Grenze International Journal of Engineering and Technology, June Issue

Camouflage Object Detection System: A Review

Siyamol Chirakkarottu[1], Arya Babu M B[2], Arunima Raj B[3], Anjali V K[4] and Aiswarya Poosari[5]
[1]Assistant Professor, Federal Institute of Science and Technology, Angamaly, Kerala
[2-5]Federal Institute of Science and Technology, Angamaly, Kerala

***Abstract*—Even for the unaided eye, it is difficult to discern camouflaged items in their natural context. CamoNet, an unique bio-inspired network that utilizes both instance segmentation and a bio-inspired attack stream, is proposed for camouflaged object segmentation. Our proposed network, unlike existing segmentation networks, contains two segmentation streams: the main stream and the bio-inspired attack stream, which correlate to the original and flipped pictures, respectively. To increase segmentation accuracy, the output of the bio-inspired assault stream is combined with the result of the mainstream for the final camouflage map. Furthermore, before using an edge detection approach, our method applies an image enhancement technique to improve the underlit and foggy input image. We're also looking for more precise characteristics than simple texture-based features to increase the accuracy of spotting disguised objects.**

***Index Terms*— Camouflage ,Bayesian, Wavelet Transformation,Isoperimetry, Anabranch, Polarization, Deep Learning.**

I. INTRODUCTION

The term "camouflage" refers to the natural, where animals used to conceal themselves by switching their body pattern, contour, or colour to avoid predators. Camouflage is the use of various materials, hues, and illumination to disguise animals or objects by making them challenging to see or making them look to be something else. A leopard's spotted coat, a modernsoldier's battle dress, and the leaf-mimic katydid's wings are all examples. A third technique is motion dazzle, which dazzles the viewer with a notable pattern, temporarily making the object recognisable but difficult to locate. Camouflaged objects, unlike salient objects, are impossible to discern in their natural state, even for humans.

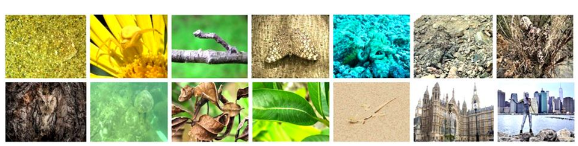

Figure 1. Examples of Camouflaged images

Inevitably detecting and segmenting camouflaged objects is a tough challenge in which discriminative features are no longer viable because we must ignore objects that attract our attention. While detecting camouflaged objects is difficult technically, it is useful in a variety of situations, such as surveillance and search-and-rescue missions. Detecting and segmenting camouflaged objects in images, for example, is undeniably beneficial in the search for camouflaged animals in order to save wild animals and discover new species.

Grenze ID: 01.GIJET.8.2.517_2

Object detection has attracted a lot of interest in recent decades as a result of it's one among the foremost underlying associated difficult issues in computer vision. the aim of visual object detection is to develop machine ways and frameworks that offer the fundamental necessary data by computer vision applications: find specific target class objects in an input image and assign a category label to each object instance.

Camouflaged object detection (COD) is a valuable but difficult task of identifying a camouflaged object against its own background [1]. Image retrieval [2],species discovery, traffic risk management, medical image analysis [3]–[5]and other real-world tasks could all benefit from seeing through camouflage. Existing deep models, on the other hand, are nonetheless not able to completely remedy the intrinsic visible resemblances among foreground and historical past objects. Current methods to address this problem extract deeper information from the situational context, together with attributes for identity [1] or classification [6], to significantly increase the underlying depictions for camouflaged item detection.

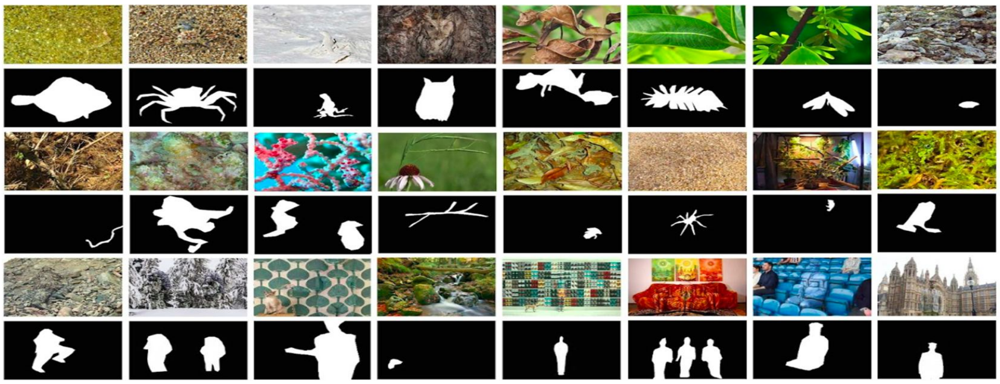

Figure 2. Few examples from Camouflaged Object (CAMO) dataset with corresponding pixel-level annotations

