A MULTICHANNEL LOCALIZATION METHOD FOR CAMOUFLAGED OBJECT DETECTION

Md. Rakibur Rahman ⋆, Md. Mafri Chowdhury ⋆, Md. Shohanur Rahaman Sarkar ⋆, Md. Shahriar Karim †

Electrical and Computer Engineering, North South University, Bangladesh
⋆ equal contribution, † corresponding author: shahriar.karim@northsouth.edu

ABSTRACT

This paper proposes a multichannel method for discriminative region localization in Camouflaged Object Detection (COD) tasks. In one channel, processing the phase and amplitude of 2-D Fourier spectra generate a modified form of the original image, used later for a pixel-wise optimal local entropy analysis. The other channel implements a class activation map (CAM) and Global Average Pooling (GAP) for object localization. We combine the channels linearly to form the final localized version of the COD images. Experimentation in multiple COD datasets demonstrates that the proposed method successfully localizes regions containing more than 80% of the camouflaged objects. Our proposed method does not require memory-intensive devices or prior training on particular image features, making it easily integrable into a resource-constrained environment. The proposed approach is also applicable to non-COD images.

***Index Terms*—** Camouflaged Object Detection, Local Entropy, Class Activation Map

1. INTRODUCTION

Object detection has revolutionized various aspects of our daily lives, ranging from safety in the transportation system [1], diagnosing diseases (such as lung infection and polyp segmentation) [2], agriculture, and arts to security and surveillance of potential threats [3]. The state-of-the-art (SOTA) deep learning models [4, 5] that work well for generic and salient object detection tasks [6] are impertinent if the objects' saliency is compromised common in camouflaged object detection (COD) tasks. Therefore, such SOTA models experience performance deterioration for Camouflaged Object Detection (COD) tasks where objects appear intrinsically similar to the image background and surroundings. Besides, image features such as brightness, tone, textures, etc., are disrupted in COD images [7], making COD tasks challenging for the existing object detection models.

The success of deep neural network (DNN) models in object detection is due to their ability to mimic human visual perception. Humans are effortlessly good at object recognition, so mimicking human visual cortex operations artificially to form the convolutional neural network (CNN) and pursue deeper image features (using DNN models) has worked well for salient and generic object detection tasks. On the contrary, camouflaged objects inherently disrupt features to deceive their appearance and detection. For instance, many animals in nature resort to disruptive coloration, dazzling coloration, disruptive contrast, and edge enhancement to avoid detection by the predators [8, 9]. So, SOTA object detection models, salient [10] or generic [11], built upon how human visual perception works, may equally be vulnerable and suffer from inadequate accuracy in COD tasks [12]. For instance, like non-COD models, the COD models developed are also mainly CNN-based, mimicking deception-prone human visual perception. So, feature extraction by the CNN-based models may exclude the necessary details for COD images' discerning features. Also, camouflaging in nature is adaptive and changes over time in coloration and luminescence [7], requiring continuous adjustment and improvement of the COD mechanisms. An intriguing question thus arises– what an optimal strategy would be to compensate for the performance degradation of COD models.

An ROI-guided localization is predominant for performance improvement in salient and generic object detection tasks. However, not much has been explored on its role in COD tasks. Earlier works attempted decoder-based object localization and subsequent ranking of objects in COD images [13]. Another work [14] relied primarily on region proposal network (RPN) [11], imposing an additional computational burden through feature training and neural network implementation. Instead, the proposed method relies primarily on exploring local-scale information of the low-level features and enhancing the COD object localization without further burdening the computationally intensive COD models. Precisely, we design a localization pipeline that, in concert with the CAM-based GAP [15], achieves camouflaged object localization up to around 80% of the object area. In summary, the multichannel pipeline successfully localizes camouflaged objects of different difficulty levels, and the method is equally effective in non-COD localization. A set of major contributions made in this paper is:

- developed a computationally less-expensive localization pipeline ($R2$) for COD tasks with no training of features nor a dependency on memory-intensive devices such as GPU/TPU.
- proposed a multichannel method for ROI detection combining $R2$ in one channel and the CAM + GAP approach in the other channel.
- performed extensive experiments for alternative COD datasets to demonstrate the efficacy of the proposed method.

