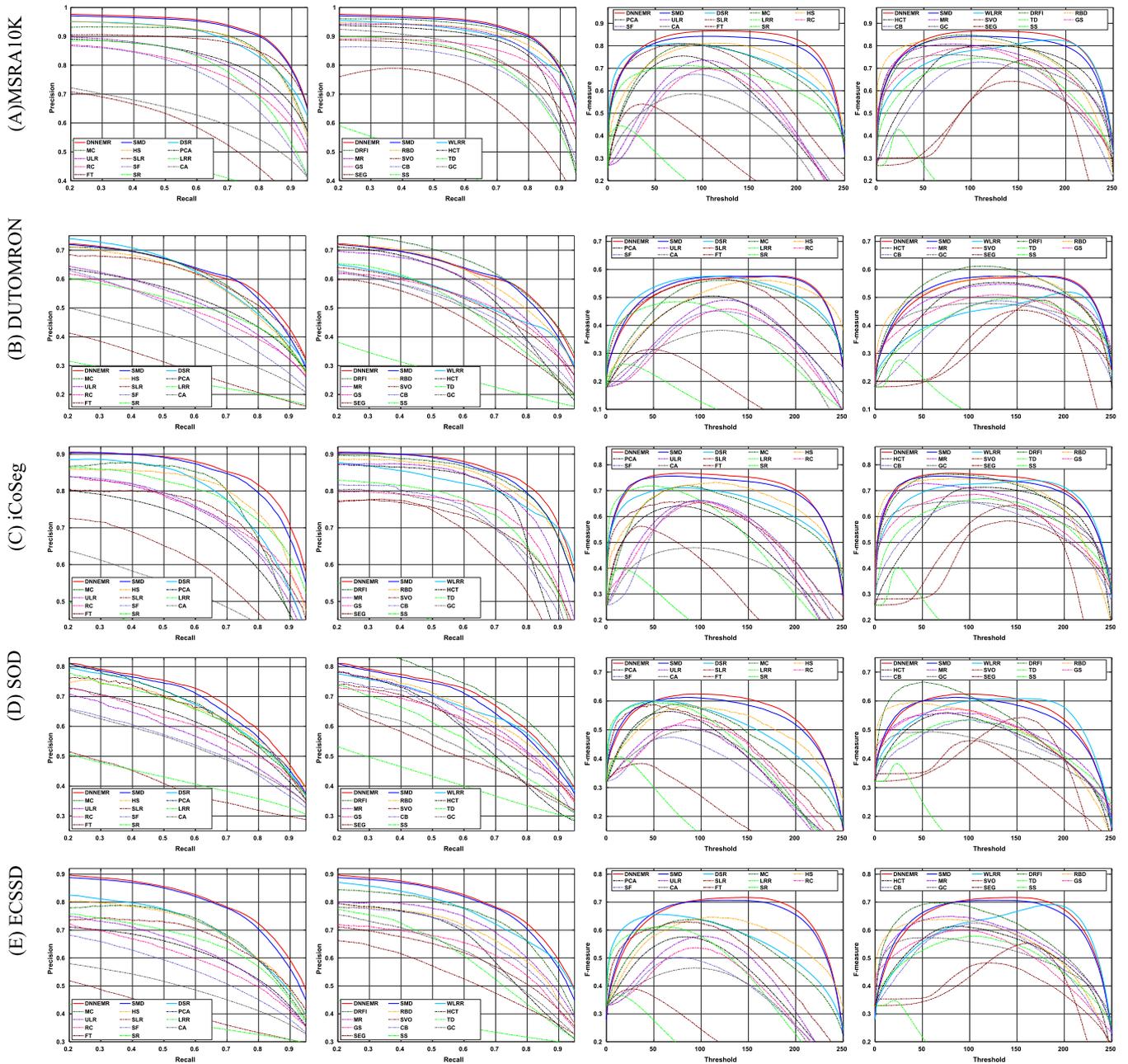


Fig. 3. Quantitative comparison on five datasets in terms of PR and F-measure curves.

curves in Fig. 3D show that the performance of our algorithms is slightly worse than that of DRFI. However, our methods outperform other algorithms in the unsupervised case. In the F-measure curves (Fig. 3D), DRFI outperforms DNN-EMR at lower threshold ranges, while our method performs better than SMD and DRFI over a wider range.

