H-V Shadow Detection Based on Electromagnetism-Like Optimization

Dimitra-Christina C. Koutsiou
Department of Computer Science and Biomedical Informatics
University of Thessaly
Lamia, Greece
dkoutsiou@uth.gr

Michalis Savelonas
General Department of Lamia
Univesity of Thessaly
Lamia, Greece
msavelonas@uth.gr

Dimitris K. Iakovidis
Department of Computer Science and Biomedical Informatics
University of Thessaly
Lamia, Greece
diakovidis@uth.gr

***Abstract*— Shadow detection is useful in a variety of image analysis applications, as it can improve scene understanding. Most of the recent shadow detection approaches use near-infrared (NIR) cameras and deep learning to provide enhanced segmentation of the shadow areas in images. In this paper a novel shadow detection method is proposed, exploiting the perceptual color representation of the HSV color space and a physics-inspired optimization algorithm for image segmentation. The comparative advantage of this method over the state-of-the-art ones is that its performance is comparable without requiring any special equipment, such as NIR cameras, while it is simpler. Quantitative and qualitative experiments on publicly available datasets in comparison with three state-of-the-art methods, validate its effectiveness.**

Keywords—Shadow Detection, Segmentation, Electromagnetism-Like Optimization, HSV.

I. INTRODUCTION

Shadow detection is a challenging problem in image processing, which has triggered intensive research efforts since the 1990s[1] . Existing works still struggle with ambiguous cases of shadowed/non-shadowed regions [2] . Currently, there are two dominant paths of research: methods based on physical observations of shadow in digital images, and methods based on neural networks. The latter, rather than following a physically interpretable approach, they rely on image data to learn recognize image regions corresponding to shadows. Several research studies on either of these paths consider imaging solutions with sensitivity beyond the visible spectrum, to derive complementary information that would enhance shadow detection.

Cameras sensitive to the near-infrared (NIR) region of the electromagnetic spectrum (700–1100 nm) provide the basis for several bimodal shadow removal methods, since the majority of objects that are dark in the visible spectrum appear brighter in NIR, whereas most of the considered illuminants in the shadow formation process have a distinct behavior in the NIR. Previous studies [3–5]combined brightness derived from RGB images with NIR for generic shadow detection. In application-oriented contexts, NIR has been employed by Zhu *et al.* [6]to detect shadows of clouds and by Fang *et al.* [7] to detect shadows of buildings in urban areas. However, NIR-based shadow detection methods can be costly and impractical for everyday applications, since they require modified or additional camera sensors, or in many cases, NIR data are simply not available, e.g. in images obtained from social media.

According to [8], [9] NIR is a useful tool for shadow detection, especially in remote sensing images. In this category are depicted mainly vegetation, clouds and soil in big distances. This is due to the fact that leaves, clouds etc. show different reflectance in NIR wavelengths because of the scattering characteristics of molecules inside in each illustrated region. The surrounded area is represented brighter in contrast to the shadows that are illustrated darker, so it is easier to be detected [10], [11]. Although NIR is a useful tool to segment big areas in remote sensing imagery, may be not always be as efficient as in natural indoor and outdoor images which have been taken under various lighting conditions. The second one contains various objects which are made from different and various materials. Consequently, it is hard to estimate the different surfaces reflectance under NIR light.[8] As a result, e.g., a black object may not always look brighter than the shadow because that fact actually depends on the reflective properties of its constructed material. Inevitably, the deep learning breakthrough in image processing inspired several shadow detection methods based on neural networks [12]. Most current such methods are focused on generative adversarial networks (GANs) [13],[14] indicated a multi-task method for shadow detection and removal using a stacked conditional adversarial network. Nguyen *et al.* [15] trained a shadow detector that corresponds to the generator of a conditional GAN, and augment its shadow detection accuracy by combining the typical GAN loss with a data loss term [16] proposed attentive recurrent GAN (ARGAN). Their method generates a shadow attention map guiding a shadow remover encoder to recover a shadow-lighter or even a shadow-free image. Beyond GANs, other deep learning strategies have also been proposed for shadow detection. Zheng *et al.* [2], proposed a network aiming at distraction-aware shadow detection by explicitly learning and integrating the semantics of visual distraction regions in an end-to-end framework. Hoseeinzadeh *et al.* [17]proposed a Convolutional Neural Network (CNN) which is used and robotic applications based on Support Vector Machine (SVM) feature extraction. Furthermore, Vicente *et al.* [18]introduced a CNN method for label recovery in mistaken annotations, utilizing in parallel shadow detection. Major drawbacks of such deep learning methods include: 1) the dependency on large datasets with sufficient variability, which

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-02070).

are not always available, and 2) the computational complexity of training a neural network, which is usually high.