2. METHODS

An image's tonal and textural information capture its spatial variation and can identify discriminative details of an image by measuring the smoothness, fineness, irregularity, etc. [16]. For instance, in radar applications, the captured images' fine and coarse texture accurately correlates to whether the images of rocks in remote regions are fine-grain or coarse-grain [17]. Besides other textural traits, the Shannon entropy [18] also predicts image background complexity.

Thanks to Professor Mark R. Bell, Dept. of ECE, Purdue University, West Lafayette, USA, for his invaluable discussion on the paper.

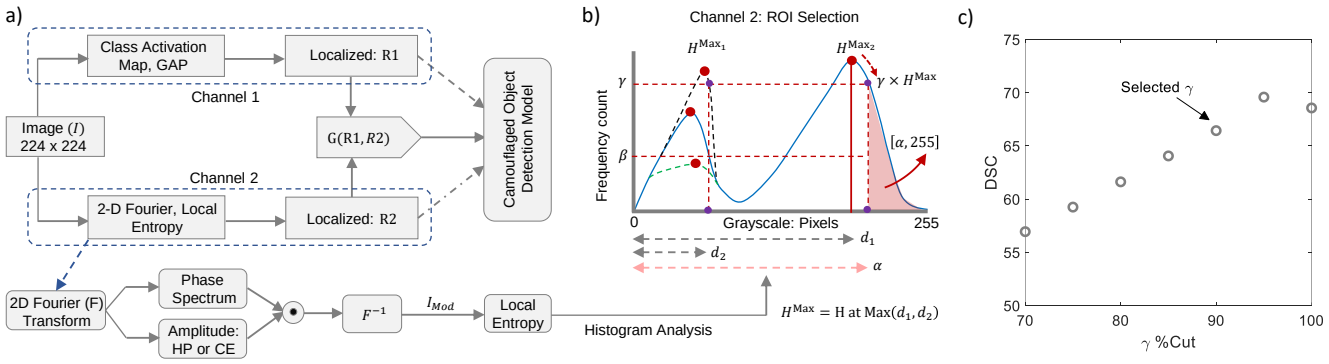

Fig. 1. a) Block diagram of the proposed multichannel localization method. b) The histogram analysis of the local entropy of the $I(a,b)_{\text{Mod}}$ is used for the ROI selection process. c) Choosing γ through a comparison object vs. surrounding pixel ratio and DSC score. As the DSC score falls after $\gamma = 0.95$, the optimal γ for the chosen 100 COD10K images is fixed at $\gamma = 0.9$.

In COD images, disruptive coloration, for instance, high contrast colors in object borders, works as concealment of the object [8], contributing to disorder in a neighborhood of image pixels. We resort to local entropy measurement as it quantifies fine-grained pixel intensity variation in a neighborhood of image pixels and provides useful textural information on the camouflaged object. Also, the phase of Fourier spectra extracts vital details on the texture in radar imaging [16], and the phase discrepancy at the pixel level possesses additional information as exploited earlier in object motion detection [19]. In short, the proposed method harnesses the ability of local entropy and Fourier phase and amplitude spectra for COD object localization. All the experiments are done on a machine with configuration as CPU - intel(R) Core (™) i5-7200U, Core - 4, Max TDP - 15W.