(3) Results on Complex Scene Images

We further compare our methods with the others on ECSSD and report the experimental results in Table 6. It is shown that DNN-EMR outperforms other algorithms in terms of WF, AUC and OR. DNN-EMR obtains the second best MAE score, with very small margins (0.002) to the best score. According to Fig. 3E, in the PR curves, DNN-EMR is best among all methods. Particularly, DNN-EMR achieves much better performance at the higher recall rates. The F-measure curves in Fig. 3E show that DNN-EMR also performs better than other algorithms, which validates that the

proposed algorithm has strong robustness against complex scenes in salient object detection.

(4) Visual Comparison

To compare visual results of those competitive methods, we select some output saliency maps in different scenarios, shown in Fig. 4. Compared with other methods on single-object images, DNN-EMR and SMD detect the salient object more accurately and have less scattered patches. By means of structured sparsity regularization and index trees, they produce identical saliency values to all superpixels within the salient objects. For images with multiple objects, MC, DSR, RBD, DRFI, SLR, ULR, PCA and GS incorrectly partition some background regions into salient objects, while MR, DSR, SMD, SLR and PCA only identify parts of the objects. By contrast, the proposed DNN-EMR method extracts all the multiple salient objects successfully. For the images with complex scenes, DNN-EMR and SMD detect the salient objects with less scattered patches, while most other methods fail to

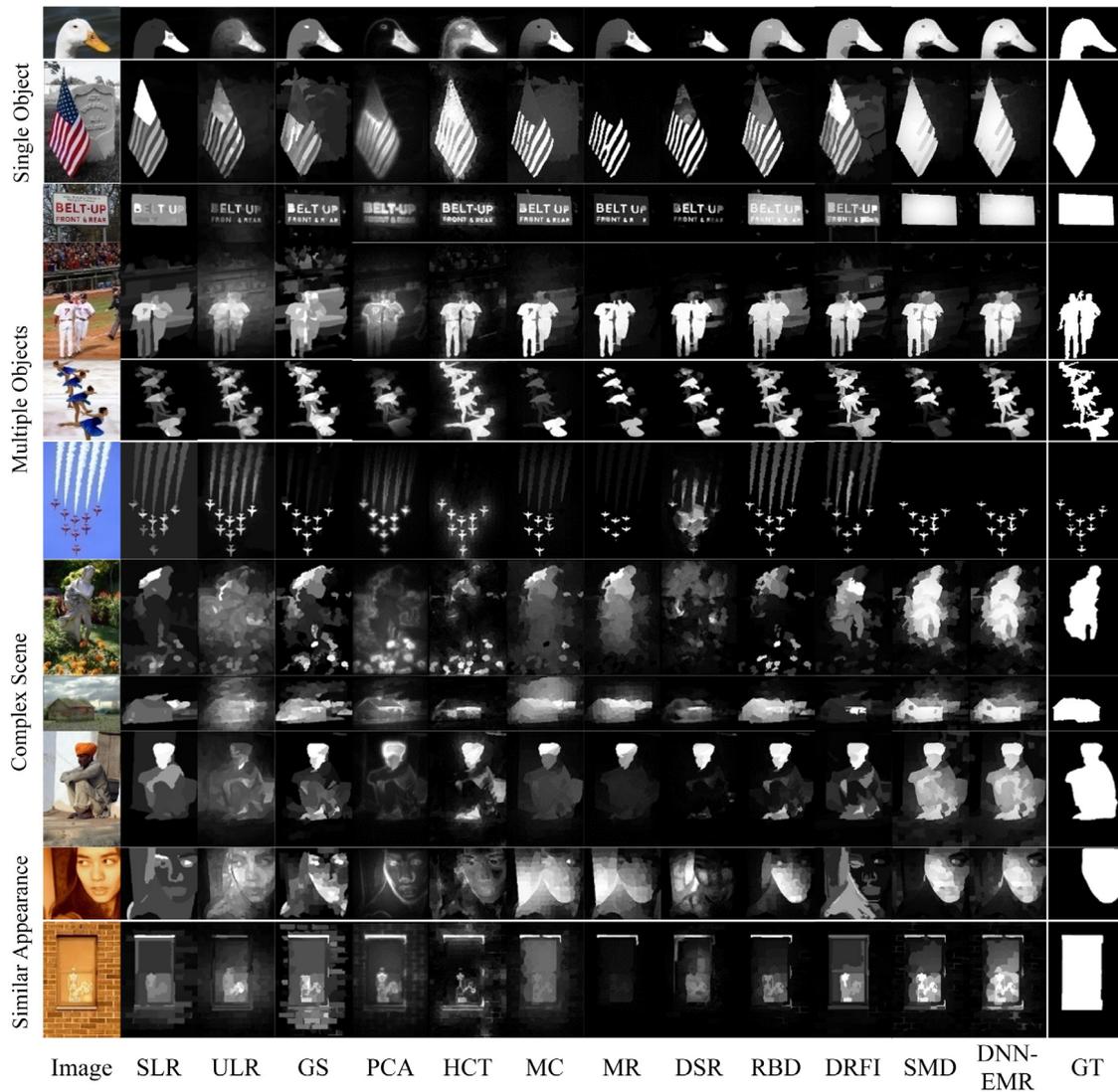


Fig. 4. Visual comparison of saliency maps of different algorithms in different scenarios.