In this paper, we challenge shadow detection by following a very simple image processing method for conventional RGB color images. The proposed method does not rely on extra sensors and it does not require any training, as neural network-based approaches do. It utilizes the perceptual representation of colors in the HSV color space, to segment shadow regions, considering only hue and brightness cues. Both stages are based on electromagnetism-like optimization (EMO)[19],[20], thus taking advantage of its heuristic search capabilities and low computational overhead of EMO to guide the maximization of variance between image regions.

The remainder of the paper is organized as follows: Section 2 describes EMO-based segmentation. Section 3 presents the proposed shadow detection method. Section 4 provides experimental results and comparisons with state-of-the-art shadow detection methods. Section 5 summarizes the main conclusions of this work.

II. EMO BASED IMAGE SEGMENTATION

Aiming to address the issue of computational complexity, several evolutionary optimization algorithms have been considered to obtain alternative balances between optimality and efficiency, including genetic algorithms [21–23], ant colony optimization (ACO) [24], simulated annealing [23] and particle swarm optimization (PSO) [25]. Oliva *et al*. [20] proposed EMO in the context of image segmentation. EMO is a global optimization algorithm that mimics the electromagnetism law of physics. It is a population-based method with an attraction–repulsion mechanism forcing "particles"-the members of the population- to move in the search space [19]. The attraction-repulsion force is calculated using the charge of each particle, based on its objective function value. In EMO, similarly to other heuristic approach, such as PSO and ACO, each particle is influenced by all other particles within the population. Despite their similarities, it has been shown that EMO outperforms PSO and ACO, with respect to precision and computational overhead [27–31].The charge q^p of each particle p can be described as:

$$q^p = exp\left(-n\frac{f(X^p)-f\left(X^{best}\right)}{\sum_{h=1}^{m}\left(f\left(X^h\right)-f\left(X^{best}\right)\right)}\right) \quad (1)$$

where X^h is the solution vector associated with the h_{th} particle, X^{best} indicates the best solution according to the objective function, n is the dimension of solutions X^h and m is the population size. The attraction/repulsion force is used in order to relegate the charges. Each particle attracts other particles with lower objective function values than it has and repels the higher ones. The equation that describes the force F^{ph} among a pair of charged particles is as follows:

$$\boldsymbol{F}^{ph} = \sum_{h\neq p}^{m}\begin{cases}\left(\boldsymbol{X}^h-\boldsymbol{X}^p\right)\dfrac{q^p q^h}{\left|\left|\boldsymbol{X}^h-\boldsymbol{X}^p\right|\right|^2}, & f\left(\boldsymbol{X}^h\right)<f\left(\boldsymbol{X}^p\right)\\ \left(\boldsymbol{X}^p-\boldsymbol{X}^h\right)\dfrac{q^p q^h}{\left|\left|\boldsymbol{X}^h-\boldsymbol{X}^p\right|\right|^2}, & f\left(\boldsymbol{X}^h\right)\geqslant f\left(\boldsymbol{X}^p\right)\end{cases} \quad (2)$$

The total force exerted on each particle p is:

$$\boldsymbol{F}^p = \sum_{i=1,\,\neq p}^{m}\boldsymbol{F}^{pi} \quad (3)$$

Each particle p moves in the search space according to:

$$\boldsymbol{X}^p \leftarrow \boldsymbol{X}^p + \lambda\frac{F^P}{\left|\left|F^p\right|\right|}(RNG) \quad (4)$$

where λ~*unif(0,1)* (*unif* represents the uniform function) for each coordinate of X^p and *RNG* denotes the allowed range of movement toward the lower or upper bound for the corresponding dimension. Following the main movement described in the previous equation, a local search is performed within a neighborhood of size δ of each particle, and the particle moves to the best solution within this neighborhood, with respect to the objective function. In each iteration, the best solution $\boldsymbol{X^{best}}$ is recalculated. The algorithm terminates either on convergence or when a pre-determined number of iterations has been reached.