Deep learning has been correctly implemented in computer vision, machine translation, speech recognition, and different fields. Deep learning is meant to enhance pictures best in low-mild situations. Chen et al. [7] proposed a low-mild picture processing pipeline targeted on end-to-end convolutional network training. However, a vast quantity of paired training pictures are required to train this network. Zhang et al. [8]

proposed a self-supervised deep learning network to enhance low-mild photographs. The training of this network calls for a smaller quantity of low-mild photographs and no paired schooling samples. As a result, we enhance and teach this self-supervised deep learning network to enhance low-mild photographs. In this way, in a low-mild environment, we are able to pick out herbal scene records even as detecting

camouflaged artificial targets. Near-ideal contrast, illumination, and occlusion are recognised by the vast bulk of available models.These conditions, on the other hand, make it hard to spot camouflaged objects. As an end result of breakthroughs in camouflage sample creation, camouflage styles are becoming higher in their adaptability to a number of herbal environments. The form of the item is disrupted, and its forte is diminished. Camouflage styles also are extensively utilized in herbal contexts like woodlands and deserts. Illumination and occlusion conditions are generally complex and poor. As an end result, fashionable item detection algorithms have the hassle of detecting the camouflaged item.

II. Camouflaged Object Detection Methods

A. Foreground Detection in Camouflaged Scenes [10]

Initially, a background model is created.It serves as a point of reference for future foreground detection operations. The background model represents a scenario in which there are no foreground objects. The input image and background model are then subjected to wavelet modifications. A picture is divided into four frequency bands by each degree of wavelet rework: LL, LH, HL, and HH. The lower frequencies estimate of the picture is represented by the LL band. In the horizontal path, LH has a better frequency approximation. In the vertical path, HL represents a better frequency estimate, while in the diagonal path, HH represents a better frequency approximation.

Then, Local Binary Pattern (LBP) is identified and calculated for each frequency band.The absolute differences of both the coefficients and also the LBP variations assessed by histogram intersection kernel are used to calculate the differences between the actual image as well as the reference image in the wavelet domain.The system is subjected to a variety of weights. To avoid the effect of gaussian noise, weight induced by noise is applied. Translation weights can be used to eliminate inaccuracies in translating images onto the wavelet domain and vice versa. Texture guided weights are used to correct the variations between textures in each frequency band. Finally, the final output is generated using just a texture guided voting mechanism with an appropriate threshold.

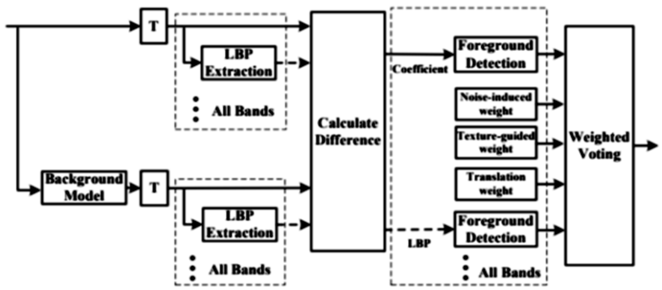

Figure 3. Architecture of the wavelet transformation method

B. A Camouflaged Object Detection Model based on DeepLearning [11]

Object detection can be carried out as a single stage or two-stage process. Preprocessing is done in the two-stage method in the same way that it is in RCNN and faster RCNN. There is no preprocessing stage in the YOLO network because it is a one stage network. The proposed technique employs YOLOv3 in conjunction with Darknet-53. YOLO divides the image into numerous regions using a single CNN, with frames and probability projected for each region. These bounding boxes are weighted using anticipated probabilities. The network examines the entire image and bases its predictions on the image's global context. As a result, one of the major benefits of this network is that it incorporates image context information, which decreases the probability of networks misunderstanding background patches in images for objects. Darknet is a 53-layer CNN that has been pre-trained with the Imagenet database. It can classify roughly 1000 different types of items. The network stabilized after 6k iterations with the image input size set to 416 × 416 pixels.

