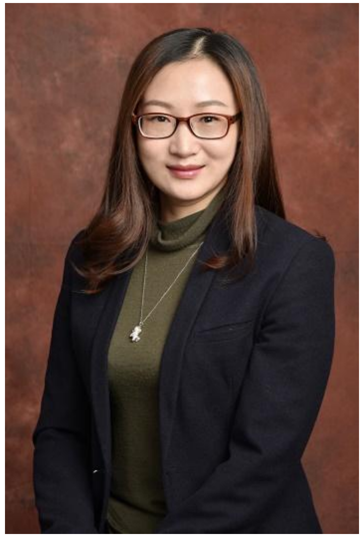
Knowledge-guided Deep Learning for Visual Understanding and Generation

Presenter: Professor Xuemiao Xu

Institute: South China University of Technology

简介

华南理工大学
South China University of Technology



徐雪妙
教授 博导
峻德书院副院长
华南理工大学

- **广东省特支人才；广州市珠江科技新星**；华南理工大学**海外高层次引进人才**；
- 2009年博士毕业于香港中文大学，长期从事计算机视觉、图形图像等方面的研究，及其在智能交通、智能安防、智能制造和智能建筑等领域应用。
- 中国计算机学会和中国图象图形学会的视觉专委会、多媒体技术专委会委员；中国图学学会图学大数据专业委员会委员；广东省计算机学会理事；
- 亚热带建筑科学国家重点实验室-大湾区智慧城市研究方向、广东省机器视觉与虚拟现实技术重点实验室的学术委员会成员，广东省计算智能与网络空间信息重点实验室的学术委员会成员等；
- 在国际重要期刊（TNNLS,TIP等）和会议（SIGGRAPH, CVPR等）发表论文60余篇,其中一作或通信论文的ESI高被引论文1篇、SCI 一区22 篇、顶级会议11篇；授权发明专利12项；
- 近五年主持项目20项，其中主持国家、省部级和国际合作项目10项，到校经费超过2千万，科研成果产业化超11亿产值。
- 获2022年中国图象图形学会科技进步二等奖（第一完成人）；2021年广东省科技进步二等奖（第一完成人）；2021中国公路学会科学技术二等奖（第四完成人）。

简介

张怀东
副教授 博导
华南理工大学
未来技术学院

教育与工作经历

2022~至今	华南理工大学	未来技术学院	副教授
2020~2022	香港理工大学	智能护理中心	博士后
2015~2020	华南理工大学	计算机科学与技术	博士
2011~2015	华南理工大学	计算机全英创新班	学士

主要贡献：

近年在国际重要期刊（TIP，TMM，TVCG，TCSVT等）和会议（CVPR, AAAI, IJCAI等）发表论文20余篇，其中第一或通信作者论文13篇，CCF A类会议5篇，SCI检索14篇。

担任奥比中光3D追光空间站联合实验室负责人

获中国图象图形学学会科技进步二等奖

Content

Background 01

Advanced Technology 02

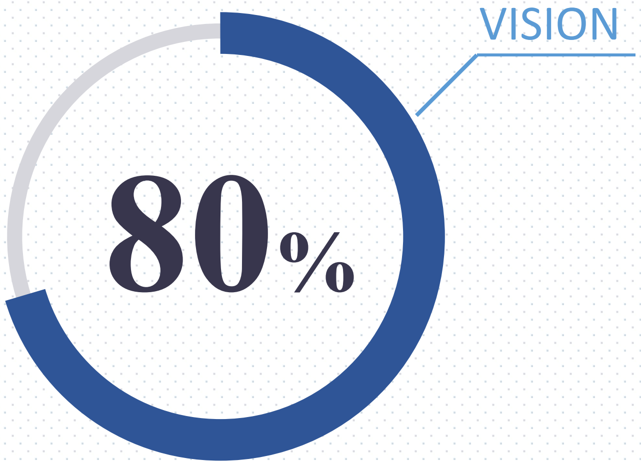
Applications 03

1 Background

1.1 Why is the vision technology important for AI?

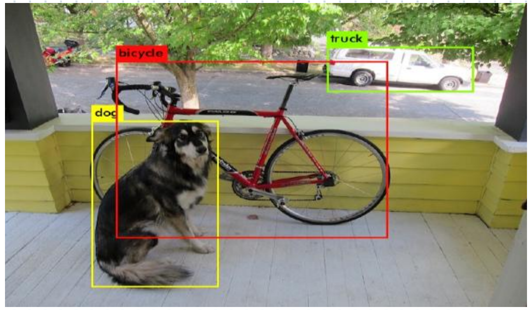
Humans perceive the world through vision, hearing, smell, taste, and touch.

The amount of information from vision accounts for at least 80% of the total amount

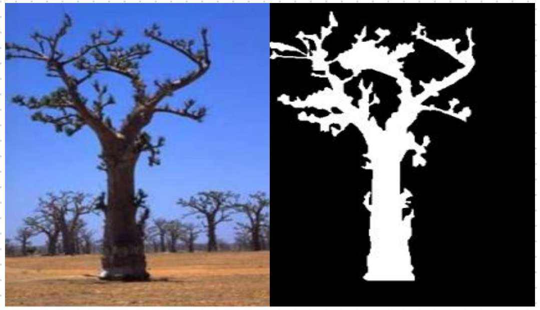


1.2 Core Vision Technology

Scene Understanding

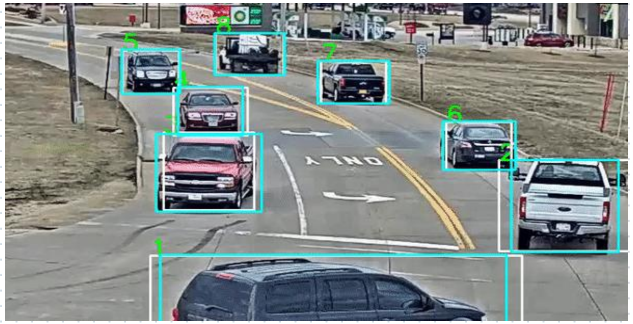


Specific Targets Detection



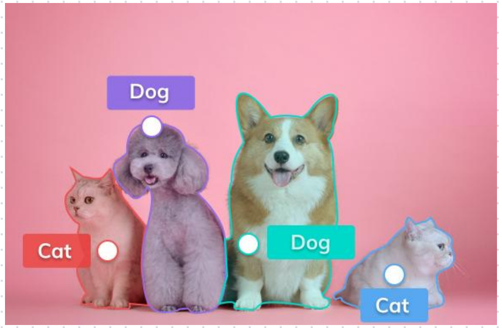
Salient Targets Detection

Classification



Tracking

Semantic Segmentation



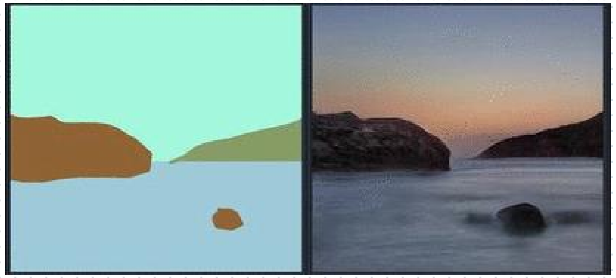
Instance Segmentation

...

Visual Generation

Image Inpainting

Style Transfer



Customized Generation

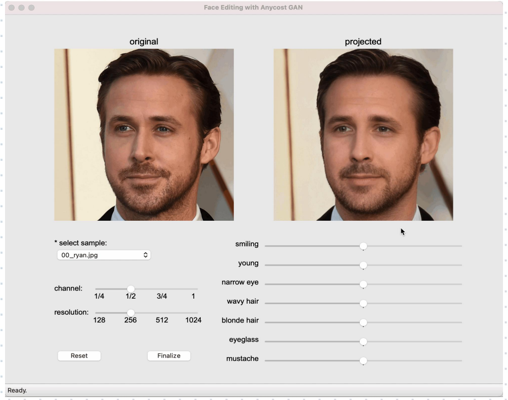
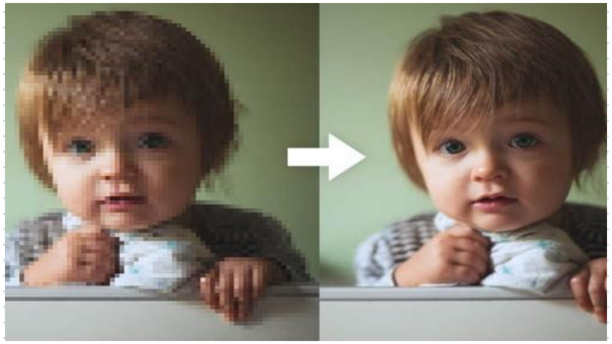


Image Editing



Super-resolution

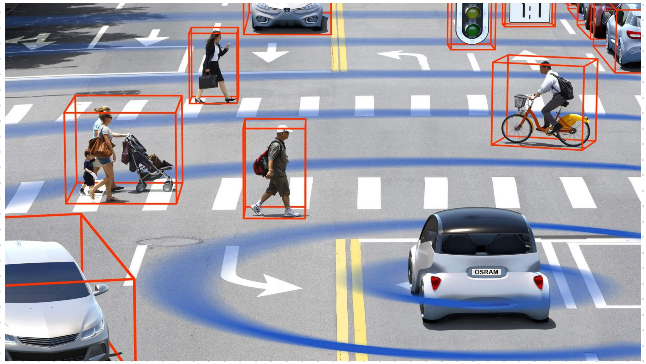
...

1.3 Applications in Industries

Artificial intelligence vision technology empowers the industries!!

Intelligent Education

Analyzing micro-expressions and concentration



Intelligent Transportation

Real-time monitoring the conditions of highway, road and sea area

Intelligent Medical Treatment

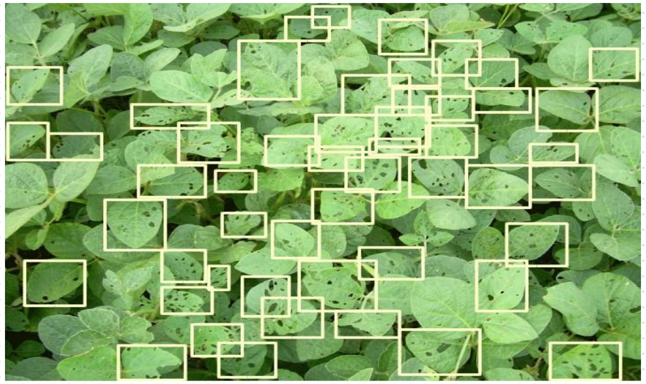
robotic surgery, medical imaging

Intelligent Manufacturing

real-time monitoring or automating the whole production process

Intelligent Monitoring

Shopping mall, schools and buildings



Intelligent Agriculture

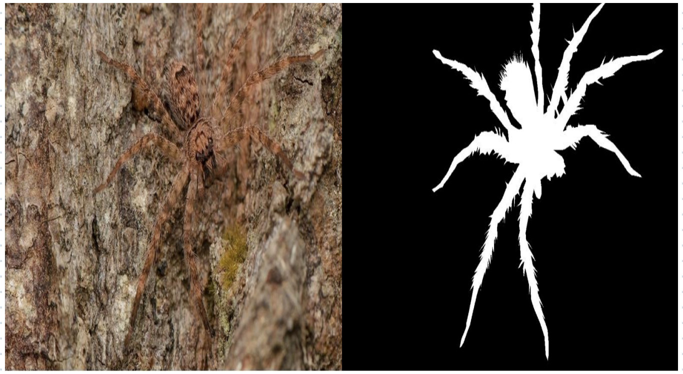
Pest detection and automatic picking

1.4 Challenges

- **Complex visual feature** in different scenes and tasks

Huge scale difference

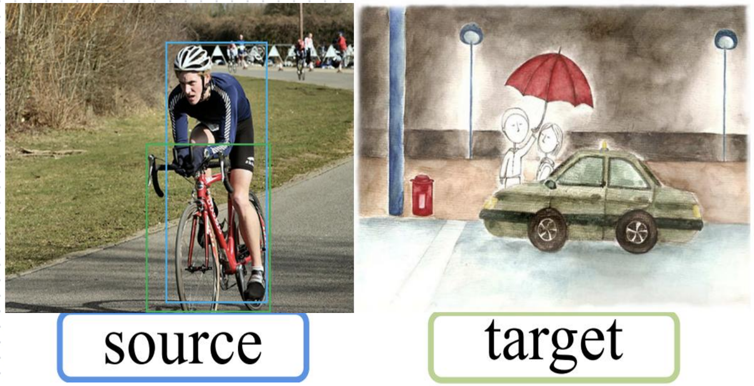
Easily Disturbed by surrounding information



Difficulties in distinguishing due to similar texture

- **Limited visual dataset**, due to the unavailability, the huge workload for labeling etc.

Anomal data missing



Cross scene object detection

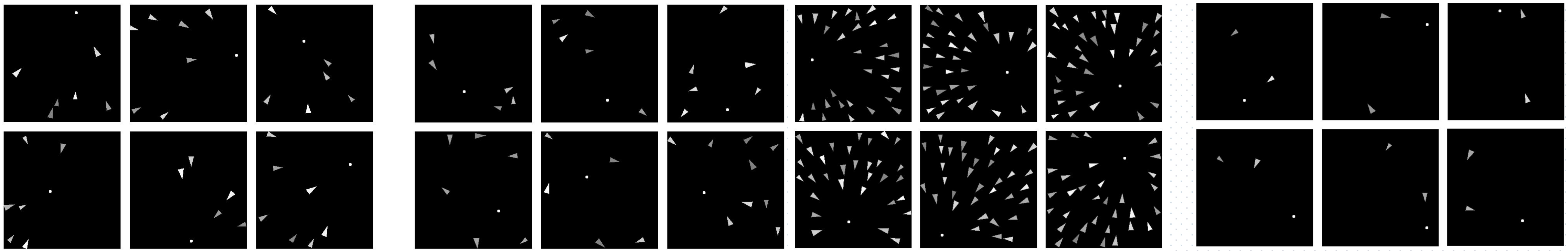
Dense targets

2 Advanced Technology

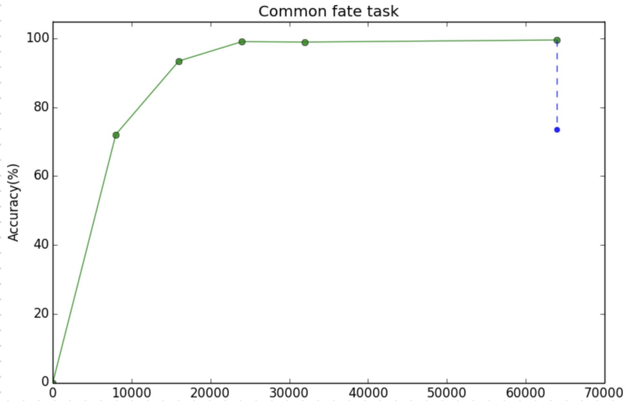
2.1 Trends in Vision Technology: Knowledge Guidance

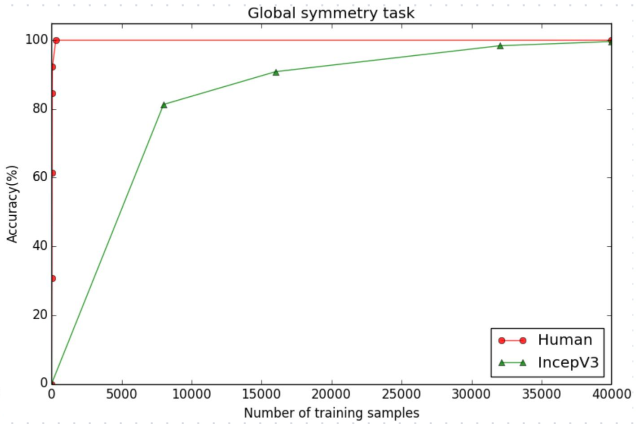
- The current success of deep learning is mainly driven by **data**, neglecting the effective utilization of **external knowledge**.

Common Fate Task[1]



Positive samples   Negative samples   Missed samples





- 20000 training samples are required to improve accuracy to over 90%
- Humans only need 20 samples to obtain 100% accuracy
- The accuracy decreases significantly when the triangle becomes larger
- Lack of knowledge leads to difficulty in learning and weak generalization

Academician Yunhe Pan:
"AI is moving towards a dual wheel drive of data and knowledge"

[1] Yan Z, Zhou X S. How intelligent are convolutional neural networks?[J]. arXiv preprint arXiv:1709.06126, 2017.