generate the saliency maps. At last, for the images with similar foreground and background appearance, DNN-EMR extracts the salient objects from the background much better than other methods.

As can be seen from Fig. 4, the output results of our method are close to ground truth, which further validates the effectiveness of our method and demonstrate that DNN-EMR has stronger robustness than those competitive methods.

(5) Analysis of the performance with ensemble manifold regularization

Traditional low-rank matrix recovery models usually introduce the manifold regularization to effectively separate the sparse matrix from the low-rank matrix. However, the cross validation is generally utilized to choose graph hyperparameters for intrinsic manifold estimation. By contrast, our method has the ability to approximate the optimal solution for intrinsic manifold estimation. To validate the effectiveness of ensemble manifold regularization, we perform the proposed algorithm based on single graph Laplacian and multiple predefined graph Laplacians, respectively. The saliency detection algorithm based on double nuclear norm maximization and standard manifold regularization is termed as DNN-MR. For DNN-MR, the hot kernel is used to construct the graph Laplacian and its hyperparameter is chosen by the cross validation. The experimental results are reported

Table 7

Comparison between DNN-MR and DNN-EMR on MSRA10K and iCoSeg.

Algorithms	MSRA10K				iCoSeg			
	MAE	WF	AUC	OR	MAE	WF	AUC	OR
DNN-MR	0.106	0.700	0.848	0.739	0.138	0.613	0.826	0.598
DNN-EMR	0.105	0.701	0.851	0.742	0.136	0.621	0.832	0.606

in Tables 7 and 8. From Tables 7 and 8, we can observe that the results of DNN-EMR are all superior to those of DNN-MR on 4 benchmark datasets, which demonstrates that DNN-EMR not only learns the composite manifold, the low-rank matrix and the sparse matrix jointly, but has the ability to automatically approximate the optimal solution. Consequently, the proposed method outperforms that based on single graph Laplacian and avoids parameter selection from discrete states in the parameter space.

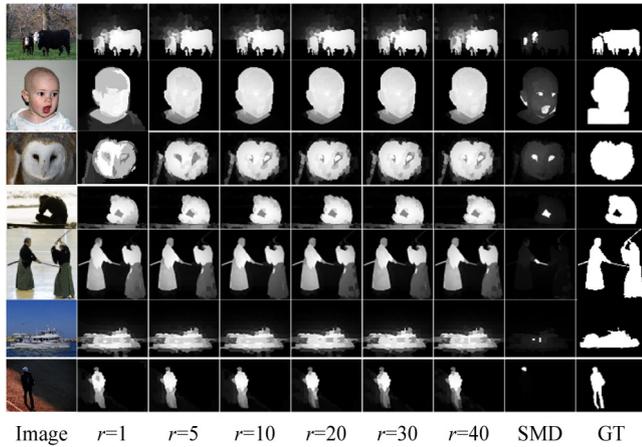
(6) Analysis of the performance with fixed ranks

We further evaluate DNN-EMR with fixed ranks instead of estimated ranks. We mainly compare our method with SMD, because the performance of SMD is closest to that of DNN-EMR and they are unsupervised methods derived from the RPCA model. In Fig. 5, we extract saliency maps with different fixed ranks using DNN-EMR. We can see that the performance of DNN-EMR

Table 8

Comparison between DNN-MR and DNN-EMR on SOD and ECSSD.