C. Passive Ground Camouflage Target Recognition based on Gray Feature and Texture Feature in SAR Images [12]

Image segmentation is carried out in two steps - Image binarization and Canny Edge Detection.

- Apply a Gaussian filter to the image to smooth it out and decrease the effect of noise.
- Identify the image's intensity gradients..
- To get rid of spurious edge detection results, utilize non-maximum suppression.
- To find probable edges, use a twofold threshold.
- Hysteresis is used to track edges, and the recognition of edges is completed by suppressing all other edges that seem to be weaker and not related to strong edges.

Then, to determine back-scattering qualities, a grey feature is extracted. A co-occurrence vector is computed using this data. Extraction of texture related features is performed on this matrix. The quantity of local variances in an image is referred to as contrast. The textural homogeneity of a picture is determined by the angular second moment. When the grey level is steady or varies on a regular basis, it reaches its maximum value. The metric of disorderliness is entropy. When all elements in the grey-level co-occurrence matrix are equal, it reaches its maximum. Correlation is a consequence of the filter's displacement, and homogeneity is greatest when the matrix's occurrences are concentrated along the major diagonal. The output is given into the k-means algorithm,

D. A Bayesian Approach for Camouflaged Moving Object Detection [13]

Camouflage modeling is used to discover camouflaged foreground pixels, while discriminative modeling is used to understand non-camouflaged shifting item pixels, consistent with Xiang Zhang et al. Discriminative function-primarily based totally modeling complements discriminative function overall performance in isolating foreground from heritage, however it tries to understand camouflage shifting objects. A widespread version is constructed for the backdrop, and an integration of widespread and neighborhood fashions is generated for the foreground. In discriminative modeling, YUV coloration traits and Spatio-temporal derivatives are used to version the heritage (DM). The heritage is simulated as a worldwide GMM within the context of camouflage modeling (CM). The CM, not like the DM, does now no longer want superior discriminative properties. For camouflage item recognition, the pixels foreground and heritage likelihoods are compared. In a Bayesian framework, DM and CM are included to discover the whole shifting item.

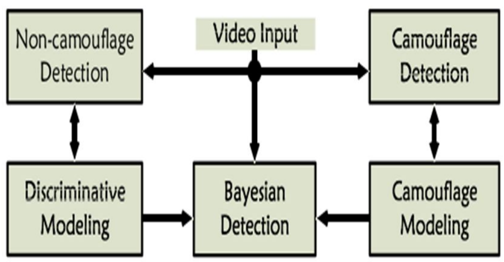

Figure 4. Block Diagram of Bayesian modeling primarily based totally camouflage shifting item detection

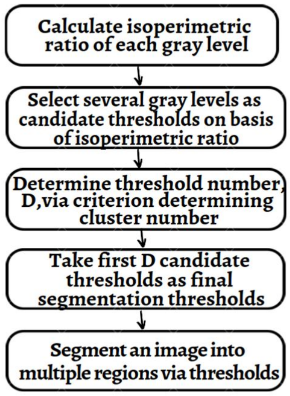

Figure. 5. Flow of the multilevel threshold segmentation algorithm based on isoperimetric theory

E. An Extraction Method for Digital Camouflage Texture Based on Human Visual Perception and Isoperimetric Theory [14]

A new extraction mathematics primarily based totally on human visible belief and isoperimetric concept is proposed to cope with the trouble of loss of subjective visible belief in texture functions exacting of virtual camouflage. The technique first creates an aspect weight feature primarily based totally on human visible belief, then makes use of isoperimetric ratio as a criterion to pick the quality candidate from a listing of candidates, and eventually makes use of an iterative scheme to pick more than one threshold for you to phase the photograph into more than one region.

In the field of digital camouflage, extracting textural features is a hot issue. At the moment, most texture feature extraction methods for digital camouflage at home are based on statistical analysis of background images and do not include human visual perception research [15].