In [14] grey-level image segmentation was considered as a constrained optimization problem for the determination of multiple grey-level thresholds, maximizing the variance between different image regions. The function to be maximized can be expressed as:

$$f(t) = max\left(\sigma_B^{2^c}(t)\right) \quad (5)$$

where $t=[t_1, t_2,\ldots t_k]$, $t\epsilon\boldsymbol{R}$, represents a vector comprising k threshold values, k is the number of image regions considered, $\sigma_B^{2^c}$ is the inter-region intensity variance, $\boldsymbol{R}$ is the bounded feasible region ($\boldsymbol{R}=\{t\epsilon \mathrm{R}^{\mathrm{k}} | 0\leq t_i \leq 255,\ i=1,2,\ldots k\}$), constrained by the interval [0, 255] in the case of 8-bit gray-level image resolution.

III. H-V SHADOW DETECTION

A previous study [26][indicates that the chromaticity component is approximately uniform in shadowed/non-shadowed regions of the same object, considering small variations in illumination. To obtain a perceptual representation of the input images, we consider the HSV color space, which provides a scalar representation of image hues. Following the RGB to HSV transformation, the proposed H-V shadow detection method is implemented in four steps:

Step 1: The EMO-based approach described in section 2, is adapted for the segmentation of the hue (H) component of the input color image. Modifications required the include extension of the search space to the interval of hue values, i.e. [0, 360] that have been normalized to [0-255], and the merging of the upper with the lower hue histogram regions. The latter modification is motivated by the fact that the both the lower and the higher values of the hue histogram represent red hues. The result of this process is a set of k images, $\tilde{h}_i, i=1,2,\ldots,k$, each of which represents approximately homogeneous regions with respect to hue. The '~' symbol corresponds to the modified component of channel H.

Step 2: Each of the k images obtained from Step 1, is subsequently multiplied by the value component (V) of the HSV image representation. Thus, a set of new images $\tilde{v}_i, i=1,2,\ldots,k$, weighted by image brightness is obtained as follows:

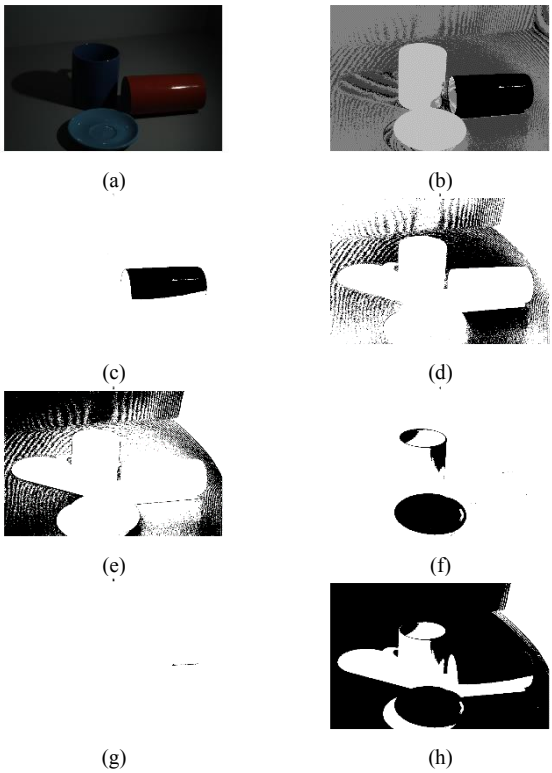


Fig. 1(a) RGB image (b) the image obtained after performing multithreshoding in channel H for 5 color regions, (c) thresholded component v_1 according to the color mask h_1, (d) thresholded component v_2 according to the color mask h_2, (e) thresholded component v_3 according to the color mask h_3, (f) thresholded component v_4 according to the color mask h_4, (g) thresholded component v_5 according to the color mask h_5 and (h) the final, aggregated segmented image.

$$\tilde{v}_i = \tilde{h}_i \cdot V \tag{6}$$

The rationale of this weighting is that the resulting images have a lower intensity in shadow regions, so that they can be more easily segmented with a bilevel thresholding operation.