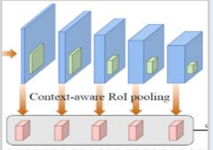
2.2 Overview of Our Work

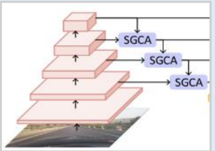
Scene Understanding

Visual Generation

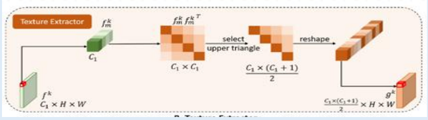
Complex visual feature

Detection



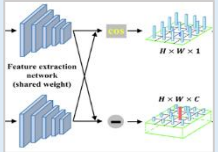


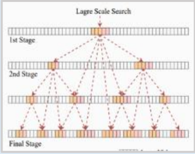
TITS 2018 *TITS 2020*



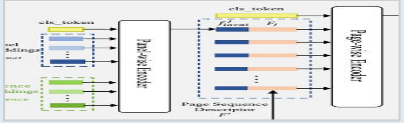
TCSVT 2021

Classification



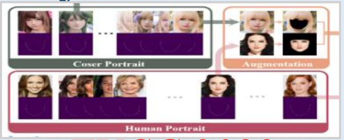


IS 2020 *CVPR 2020*



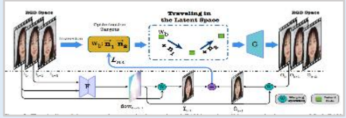
TIP 2023

Style Transfer



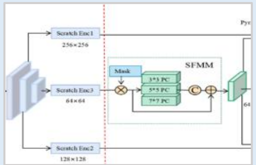
TVCG 2022

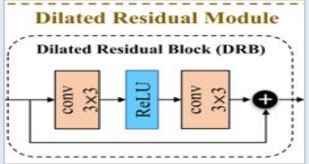
Face Editing



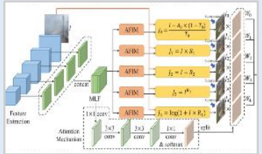
ICCV 2021

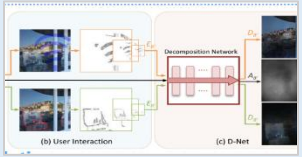
Image Enhancement





TCSVT 2023 *TCSVT 2021*

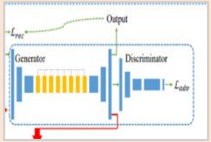


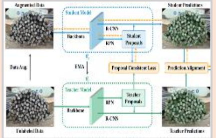


ICCV 2019 *TMM 2019*

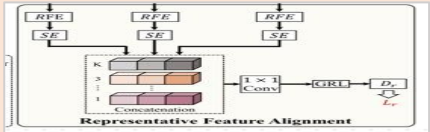
Limited visual dataset

Detection



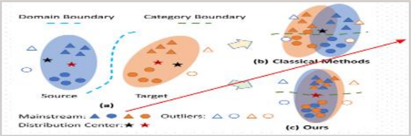


AAAI 2021 *IJCAI 2021*

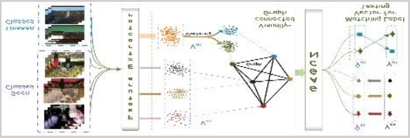


TCSVT 2023

Classification

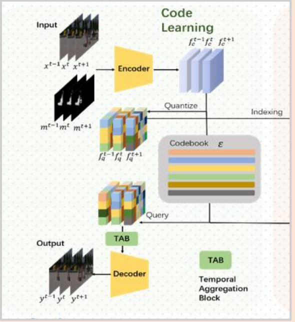


TCSVT 2023



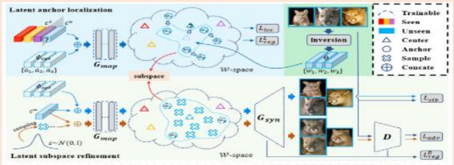
TNNLS 2020

Video Inpainting

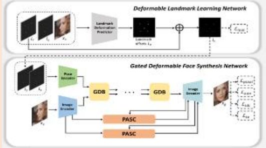


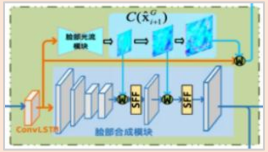
CVPR 2022

Image Generation



CVPR 2023





TNNLS 2022 *TIP 2021*

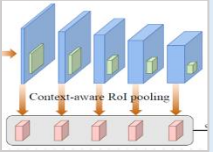
2.2 Overview of Our Work

Scene Understanding

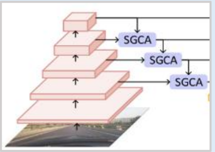
Visual Generation

Complex visual feature

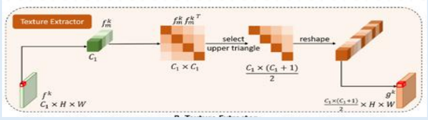
Detection



TITS 2018

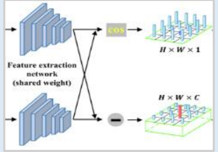


TITS 2020

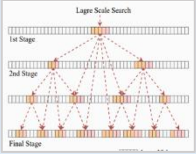


TCSVT 2021

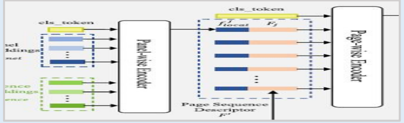
Classification



IS 2020



CVPR 2020



TIP 2023

Style Transfer

TVCG 2022

Face Editing

ICCV 2021

Image Enhancement

TCSVT 2023

TCSVT 2021

ICCV 2019

TMM 2019

Limited visual dataset

Detection

AAAI 2021

IJCAI 2021

TCSVT 2023

Classification

TCSVT 2023

TNNLS 2020

Video Inpainting

CVPR 2022

Image Generation

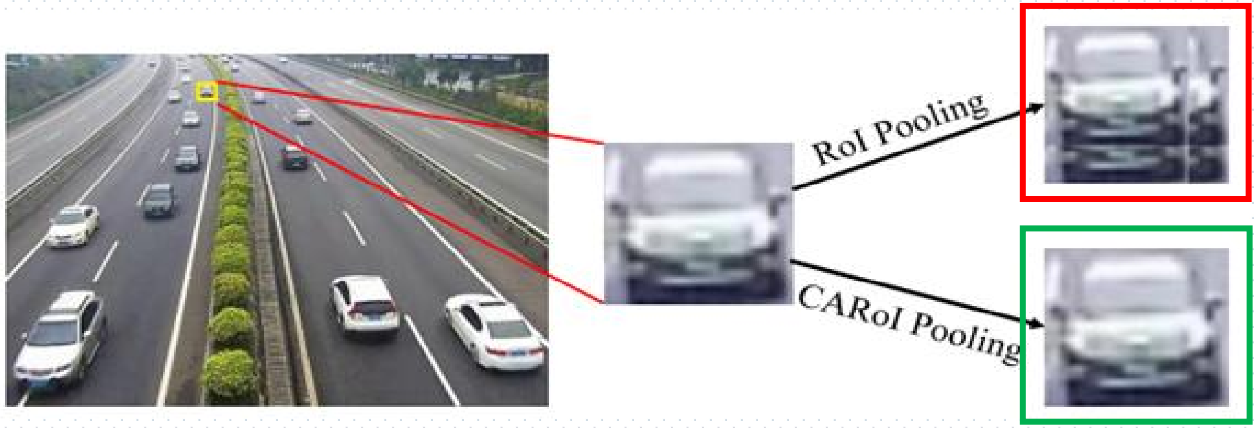
CVPR 2023

TNNLS 2022

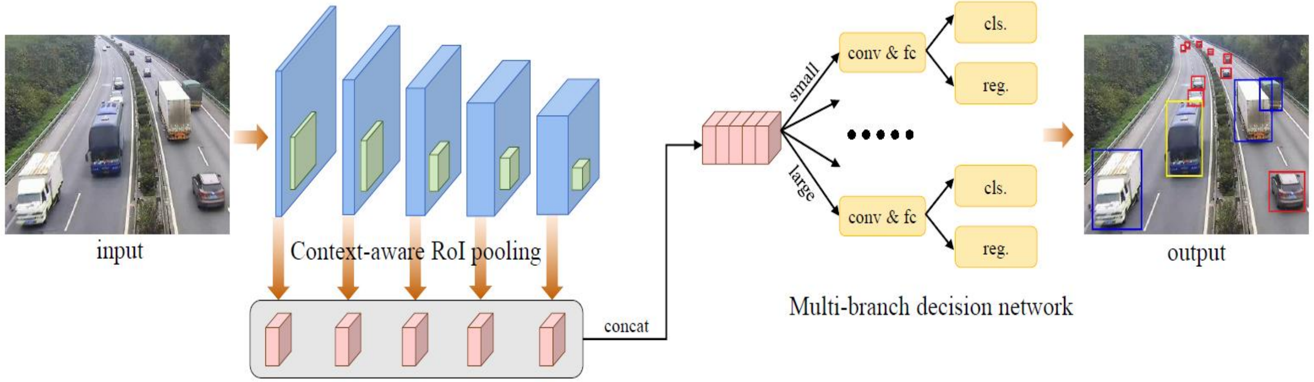
TIP 2021

Detecting Targets with Very Large Scale Difference using Scale Prior Knowledge

Xiaowei Hu, **Xuemiao Xu***, Jing Qin, and Pheng-Ann Heng, SINet: A Scale-insensitive Convolutional Neural Network for Fast Vehicle Detection, *IEEE Transactions on Intelligent Transportation Systems* (***TITS***), 2018. **[IF=8.50, ESI Highly Cited, Google Scholar 237 times]**



Difficulty: **The scale of vehicles changes greatly, but the traditional max pooling mechanism cannot maintain the original structure for the ultra-small targets.**



Our Solution: **Scale prior knowledge, that is, context-aware pooling mechanism, is introduced. For ultra-small targets, bilinear kernel deconvolution operation is used to expand candidate regions without damaging the surrounding structure.**

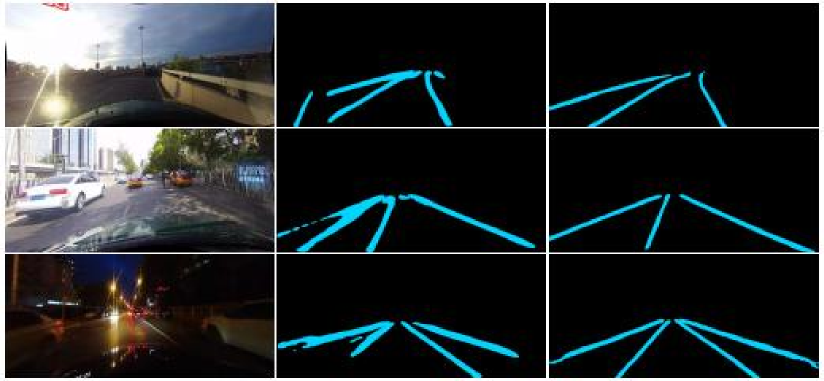
Model	Time/Image	Strategy 1						
		Mean	Sparse			Crowded		
			Car	Bus	Van	Car	Bus	Van
SINet_VGG (ours)	0.20s	**70.17**	**81.82**	**85.60**	**78.65**	**56.80**	**55.78**	62.38
SINet_PVA (ours)	**0.08s**	70.04	81.40	84.39	77.39	53.76	54.06	**69.23**
MS-CNN [28]	0.23s	63.23	79.94	83.71	76.79	51.74	32.95	54.26
Faster RCNN [27]	0.31s	46.44	60.93	66.68	60.14	26.08	24.55	40.24
YOLOv2 [45]	**0.03s**	43.82	59.71	65.51	58.35	17.39	21.55	40.42
YOLO [44]	**0.03s**	16.53	23.06	31.13	22.44	3.87	8.35	10.32

Quantitative Evaluation: **Comparison on our highway dataset**

Visual evaluation: **Results on our highway dataset**

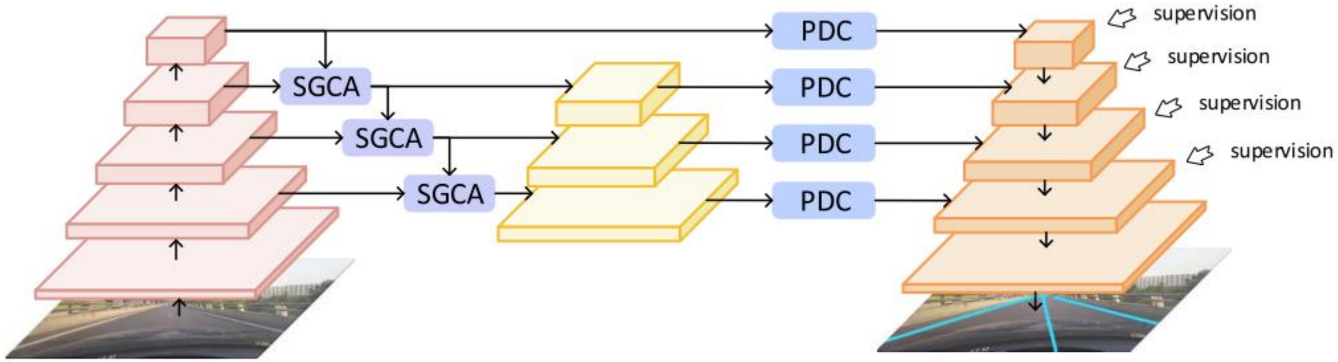
Detecting the Lane Marking with Structure Prior Knowledge

Xu, Xuemiao; Yu, Tianfei; Hu, Xiaowei; Ng, Wing; Heng, Pheng-Ann, SALMNet: A Structure-Aware Lane Marking Detection Network, *IEEE Intelligent Transportation Systems Transactions (**TITS**)*, 2020. **[IF=8.50]**



(a) Input images (b) SAD [2] (c) Our results

Difficulty: **Complex and uncontrollable driving environment and discontinuous lane markings lead to the failure of detection.**

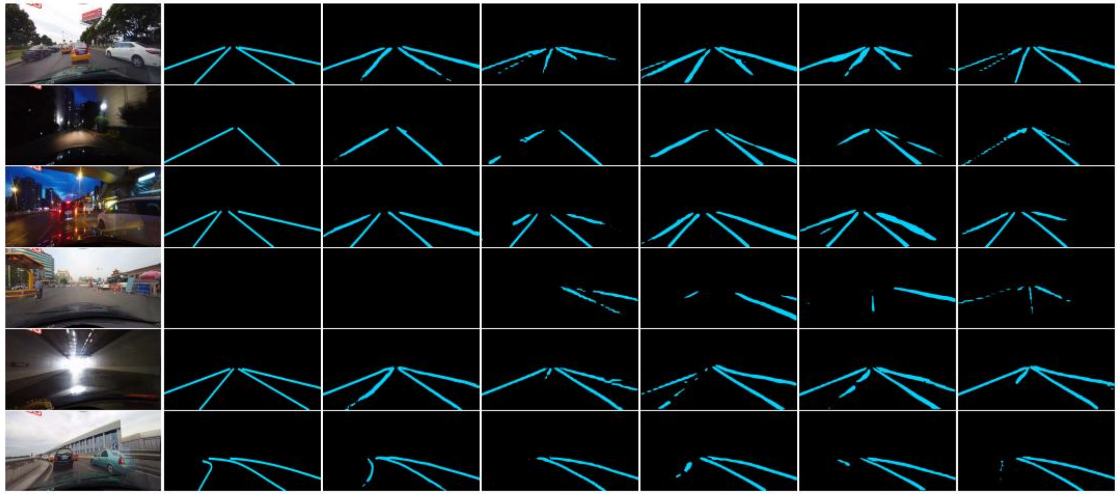


Our Solution: **Structure prior knowledge , that is, the semantic-guided channel attention module and the pyramid deformable convolution module are introduced.**

Method \ Scene	Total	Normal	Crowded	Dazzle light	Shadow	No line	Arrow	Curve	Night	Crossroad
VGG16 [37]	63.2	83.1	61.0	49.9	54.7	34.0	74.0	61.0	56.9	2060
SCNN [12]	71.6	90.6	69.7	58.5	66.9	43.4	84.1	64.4	66.1	**1990**
Stripnet-SCNN [10]	72.2	90.8	69.9	60.0	69.7	44.5	**85.3**	66.1	66.9	2020
SALMNet-VGG16	**72.7**	**91.3**	**71.4**	**64.5**	**70.8**	**45.1**	84.0	**70.3**	**68.1**	3015