Algorithms	SOD				ECSSD			
	MAE	WF	AUC	OR	MAE	WF	AUC	OR
DNN-MR	0.236	0.457	0.739	0.421	0.177	0.539	0.816	0.599
DNN-EMR	0.234	0.465	0.744	0.425	0.175	0.544	0.819	0.563

**Fig. 5.** Visual comparison of saliency maps of SMD and DNN-EMR.

tends to be stable as the rank r increases. When r is larger than 10, the saliency maps of DNN-EMR remain almost unchanged. This is due to the fact that about 90 percent of these matrices can be approximated by a matrix with rank no greater than 10 according to the statistics in [32]. In terms of Lemma 3 in [34], if we use any right upper bound of the rank of the matrix L , the low-rank matrix can be accurately generated via the Schatten-1/2 quasi-norm. Consequently, the experimental results are consistent with the theoretical analysis. As can be seen from Fig. 5, DNN-EMR still performs better than SMD if a proper r is selected. In Tables 2 to 6, we estimate an appropriate upper bound of the rank and achieve better performance than SMD. These experimental results demonstrate that the proposed methods are more effective in real applications.

5. Conclusion

In this paper, we have presented a novel structured matrix decomposition model for salient object detection. The ADM based iterative algorithm was developed, which not only exploits double nuclear norm penalty to fit the heavy-tailed distribution of singular values of low-rank matrices, but introduces ensemble manifold regularization to automatically learn the optimal intrinsic manifold from a set of predefined graph Laplacians. In addition, structured sparsity regularization and high-level priors are incorporated into our model to achieve favorable performance. Experimental results on five challenging datasets validate the efficiency and effectiveness of our method in different scenarios.

For future work, we will study how to impose high-level priors on the feature matrix and low-rank matrix. In addition, we will combine some probabilistic graphical models to enhance the performance of our models further.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (2017XKQY082), China. The authors would like to thank the anonymous reviewers and the associate editor for their valuable comments.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.knosys.2019.07.021>.

References

- [1] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (11) (1998) 1254–1259.
- [2] H. Li, F. Meng, K.N. Ngan, Co-salient object detection from multiple images, *IEEE Trans. Multimedia* 15 (8) (2013) 1896–1909.
- [3] B. Alexe, T. Deselaers, V. Ferrari, Measuring the objectness of image windows, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (11) (2012) 2189–2202.
- [4] Z. Liu, R. Shi, L. Shen, et al., Unsupervised salient object segmentation based on kernel density estimation and two-phase graph cut, *IEEE Trans. Multimedia* 14 (4) (2012) 1275–1289.
- [5] P. Wang, J. Wang, G. Zeng, J. Feng, H. Zha, S. Li, Salient object detection for searched web images via global saliency, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2012, pp. 3194–3201.
- [6] L. Itti, Automatic foveation for video compression using a neurobiological model of visual attention, *IEEE Trans. Image Process* 13 (10) (2004) 1304–1318.
- [7] L. Marchesotti, C. Cifarelli, G. Csurka, A framework for visual saliency detection with applications to image thumbnailing, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 2232–2239.
- [8] Z. Wang, D. Xiang, S. Hou, et al., Background-driven salient object detection, *IEEE Trans. Multimed.* 19 (4) (2017) 750–762.
- [9] S. He, R.W. Lau, W. Liu, Z. Huang, Q. Yang, SuperCNN: A super pixelwise convolutional neural network for salient object detection, *Int. J. Comput. Vis.* 115 (3) (2015) 330–344.
- [10] N. Liu, J. Han, Dhsnet: Deep hierarchical saliency network for salient object detection, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2016, pp. 678–686.
- [11] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, 2009, pp. 1597–1604.
- [12] S. Goferman, L.Z. Manor, A. Tal, Context-aware saliency detection, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2010, pp. 1915–1926.
- [13] S. Huo, Y. Zhou, J. Lei, et al., Linear feedback control system based salient object detection, *IEEE Trans. Multimedia* 20 (6) (2018) 1350–1364.
- [14] M. Cheng, G. Zhang, N.J. Mitra, X. Huang, S. Hu, Global contrast based salient region detection, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2011, pp. 409–416.
- [15] E. Rahtu, J. Kannala, M. Salo, J. Heikkilä, Segmenting salient objects from images and videos, in: *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 366–379.
- [16] X. Hou, L. Zhang, Saliency detection: A spectral residual approach, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2007, pp. 1–8.
- [17] X. Hou, J. Harel, C. Koch, Image signature: Highlighting sparse salient regions, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (1) (2012) 194–201.
- [18] Y. Zhai, M. Shah, Visual attention detection in video sequences using spatiotemporal cues, in: *Proc. 14th ACM Int. Conf. Multimedia*, 2006, pp. 815–824.
- [19] J. Kim, D. Han, Y. Tai, J. Kim, Salient region detection via high-dimensional color transform, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2014, pp. 883–890.
- [20] K.-Y. Chang, T.-L. Liu, H.-T. Chen, S.-H. Lai, Fusing generic objectness and visual saliency for salient object detection, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 914–921.
- [21] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, S. Li, Automatic salient object segmentation based on context and shape prior, in: *Proc. Brit. Mach. Vis. Conf.*, 2011, pp. 1–12.
- [22] C. Scharfenberger, A. Wong, K. Fergani, J.S. Zelek, D.A. Clausi, Statistical textural distinctiveness for salient region detection in natural images, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 979–986.
- [23] Q. Yan, L. Xu, J. Shi, J. Jia, Hierarchical saliency detection, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 1155–1162.
- [24] R. Margolin, A. Tal, L. Zelnic-Manor, What makes a patch distinct? in: *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 1139–1146.
- [25] M. Cheng, J. Warrell, W. Lin, S. Zheng, V. Vineet, N. Crook, Efficient salient region detection with soft image abstraction, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 1529–1536.
- [26] Y. Wei, F. Wen, W. Zhu, J. Sun, Geodesic saliency using background priors, in: *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 29–42.
- [27] J. Yan, M. Zhu, H. Liu, Y. Liu, Visual saliency detection via sparsity pursuit, *IEEE Signal Process Lett.* 17 (8) (2010) 739–742.