Step 3: Bilevel thresholding on each of $\tilde{v}_i, i=1,2,\ldots,k$ is performed by EMO segmentation as described in section 2, with only a single threshold to be optimized. The result of this process is a set of k binary images b_i , i=1,2, …, k, in which the pixels corresponding to lower intensities (possible shadow regions) are set to white and the rest are set to black.

Step 4: In the final step of the proposed method, the binary masks b_i, i=1,2, …, k, obtained per input image are aggregated as follows, to generate a binary image representing a mask shadow of the shadow regions within the input image:

$$B = \bigwedge_{i=1}^{K} b_i \tag{7}$$

where **Λ** is the logical operator AND. Figure 1 illustrates the steps of the proposed two-stage segmentation method for shadow detection, virtually.

IV. Results

A. Datasets

The proposed method has been evaluated and compared to state-of-the-art shadow detection methods using two popular datasets. The first dataset has been created by Ruefenacht *et al*. [4] and is the only one that contains both RGB and NIR images combining its ground truth shadow masks. It consists of 74 images, from which 42 are outdoor and 32 are indoor. Indoor images are further separated in two sub-categories according to illumination: 16 images have been acquired under uncontrolled illumination and 16 images have been taken using either fluorescent flash (8 images) or incandescent flash (8 images). All ground truth masks have been created manually. The second dataset is a combination of UIUC , UCF and SBU [27], [28][18] shadow datasets. UCF consists of 355 images and manually created ground truth masks. A small part of them are remote sensing images. UIUC is composed of 108 images and ground truth images created by the subtraction of shadow and non-shadow images in channel V.SBU-Train-Recover Dataset which consists of 4085 annotated images collected from the web, and a part of them collected from the MS COCO Dataset [29]. Therefore, the final UIUC+UCF+SBU dataset consists of a total of 4548 images along with their ground truth masks.

B. Comparative Evaluation

Shadow detection quality has been quantified by comparing the extracted mask with the ground truth mask. Two widely known metrics in statistics are accuracy and Matthews Correlation Coefficient (*MCC*), which are defined as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

$$MCC = \frac{(TP \cdot TN - FP \cdot FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{9}$$

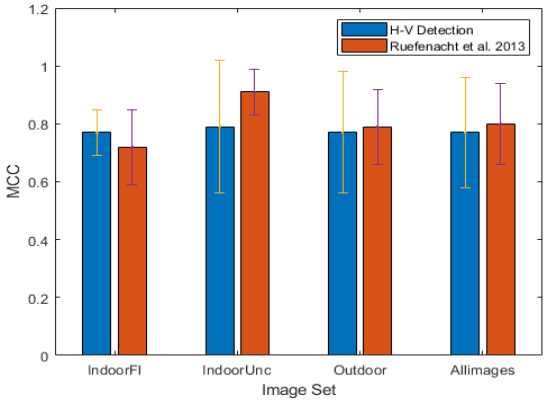
where, *TP = True positive*; *FP = False positive*; *TN = True negative*; *FN = False negative*. The experiments were carried out for five color regions [14] in Hue, for all images. The results of the proposed method are summarized in Table I in comparison to the state-of-the-art method *of* Ruefenacht *et al.* [4] on the dataset created in that study. For visualization purposes we also provide Fig. 2, which makes clear that the differences between the results obtained with the compared methods are small, and the estimated errors are overlapping. In the Indoor Flash image set, the proposed method is more effective, as indicated by both metrics.

TABLE I. Comparative H-V Shadow Detection Results on Ruefenacht's *et al*. [4], Dataset.

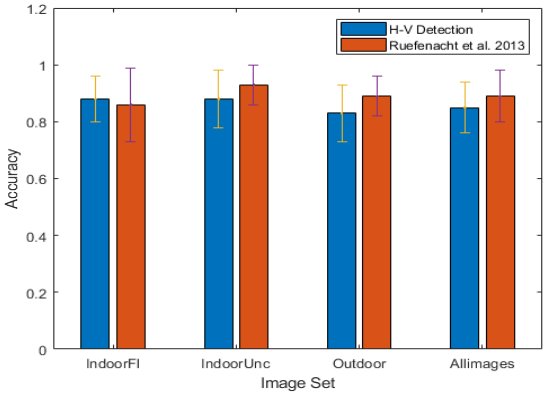
Image Set	***Metric***	***Proposed Method***	***Ruefenacht et al.* [4]**
Indoor Flash	Accuracy	0.88±0.03	0.86±0.13
	MCC	0.77±0.08	0.72±0.13
Indoor Uncontrolled	Accuracy	0.88±0.10	0.93±0.07
	MCC	0.79±0.23	0.91±0.08
Outdoor	Accuracy	0.83±0.10	0.89±0.07
	MCC	0.77±0.21	0.79±0.13
All images	Accuracy	0.85±0.09	0.90±0.09
	MCC	0.77±0.19	0.80±0.14

TABLE II. Comparative H-V Shadow Detection Results on The UIUC+UCF+SBU Dataset.