Method \ Scene	Total	Normal	Crowded	Dazzle light	Shadow	No line	Arrow	Curve	Night	Crossroad
ResNet50 [36]	66.7	87.4	64.1	54.1	60.7	38.1	79.0	59.8	60.6	2505
Stripnet-ResNet50 [10]	67.4	86.7	65.3	55.5	66.6	39.2	79.7	63.9	61.4	**2468**
FastDraw [40]	-	85.9	63.6	57.0	59.9	40.6	79.4	65.2	57.8	7013
SALMNet-ResNet50	**72.9**	**91.5**	**71.0**	**64.0**	**66.8**	**46.6**	**85.9**	**70.9**	**67.8**	2708

Quantitative Evaluation: **Comparison on public datasets of lane detection**



(a) Input images (b) Ground truth (c) Our results (d) ResNet101 [36] (e) SCNN [12] (f) SAD [2] (g) PSPNet [38]

Visual evaluation: **Results on public datasets of lane detection**

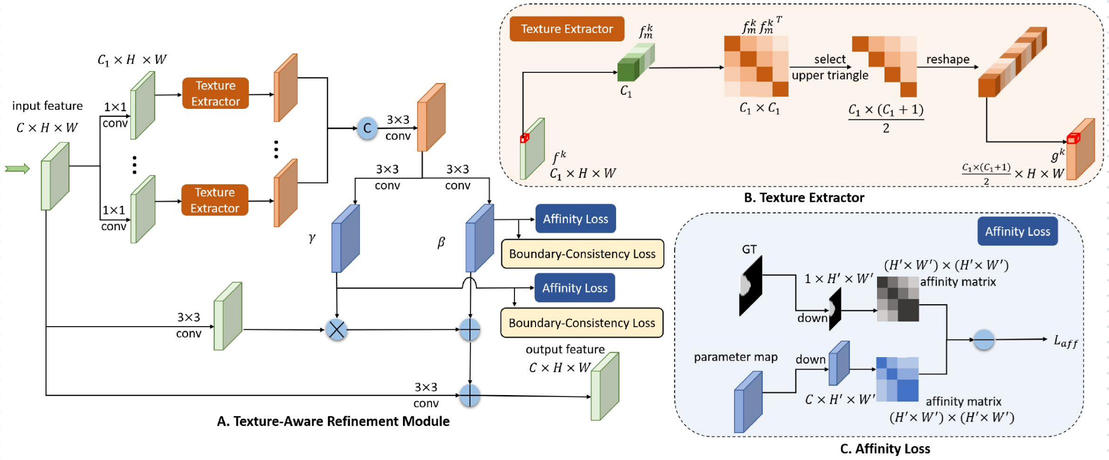
Camouflaged Object Detection with Texture Similarity Priors

Jingjing Ren, Xiaowei Hu, Lei Zhu, **Xuemiao Xu***, Yangyang Xu, Weiming Wang, Zijun Deng, and Pheng-Ann Heng, Deep Texture-aware Features for Camouflaged Object Detection, *IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**)*, 2021.

[IF=8.40, Google Scholar 41times]

(a) input image (b) ground truth (c) convolutional feature (d) texture-aware feature

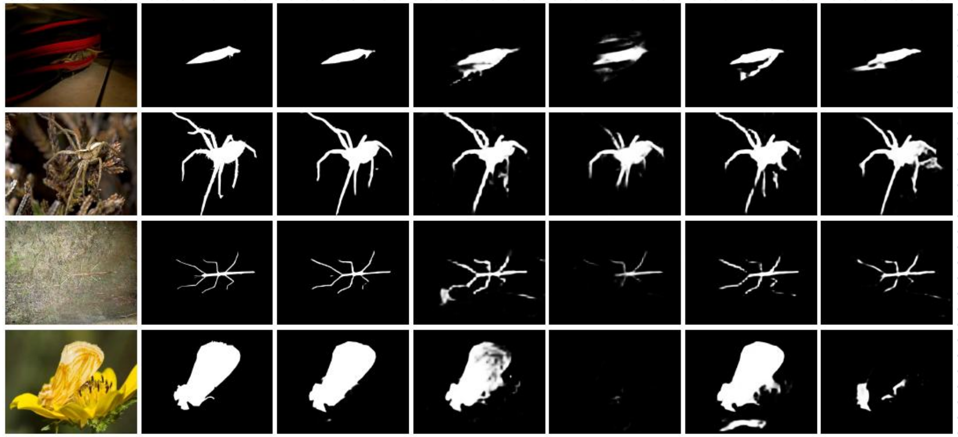
Difficulty: **Camouflaged objects are easily classified as background due to their similar textures with their surroundings.**



Our Solution: **Introducing the texture similarity prior knowledge, the texture difference between camouflaged objects and the background is amplified by learning texture-aware features from the deep neural network.**

Method	Year	CHAMELEON				CAMO-Test				COD10K-Test			
		$S_\alpha \uparrow$	$E_\phi \uparrow$	$F_\beta^w \uparrow$	M $\downarrow$	$S_\alpha \uparrow$	$E_\phi \uparrow$	$F_\beta^w \uparrow$	M $\downarrow$	$S_\alpha \uparrow$	$E_\phi \uparrow$	$F_\beta^w \uparrow$	M $\downarrow$
FPN [30]	2017	0.794	0.783	0.590	0.075	0.684	0.677	0.483	0.131	0.697	0.691	0.411	0.075
MaskRCNN [15]	2017	0.643	0.778	0.518	0.099	0.574	0.715	0.430	0.151	0.613	0.748	0.402	0.080
PSPNet [61]	2017	0.773	0.758	0.555	0.085	0.663	0.659	0.455	0.139	0.678	0.680	0.377	0.080
UNet++ [64]	2018	0.695	0.762	0.501	0.094	0.599	0.653	0.392	0.149	0.623	0.672	0.350	0.086
PiCANet [32]	2018	0.769	0.749	0.536	0.085	0.609	0.584	0.356	0.156	0.649	0.643	0.322	0.090
MSRCNN [20]	2019	0.637	0.686	0.443	0.091	0.617	0.669	0.454	0.133	0.641	0.706	0.419	0.073
BASNet [42]	2019	0.687	0.721	0.474	0.118	0.618	0.661	0.413	0.159	0.634	0.678	0.365	0.105
PFANet [63]	2019	0.679	0.648	0.378	0.144	0.659	0.622	0.391	0.172	0.636	0.618	0.286	0.128
CPD [55]	2019	0.853	0.866	0.706	0.052	0.726	0.729	0.550	0.115	0.747	0.770	0.508	0.059
HTC [2]	2019	0.517	0.489	0.204	0.129	0.476	0.442	0.174	0.172	0.548	0.520	0.221	0.088
EGNet [62]	2019	0.848	0.870	0.702	0.050	0.732	0.768	0.583	0.104	0.737	0.779	0.509	0.056
ANet-SRM [26]	2019	‡	‡	‡	‡	0.682	0.685	0.484	0.126	‡	‡	‡	‡
MirrorNet [56]	2020	‡	‡	‡	‡	0.741	0.804	0.652	0.100	‡	‡	‡	‡
SINet-v1 [9]	2020	0.872	0.936	0.827	0.034	0.745	0.804	0.702	0.092	0.776	0.864	0.679	0.043
TINet [65]	2021	0.874	0.916	0.783	0.038	0.781	0.847	0.678	0.087	0.793	0.848	0.635	0.043
Rank-Net [35]	2021	0.893	0.938	0.822	0.033	0.793	0.826	0.696	0.085	0.793	0.868	0.673	0.041
MGL [59]	2021	0.893	0.923	0.813	0.030	0.775	0.847	0.673	0.088	0.814	0.865	0.666	0.035
PFNet [38]	2021	0.882	0.942	0.810	0.033	0.782	0.852	0.695	0.085	0.800	0.868	0.660	0.040
C^2F-Net [50]	2021	0.888	0.935	0.828	0.032	0.796	0.854	0.719	0.080	0.813	0.890	0.686	0.036
SINet-v2 [9]	2021	0.888	0.942	0.816	0.030	0.820	0.882	0.743	0.070	0.815	0.887	0.680	0.037
TANet (ours)	-	**0.903**	**0.963**	**0.862**	**0.023**	**0.823**	**0.884**	**0.763**	**0.066**	**0.829**	**0.902**	**0.725**	**0.030**

Quantitative Evaluation: **Comparison on camouflaged object datasets**

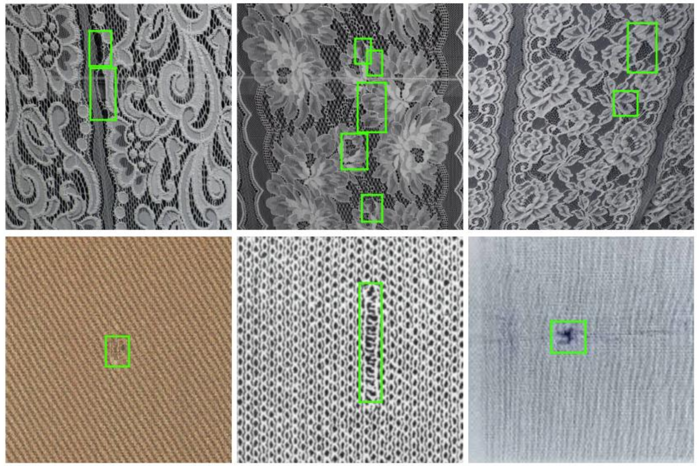


(a) Input (b) GT (c) Ours (d) SINet-V2 (e) MGL (f) C^2F-Net (g) PFNet

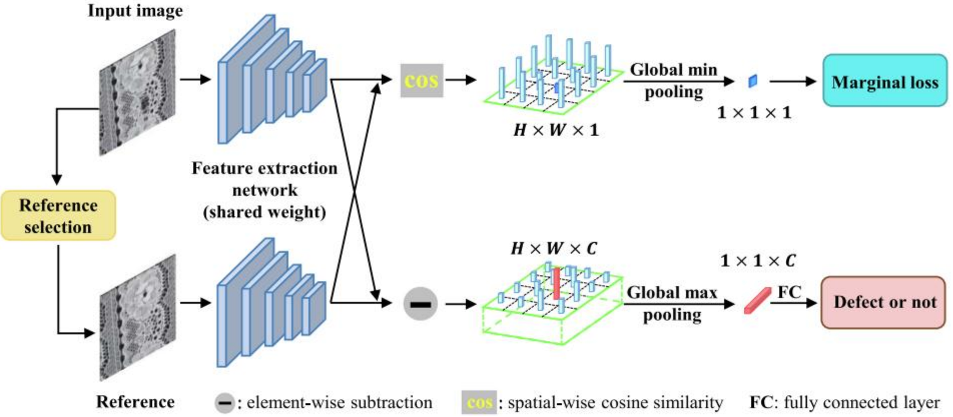
Visual evaluation: **Results on camouflaged object datasets**

Defect Classification for Non-rigid Products with Large Pattern using Deformation Priors

Xuemiao Xu, Jiaxing Chen, Huaidong Zhang, Wing W. Y. Ng, D4Net: De-Deformation Defect Detection Network for Non-Rigid Products with Large Patterns, *Information Science (**IS**)*, 2020. **[IF=8.10]**



Difficulty: **Non-rigid products suffer from large and uncontrollable deformation, thus it is difficult to distinguish whether the difference is due to the defect or deformation.**

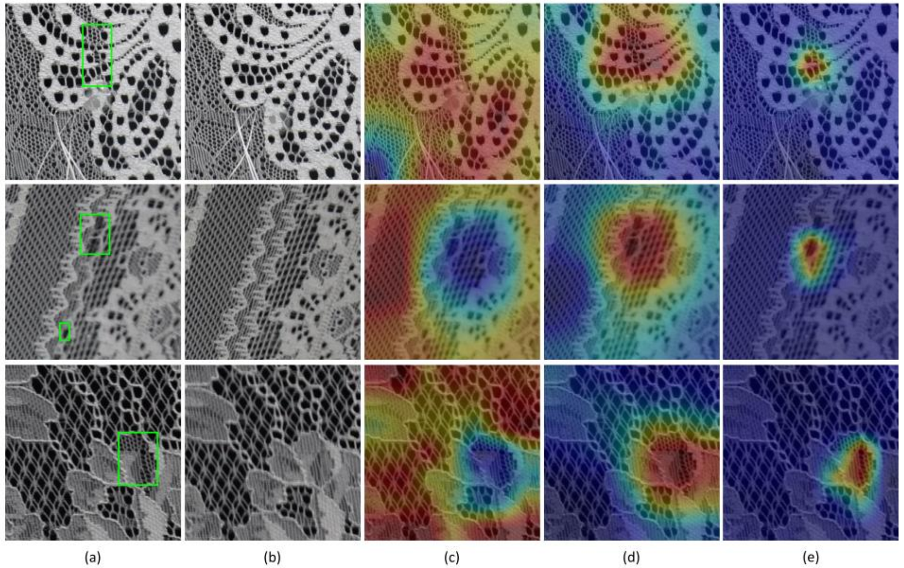


Our Solution: **Introduce the deformation prior knowledge, enforcing the network to distinguish defects and acceptable deformations by integrating an extra task, thereby reducing the impact of local non rigid deformations.**

Method	pre-train	Accuracy	Precision	Recall	F-measure
ResNet [7]	×	0.845	0.880	0.623	0.683
	√	0.906	0.910	0.838	0.855
ResNeXt [26]	×	0.863	0.865	0.681	0.722
	√	0.922	0.906	0.783	0.828
DenseNet [9]	×	0.870	0.854	0.664	0.696
	√	0.918	0.865	0.860	0.847
WACNet [17]	×	0.881	0.753	0.687	0.699
	√	0.932	0.848	0.765	0.782
MICNN [23]	×	0.863	0.894	0.666	0.717
	√	0.927	0.872	0.886	0.870
D4Net	×	0.914	0.945	0.761	0.807
	√	**0.969**	**0.938**	**0.903**	**0.917**

The highest scores are marked with bold.