The proposed method of binarization based on isoperimetric theory can be extended to multilevel thresholding segmenta- tion in the practical application of digital camouflage. In order to segment an image into multi-regions, the proposed method uses an iterative scheme to select multiple thresholds.

F. Anabranch Network for Camouflaged Object Segmentation [17]

Not the same as existing organizations for division, the proposed network has the second branch for order to anticipate the likelihood of containing disguised object(s) in a picture, which is then intertwined into the fundamental branch for division to help up the division exactness. Broad investigations led on the recently fabricated dataset exhibit the viability of our organization utilizing different completely convolutional networks.

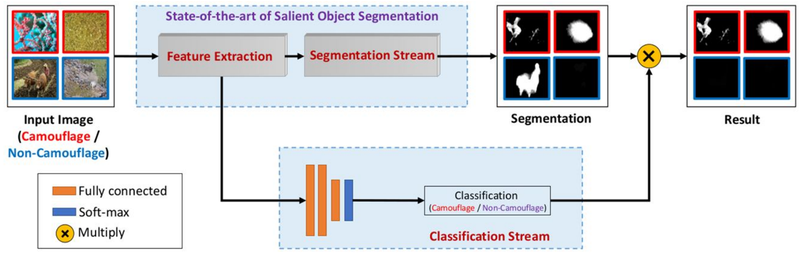

Figure 6. Overview of Anabranch Network (ANet)

The suggested ANet is a generic network for segmenting camouflaged objects. The ANet combines two network models: a classification network and a segmentation network. We'd want to train the CNN model to recognise two types of images: camouflaged and non-camouflaged.

G. Mutual Graph Learning for Camouflaged Object Detection [19]

For camouflage object detection,a perfect model is proposed which is to design a novel Mutual Graph model(MGL). By MGL, there are two task-specific feature maps that are decoupled. The first one for roughly locating the target and the next for accurately capturing its boundary details, and recurrently maintaining their high-order references through graphs and completely uses the mutual advantages. In Figure 7 Given an image as input. For the extraction of the two task-specific features for the camouflaged object detection (COD) and camouflaged object-aware edge extraction (COEE), here uses a ResNet-FCN .So it is called the backbone. With the incorporation of the RIGR, that is Region-Induced Graph Reasoning module and the Edge Constricted Graph Reasoning ECGR and RIGR , that is Region-Induced Graph Reasoning module, and manipulate the mutual uses from both these tasks by reasoning regarding their mutual relations in a recurrent manner. At last these developed characteristics are mapped into the respective results.

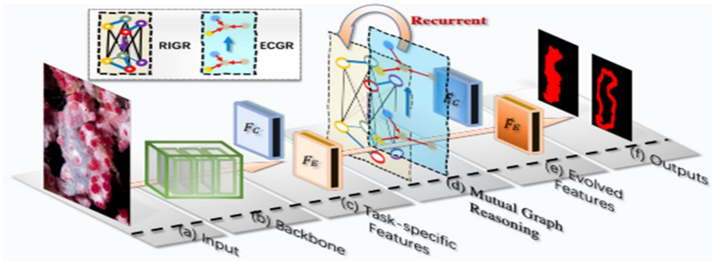

Figure 7. Illustration of the proposed MGL

H. Rapid Detection of Camouflaged Artificial Target Based on Polarization Imaging and Deep Learning [20]

To virtually detect the synthetic targets camouflaged in natural scenes under normal and low illumination conditions, the detection technique was proposed by integrating deep learning and polarization imaging. This technique mainly comprises of 4 steps:

1. Original polarization image preprocessing: Here DoFP polarization camera is employed to detect the initial polarization image and it's split into four intensity images(I). These intensity images include four different polarization directions. Using this intensity Stokes images and also the proposed IS parameter image are often estimated.
2. Polarization signature extraction: For extraction here the Otsu algorithm is applied to the IS parameter image of the target, and morphological operations are wont to make the extraction results more authentic.
3. Low-light image enhancement: A self-supervised deep learning network is employed that's trained to boost low-light S0 Stokes image.
4. Highlighting target location: For highlighting the target location extracted polarization signatures are employed in the polarization intensity image.
   Here for emphasizing the S0 Stokes pictures that's currently in the conversion of RGB to HSV color space and which are H (hue), S (saturation), and V (value) channel images. Concurrently that requires marking whether a pixel within the S0 Stokes image is into the hidden artificial target or the natural environment. For this determination, IE images are used.