2.1. Off-the-shelf Approach

We revisited several image segmentation techniques, alone or in combination (see Table 3), to identify an optimal stack of Off-the-shelf segmentation techniques applicable for COD tasks [20]. The screen used the Dice Similarity Coefficient (DSC) and Structural Similarity Index Measure (SSIM) of a few COD images to differentiate between the competing candidates or the stack of multiple methods, as in Table 3. As found, the stack *Fourier Transform* $\rightarrow$ *Denoiser* $\rightarrow$ *(Phase spectra, High-pass of amplitude spectra)* $\rightarrow$ *Local Entropy* pipeline is superior to others on the COD tasks. As observed, the Fourier phase of the original image improves both DSC and SSIM scores (rows 3 and 5), reiterating the intriguing role of phase alone in image retrievability. The inclusion of High-pass (HP) filter for the amplitude spectra preserves (Canny Edge: CE improves further) the object's edge, which is often high-contrasted as an effective camouflaging strategy [7]. Analytically, the proposed $R2$ pipeline proceeds as outlined in the subsequent sections.

2.2. Processing Phase and Amplitude Spectra

Consider $I(a,b)$ is an image with width of M and N pixels, where $a = 0, 1, \ldots M-1$ and $b = 0, 1, \ldots N-1$, respectively. The 2D discrete Fourier transform (DFT) of the image $I(a,b)$, denoted as $F(u,v)$, is:

$$\begin{aligned} F(u,v) &= \sum_{a=0}^{M-1}\sum_{b=0}^{N-1} I(a,b)e^{-j2\pi ua/M}e^{-j2\pi vb/N} \\ &= |F(u,v)|e^{-j\phi(u,v)}, \text{ in polar form} \end{aligned} \tag{1}$$

where $|F(u,v)|$ and $\phi(u,v)$ are the amplitude spectra and phase spectra, respectively. The object localization approach, as in Channel 2, uses the 2-D DFT (Eq. 1) and produces the modified version $I(a,b)_{\text{Mod}}$ of the original grayscale image $I(a,b)$. Precisely, we obtain $I(a,b)_{\text{Mod}}$ by performing an Inverse Discrete Fourier Transform (IDFT) of the point-wise product of the $|F(u,v)|_{\text{High}}$ and $\phi(u,v)$ both:

$$I(a,b)_{\text{Mod}} = F^{-1}(|F(u,v)|_{\text{High}} \odot e^{j(\arctan(\phi(u,v)))}) \tag{2}$$

where $|F(u,v)|_{\text{High}}$ represents the High-pass refinement of the Fourier amplitude spectra $I(a,b)$ evaluated using Eq. 1.

2.3. Local Entropy Analysis of $I(a,b)_{\text{Mod}}$

The local entropy calculation at any pixel (a,b) of the modified grayscale image $I(a,b)_{\text{Mod}}$ uses the Shannon entropy [21]. Suppose, number of pixels in window W of gray scale intensity k is n_k, so, the probability p_k becomes $p_k = n_k/N^2$, with $k = 0, 1, \ldots 255$. The local entropy $H(I_{a,b})$ formulation of an image pixel $I(a,b)$ over a window W of dimension $N \times N$ is

$$H(I_{a,b}) = \sum_{k=0}^{N-1} p_k log_2 \frac{1}{p_k} = \sum_{k=0}^{N-1} n_k/N^2 log_2(N^2/n_k) \tag{3}$$

where p_k is the probability of k^{th} gray scale intensity in the neighborhood window W. The window W slides over a stride of a single pixel size to produce a fine-grained local entropy map of the entire I_{Mod}. Later, a histogram analysis over 0 to 255 bins produces the entropy distribution and identifies the decision boundary for the region of interest (ROI) selection in channel 2 (shown in Figure 1b). A low value of $H(I_{a,b})$ in local entropy analysis represents less textural variability in a neighborhood, which suggests COD objects' presence in the surroundings to be less likely.

Table 1. Point-wise linear and nonlinear merge of the multichannel outputs for 100 COD10K images.