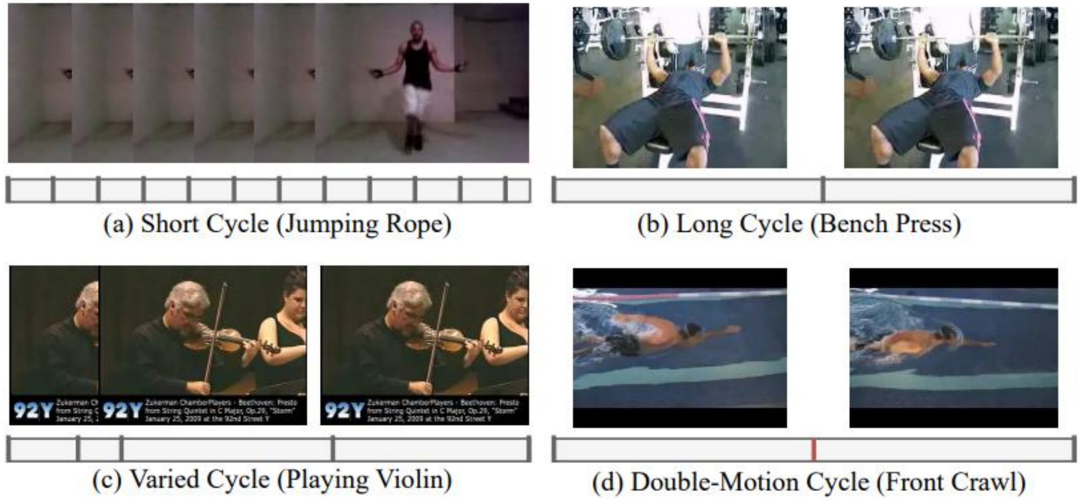
Quantitative Evaluation: **Comparison on our large patterned lace fabrics dataset**



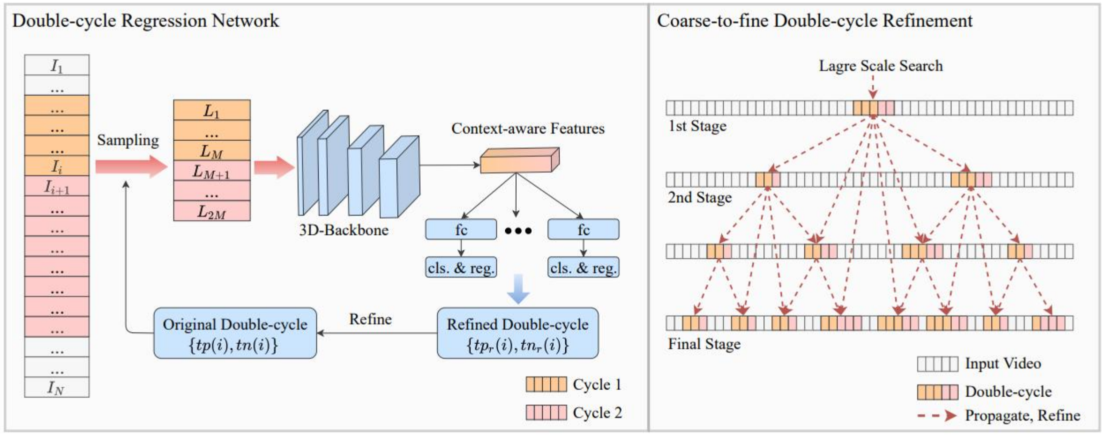
Visual evaluation: **Results on our large patterned lace fabrics dataset**

Counting Temporal Repetitive Patterns with Dual-Cycle Priors

Zhang, Huaidong, **Xu, Xuemiao***, Han, Guoqiang, and He, Shengfeng. Context-aware and Scale-insensitive Temporal Repetition Counting , *IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2020. **[CCF A]**



Difficulty: **Repetitive patterns are difficult to be identified when the scales in temporal for repetitive patterns are changed largely.**



Our Solution: **Introduce dual-cycle prior knowledge, that is a dual-cycle temporal scale regression mechanism. By utilizing a dual-cycle perception-based regression and correction network, repetitive patterns can be recognized at various temporal scales.**

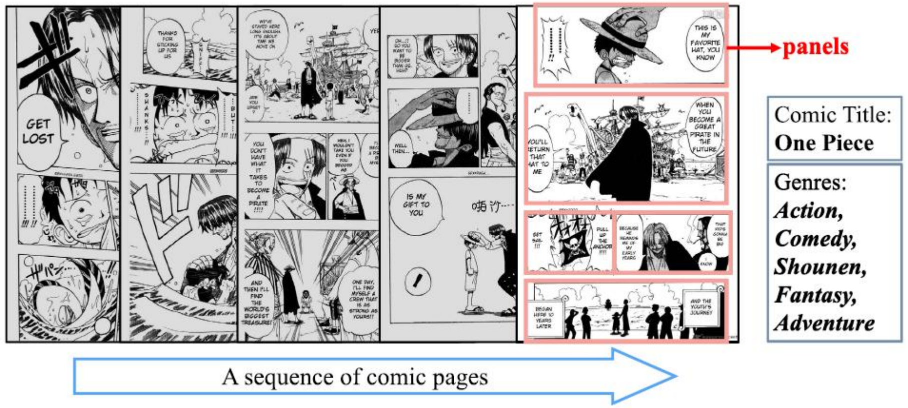
Method	*QUVA Repetition* [25]		*YTsegments* [14]	
	MAE↓	OBOA↑	MAE↓	OBOA↑
Pogalin *et al.* [21]	0.385 ± 0.376	0.49	0.219 ± 0.301	0.68
Levy and Wolf [14]	0.482 ± 0.615	0.45	0.065 ± 0.092	0.90
Levy and Wolf* [14]	0.237 ± 0.339	0.52	0.142 ± 0.231	0.73
Runia *et al.* [25]	0.232 ± 0.344	0.62	0.103 ± 0.198	0.89
Runia *et al.* [26]	0.261 ± 0.396	0.62	0.094 ± 0.174	0.89
Ours-Resnet18	0.190 ± 0.327	0.70	0.062 ± 0.125	0.91
Ours-Resnet50	0.167 ± 0.293	0.75	0.081 ± 0.261	0.94
Ours-Resnet101	$\mathbf{0.148 \pm 0.290}$	0.75	0.066 ± 0.170	0.94
Ours-Resnext101	0.163 ± 0.311	**0.76**	$\mathbf{0.053 \pm 0.115}$	**0.95**

Quantitative Evaluation: **Comparison on QUVA and Ytsegments datasets**

Visual evaluation: **Results on our UCFRep dataset**

Comic Books Classification with Panel-page Priors

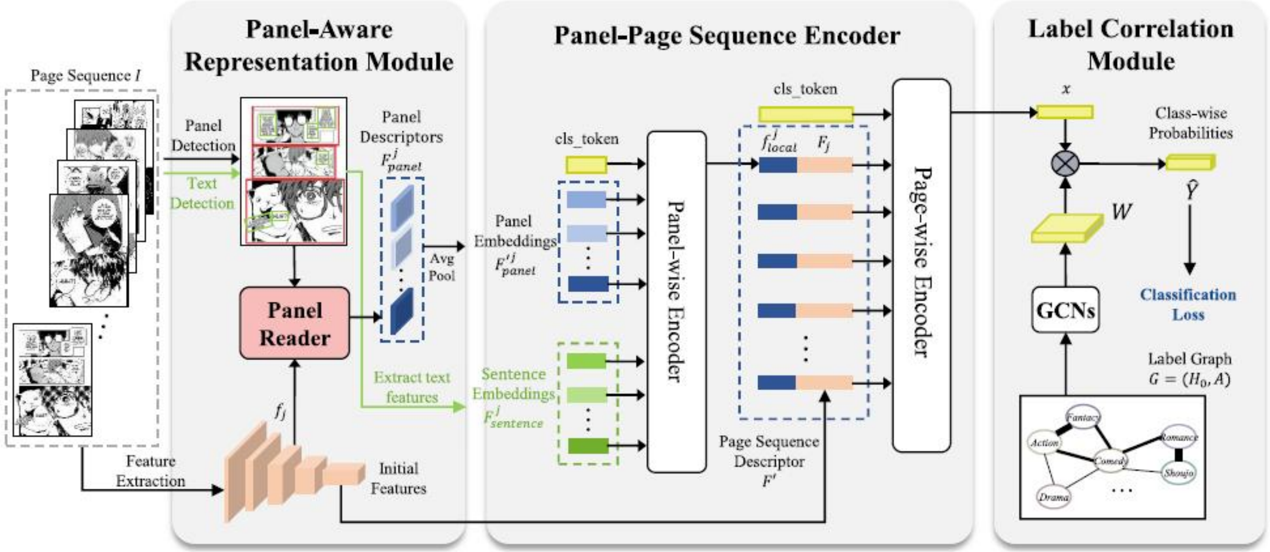
Chenshu Xu, **Xuemiao Xu***, Nanxuan Zhao, Weiwei Cai, Huaidong Zhang*, Chengze Li, Xueting Liu, Panel-Page-Aware Comic Genre Understanding, *IEEE Transactions on Image Processing* (***TIP***), 2023. **[IF=10.60]**



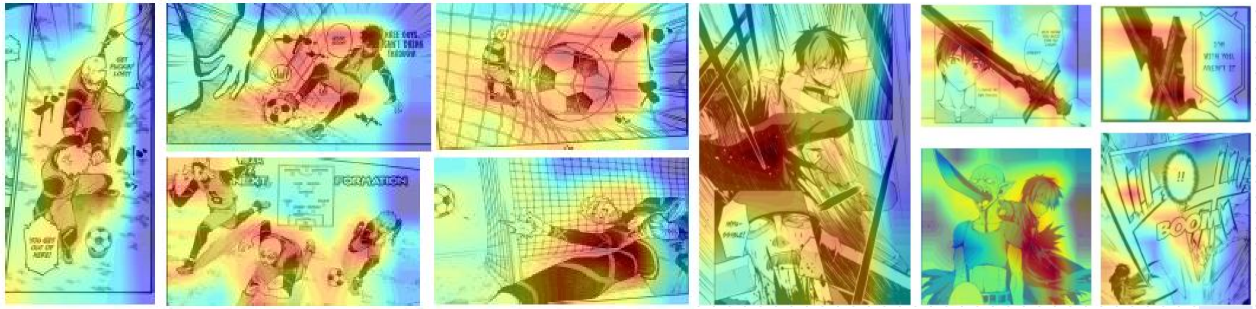
Difficulty: **Different from traditional videos with consistent frames, the paneled nature of comics makes a single page depict completely different scenes for storytelling and voids the assumption of dense pixel-level correspondence.**

Method		Sequence-Level mAP_{Macro}	Sequence-Level mAP_{Micro}	Book-Level mAP_{Macro}	Book-Level mAP_{Micro}
Video-based	I3D (224) [5]	44.18	54.74	50.87	58.98
	I3D (224) [5]+FL+LS	44.60	54.28	52.72	59.77
	TRN (224) [9]	48.17	56.03	53.04	58.38
	TRN (224) [9]+FL+LS	49.92	57.73	54.79	59.38
	SlowFast (224) [10]	40.12	51.50	43.74	53.49
	SlowFast (224) [10]+FL+LS	41.07	48.94	47.05	53.89
	X3D (224) [8]	42.02	51.52	45.98	53.55
	X3D (224) [8]+FL+LS	38.91	49.77	43.76	53.45
	TimeSFormer (224) [33]	37.36	49.33	40.75	50.56
	TimeSFormer (224) [33]+FL+LS	35.89	47.24	39.80	48.07
Image-based	ASL (TResNet-M)(224) [16]	41.01	52.36	51.52	63.35
	ASL (TResNet-L)(448) [16]	37.11	45.79	51.09	59.46
	ML-GCN (224) [13]	38.35	50.78	46.95	59.09
	ML-GCN (448) [13]	39.51	51.22	47.94	59.82
	SSGRL (640) [14]	45.64	52.69	57.62	60.42
	Ours (224)	53.38	62.35	61.12	**67.74**
	Ours (448)	**57.07**	**63.46**	**63.68**	67.59

Quantitative Evaluation: **Comparison with related video-based and image-based works**



Our Solution: **Based on prior knowledge of the panel-page layout of comic books, a framework, P^2Comic, is proposed to understand comic sequences, which is in a panel-to-page manner of feature and semantic understanding and processing.**



Visualization: **Visualization of the attention heatmaps of Intra-Panel Attention**

2.2 Overview of Our Work

Scene Understanding | **Visual Generation**

Complex visual feature

Detection

TITS 2018 *TITS 2020*

TCSVT 2021

Classification

IS 2020 *CVPR 2020*

TIP 2023

Style Transfer

TVCG 2022

Face Editing

ICCV 2021

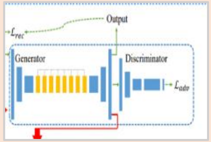
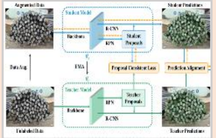
Image Enhancement

TCSVT 2023 *TCSVT 2021*

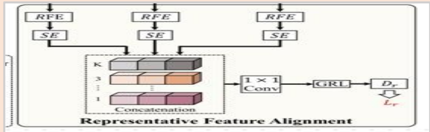
ICCV 2019 *TMM 2019*

Limited visual dataset

Detection

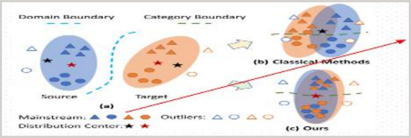



AAAI 2021 *IJCAI 2021*

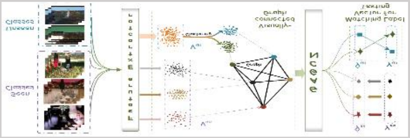


TCSVT 2023

Classification



TCSVT 2023



TNNLS 2020

Video Inpainting

CVPR 2022

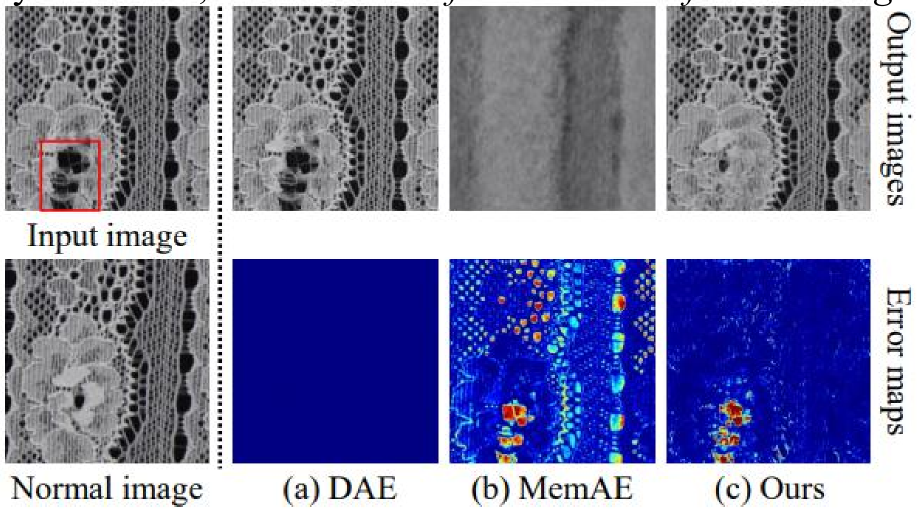
Image Generation

CVPR 2023

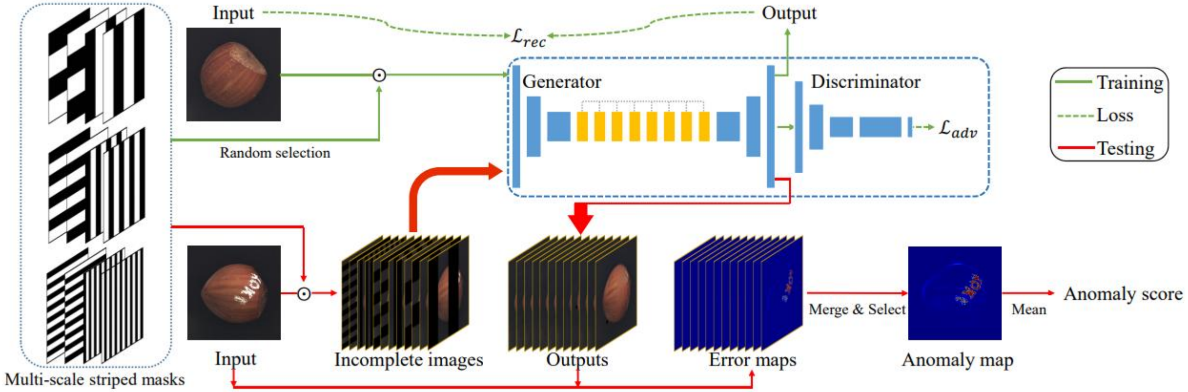
TNNLS 2022 *TIP 2021*

Unsupervised Anomaly Detection with Contextual Semantic Consistency Priors

Xudong Yan, Huaidong Zhang, **Xuemiao Xu***, Xiaowei Hu, Pheng-Ann Heng, Learning Semantic Context from Normal Samples for Unsupervised Anomaly Detection, *The AAAI Conference on Artificial Intelligence (**AAAI**),* 2021. **[CCF A, Google Scholar 46 times]**



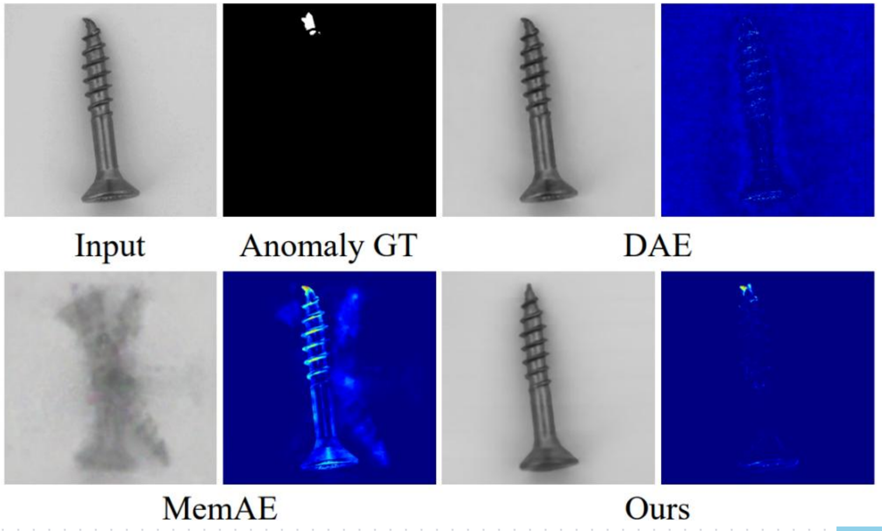
Difficulty: **Anomaly data is difficult to be collected. However, the normal data is easily obtained. We propose to detect the anomaly cases based on the normal data, instead of the anomaly data.**