- [28] X. Shen, Y. Wu, A unified approach to salient object detection via low rank matrix recovery, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2012, pp. 2296–2303.
- [29] C. Lang, G. Liu, J. Yu, S. Yan, Saliency detection by multitask sparsity pursuit, *IEEE Trans. Image Process* 21 (3) (2012) 1327–1338.
- [30] W. Zou, K. Kpalma, Z. Liu, J. Ronsin, Segmentation driven low-rank matrix recovery for saliency detection, in: Proc. Brit. Mach. Vis. Conf., 2013, pp. 1–13.
- [31] C. Tang, P. Wang, C. Zhang, W. Li, Salient object detection via weighted low rank matrix recovery, *IEEE Signal Process. Lett.* 24 (4) (2017) 490–494.
- [32] H. Peng, B. Li, H. Ling, W. Hu, W. Xiong, S.J. Maybank, Salient object detection via structured matrix decomposition, *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4) (2017) 818–832.
- [33] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, H. Shum, Learning to detect a salient object, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (2) (2011) 353–367.
- [34] N. Imamoglu, W. Lin, Y. Fang, A saliency detection model using low-level features based on wavelet transform, *IEEE Trans. Multimedia* 15 (1) (2013) 96–105.
- [35] Y. Fang et al, Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum, *IEEE Trans. Multimedia* 14 (1) (2012) 187–198.
- [36] F. Shang, J. Cheng, Y. Liu, Z. Luo, Z. Lin, Bilinear factor matrix norm minimization for robust PCA: Algorithms and applications, *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (9) (2018) 2066–2080.
- [37] J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its Oracle properties, *J. Amer. Statist. Assoc.* 96 (456) (2001) 1348–1360.
- [38] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (1) (2013) 171–184.
- [39] F. Nie, H. Wang, X. Cai, H. Huang, C. Ding, Robust matrix completion via joint Schatten p-norm and Lp-norm minimization, in: Proc. IEEE Int. Conf. Data Mining, 2012, pp. 566–574.
- [40] M. Lai, Y. Xu, W. Yin, Improved iteratively reweighted least squares for unconstrained smoothed p minimization, *SIAM J. Numer. Anal.* 51 (2) (2013) 927–957.
- [41] T. Oh, Y. Matsushita, Y. Tai, I. Kweon, Fast randomized singular value thresholding for nuclear norm minimization, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2015, pp. 4484–4493.
- [42] J. Cai, S. Osher, Fast singular value thresholding without singular value decomposition, *Meth. Anal. Appl.* 20 (4) (2013) 335–352.
- [43] Y. Shen, Z. Wen, Y. Zhang, Augmented Lagrangian alternating direction method for matrix separation based on low-rank factorization, *Optim. Methods Softw.* 29 (2) (2014) 239–263.
- [44] F. Jiang, M. Oskarsson, K. Astrom, On the minimal problems of low-rank matrix factorization, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2015, pp. 2549–2557.
- [45] X. Shu, F. Porikli, N. Ahuja, Robust orthonormal subspace learning: Efficient recovery of corrupted low-rank matrices, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2014, pp. 3874–3881.
- [46] G. Liu, S. Yan, Active subspace: Toward scalable low-rank learning, *Neural Comput.* 24 (12) (2012) 3371–3394.
- [47] Y. Liu, L. Jiao, F. Shang, A fast tri-factorization method for low-rank matrix recovery and completion, *Pattern Recog.* 46 (1) (2013) 163–173.
- [48] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, SLIC Superpixels compared to state-of-the-art superpixel methods, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (11) (2012) 2274–2282.
- [49] J. Liu, J. Ye, Moreau–Yosida regularization for grouped tree structure learning, in: Proc. Adv. Neural Inf. Process. Syst., 2010, pp. 1459–1467.
- [50] P. Felzenszwalb, D. Huttenlocher, Efficient graph-based image segmentation, *Int. J. Comput. Vis.* 59 (2) (2004) 167–181.
- [51] W. Zhu, S. Liang, Y. Wei, J. Sun, Saliency optimization from robust background detection, in: Proc. IEEE Comput. Vis. Pattern Recog., 2014, pp. 2814–2821.