Point-wise	Formulation	DSC(%): HP	DSC(%): Canny Edge
Addition	$R1 + R2$	83.0	87.0
Product	$R1 \odot R2$	40.0	45.0
Minimum	$\min(R1, R2)$	40.0	45.0
Maximum	$\max(R1, R2)$	83.0	87.0

Local entropy calculation in a given neighborhood depends on the chosen window size (W) around the neighborhood [22]. So, it is imperative to identify an optimal window size that maximizes mean local entropy and COD area localization. To resolve the optimality choice, we screened window size variations with a range between 3 to 27, as demonstrated in Fig. 2.

2.4. Region of Interest (ROI) Selection

The ROI selection process uses the normalized histogram of the local entropy analysis of $I(a,b)_{\text{Mod}}$ at the neighborhood of window W. A perpendicular line on the x-axis drawn at γH_{Max} intersects at α, corresponding to the starting point of the local entropy of the ROI region. The ROI retrieval steps uses the pixel range from α to 255, as schematically shown in Fig. 3b. For a multimodal local entropy distribution, as for the green, blue, and black peaks at the left, the α calculation considers multimodality if the second peak is larger than a threshold β. For a multimodal distribution, among the competing local-entropy peak, for instance, H_{Max_1} and H_{Max_2}, the local-entropy peaks producing a greater distance $\max(d_1, d_2)$ from the zero-intensity pixel selects the position of H_{Max}. Subsequently, a perpendicular line drawn on the x-axis for γH_{max} decides α for the ROI selection. The ROI process is schematically shown in Fig. 1b. Here, β and γ are optimizable and are fixed at 0.4 and 0.9, respectively. Algorithm 1 summarizes the underlying steps [1].

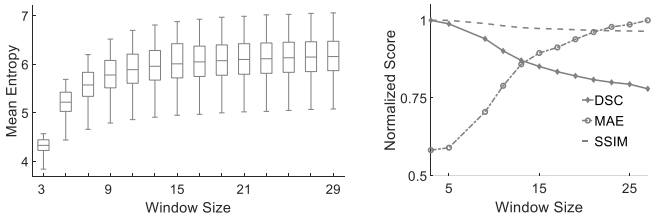

Fig. 2. a) Mean entropy vs. window size for 20 images from COD10K dataset. b) Window size vs. normalized SSIM, DSC, and MAE scores.

2.5. Combining the Multichannel Outputs

This study applies both linear and non-linear merging of the channel outputs, considering that optimal merging often appears highly context-dependent in other areas [23]. For instance, a generic merger $G(R1, R2)$, as in Fig. 1, merges the two localized image versions R_1 and R_2 obtained from the alternative localization channels. The merging processor $G(R1, R2)$ considers point-wise addition, multiplication, minimum, and maximum between pixels.

2.6. Datasets and Metrics

We consider COD10K [24] datasets that include CHAMELEON, CAMO, and NC4K. The DSC score, used as a metric, assesses the segmentation quality by measuring the overlap between the predicted (P) and ground (G) truth regions as $DSC = 2(P \cap G)/P + G$. The DSC value ranges between 0 and 1, and values close to 1 represent better localization of the COD object. Another metric used for comparison is the widely used mean absolute error (MAE) that assesses the pixel-level differences between the predicted (P) and ground-truth (G) map. The other metric used here is the SSIM, ranging between 1 (very similar) and -1 (very dissimilar).

Algorithm 1 ROI Selection Steps for COD images

N images: $\{I_{224\times224}\}$, γ, Metrics = {DSC, SSIM, MAE }, W_{opt}
while *($i \leq N$)* **do**

- Calculate CAM + GAP for image I_i as in [15] for channel R1
- Evaluate 2D-DFT of I_i using Eq. 1
- Filter $|F(u,v)|$ using either HP (or CE) and obtain $|F(u,v)|_{\text{High}}$
- Perform point-wise merging, IDFT (Eq. 2) to obtain $I(a,b)_{\text{Mod}}$
- Calculate local entropy $H(I_{a,b})$ of $I(a,b)_{\text{Mod}}$ using Eq. 3 for W_{opt} evaluated using Metrics.
- Do a histogram analysis of $H(I_{a,b})$ on grayscale and evaluate α for $\gamma = 0.9$ for ROI range $[\alpha, 255]$ in R2 channel
- Combine R1 and R2 as R1 + R2: Final ROI
- Use Final ROI to produce the Image-mask and cut I_i