Our Solution: **Introduce contextual semantic consistency prior knowledge , design a context-aware mask reconstruction mechanism, force the network to learning and reconstruct the accurate normal patterns, so that the anomaly patterns can be detected by comparing the reconstructed pattern and the query pattern.**

Pattern	ALOCC D	ALOCC DR	DAE	OCGAN	MemAE	Ours
1	0.929	**0.986**	0.604	0.648	0.646	0.975
2	**0.847**	0.825	0.638	0.725	0.833	0.843
3	0.870	0.635	0.878	0.380	0.969	**0.984**
4	0.924	0.829	0.810	0.648	0.921	**0.956**
5	0.470	0.367	0.583	0.689	0.810	**0.927**
6	0.821	0.702	0.930	0.898	0.811	**0.949**
7	0.460	0.677	0.647	0.578	0.604	**0.731**
8	0.730	**0.730**	0.478	0.415	0.500	0.660
9	0.807	0.717	0.746	0.671	0.806	**0.935**
10	0.886	0.938	**1.000**	0.937	0.821	**1.000**
11	0.106	0.950	**0.992**	0.935	0.956	0.987
12	0.576	0.499	0.568	0.443	0.648	**0.762**
13	0.799	0.708	0.898	0.840	0.834	**0.998**
14	0.618	0.637	0.686	0.488	0.591	**0.800**
15	0.511	0.791	0.734	0.733	0.729	**0.889**
16	0.930	0.965	0.825	0.816	0.921	**0.998**
17	0.503	0.478	0.761	0.329	0.460	**0.866**
Mean	0.693	0.731	0.752	0.657	0.756	**0.898**

Quantitative Evaluation: **Comparison on our LaceAD datasets**

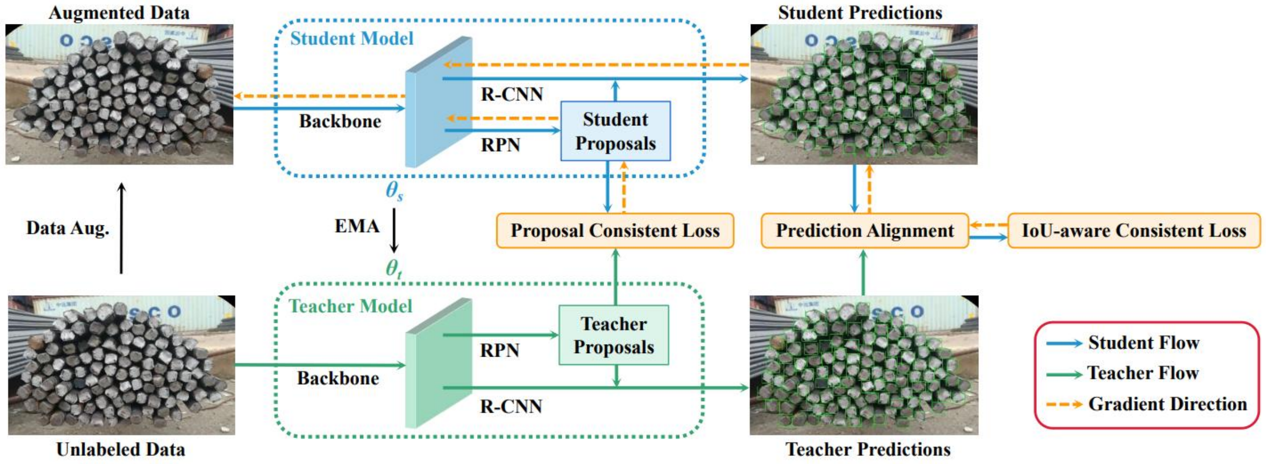


Visual evaluation: **Results on MVTEC AD datasets**

Semi-Supervised Dense Object Detection with Spatial Localization Consistency Priors

Chao Ye, Huaidong Zhang, **Xuemiao Xu***, Weiwei Cai, Jing Qin, Kup-Sze Choi, Object Detection in Densely Packed Scenes via Semi-Supervised Learning with Dual Consistency, *International Joint Conference on Artificial Intelligence* ***(IJCAI)***, 2021. **[CCF A]**

Difficulty: **Dense object detection scenarios involve many objects, making manual annotation costly.**



Our Solution: **Introduce the spatial localization consistency prior knowledge. By incorporating candidate IoU consistency into the teacher-student framework, it enables the student network to learn spatial localization knowledge from the teacher network.**

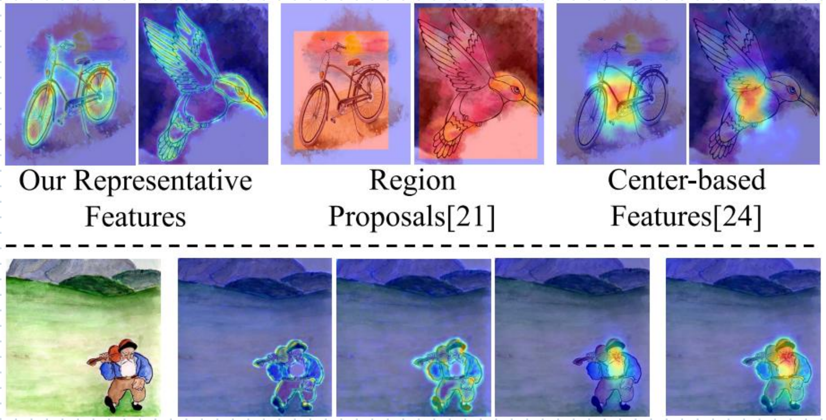
Method	10% AP	10% $AP^{.75}$	10% AR^{300}	20% AP	20% $AP^{.75}$	20% AR^{300}
Backbone [Ren *et al.*, 2016]	57.3	67.7	63.9	59.4	71.2	66.4
CSD [Jeong *et al.*, 2019]	56.7	66.2	62.5	60.6	72.8	65.2
STAC [Sohn *et al.*, 2020b]	57.3	67.4	63.5	59.6	71.2	66.0
SESS [Zhao *et al.*, 2020]	57.2	67.1	63.1	59.3	71.3	65.2
Ours	**59.7**	**70.5**	**66.6**	**62.7**	**76.2**	**68.2**
Gain	**2.4**	**2.8**	**2.7**	**3.3**	**5.0**	**1.8**

Quantitative Evaluation: **Comparison on our steel bar datasets**

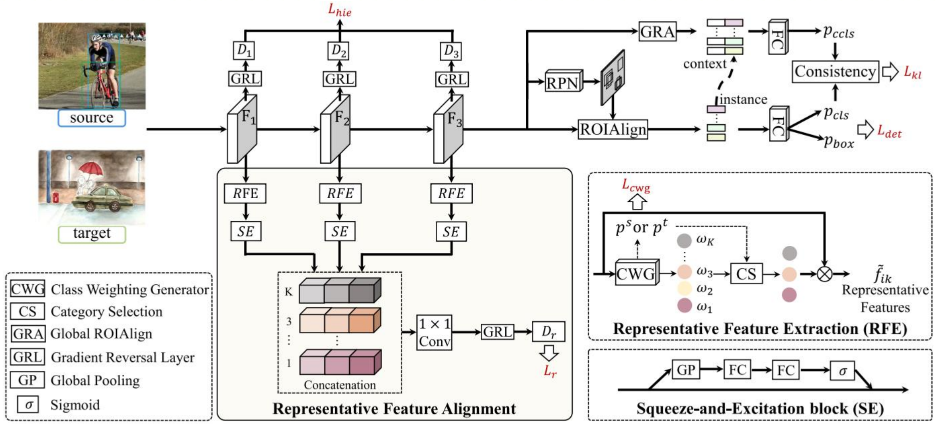
Visual evaluation: **Results on steel bar and SKU-110K datasets**

Adaptive Object Detection with Representative Feature Priors

Shan Xu , Huaidong Zhang , **Xuemiao Xu*** , Xiaowei Hu , Yangyang Xu , Liangui Dai, Kup-Sze Choi , and Pheng-Ann Heng, Representative Feature Alignment for Adaptive Object Detection. *IEEE Transactions on Circuits and Systems for Video Technology* (***TCSVT***), 2023. **[IF=8.40]**



Difficulty: **There is a big gap from the real to art domain. Moreover, the proposal region contains redundant background features, resulting in negative transfer.**



Our Solution: **Introduce representative feature prior knowledge, design the representative feature alignment module to activate discriminative features such as contours and textures while suppressing redundant background features.**

Method	KITTI → City	City → KITTI
Source-only	30.2	53.5
DAF [14]	38.5	64.1
MAF [64]	41.0	72.1
SWDA [15]	37.9	71.0
ATF [56]	42.1	73.5
RFA (ours)	**43.0**	**75.7**
Oracle	45.7	84.5

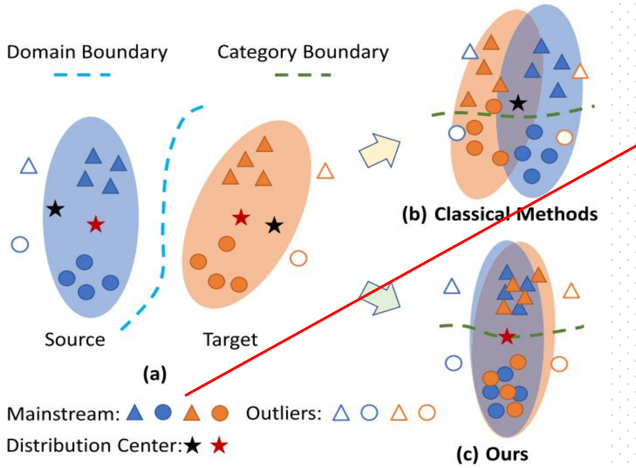
Quantitative Evaluation: **Comparison of cross-domain detection on cityscape datasets**

Ground True **Ours**

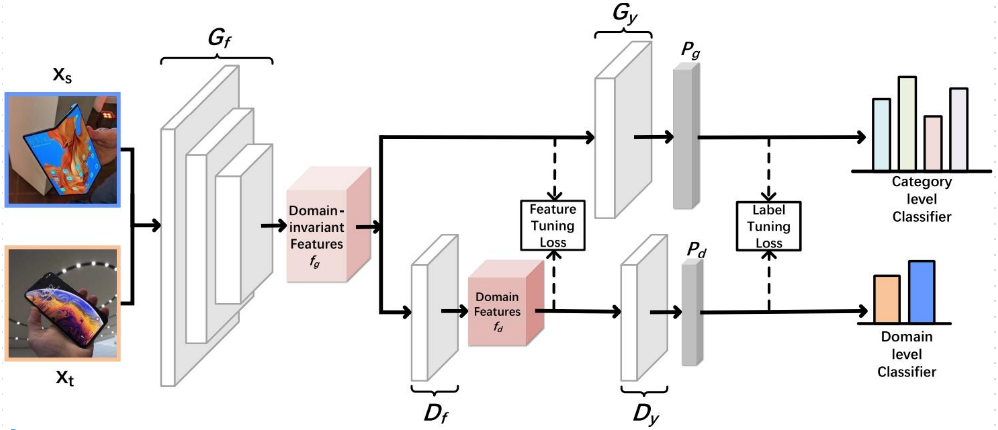
Visual evaluation: **Results on cross-domain detection datasets**

Cross-domain Classification with Importance Sampling Priors

Xuemiao Xu, Hai He , Huaidong Zhang , Yangyang Xu, and Shengfeng He, Unsupervised domain adaptation via importance sampling. *IEEE Transactions on Circuits and Systems for Video Technology* (***TCSVT***), 2020. **[IF=8.40, Google Scholar 32 times]**



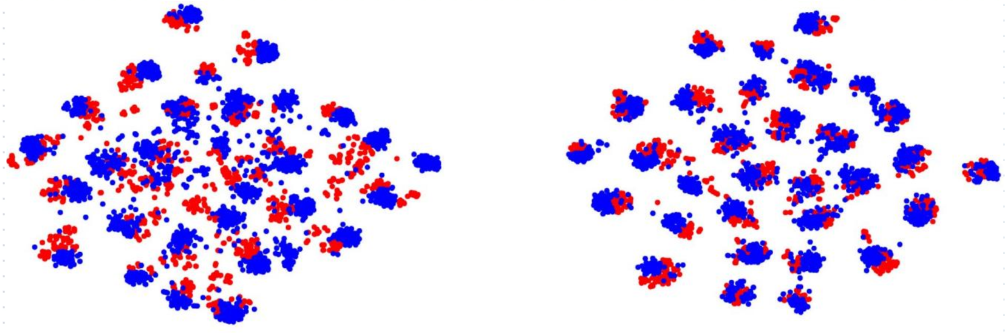
Difficulty: **Traditional domain adaptation methods fail to solve the negative transfer effects due to some outliers in the target domain.**



Our Solution: **Introduce the importance sampling prior. By incorporating the feature tuning loss and the label tuning loss, it makes learning more rely on the normal samples.**

Method	$Ar \rightarrow Cl$	$Ar \rightarrow Pr$	$Ar \rightarrow Rw$	$Cl \rightarrow Ar$	$Cl \rightarrow Pr$	$Cl \rightarrow Rw$	$Pr \rightarrow Ar$	$Pr \rightarrow Cl$	$Pr \rightarrow Rw$	$Rw \rightarrow Ar$	$Rw \rightarrow Cl$	$Rw \rightarrow Pr$	Avg
ResNet-50 [58]	38.57	60.78	75.21	39.94	48.12	52.90	49.68	30.91	70.79	65.38	41.79	70.42	53.71
DAN [6]	44.36	61.79	74.49	41.78	45.21	54.11	46.92	38.14	68.42	64.37	45.37	68.85	54.48
RTN [7]	49.37	64.33	76.19	47.56	51.74	57.67	50.38	41.45	75.53	70.17	51.82	74.78	59.25
RevGrad [5]	44.89	54.06	68.97	36.27	34.34	45.22	44.08	38.03	68.69	52.98	34.68	46.50	47.39
PADA [11]	51.95	67	78.74	52.16	53.78	59.03	52.61	43.22	78.79	73.73	56.6	77.09	62.06
ETN [63]	**59.24**	77.03	79.54	62.92	65.73	75.01	68.29	**55.37**	**84.37**	75.72	57.66	**84.54**	70.45
Ours-F	48.78	67.84	79.46	62.63	59.33	67.92	63.27	44.42	75.48	70.43	49.79	76.08	63.79
Ours-L	51.70	73.28	80.56	**71.99**	66.11	**75.87**	**75.94**	53.79	83.66	**78.70**	56.48	82.07	70.80
Ours-FL	56.06	**77.65**	**83.88**	69.88	**69.64**	73.83	73.37	52.84	82.50	78.42	**58.03**	81.96	**71.51**

Quantitative Evaluation: **Comparison on common indoor object classification benchmark OfficeHome**

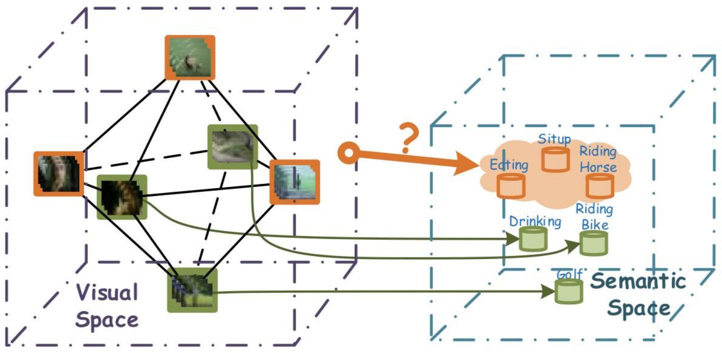


(a) (b)