I. Detection of People With Camouflage Pattern Via Dense Deconvolution Network [21]

Individuals can be glimpsed by camouflage patterns operating dense deconvolution networks. The camouflaged people detection network using the fully convolutional system brings the entire raw image and the corresponding ground-truth map as the input and output respectively. In the training process, the proposed network directly achieves the learning and modeling of the semantic information. For the detection of people using the camouflage pattern the VGG16 network is altered into a totally convolutional structure. "C1" and "C2" layers in VGG mainly extract the colour and texture features. The "C3","C4", and "C5" layers outputs contain more abundant semantic features. To incorporate multiscale features effectively, all classes of features, omitting the features in "C3" are upsampled and merged into the shallower feature maps. Like that of the input image the four classes of feature maps are then upsampled to the identical measurement. In the end, we fuse all deconvolution semantic maps to emanate the last detection map. The superpixel segmentation and spatial smoothness constraints are applied to the improvement of the detection results.

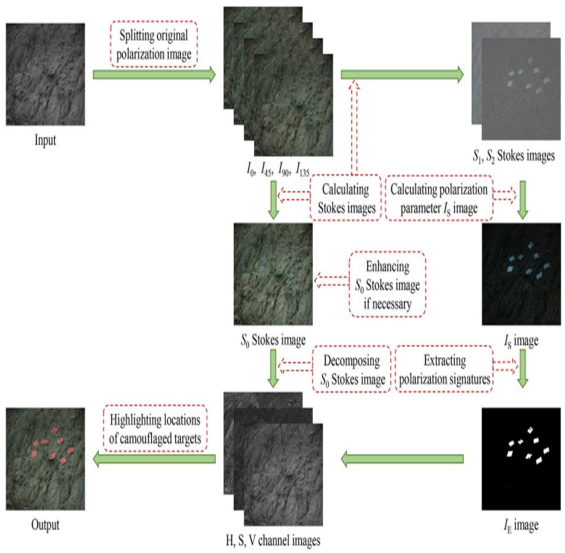

Figure 8. The proposed camouflaged artificial target detection method

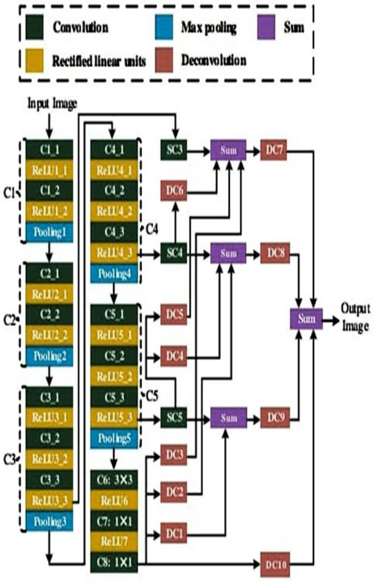

Figure 9. Architecture of DDCN-4C

J. Camouflaged Object Detection [22]

The huge COD study aids in the evaluation of 13 cutting-edge models, revealing some intriguing findings and demonstrating a variety of potential applications. COD10K is a dataset containing 10,000 pictures of camouflaged entities in different natural conditions, split into 78 entity types. To deliver a significant portion of training data to deep learning models COD10K is broken into 6000 and 4000 pictures for training and testing. The receptive field(RF) and the partial decoder(PDC) are two primary parts of SINet framework. The RF is used to replicate the pattern of RFs in the visual system of humans. The PDC simulates the steps of animal predation's search and identification. The collected features are classified as low X0 X1 middle level X2 high-level X3 X4 and combined using concatenation up-sampling, and down-sampling procedures due to this fundamental nature of neural networks. SINet is a competitive system that makes more aesthetically attractive results.