3. EXPERIMENTAL RESULTS

3.1. Optimal Window Size Improves Localization Accuracy

We hypothesize that localizing COD objects is facilitated by a statistical measure of disorder in a local region. The high contrast in the object-boundary neighborhood is typical for a camouflaging strategy. Hence, a comparatively higher local entropy is relatable to the existence of a COD object in a given locality. Interestingly, our analysis reveals that neighborhood size affects the local entropy calculation and the localization performance. As shown in Fig. 2b, a higher DSC score occurs for smaller window sizes (say, 3×3, 5×5), whereas the mean entropy is the smallest for the window size 3×3 (see Fig. 2a). The proposed method performs a histogram analysis of the local entropy of the modified grayscale image $I(a,b)_{\text{Mod}}$ and uses local entropy-based thresholding to determine the ROI. A higher local entropy threshold may discard the object area, whereas a comparatively smaller local entropy-based threshold may include a large unwanted area reducing the DSC score. Also, neither a too-small W nor a too-large W is optimal for a local entropy analysis. [25]. Thus, an optimal window size is set at 5×5 for the efficacy study of the proposed method.

3.2. Localization Performance in COD Datasets

A localization scheme's effectiveness depends on whether the localized area accurately includes the COD object. The inadequacy of

[1] https://github.com/MohammadRakiburRahman/Multichannel-Localization-.git

Table 2. Performance evaluation of the multichannel localization for CHAMELEON, CAMO, NC4K, and COD10K standard COD datasets using DSC, MAE, and SSIM scores.

Method	Performance Evaluation: Datasets											
	CHAMELEON (75)			NC4K (106)			CAMO (50)			COD10K (500)		
	DSC% (↑)	MAE (↓)	SSIM (% ↑)	DSC	MAE	SSIM	DSC	MAE	SSIM	DSC	MAE	SSIM
R1	0.43	23.53	0.88	0.37	31.26	0.85	0.50	28.87	0.86	0.57	22.65	0.89
R2	0.69	15.07	0.91	0.67	20.18	0.88	0.75	16.15	0.90	0.70	16.53	0.90
R1 + R2	0.73	14.69	0.92	0.71	19.93	0.89	0.81	15.57	0.92	0.83	11.88	0.93

Table 3. Off-the-shelf techniques for 10 COD10K images: Sobel (S), FT: Fourier Transform, D: Denoiser, HP: Highpass, CE: Canny Edge, E: Entropy. HP + S, S + Phase, and others do not produce useful information (data not shown).

Method	DSC	SSIM
1. S → HP → E	0.59	0.81
2. S → LP → E	0.41	0.79
3. S → D → HP → E	0.63	0.89
4. S → D → LP → E	0.44	0.84
5. FT → D → (Phase, HP) → E	**0.70**	0.90
6. FT → D → (Phase, LP) → E	0.69	0.89
For 100 Images		
5A. FT → D → (Phase, HP) → E	**0.67**	0.90
6A. FT → D → (Phase, LP) → E	0.64	0.89
7A. FT → D → (Phase, CE) → E	**0.76**	0.92

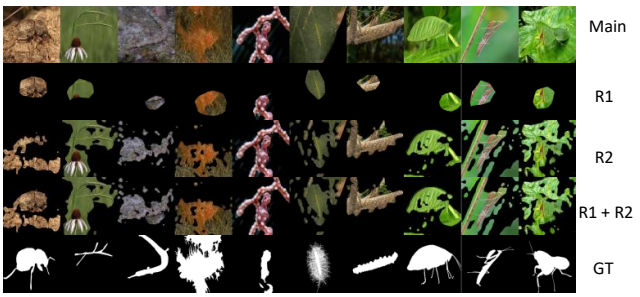

Fig. 3. Visual inspection of the performance of the localization channels for a few COD10K images.