Metric	***Proposed Method***	***Vicente et al. 2016***	***Hosseinzadeh et al. 2018***
Accuracy	0.82±0.13	0.86±0.18	0.85±0.20

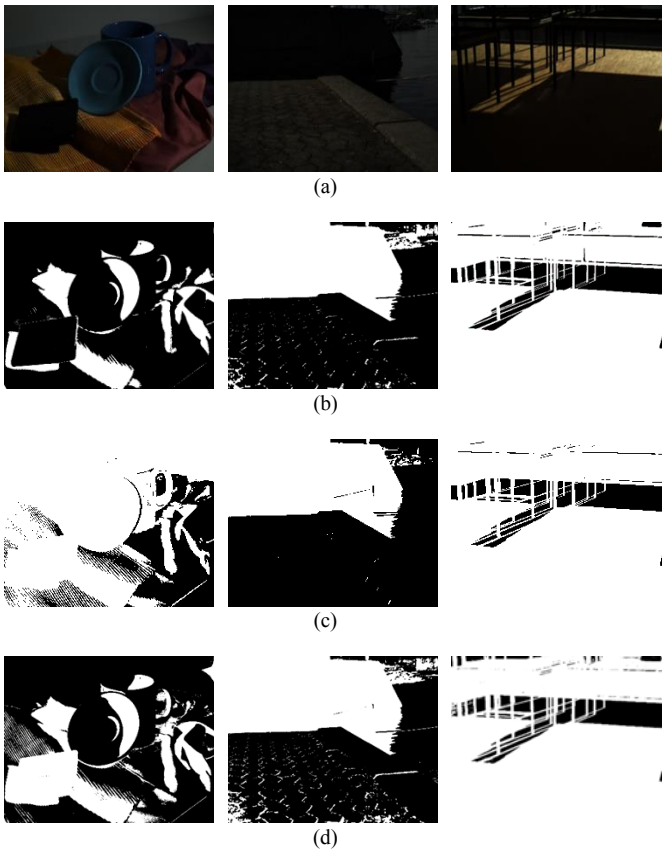


(a)



(b)

Fig. 2 Visualization of the results presented in Table I



(a) (b) (c) (d)

Fig. 3 (a), RGB images from Ruefenacht's dataset (b), ground truth images (c), Ruefenacht's results and (d) the images obtained from the proposed method

In the Indoor Uncontrolled and Outdoor image sets, the proposed method obtains lower *MCC* but remains comparable. Overall, the results of the proposed method are comparable to the method of Ruefenacht *et al*. [4] without making use of an infrared camera (highly overlapping error bars). The observed increase in Indoor Flash image set could be attributed to the illumination conditions [8]. In the last two categories solar radiation is the main light source that is composed of 44% visible and 53% infrared radiation. Thus, a possible explanation for the leveling results in our study and Ruefenacht's one for Indoor Uncontrolled and Outdoor image is the fact that sunlight reflection is quite high in NIR cameras. However, fluorescent and incandescent light do not respond well enough in NIR. In this way, one can find out that NIR is not always necessary for shadow detection, especially in natural environment images.

Qualitatively, indicative results of the proposed method in comparison to the method of Ruefenacht *et al.* [4] are illustrated in Fig. 3. It can be observed that in Indoor Flash image set the results are better than those of [4].This is due to the fact that the last technique misconstrues most of the non-shadow regions as shadows. In addition, in Outdoor set, the image obtained from the proposed method is again finer than Ruefenacht's for the reason that here shadow regions are detected thoroughly such as the slits on the pavement according to the ground truth mask. In Indoor Uncontrolled set the segmented images came up from Ruefenacht's and the proposed method are comparable to each other, although the proposed one does not require NIR data.