- [52] T. Oh, Y. Tai, J. Bazin, H. Kim, I. Kweon, Partial sum minimization of singular values in robust PCA: Algorithm and applications, *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (4) (2016) 744–758.
- [53] K. Jia, T. Chan, Y. Ma, Robust and practical face recognition via structured sparsity, in: Proc. Eur. Conf. Comput. Vis., 2012, pp. 331–344.
- [54] D. Cai, X. He, J. Han, T.S. Huang, Graph regularized nonnegative matrix factorization for data representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (8) (2011) 1548–1560.
- [55] Jingjun Bi, Chongsheng Zhang, An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme, *Knowl.-Based Syst.* 158 (2018) 81–93.
- [56] J. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, *SIAM J. Optim.* 20 (4) (2010) 1956–1982.
- [57] R. Jenatton, J. Mairal, G. Obozinski, F. Bach, Proximal methods for hierarchical sparse coding, *J. Mach. Learn. Res.* 12 (2011) 7–24.
- [58] E. Kim, M. Lee, S. Oh, Elastic-net regularization of singular values for robust subspace learning, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2015, pp. 915–923.
- [59] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, M. Okutomi, Practical low-rank matrix approximation under robust L1-norm, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2012, pp. 1410–1417.
- [60] R. Cabral, F. De la Torre, J. Costeira, A. Bernardino, Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition, in: Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 2488–2495.
- [61] X. Li, H. Lu, L. Zhang, X. Ruan, M.-H. Yang, Saliency detection via dense and sparse reconstruction, in: Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 2976–2983.
- [62] M. Cheng, N.J. Mitra, X. Huang, P.H. Torr, S. Hu, Global contrast based salient region detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (3) (2015) 569–582.
- [63] C. Yang, L. Zhang, H. Lu, X. Ruan, M.-H. Yang, Saliency detection via graph-based manifold ranking, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2013, pp. 3166–3173.
- [64] D. Batra, A. Kowdle, D. Parikh, J. Luo, T. Chen, Interactively co-segmenting topically related images with intelligent scribble guidance, *Int. J. Comput. Vis.* 93 (3) (2011) 273–292.
- [65] V. Movahedi, J.H. Elder, Design and perceptual validation of performance measures for salient object segmentation, in: Proc. IEEE Comput. Vis. Pattern Recog. Workshop, 2010, pp. 49–56.
- [66] J. Shi, Q. Yan, L. Xu, J. Jia, Hierarchical image saliency detection on extended CSSD, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (4) (2015) 717–729.
- [67] R. Achanta, S. Hemami, F. Estrada, S. Süsstrunk, Frequency-tuned salient region detection, in: Proc. IEEE Comput. Vis. Pattern Recog., 2009, pp. 1597–1604.
- [68] F. Perazzi, P. Krähenbühl, Y. Pritch, A. Hornung, Saliency filters: Contrast based filtering for salient region detection, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2012, pp. 733–740.
- [69] B. Jiang, L. Zhang, H. Lu, C. Yang, M.-H. Yang, Saliency detection via absorbing Markov chain, in: Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 65–72.
- [70] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, S. Li, Salient object detection: A discriminative regional feature integration approach, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2013, pp. 1–8.
- [71] J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization, in: Proc. Adv. Neural Inf. Process. Syst., 2009, pp. 2080–2088.
- [72] X. Li, Y. Li, C. Shen, A.R. Dick, A. van den Hengel, Contextual hypergraph modeling for salient object detection, in: Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 3328–3335.
- [73] R. Margolin, L. Zelnik-Manor, A. Tal, How to evaluate foreground maps, in: Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2014, pp. 248–255.