the salient and generic object detection localization approach is evident from the low DSC scores obtained over 500 COD10K images. The CAM + GAP method ($R1$) captures about 57% of the ground truth (GT) object area (Table 2), leaving a significant part of the COD object. In comparison, about 83% (see Table 2) of the GT area of the COD object falls within the localized region obtained via the multichannel (R1 + R2) approach, whereas the proposed R2 channel alone localizes about 70% of COD object. The superior performance of R2 over R1 in COD tasks remains preserved for other datasets (CHAMELEON, NC4K, CAMO). The experimentation considers four alternative datasets measuring three performance metrics MAE, SSIM, and DSC. As in Table 2, the proposed localization channel R2 achieves better DSC, MAE, and SSIM scores than the salient object localization approach implemented in R1 Channel (Fig. 1a). As observed, the multichannel system incorporating R1 and R2 for COD object localization further enhances the DSC, MAE, and SSIM scores. We also estimated energy and carbon emission calculation using the formula available in [26] and compared the proposed R2 channel with other alternative approaches, such as the CAM + GAP [15] (pre-trained) and SINet [12] (requires feature training). The R2 channel is independent of any feature training through a neural network or other means, so it appears lightweight against pre-trained R2 and SINet (see Table 4).

3.3. Ablation Study of the Proposed Method

The impact of the proposed R2 localization channel is evident from an exhaustive comparison done in the ablation study (shown in Table 2). The R2 channel considerably improves DSC, MAE, and SSIM metrics if compared to the localization performance of R1 alone. Also, the image portion produced from the localization reflects its ability to include the COD object in challenging surroundings (see Fig. 3). Besides, ablation of the submodule performing the local entropy of $I(a,b)_{\text{Mod}}$ reduces the DSC score, as evident from a reduction of DSC 0.70 to 0.48 and the SSIM from 0.90 to 0.87 for a randomly chosen ten images (data not shown).

Table 4. Estimated energy and carbon emission comparison: SINet cost is calculated for 6K COD10K images. The R1, R2, and R1 + R2 cost were initially calculated for 100 images and later linearly extrapolated for 6K images.

Method	Total Power(kWh)	CO_2e
R1	0.36	0.34
R2	**0.03**	**0.03**
R1 + R2	0.39	0.37
SINet [12]	1869	1783

4. CONCLUSION

The proposed method requires no training of features yet can localize camouflaged objects with comparable accuracy and reduces the searchable area of the original COD images. It localizes about 80% of the total GT of the COD object. Many COD mechanisms reported in earlier works are context-dependent, whereas the proposed approach equally applies to non-COD images. Moreover, the proposed is amenable further to enhance its performance, for instance, by optimizing how disruptive coloration can be tackled and how they interrelate with other camouflaging strategies (recursive canny-edge, Gabor filter, etc.) [27]. Our immediate improvement study demonstrates potential accuracy improvement by adding a canny-edge instead of an edge-preserving high-pass filter (Fig.1a) as reported in Table. 3. Intriguingly, the proposed work requires no memory-intensive devices common for other mechanisms of COD tasks, staying computationally inexpensive.

5. REFERENCES

[1] Shuo Yang, Huimin Lu, and Jianru Li, "Multifeature fusion-based object detection for intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[2] Deng-Ping Fan, Ge-Peng Ji, Tao Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao, "Pranet: Parallel reverse attention network for polyp segmentation," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VI 23*. Springer, 2020, pp. 263–273.

[3] Carlos R Del-Blanco, Fernando Jaureguizar, and Narciso García, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 857–862, 2012.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[5] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.

[6] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár, "Panoptic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404–9413.