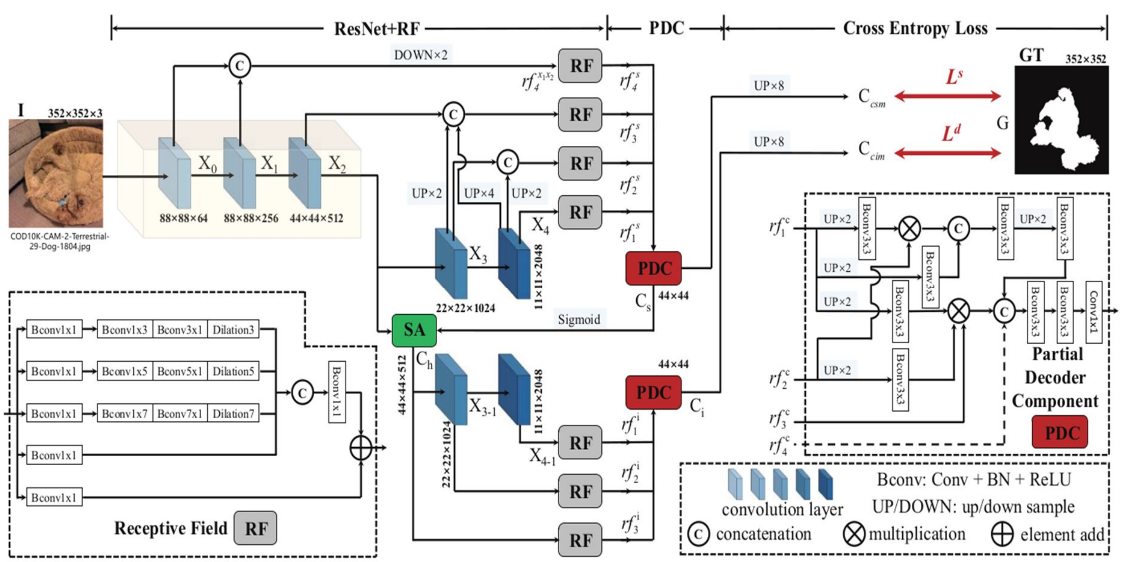

Figure 10. Overview of SINet framework

III. Comparison Study

TABLE I. COMPARATIVE STUDY

Reference Paper	Methodology	Advantage	Disadvantage
Foreground detection in camouflaged scenes	Wavelet transformation	Superior performance and high recall and F-measure	High computational complexity
A camouflaged object detection model based on deep learning	YOLO Object detection	Detecting camouflage level of low,medium and high	Small number of images in dataset
Passive Ground Camouflage Target Recognition based on Gray Feature and Texture Feature in SAR Images	Image segmentation, texture feature extraction, k-means algorithm	Detection of real and passive camouflage targets	Effect of noise can affect efficiency of edge detection
A Bayesian Approach for Camo-uflaged Moving Object Detection	Bayesian detection	A unified Bayesian framework shows superior performance in terms of complete object detection compared with other relevant algorithms.	It is difficult to detect the defective part if different types of defect are present in a frame.
An Extraction Method for Digital Camouflage Texture Based on Human Visual Perception and Isoperimetric Theory	Perception and Isoperimetric Theory	more effective than current threshold methods in segmentation quality	The algorithm is working for partially camouflage image but not for the fully camouflaged image
Anabranch Network for Camouflaged Object Segmentation	Anabranch Network	Deep features replace low-level features even for camouflaged object segmentation by incorporating them into a new network	Extraction and Segmentation cannot be performed simultaneously
Mutual graph learning for camouflaged object detection	Mutual graph learning model	Automatically detecting the objects that blend in with their surroundings	Heavy occlusion and indefinable boundaries affect the model.
Rapid detection of camouflaged artificial target based on polarization imaging and deep learning	Otsu segmentation and morphological operation	Rapid detection of artificial targets camouflaged in natural scenes	Some observational geometries cannot detect artificial targets
Detection of People With Camouflage Pattern Via Dense Deconvolution Network	DDCN-4C	Withstand the influence of camouflage patterns and strong background noise	High computational complexity
Camouflaged object detection	SINet framework	Gives more efficiency for the Medical image segmentation and Search engins.	Small number of images in dataset

IV. Conclusion

In the actual world, camouflaged object detection has a broad array of applications, especially military camouflage detection. The many detection strategies that have been used to detect disguised targets are discussed. Every technique has its own arrangement of advantages and downsides. In this discipline, detecting targets in low light and cloudy images has been a big issue. In the case of camouflaged object segmentation, texture-based features are also more deceiving. As a result, a new and more efficient technique for detecting camouflaged objects is required.

References

[1] D.-P. Fan, G.-P. Ji, G. Sun, M.-M. Cheng, J. Shen, and L. Shao, “Camouflaged object detection,” in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2777–2787, 2020.

