A MULTICHANNEL LOCALIZATION METHOD FOR CAMOUFLAGED OBJECT DETECTION (COD)

by

Md. Rakibur Rahman
Department of Electrical and Computer Engineering
North South University, Dhaka, Bangladesh

IEEE ICIP 2023

Camouflaged Object Detection (COD)

- Why is it difficult to detect camouflaged objects?
  - Camouflaged objects blend into their surroundings, so objects and background appear astonishingly similar
  - They have variability in patterns and lighting conditions
  - They employ occlusion technique to create challenges or obstacles
  - They lack distinctive features for detection algorithms
  - Scarce training data available for deep learning models

Fig: Camouflaged objects from COD10K Dataset

Related Works – Sate-of-the-art (SOTA) Methods

- Search Identification Network - SINet [1]:
  SiNet uses search and identification module to find the exact location of camouflaged objects and identify them.

- Texture-Aware Network- TANet [2]:
  TANet focuses on a texture-aware refinement module(TARM) to extract the texture information and amplifies the texture difference and predicts a detection map. It uses deep learning model.

- Boundary-Guided Network- BGNet [3]:
  BGNet uses boundary-guided network that forces the deep learning model to generate features that highlight object structure promoting COD of accurate boundary localization.

Motivation For The Work

- Challenges in COD with SOTA methods:
  - Memory intensive process**:**
    - Requires higher computation power and time, leading to high power consumption and CO_2 emission.
  - Requires large data for mode of training:
    - For instance, SiNet uses Intel i9-9820X CPU @3.30GHz × 20 and TI-TAN RTX, with a training time around 70 minutes

Proposed Solution

- To reduce the enormous computational burden, we design a multichannel image analysis pipeline to localize camouflaged objects in an image

- Our specific contributions are made by

  - Algorithm design:
    Developing an algorithm tuned to localize the camouflaged objects, capable of narrowing down the region of interest (ROI) in the image

  - Maximizing efficiency**:**
    Comparatively, a fewer pixels are needed to be analyzed, resulting in a less memory and computation requirement

Models and Methods

Off-the-self Approach

- Initially, we screened 28 different filters and combinations to assess their roles in COD tasks:
  - 2D Fourier transform (FT)
  - High pass (HP)
  - Low pass (LP)
  - Canny edge (CE)
  - Hough transformation
  - Local Shannon entropy (E)
  - Gaussian blur as de-nosier (D)
  - Laplacian
  - Thinning
  - Saliency
  - Sobel (S)
  - Water-shedding
- Metrics used: Dice Similarity Coefficient (DSC), Structural Similarity Index Measure (SSIM)

Note: The amplitude of HP and LP is used

Method	DSC	SSIM
S →HP→E	0.59	0.81
S→LP→E	0.41	0.79
S→D→HP→E	0.63	0.89
S→D→LP→E	0.44	0.84
FT→D→(Phase, HP)→E	**0.70**	**0.90**
FT→D→(Phase, LP)→E	0.69	0.89

Table**:** Combination of the filters and relevant DSC and SSIM scores.

Off-the-self Approach

- FT→D→(Phase, Amplitude of HP)→E combination showed prominent results.
- So, we consider this for subsequent analysis found FT→D→(Phase, Amplitude of CE) →E shows better result
- Replacing HP with the Amplitude of CE further improves the performance

Method	DSC	SSIM
FT→D→ (Phase, HP) →E	**0.67**	**0.90**
FT→D→ (Phase, LP) →E	0.64	0.89
FT→D→ (Phase, CE) →E	**0.76**	**0.92**

Table**:** Combination of the selected filters and relevant DSC and SSIM scores.