[7] D Osorio and MV Srinivasan, "Camouflage by edge enhancement in animal coloration patterns and its implications for visual mechanisms," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 244, no. 1310, pp. 81–85, 1991.

[8] Innes C Cuthill, Martin Stevens, Jenna Sheppard, Tracey Maddocks, C Alejandro Párraga, and Tom S Troscianko, "Disruptive coloration and background pattern matching," *Nature*, vol. 434, no. 7029, pp. 72–74, 2005.

[9] Martin Stevens and Sami Merilaita, "Defining disruptive coloration and distinguishing its functions," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1516, pp. 481–488, 2009.

[10] Xiaodi Hou and Liqing Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on computer vision and pattern recognition*. Ieee, 2007, pp. 1–8.

[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[12] Xiaowei Hu, Xuemiao Xu, Yongjie Xiao, Hao Chen, Shengfeng He, Jing Qin, and Pheng-Ann Heng, "Sinet: A scale-insensitive convolutional neural network for fast vehicle detection," *IEEE transactions on intelligent transportation systems*, vol. 20, no. 3, pp. 1010–1019, 2018.

[13] Yunqiu Lv, Jing Zhang, Yuchao Dai, Aixuan Li, Bowen Liu, Nick Barnes, and Deng-Ping Fan, "Simultaneously localize, segment and rank the camouflaged objects," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 11591–11601.

[14] Jinnan Yan, Trung-Nghia Le, Khanh-Duy Nguyen, Minh-Triet Tran, Thanh-Toan Do, and Tam V Nguyen, "Mirrornet: Bio-inspired camouflaged object segmentation," *IEEE Access*, vol. 9, pp. 43290–43300, 2021.

[15] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.

[16] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, , no. 6, pp. 610–621, 1973.

[17] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, , no. 6, pp. 610–621, 1973.

[18] Yue Wu, Yicong Zhou, George Saveriades, Sos Agaian, Joseph P Noonan, and Premkumar Natarajan, "Local shannon entropy measure with statistical tests for image randomness," *Information Sciences*, vol. 222, pp. 323–342, 2013.

[19] Bolei Zhou, Xiaodi Hou, and Liqing Zhang, "A phase discrepancy analysis of object motion," in *Computer Vision–ACCV 2010: 10th Asian Conference on Computer Vision, Queenstown, New Zealand, November 8-12, 2010, Revised Selected Papers, Part III 10*. Springer, 2011, pp. 225–238.

[20] Xi Li and Hichem Sahbi, "Superpixel-based object class segmentation using conditional random fields," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 1101–1104.

[21] Thomas M Cover and Joy A Thomas, "Information theory and statistics," *Elements of information theory*, vol. 1, no. 1, pp. 279–335, 1991.

[22] Chengxin Yan, Nong Sang, and Tianxu Zhang, "Local entropy-based transition region extraction and thresholding," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 2935–2941, 2003.

[23] Shahzada B Rasool and Mark R Bell, "Biologically inspired processing of radar waveforms for enhanced delay-doppler resolution," *IEEE transactions on signal processing*, vol. 59, no. 6, pp. 2698–2709, 2011.

[24] Deng-Ping Fan, Ge-Peng Ji, Guolei Sun, Ming-Ming Cheng, Jianbing Shen, and Ling Shao, "Camouflaged object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2777–2787.

[25] Jacopo Grazzini, Antonio Turiel, and Hussein Yahia, "Entropy estimation and multiscale processing in meteorological satellite images," in *2002 International Conference on Pattern Recognition*. IEEE, 2002, vol. 3, pp. 764–767.

[26] Emma Strubell, Ananya Ganesh, and Andrew McCallum, "Energy and policy considerations for deep learning in nlp," *arXiv preprint arXiv:1906.02243*, 2019.

[27] Jolyon Troscianko, John Skelhorn, and Martin Stevens, "Quantifying camouflage: how to predict detectability from appearance," *BMC evolutionary biology*, vol. 17, no. 1, pp. 1–13, 2017.