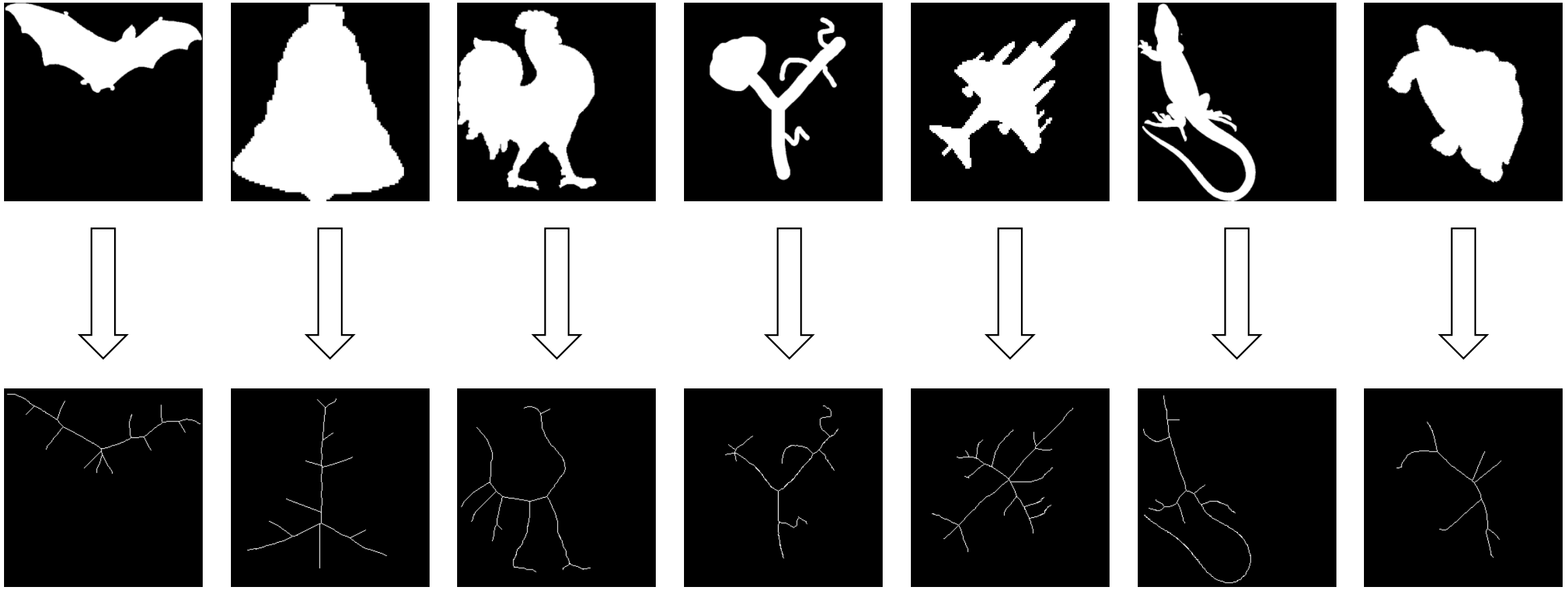
山东科技大学
SHANDONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