Block Diagram of the Proposed Multi Channel Localization Method

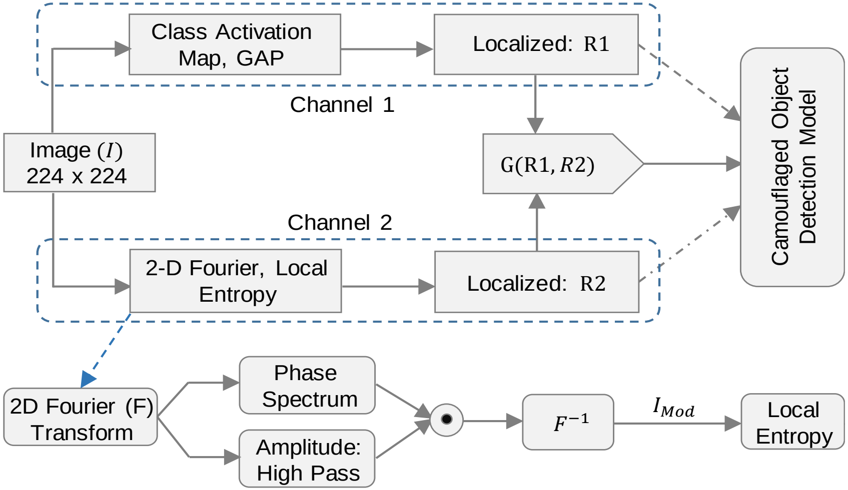

Fig: Block diagram of the proposed multi channel localization method

- Applied global average pooling layer in ResNet50 model to localize object(R1)
- 2D Fourier transform → (pixel-wise addition of phase Spectrum and High Pass) → Local Entropy, to localize camouflaged object
- Merge R1 + R2 to identify the ROI

Window Size Selection for Local Entropy Calculation

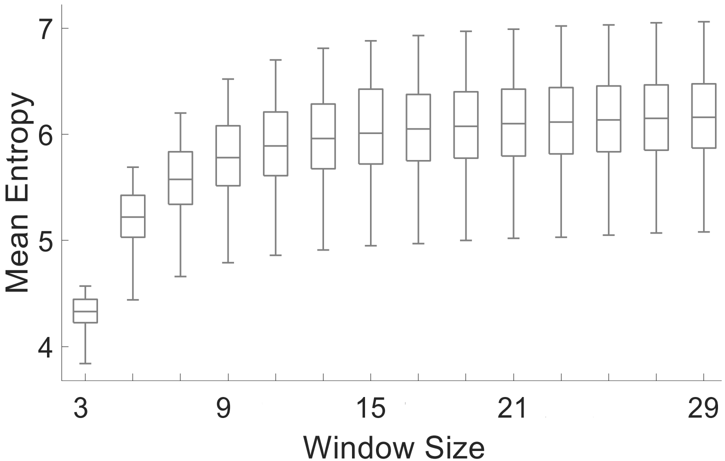

Fig: Mean entropy vs. window size for 20 images from COD10K dataset.

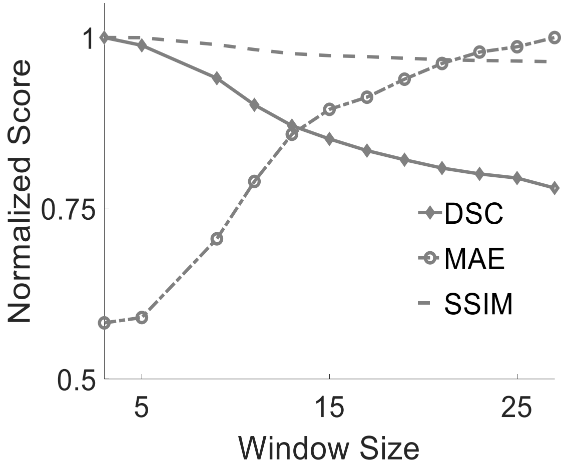

Fig: Window size vs. normalized SSIM, DSC, and MAE scores

- A higher DSC score occurs for smaller window sizes (such as, 3 × 3, 5 × 5), whereas the mean entropy is the smallest for the window size 3 × 3.