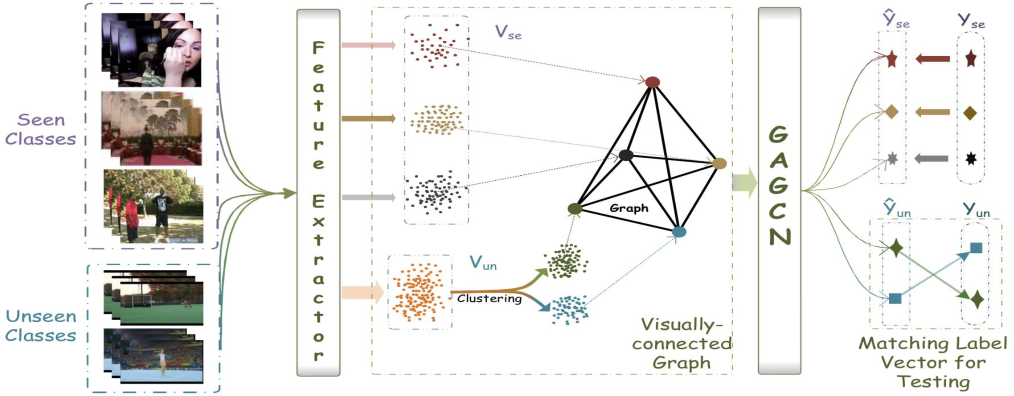
Visual evaluation: **Comparison of discriminative features with state-of-the-art, Fig.b is our results**

Zero-shot Action Recognition with Visual Similarity Prior

Yangyang Xu, Chu Han , Jing Qin, **Xuemiao Xu*,** Guoqiang Han and Shengfeng He,Transductive Zero-Shot Action Recognition via Visually Connected Graph Convolutional Networks, *IEEE Transaction on Neural Network (TNNLS),* 2020. **[CCF A]**



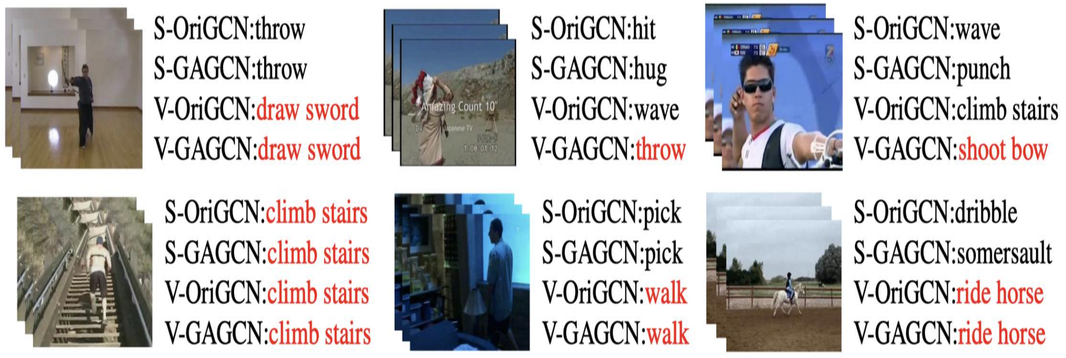
Difficulty: Domain Gap between seen and unseen classes can not be simply bridged by text semantic similarity. For action videos, visual similarity prior is more important as video contains both spatial information and temporal information.



Our Solution: Introduce the Visual Similarity Prior, propose a grouped attention graph convolutional networks (GAGCNs) to propagate the visual–semantic connections from seen actions to unseen ones.

Methods	Visual	Semantic	HMDB51	UCF101
Random Guess	-	-	4.0	2.0
Baseline	D	WV	14.7±2.3	13.7±1.2
RSC [17]	L	WV+Att	-	14.0±1.8
SER [46]	L	WV	21.2±3.0	18.6±2.2
MR [48]	L	WV	24.1±3.8	22.1±2.5
PDA [47]	L	WV	24.8±2.2	22.9±3.3
TOM [20]	D	WV	-	26.8±4.4
ZSECOC [33]	L	ECOC	22.6±1.2	15.1±1.7
VDS [52]	D	WV	25.3±4.5	28.8±5.7
BiDiLEL [43]	D+L	WV	22.3±1.1	23.0±0.9
Ours	D	WV	**29.8±2.2**	**30.0±1.8**

Quantitative Evaluation: Comparison on HMDB51 and UCF101 dataset.



Visual evaluation: Comparison with different methods on HMDB51 dataset

2.2 Overview of Our Work

Scene Understanding

Visual Generation

Complex visual feature

Detection

TITS 2018 *TITS 2020*

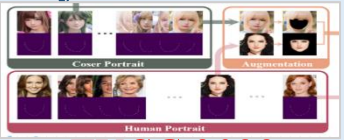
TCSVT 2021

Classification

IS 2020 *CVPR 2020*

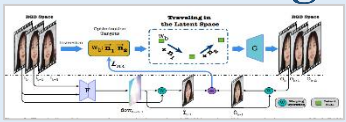
TIP 2023

Style Transfer



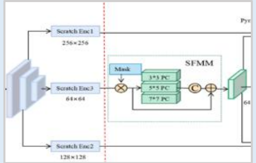
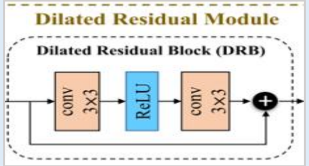
TVCG 2022

Face Editing

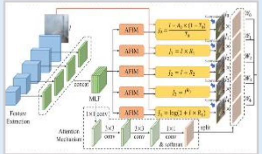
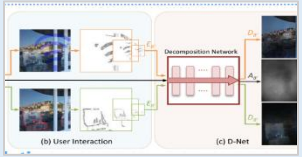


ICCV 2021

Image Enhancement

TCSVT 2023 *TCSVT 2021*

ICCV 2019 *TMM 2019*

Limited visual dataset

Detection

AAAI 2021 *IJCAI 2021*

TCSVT 2023

Classification

TCSVT 2023

TNNLS 2020

Video Inpainting

CVPR 2022

Image Generation

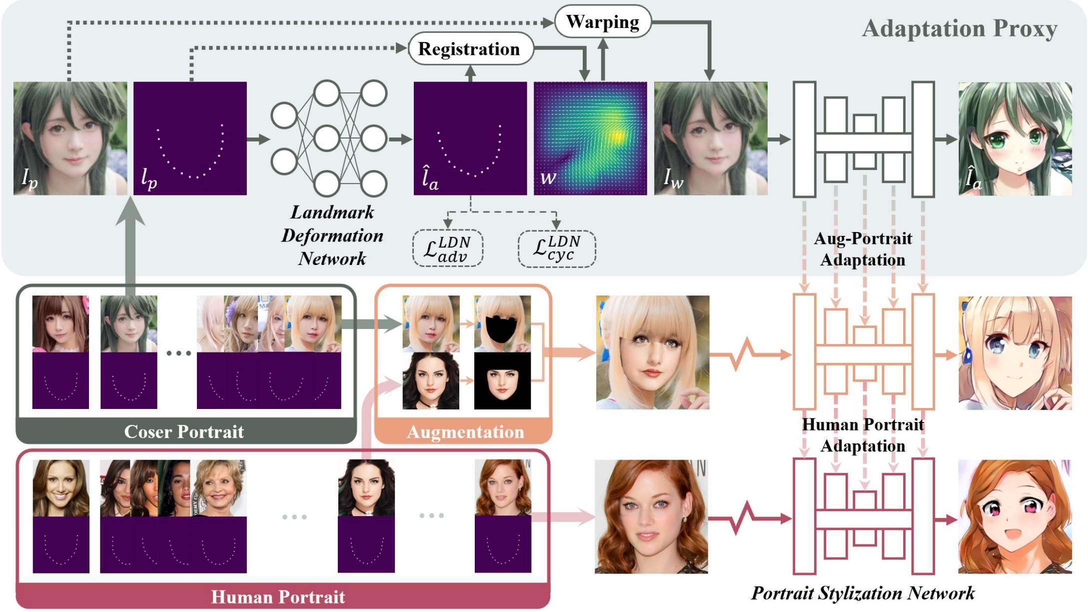
CVPR 2023

TNNLS 2022 *TIP 2021*

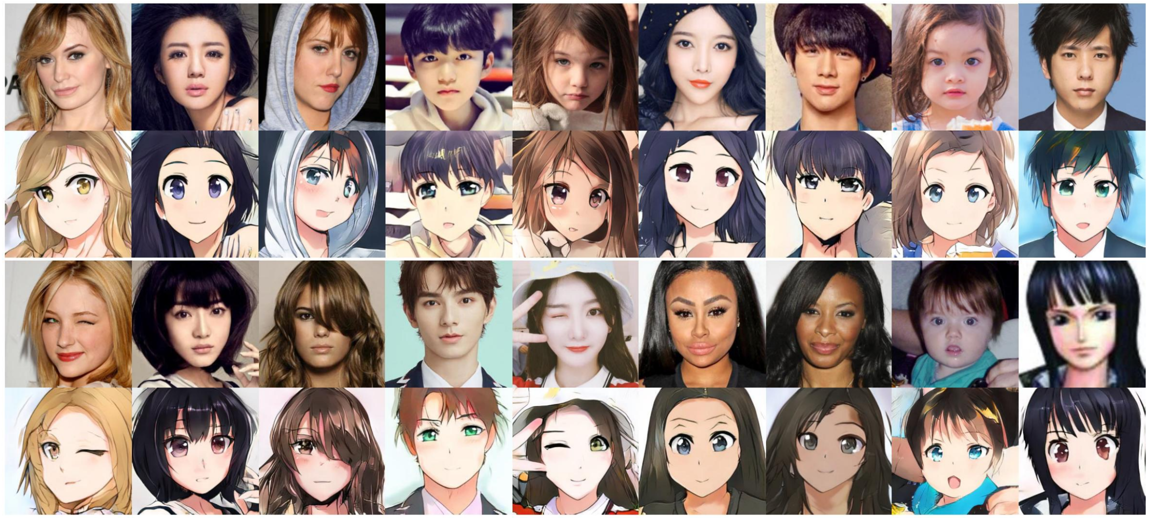
Portrait-to-anime Translation with Coser Proxy Priors

Wenpeng Xiao, Cheng Xu, Jiajie Mai, **Xuemiao Xu***, Yue Li, Chengze Li, Xueting Liu, and Shengfeng He, Appearance-preserved Portrait-to-anime Translation via Proxy-guided Domain Adaptation，*IEEE Transaction on Visualization and Computer Graphics (**TVCG**)*, 2022. **[IF=5.20]**

Difficulty: Due to the huge domain gap and distribution difference in facial appearance between portrait and anime domains, existing methods fail to handle the learning ambiguity, thus yield results with severe appearance changes.



Our Solution: Introduce the Coser-proxy prior knowledge. By exploring the correlation among *Coser*, *portraits* and *anime*, a 3-stage proxy-guided domain adaptation learning scheme is proposed to achieve appearance-preserved portrait-to-anime translation.



Visual Evaluation: Results on various real-life portraits

Method	LPIPS ↓		Hue-HistD ↓	
	CelebA-HQ	Selfie	CelebA-HQ	Selfie
UNIT [2]	0.5670	0.6780	0.4535	0.5172
MUNIT [19]	0.5597	0.7029	0.4902	0.5579
CycleGAN [1]	0.5101	0.5303	0.3335	0.4024
U-GAT-IT [3]	0.5487	0.6179	0.4017	0.5352
Toonify [33]	0.4229	0.6860	0.4058	0.4748
Cartoon-StyleGAN [35]	0.6174	0.6822	0.4206	0.4658
UI2I-via-StyleGAN2 [34]	0.6199	0.6789	0.3671	0.4772
Ours	**0.4079**	**0.5137**	**0.1760**	**0.3889**

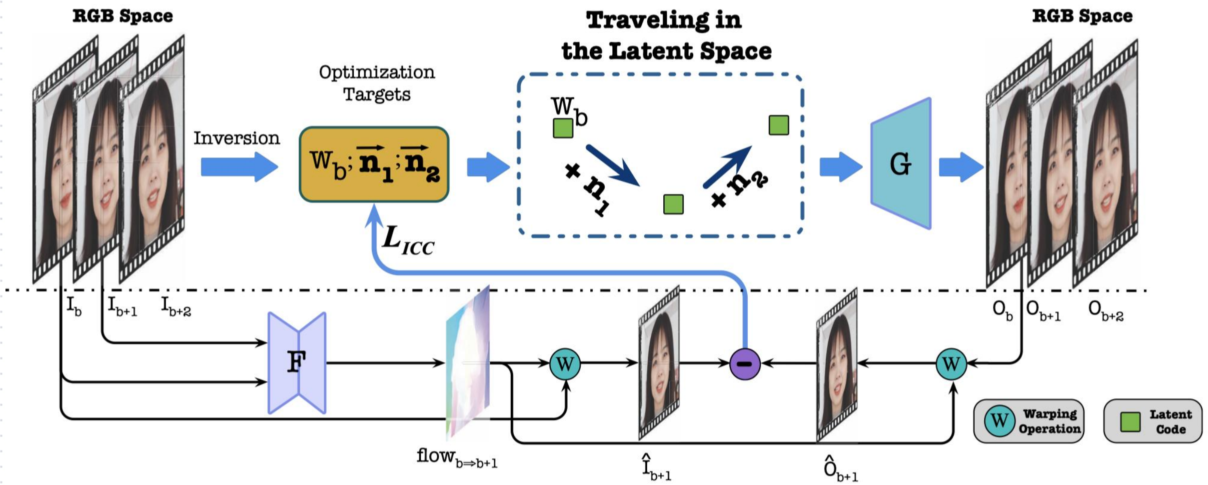
Quantitative Evaluation: Comparison to state-of-the-art methods

GAN Inversion with Consecutive Images Priors

Yangyang Xu, Yong Du, Wenpeng Xiao, **Xuemiao Xu***, Shengfeng He, From Continuity to Editability: Inverting GANs with Consecutive Images, *IEEE International Conference on Computer Vision (ICCV)*, 2021. **[CCF A]**

(a) Inputs (b) I2S [1] (c) I2S++ [2] (d) pSp [18] (e) InD [24] (f) Ours

Difficulty: **Invert a real image into latent space of GANs cannot achieve the fidelity of reconstructed images and the editability of latent codes simultaneously .**



Our Solution: **Introduce consecutive images prior for GAN inversion, enforce the latent codes of consecutive images can be transformed in the latent space to constrain the editability.**

Methods \ Metrics	RAVDESS-12 Dataset NIQE↓	FID↓	LPIPS↓	MSE↓(×e-3)	Synthesized Dataset NIQE↓	FID↓	LPIPS↓	MSE↓(×e-3)
I2S [1]	3.770	16.284	0.162	8.791	3.374	48.909	0.271	35.011
pSp [18]	3.668	29.701	0.202	22.337	3.910	84.355	0.391	46.244
InD [24]	3.765	18.135	0.193	9.963	3.152	42.773	0.352	44.645
Ours	**3.596**	**13.136**	**0.148**	**5.972**	**2.807**	**37.225**	**0.250**	**24.395**
I2S++ [2]	3.358	0.320	**0.003**	0.174	2.644	2.967	**0.014**	1.458
Ours++	**3.352**	**0.311**	**0.003**	**0.165**	**2.597**	**2.897**	**0.014**	**1.432**

Quantitative Evaluation: **Comparison with state-of-the-art methods for image restoration on RVDESS-12 and Synthesized datasets**

(a) GT (b) I2S (c) Ind (d) pSp (e) Ours

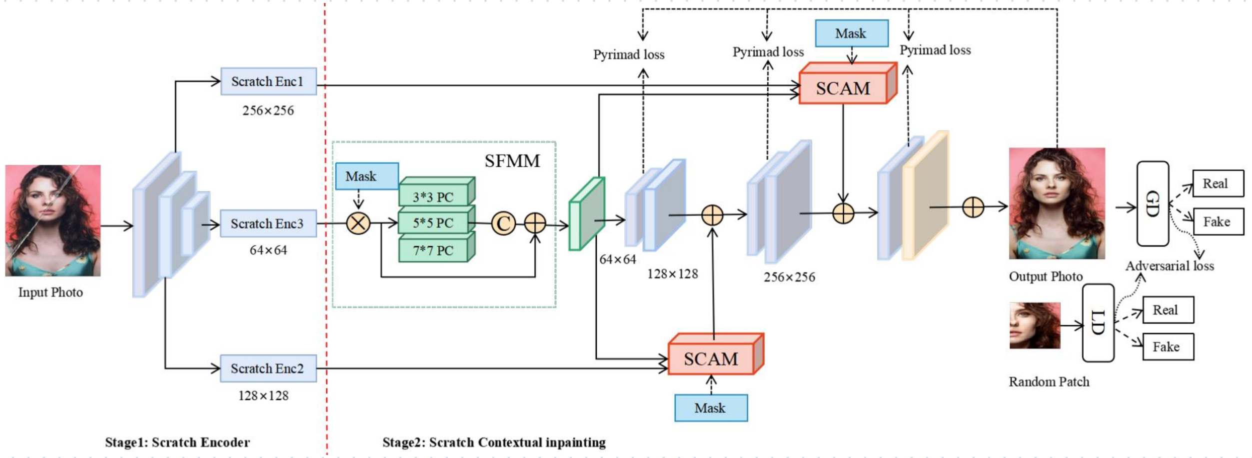
Visual evaluation: **Comparison of different methods for image editing.**