CAMION: Cascade Multi-input Multi-output Network for Skeleton Extraction

Sheng Fang, Kaiyu Li, Zhe Li

Deep Learning for **Geometric Computing**
CVPR 2022 Workshop and Challenge

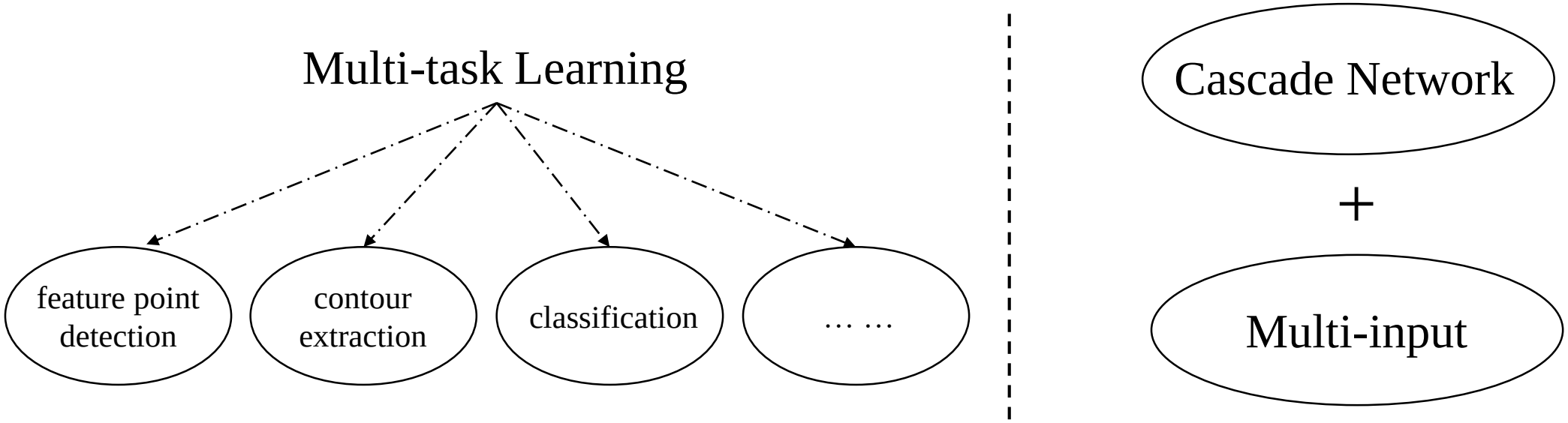
Pixel Skeletonization: Extract skeleton pixels from a binary shape image



Motivation

- Although the shape of the object directly contains the geometric information of a real object, it contains limited semantic information, i.e., without the texture features of the object.

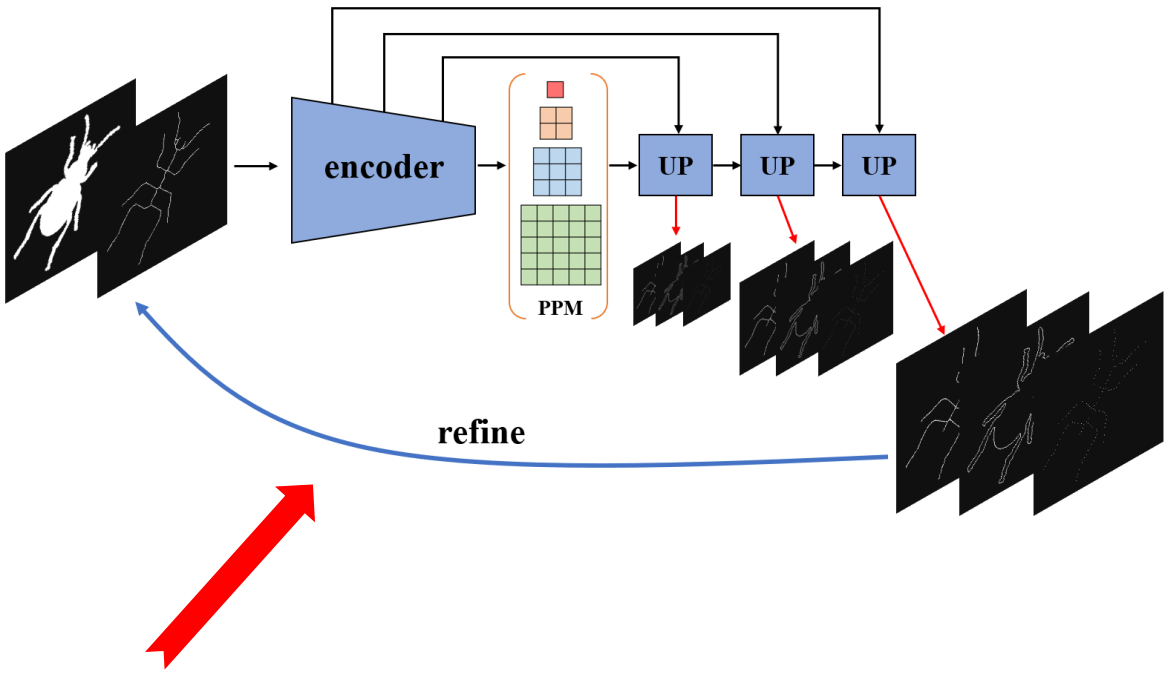
- Some skeletal lines extracted using deep learning methods are discontinuous.