Histogram Analysis of Local Entropy

- **γ** - 90% of the maximum pixel intensity level of grayscale local Shannon entropy of their moditual image $I_{(a,\ b)Mod}$
- **β** - 40% of the max pixel intensity level of grayscale local Shannon entropy of their moditual image $I_{(a,\ b)Mod}$
- Do not count pixels less than value of β
- Now, select all the pixels of intensity from **γ** to 255

Channel 2: ROI Selection

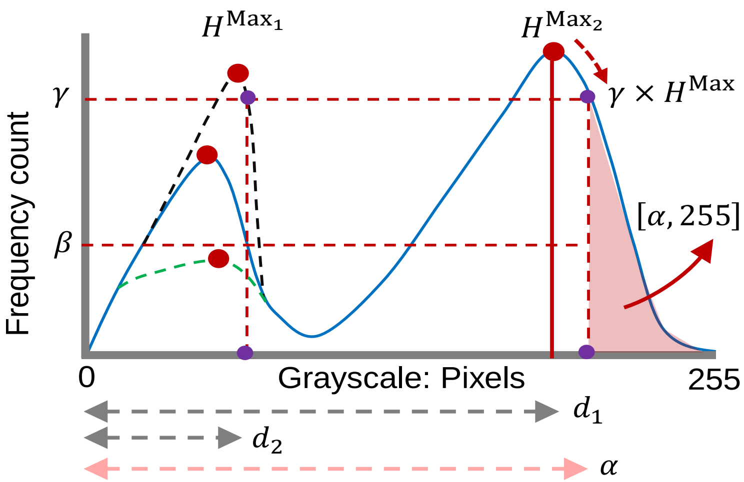

Fig: The histogram analysis of the local entropy of the $I_{(a,\ b)Mod}$ is used for the ROI selection process.

Histogram Analysis of Local Entropy

- Process for selecting the value of γ%
  - To find the best γ% cut we plot DSC vs γ% cut.
  - The plot shows the score starts to drop after 95% so we took 90%

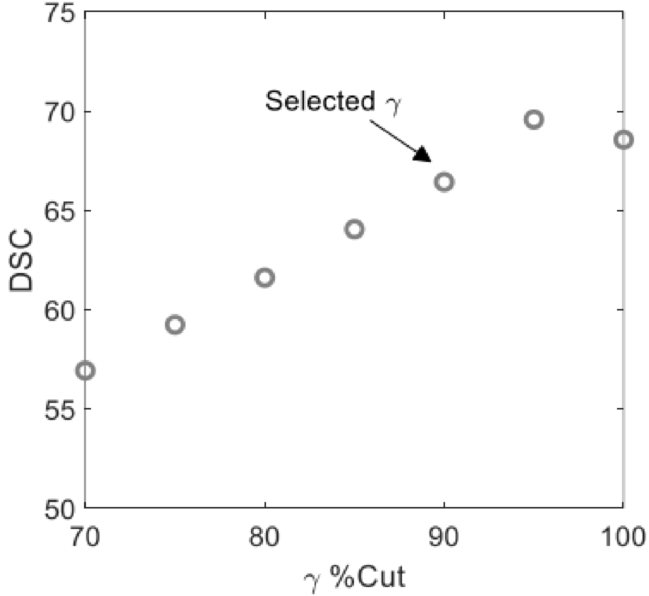

Fig: DSC vs γ% cut

Generic Merger: Alternative approaches

Point-wise	Formulation	DSE(%): 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

Table: Alternative merging techniques of R1 and R2 channel

- Point-wise addition and maximum (R1, R2) appear as better performing techniques

ROI Selection Method

Implemented Steps for ROI selection

1. N = number of images $\{I_{224\times224}\}$ γ, Metrics = {DSC, SSIM, MAE}
2. while ($i \leq N$) do
3. Calculate CAM+GAP for image I_i as for channel R1
4. Evaluate 2D-DFT of I_i Filter $|F(u, v)|$ using either HP (or CE) and obtain $|F(u,v)|High$
5. Perform point-wise merging, IDFT to obtain $I(a, b)_{Mod}$
6. Calculate local entropy $H(I_{a\ b})$ of $I(a, b)_{Mod}$ using W_{opt}
7. Perform a histogram of $H(I_{a\ b})$ on grayscale and evaluate $\boldsymbol{\alpha}$ for γ = 0.9 and identify ROI range [$\boldsymbol{\alpha}$, 225] in R2 channel
8. Combine R1 and R2 as R1+ R2: Final ROI
9. Use Final ROI to produce the Image-mask and cut I_i