Scratched Photo Restoration assisted by Contextual Priors

Weiwei Cai, Huaidong Zhang, **Xuemiao Xu***, Shengfeng He*, Contextual-assisted Scratched Photo Restoration, *IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**)*, 2023. **[IF=8.40]**

Input Ours OPBL

Difficulty: **The existing methods have not fully utilized the contextual information of the scratches, semantically inconsistent and blurry artifacts can easily occur.**



Our Solution: **Introduce the scratch contextual information as prior knowledge. We propose a two-stage scratched photo inpainting network to leverage the scratch context and the background context, and finally generate a context-consistent, scratch-free photo.**

	Method	PSNR ↑	SSIM ↑	LPIPS ↓	FID ↓
OSPD	CYG [45]	23.6028	0.8177	0.2262	140.3986
	MFE [36]	27.8243	0.8414	0.1822	122.4771
	RFR [38]	28.7103	0.8416	0.1914	124.5095
	OPBL [6]	23.0001	0.8038	0.2051	115.5456
	OPBL-Re [6]	22.1193	0.7863	0.2839	137.0829
	SPL [39]	21.8400	0.7652	0.2908	234.3566
	Ours	**29.4798**	**0.8715**	**0.1438**	**81.2526**
MSPD	CYG [45]	23.7442	0.7960	0.1185	99.0443
	MFE [36]	25.2060	0.8043	0.0790	74.2280
	RFR [38]	24.1487	0.7925	0.1096	84.5165
	OPBL [6]	23.1347	0.7793	0.1226	87.1060
	OPBL-Re [6]	24.3946	0.8063	0.1548	102.5973
	SPL [39]	20.7545	0.7263	0.2307	145.0990
	Ours	**25.5441**	**0.8294**	**0.0537**	**41.3191**

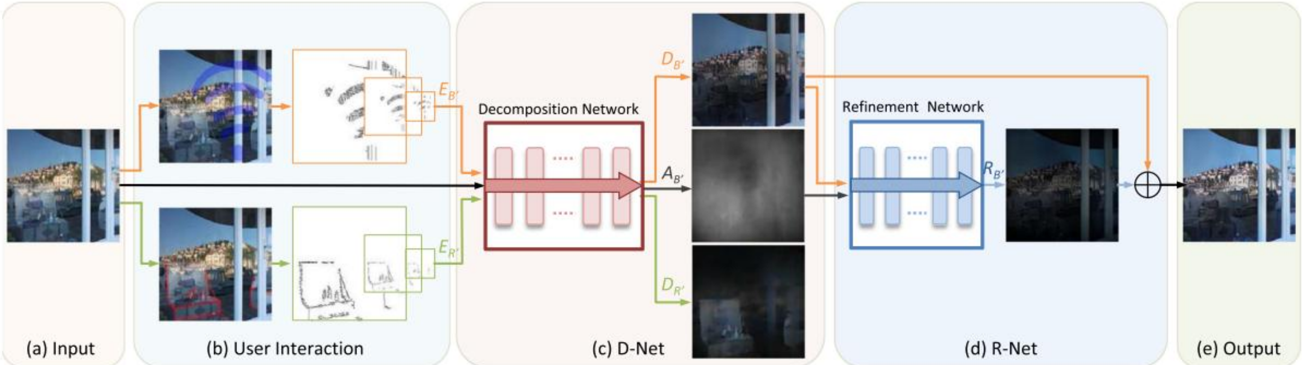
Quantitative Evaluation: **Comparison on our OSPD and MSPD datasets**

Visual evaluation: **Results on our OSPD and MSPD datasets**

Single Image Reflection Removal with User Guidance Priors

Huaidong Zhang, **Xuemiao Xu***, Hai He, Le Zhou, Shengfeng He, Guoqiang Han, Jing Qin, DapengWu. Fast User-Guided Single Image Reflection Removal via Edge-aware Cascaded Networks. *IEEE Transactions on Multimedia (**TMM**)*, 2019. **[IF=7.30]**

Difficulty: **It is an NP hard problem to separate reflection layer from a single image.**



Our Solution: **Introduce the user guidance priors, By incorporating information from user interaction, our method accurately separates the specular reflection layer and background layer.**

Guidance	Method	SIR^2 sLMSE	NCC	SSIM	SI
NG	Li and Brown [7]	0.970	0.960	0.740	0.856
	Fan *et al.* [17]	0.982	0.969	0.825	0.896
	Arvanitopoulos *et al.* [9]	**0.989**	**0.978**	**0.862**	0.911
	Wan *et al.* [22]	0.987	**0.978**	0.815	0.875
	Zhang *et al.* [21]	0.984	0.971	0.854	0.914
	Levin and Weiss [18]	0.879	0.858	0.614	0.857
	Ours(w/o R)	0.987	0.969	0.850	0.908
	Ours	0.988	0.973	0.855	**0.915**
SG	Levin and Weiss [18]	0.893	0.880	0.617	0.842
	Ours(w/o R)	0.986	0.976	0.852	0.908
	Ours	**0.988**	**0.982**	**0.865**	**0.916**
DG	Levin and Weiss [18]	0.973	0.968	0.798	**0.942**
	Ours(w/o R)	0.988	0.979	0.861	0.919
	Ours	**0.991**	**0.986**	**0.875**	0.923

Quantitative Evaluation: **Comparison on SIR2 datasets**

Visual evaluation: **Results on our datasets**

2.2 Overview of Our Work

Scene Understanding

Visual Generation

Complex visual feature

Detection

TITS 2018 TITS 2020

TCSVT 2021

Classification

IS 2020 CVPR 2020

TIP 2023

Style Transfer

TVCG 2022

Face Editing

ICCV 2021

Image Enhancement

TCSVT 2023 TCSVT 2021

ICCV 2019 TMM 2019

Limited visual dataset

Detection

AAAI 2021 IJCAI 2021

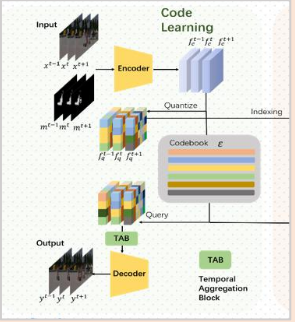
TCSVT 2023

Classification

TCSVT 2023

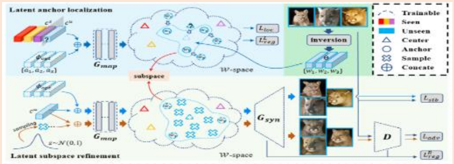
TNNLS 2020

Video Inpainting

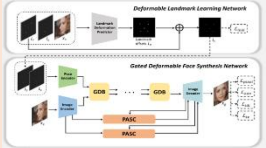
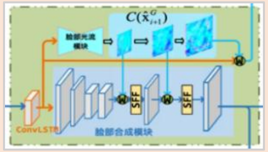


CVPR 2022

Image Generation



CVPR 2023

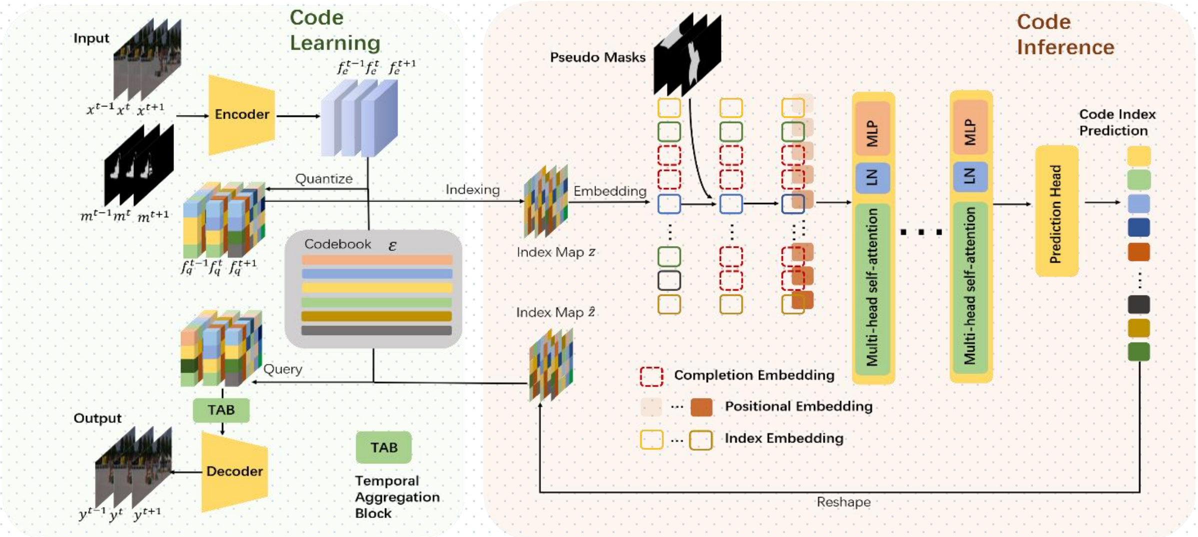
TNNLS 2022 TIP 2021

Video Inpainting of Internal Priors in Discrete Latent Space

Jingjing Ren, QingqingZheng, Yuanyuan Zhao, **Xuemiao Xu***, Chen Li, DLFormer:Discrete Latent Transformer for Video Inpainting, *IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2022. **[CCF A]**

(a) Input (b) VINet (c) STTN (d) Ours

Difficulty: **Previous methods , e.g. (b) and (c) formulate in a continuous feature space based on external knowledge and usually produce artifacts and blurry results around the occluded bars and background.**



Our Solution: **Internal prior knowledge in discrete latent space. Learn a compact discrete codebook to represent the target video. Employs a self-attention mechanism to infer appropriate codes for unknown regions, so that resulting in the generation of fine-grained content with spatial temporal consistency.**

Method	Youtube-VOS PSNR ↑	Youtube-VOS SSIM ↑	Youtube-VOS VFID ↓	DAVIS PSNR ↑	DAVIS SSIM ↑	DAVIS VFID ↓
VINet [33]	29.72	0.953	0.111	32.38	0.967	0.105
FFVI [2]	33.39	0.968	0.119	31.13	0.972	0.087
CPNet [13]	30.21	0.957	0.117	29.57	0.955	0.147
STTN [35]	33.67	0.965	0.087	33.07	0.976	0.071
FuseFormer [17]	33.26	0.968	0.089	33.45	**0.979**	0.074
DLFomer (ours)	**33.95**	**0.970**	**0.082**	**34.22**	0.977	**0.062**

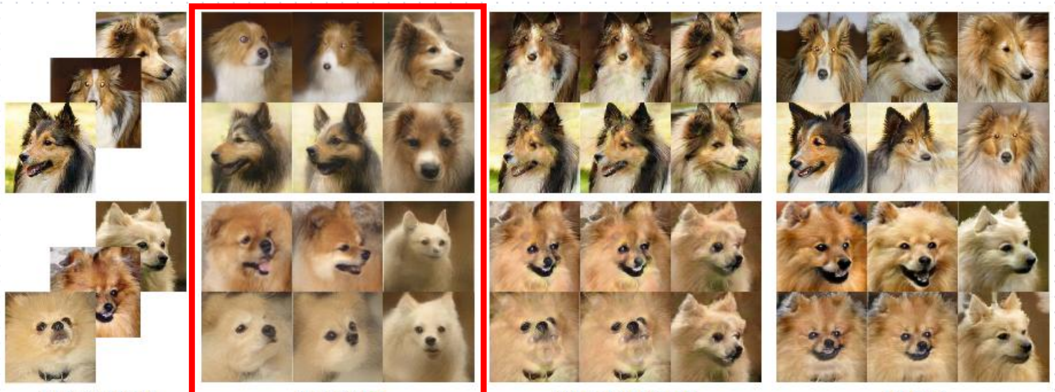
Quantitative Evaluation: **Comparison with state-of-the-art methods for video restoration on Youtube-VOS and DAVIS datasets**

(a) Input (b) VINet (c) STTN (d) Fuseformer (e) ILVI (f) Ours

Visual evaluation: **Comparison of different methods for object removal.**

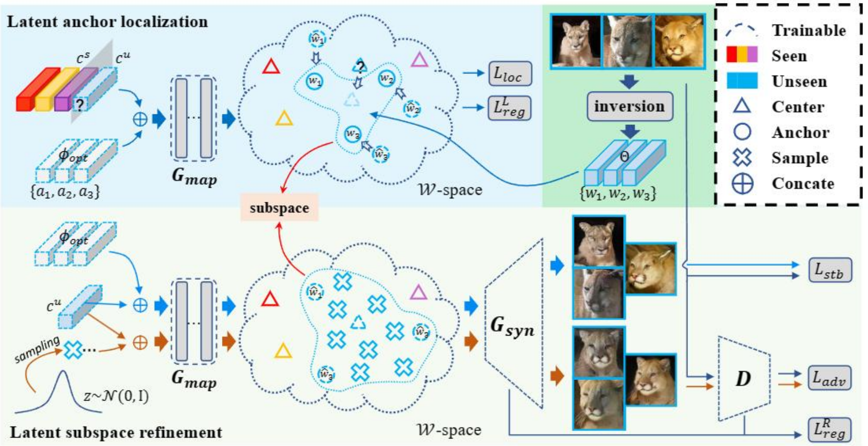
Few-shot Image Generation with Knowledge Collaborative Mechanism

Chenxi Zheng, Bangzhen Liu, Huaidong Zhang, **Xuemiao Xu***, Shengfeng He， Where is My Spot? Few-shot Image Generation via Latent Subspace Optimization， *IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, 2023. **[CCF A]**



(a) 3-shot Input (b) AGE [7] (c) WaveGAN [44] (d) Ours

Difficulty: The prior knowledge can not be well adapted to the unseen category, a unified framework to effectively leverage prior and posterior knowledge is needed.



Our Solution: Knowledge Collaborative Mechanism, that is a two-stage optimization framework. First, we localize the subspace for the target category using priors. Next, we further optimize the generator with the images of target images to inject the low-level features.