Network Architecture

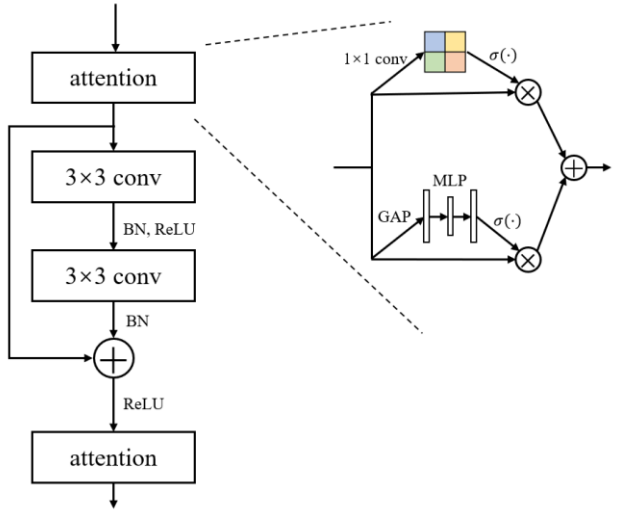
Macro

Illustration of the CAMION.

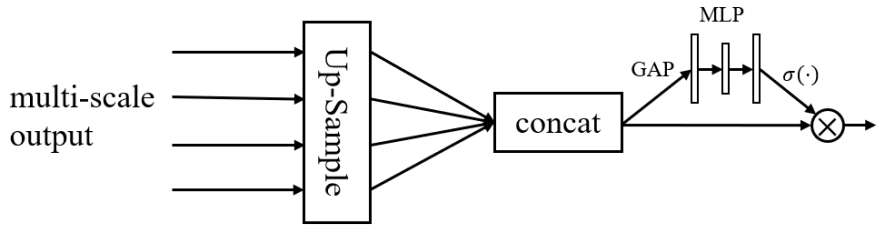


Micro

The basic block of CAMION.



Ensemble attention module.



Loss Function

$$\mathcal{L} = \mathcal{L}_s + \lambda_p \mathcal{L}_p + \lambda_c \mathcal{L}_c \quad (1)$$

$$\mathcal{L}_s = \mathcal{L}_{s\_focal} + \mathcal{L}_{s\_dice} \quad (2)$$

$$\mathcal{L}_p = \mathcal{L}_{p\_focal} + \mathcal{L}_{p\_dice}$$

$$\mathcal{L}_c = \mathcal{L}_{c\_focal} + \mathcal{L}_{c\_dice}$$

Data Augmentations

- All possible flips and rotations multiple 90°.
- Randomly remove some lines from the input skeletons in order to better enable the cascade network to learn the potential task of ”connecting discontinuous skeletons”.

Experiments

PPM	scSE	ensemble att.	F1-score
			0.7903
✓			0.7951
✓	✓		0.8023
✓	✓	✓	**0.8051**