Table II summarizes the results obtained with the dataset UIUC+UCF+SBU in comparison to two deep learning-based state-of-the-art shadow detection methods Vicente *et al*. 2016 [18] and Hosseinzadeh *et al*. 2018 [17]. It can be noticed that on average the performance of the proposed method is slightly lower; however, it can be considered comparable, since the errors are overlapping. Owing the fact that the deviation in accuracy scores among the various methodologies is less than 5%, makes the suggested algorithm enough strong and efficient against to other state-of-the art techniques using learning. This is a comparative advantage of the proposed approach, since it is fast, simple, and unsupervised, without requiring computationally complex computations as deep learning approaches do.

Figure 4 illustrates shadow segmentation examples using images from the UIUC+UCF+SBU dataset. One can see that using Stacked-CNN method [18] misreads the column as a shadow. Also, the H-V Shadow Detection method gives almost the same quantitative results as [17] by utilizing an effective and simpler algorithm. Furthermore, in order to verify that the results obtained by the proposed method, which does not use NIR, are comparable to those obtained with the state-of-the-art methods that use NIR, we performed a statistical *p*-value test [30]. Let us assume the null hypothesis H_0 that there is no difference between the experiments with and without using NIR and the alternative hypothesis H_1. Performing the test, we obtained a *p*-value=0.48. Since the *p*-value is higher than significance level of 0.05 (5%) the null hypothesis H_0 can be accepted. Therefore, the difference of the results obtained either using or not using NIR images for shadow detection problems is insignificant.

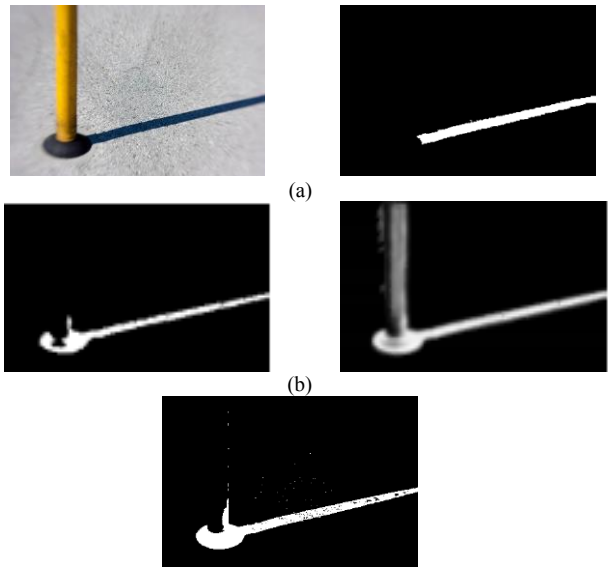


Fig. 4 (a) RGB image and its ground truth from UIUC+UCF+SBU dataset (b) the segmented images obtained using the method of Hosseinzadeh *et al.* [17] (on the left) and using the Stacked-CNN [18] (on the right), (c) the segmented image that obtained via H-V Shadow Detection method.

V. Conclusions

This work presents a novel shadow detection method in H-V, based on EMO. The proposed method has a simplified design in comparison to other state-of-the-art shadow detection methods using CNNs and other deep learning strategies. Experimental comparisons with state-of-the-art methods have been performed on two popular shadow detection benchmark datasets. The results of the experimental evaluation lead to the conclusions that the proposed method leads to a shadow detection that is comparable with the one obtained using NIR-based methods, without using NIR camera, more efficiently. In future research we intend to investigate shadow detection methods based on other physics-inspired optimization algorithms and shadow removal techniques.

Acknowledgment

The authors would like to thank Dr. Dominic Ruefenacht and Prof. Derek Hoiem for datasets provision.

References

[1] P. L. Rosin and T. J. Ellis, "Image difference threshold strategies and shadow detection.," in *BMVC*, 1995, vol. 95, pp. 347–356.

[2] Q. Zheng, X. Qiao, Y. Cao, and R. W. Lau, "Distraction-aware shadow detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5167–5176.

[3] N. Salamati, A. Germain, and S. Siisstrunk, "Removing shadows from images using color and near-infrared," in *2011 18th IEEE International Conference on Image Processing*, 2011, pp. 1713–1716.