Performance Evaluation on Datasets

COD10K (500 images)

Method	DSC% (↑)	MAE(↓)	SSIM%(↑)
R1	0.57	22.65	0.89
R2	0.70	16.53	0.90
R1+R2(ROI)	0.83	11.88	0.93

NC4K (106 images)

Method	DSC% (↑)	MAE(↓)	SSIM%(↑)
R1	0.37	31.26	0.85
R2	**0.67**	**20.18**	**0.88**
R1+R2 (ROI)	**0.71**	**19.93**	**0.89**

- Similar results obtained for CHAMELEON (75 images) and CAMO (50 images)
- R2 increases the overall performance of R1 after merging

Results

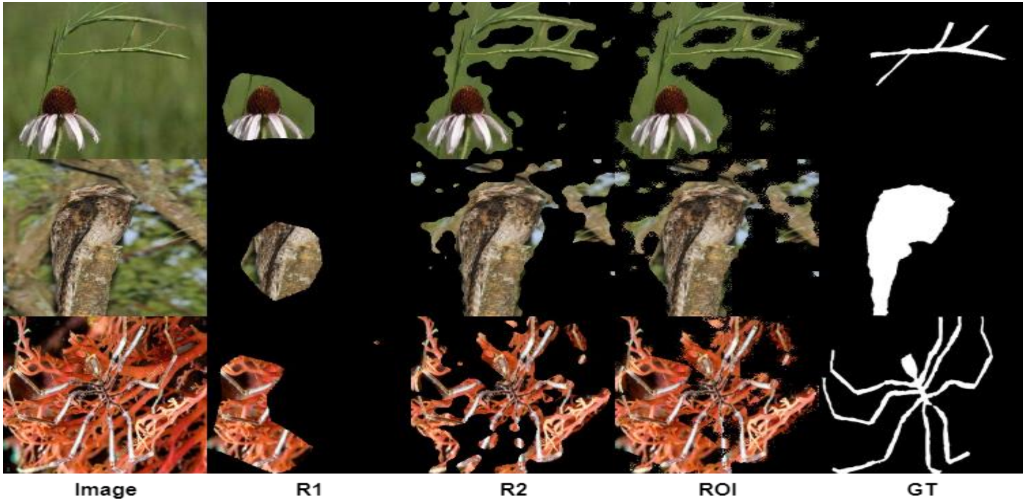

Fig: Results of R1, R2 and ROI

- R1 fails to detect a camouflaged object, and the proposed R2 by an improvement in performance by R2.
- The identified ROI includes the object of interest, thereby reducing the search area.

Power Consumption and Carbon Emission

- Power consumption:

$$p_t = \frac{1.58t(p_c + p_r + gp_g)}{1000} \quad [4]$$

- Carbon emission.

$$CO_2e = 0.954\, p_t \quad [4]$$

- The proposed R2 has a low carbon footprint.

Method	Total Power(KWh)	CO_2e
R1	0.36	0.34
R2	**0.03**	**0.03**
R1+R2	0.39	0.37
SINet	1869	1783

Table: Power consumption and CO_2 emission comparison

Conclusion

- The proposed COD localization approach
  - reduces the search area in COD images by about 80%
  - can localize both COD and non-COD images
  - is computationally efficient and does not require any memory-intensive devices such as GPU, TPU, etc.
  - results in less CO_2 emission

Reference

[1] 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.

[2] J. Ren et al., "Deep Texture-Aware Features for Camouflaged Object Detection," in IEEE Transactions on Circuits and Systems for Video Technology, doi: 10.1109/TCSVT.2021.3126591.

[3] Yijie Zhong, Bo Li, Lv Tang, Senyun Kuang, Shuang Wu, Shouhong Ding; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 4504-4513

[4] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv:1906.02243*.

[5] Lin, M., Chen, Q., & Yan, S. (2013). Network in network. *arXiv preprint arXiv:1312.4400*.

Thank You

Questions?

mohammad.rahman2@northsouth.edu