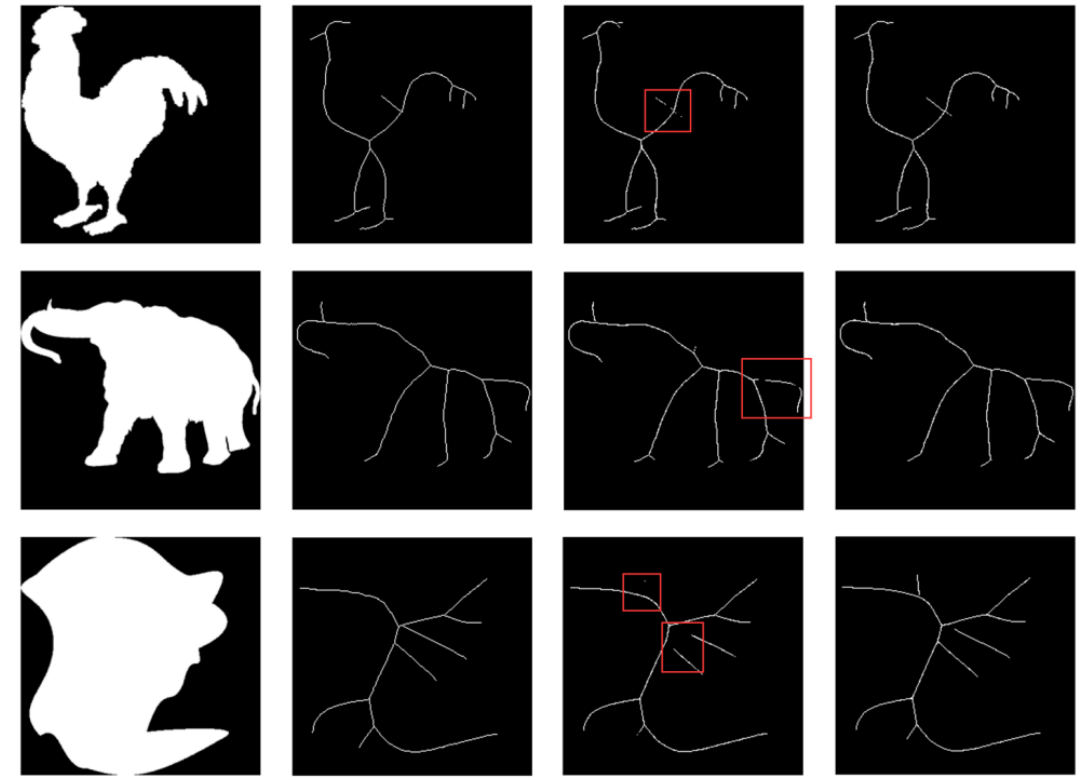
Table 1. Ablation study of several modules of basic network.

skeleton	feature point	contour	cascade	F1-score
✓				0.8051
✓	✓			0.8162
✓	✓	✓		0.8188
✓	✓	✓	✓	**0.8285**

Table 2. Ablation study of multi-task learning and cascade manner.

Mehond	Validation (242)	Test (266)
SkeletonNet [15]	0.7480	0.7711
Panichev et. al. [17]	0.7500	0.7846
Subpixel [5]	0.7708	-
CAMION	-	0.8289

Table 3. Comparison of results on Pixel SkelNetOn validation and testing data.



Visualization results of basic U-Net and CAMION. The first and second columns are the shapes and skeletons provided in Pixel SkelNetOn dataset. The third column is the results generated by our basic U-Net. The fourth column is the results generated by CAMION. Some discontinuous lines and false-negative pixels are indicated by red frames.

Future work

- more relevant tasks
- cascade manner for better models
- … …

References

[1] Sabari Nathan and Priya Kansal. Skeletonnet: Shape pixel to skeleton pixel. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 0–0, 2019. 5
[2] Oleg Panichev and Alona Voloshyna. U-net based convolutional neural network for skeleton extraction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 0–0, 2019. 1, 5
[3] Sohom Dey. Subpixel dense refinement network for skeletonization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 258–259, 2020. 5
[4] Jianbo Shi et al. Good features to track. In 1994 Proceedings of IEEE conference on computer vision and pattern recognition, pages 593–600. IEEE, 1994. 2, 3
[5] Satoshi Suzuki et al. Topological structural analysis of digitized binary images by border following. Computer vision, graphics, and image processing, 30(1):32–46, 1985. 3
[6] Tongjie Y Zhang and Ching Y. Suen. A fast parallel algorithm for thinning digital patterns. Communications of the ACM, 27(3):236–239, 1984. 1, 3

Thanks