[4] D. Rüfenacht, C. Fredembach, and S. Süsstrunk, "Automatic and accurate shadow detection using near-infrared information," *IEEE transactions on patt. anal. and machine intelligence*, vol. 36, no. 8, pp. 1672–1678, 2013.

[5] H. Hiary, R. Zaghloul, and M. B. Al-Zoubi, "Single-image shadow detection using quaternion cues," *The Computer Journal*, vol. 61, no. 3, pp. 459–468, 2018.

[6] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Rem. sens. env.,* vol. 118, pp. 83–94, 2012.

[7] H. Fang, Y. Wei, H. Luo, and Q. Hu, "Detection of Building Shadow in Remote Sensing Imagery of Urban Areas With Fine Spatial Resolution Based on Saturation and Near-Infrared Information," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2695–2706, 2019.

[8] C. Fredembach and S. Susstrunk, "Illuminant estimation and detection using near-infrared," in *Digital Photography V*, 2009, vol. 7250, p. 72500E.

[9] H. Jiang, S. Wang, X. Cao, C. Yang, Z. Zhang, and X. Wang, "A shadow-eliminated vegetation index (SEVI) for removal of self and cast shadow effects on vegetation in rugged terrains," *International Journal of Digital Earth*, vol. 12, no. 9, pp. 1013–1029, 2019.

[10] M. Vollmer, K.-P. Möllmann, and J. A. Shaw, "The optics and physics of near infrared imaging," in *Education and Training in Optics and Photonics*, 2015, p. TPE09.

[11] K. Mangold, J. A. Shaw, and M. Vollmer, "The physics of near-infrared photography," *European Journal of physics*, vol. 34, no. 6, p. S51, 2013.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[14] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1788–1797.

[15] V. Nguyen, Y. Vicente, F. Tomas, M. Zhao, M. Hoai, and D. Samaras, "Shadow detection with conditional generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4510–4518.

[16] B. Ding, C. Long, L. Zhang, and C. Xiao, "ARGAN: attentive recurrent generative adversarial network for shadow detection and removal," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 10213–10222.

[17] S. Hosseinzadeh, M. Shakeri, and H. Zhang, "Fast shadow detection from a single image using a patched convolutional neural network," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 3124–3129.

[18] T. F. Y. Vicente, L. Hou, C.-P. Yu, M. Hoai, and D. Samaras, "Large-scale training of shadow detectors with noisily-annotated shadow examples," in *European Conference on Computer Vision*, 2016, pp. 816–832.

[19] S. lker Birbil and S.-C. Fang, "An electromagnetism-like mechanism for global optimization," *J. glob. opt.*, vol. 25, no. 3, pp. 263–282, 2003.

[20] D. Oliva, E. Cuevas, G. Pajares, D. Zaldivar, and V. Osuna, "A multilevel thresholding algorithm using electromagnetism optimization," *Neurocomputing*, vol. 139, pp. 357–381, 2014.

[21] K. Hammouche, M. Diaf, and P. Siarry, "A comparative study of various meta-heuristic techniques applied to the multilevel thresholding problem," *Eng. appl.of artificial intelligence*, vol. 23, no. 5, pp. 676–688, 2010.

[22] D. Goldberg, "Genetic Algorithms in Search, Optimization & Machine Learning (p. 126) Addison-Wesley," *Reading, Mass*, 1989.

[23] P.-Y. Yin, "A fast scheme for optimal thresholding using genetic algorithms," *Signal processing*, vol. 72, no. 2, pp. 85–95, 1999.

[24] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE computational intelligence magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[25] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948.

[26] E. Salvador, A. Cavallaro, and T. Ebrahimi, "Cast shadow segmentation using invariant color features," *Computer vision and image understanding*, vol. 95, no. 2, pp. 238–259, 2004.

[27] R. Guo, Q. Dai, and D. Hoiem, "Paired regions for shadow detection and removal," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 12, pp. 2956–2967, 2012.

[28] J. Zhu, K. G. Samuel, S. Z. Masood, and M. F. Tappen, "Learning to recognize shadows in monochromatic natural images," in *2010 IEEE conf on computer vision and pattern recognition*, 2010, pp. 223–230.

[29] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*, 2014, pp. 740–755.

[30] W. R. Rice, "A consensus combined P-value test and the family-wide significance of component tests," *Biometrics*, pp. 303–308, 1990.