[2] G. Liu and D. Fan, “A model of visual attention for natural image retrieval,” in 2013 International Conference on Information Science and Cloud Computing Companion, pp. 728–733, IEEE, 2013.

[3] D.-P. Fan, G.-P. Ji, T. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao,“Pranet: Parallel reverse attention network for polyp segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 263–273, Springer, 2020.

[4] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2626–2637, 2020.

[5] Y.-H. Wu, S.-H. Gao, J. Mei, J. Xu, D.-P. Fan, R.-G. Zhang, and M.-M. Cheng, "Jcs: An explainable covid-19 diagnosis system by joint classification and segmentation," IEEE Transactions on Image Processing, vol. 30, pp. 3113–3126, 2021.

[6] T.-N. Le, T. V. Nguyen, Z. Nie, M.-T. Tran, and A. Sugimoto, "Anabranch network for camouflaged object segmentation," Computer Vision and Image Understanding, vol. 184, pp. 45–56, 2019.

[7] A. Boiko and R. Malashin, "Single-frame noise2noise: method of training a neural network without using reference data for video sequence image enhancement," Journal of Optical Technology, vol. 87, no. 10, pp. 567–573, 2020.

[8] Y. Zhang, X. Di, B. Zhang, and C. Wang, "Self-supervised image enhancement network: Training with low light images only," arXiv preprint arXiv:2002.11300, 2020.

[9] A. Tankus and Y. Yeshurun, "Convexity-based visual camouflage breaking," Computer Vision and Image Understanding, vol. 82, no. 3, pp. 208–237, 2001.

[10] Y. Li and A. Gupta, "Beyond grids: Learning graph representations for visual recognition," Advances in Neural Information Processing Systems, vol. 31, pp. 9225–9235, 2018.

[11] Y. Wang, L. Li, X. Yang, X. Wang, and H. Liu, "A camouflaged object detection model based on deep learning," in 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS), pp. 150–153, IEEE, 2020.

[12] M. Huang, W. Xia, L. Huang, Y. Zhou, and Y. Pan, "Passive ground camouflage target recognition based on gray feature and texture feature in sar images," in 2016 CIE International Conference on Radar (RADAR), pp. 1–4, IEEE, 2016.

[13] X. Zhang, C. Zhu, S. Wang, Y. Liu, and M. Ye, "A bayesian approach to camouflaged moving object detection," IEEE transactions on circuits and systems for video technology, vol. 27, no. 9, pp. 2001–2013, 2016.

[14] C. Miao and T. Shaohui, "An extraction method for digital camouflage texture based on human visual perception and isoperimetric theory," in 2017 2nd International Conference on Image, Vision and Computing (ICIVC), pp. 158–162, IEEE, 2017.

[15] S. Brutzer, B. Höferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance," in CVPR 2011, pp. 1937–1944, IEEE, 2011.

[16] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 8, pp. 747–757, 2000.

[17] T.-N. Le, T. V. Nguyen, Z. Nie, M.-T. Tran, and A. Sugimoto, "Anabranch network for camouflaged object segmentation," Computer Vision and Image Understanding, vol. 184, pp. 45–56, 2019.

[18] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Object detectors emerge in deep scene cnns," arXiv preprint arXiv:1412.6856, 2014.

[19] Q. Zhai, X. Li, F. Yang, C. Chen, H. Cheng, and D.-P. Fan, "Mutual graph learning for camouflaged object detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12997–13007, 2021.

[20] Y. Shen, W. Lin, Z. Wang, J. Li, X. Sun, X. Wu, S. Wang, and F. Huang, "Rapid detection of camouflaged artificial target based on polarization imaging and deep learning," IEEE Photonics Journal, vol. 13, no. 4, pp. 1–9, 2021.

[21] Y. Zheng, X. Zhang, F. Wang, T. Cao, M. Sun, and X. Wang, "Detection of people with camouflage pattern via dense deconvolution network," IEEE Signal Processing Letters, vol. 26, no. 1, pp. 29–33, 2018.

[22] D.-P. Fan, G.-P. Ji, G. Sun, M.-M. Cheng, J. Shen, and L. Shao, "Camouflaged object detection," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2777–2787, 2020.