Methods	k-shot	Flowers		Animal Faces		VGG Faces	
		FID↓	LPIPS↑	FID↓	LPIPS↑	FID↓	LPIPS↑
DAGAN [1]	1	179.59	0.0496	185.54	0.0687	134.28	0.0608
DeltaGAN [13]	1	109.78	0.3912	89.81	0.4418	80.12	0.3146
AGE [7]	1	45.96	0.4305	28.04	**0.5575**	34.86	0.3294
Ours	1	**35.87**	**0.4338**	**27.20**	0.5382	**4.15**	**0.3834**
FIGR [6]	3	190.12	0.0634	211.54	0.0756	139.83	0.0834
DAWSON [25]	3	188.96	0.0583	208.68	0.0642	137.82	0.0769
GMN [2]	3	200.11	0.0743	220.45	0.0868	136.21	0.0902
MatchingGAN [12]	3	143.35	0.1627	148.52	0.1514	118.62	0.1695
F2GAN [14]	3	120.48	0.2172	117.74	0.1831	109.16	0.2125
LoFGAN [10]	3	79.33	0.3862	112.81	0.4964	20.31	0.2869
WaveGAN [44]	3	42.17	0.3868	30.35	0.5076	4.96	0.3255
Ours	3	**34.59**	**0.3914**	**23.67**	**0.5198**	**3.98**	**0.3344**

Quantitative Evaluation: Comparison with existing methods

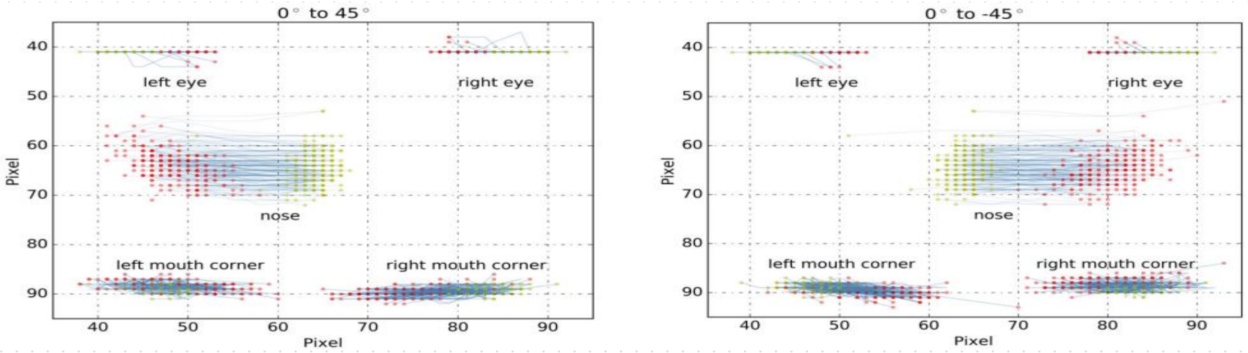
Input AGE [7] Ours Input WaveGAN [44] Ours

Visual evaluation: Comparison with state-of-the-art methods under 1-shot and 3-shot setting

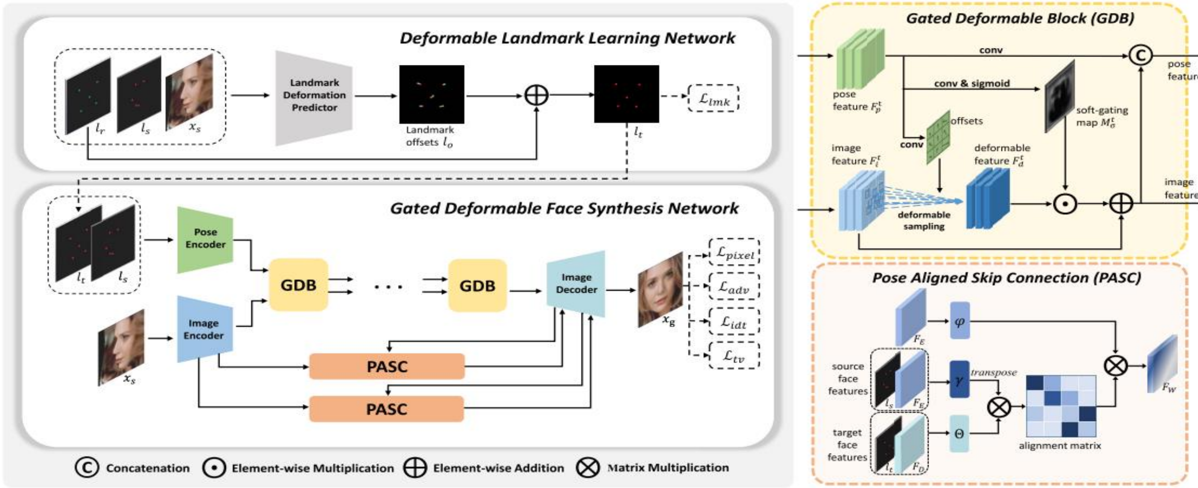
Deformation Priors for Multiview Face Image Synthesis

Cheng Xu, Keke Li, Xuandi Luo, **Xuemiao Xu***, Shengfeng He, and Kun Zhang, Fully Deformable Network for Multiview Face Image Synthesis, *IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**)*, 2022. **[IF=10.40]**

Xuemiao Xu, Keke Li, Cheng Xu, Shengfeng He, GDFace: Gated Deformation for Multi-View Face Image Synthesis, *The AAAI Conference on Artificial Intelligence (**AAAI**)*, 2020. **[CCF A]**



Difficulty: **The face deformation pattern caused by large pose variation is too complex to model , current methods fail to model complex face deformation.**



Our Solution : **Face deformation prior knowledge** **is introduced. By proposing the personalized landmark learning module and gated deformation face synthesis module, the complex face deformation pattern can be explicitly modeled for better synthesis.**

Methods	±90°	±75°	±60°	±45°	±30°	±15°
FIP+LDA [67]	-	-	45.90	64.10	80.70	90.70
MVP+LDA [29]	-	-	60.10	72.90	83.70	92.80
CPF [12]	-	-	61.90	79.90	88.50	95.00
DR-GAN [8]	-	-	83.20	86.20	90.10	94.00
A3F-CNN [20]	-	-	92.70	95.80	98.90	98.70
Light CNN [56]	5.51	24.18	62.09	92.13	97.38	98.59
FF-GAN [16]	61.20	77.20	85.20	89.70	92.50	94.60
TP-GAN [6]	64.64	77.43	87.72	95.38	98.06	98.68
CAPG-GAN [4]	66.05	83.05	90.63	97.33	99.56	99.82
PIM [64]	86.50	95.00	98.10	98.50	99.00	99.30
3D-PIM [17]	86.73	95.21	98.37	98.81	99.48	99.64
AD-GAN [19]	89.70	95.30	98.80	99.50	99.70	99.80
HF-PIM [18]	92.32	96.40	99.14	**99.88**	99.98	99.99
FFlowGAN [23]	93.01	97.00	99.17	99.83	99.99	99.99
Ours	**93.60**	**97.43**	**99.31**	**99.88**	**100**	**100**

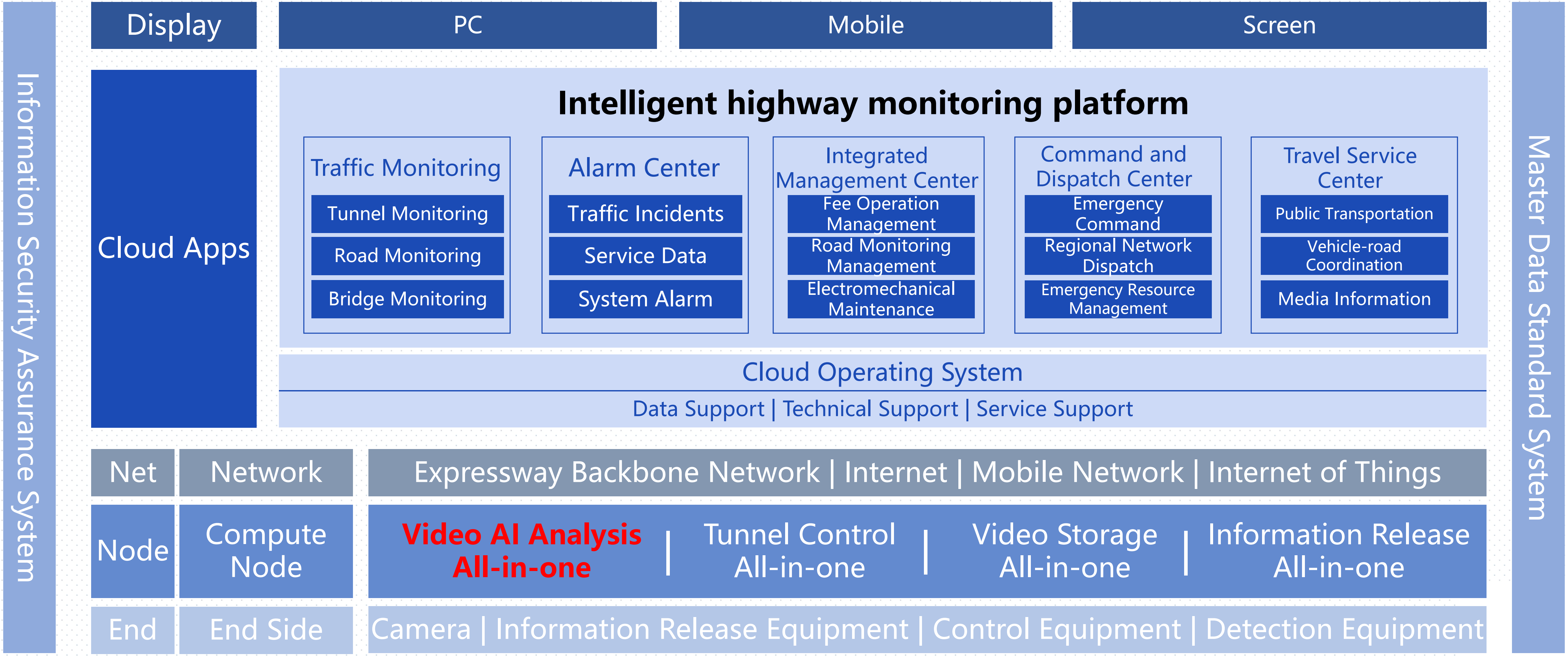
Quantitative Evaluation: **Comparison on Multi-PIE**

Visual evaluation: **Comparison on CelebA**

3 Applications

Application 1: Intelligent Highway Monitoring Platform

- **Cooperate with Guangdong Transportation Group (TOP 1 in Guangdong Province)**
- This platform have been **widely deployed** in the Guangdong province highways.

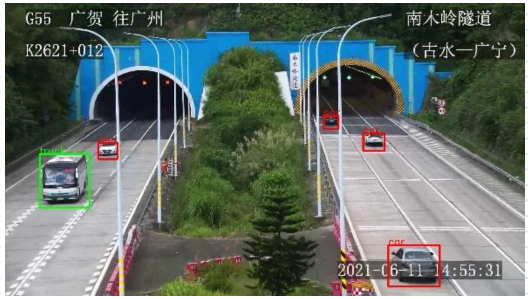


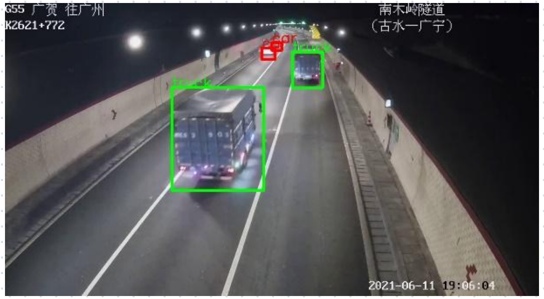
Application 1: Intelligent Highway Monitoring Platform

Video AI Analysis All-in-one

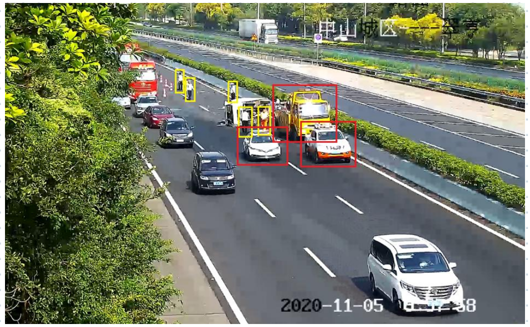
Video AI Analysis All-in-one

- Accurate detecting, tracking and identification of vehicles, pedestrians, and traffic signs
- Statistics of traffic flow and average speed





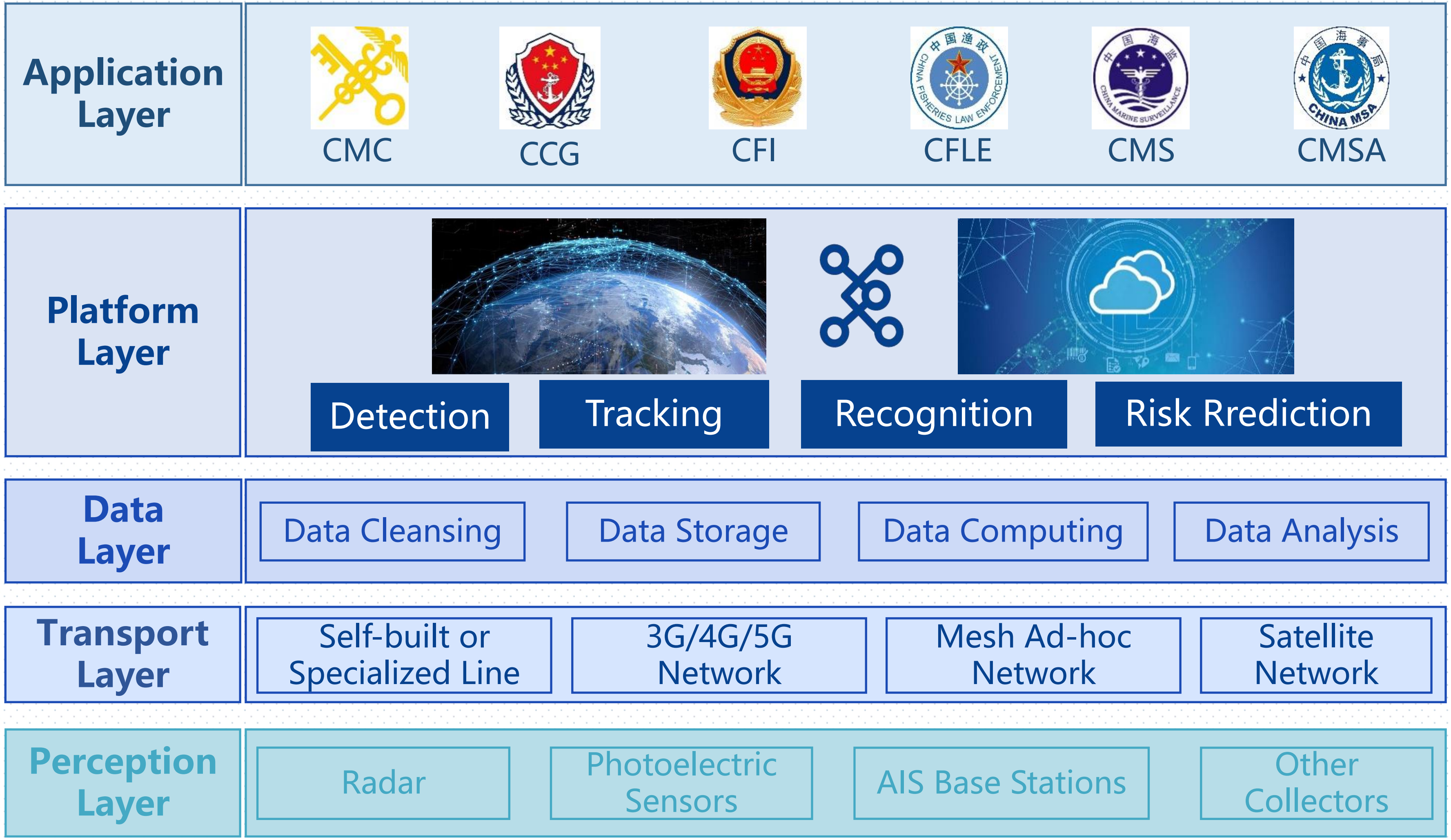
- Real-time detection of abnormal events (abnormal parking/road work/throwing objects/retrograde, etc.)



- Video quality assessment
- Surveillance video restoration under bad weathers

Application 2: Intelligent Coastal Ship Monitoring and Risk Control Platform

- **Cooperate with South Coast Company (Only company permitted to access the shipping data at Pearl River Estuary)**
- **Serve more than 5 port supervision departments, including customs, maritime, coastal defense, etc.**

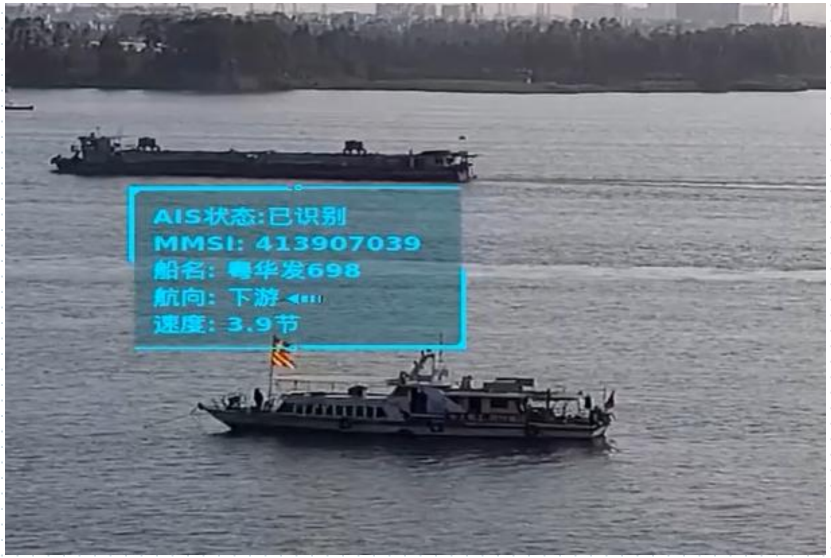


Application 2: Intelligent Coastal Ship Monitoring and Risk Control Platform

Key technologies

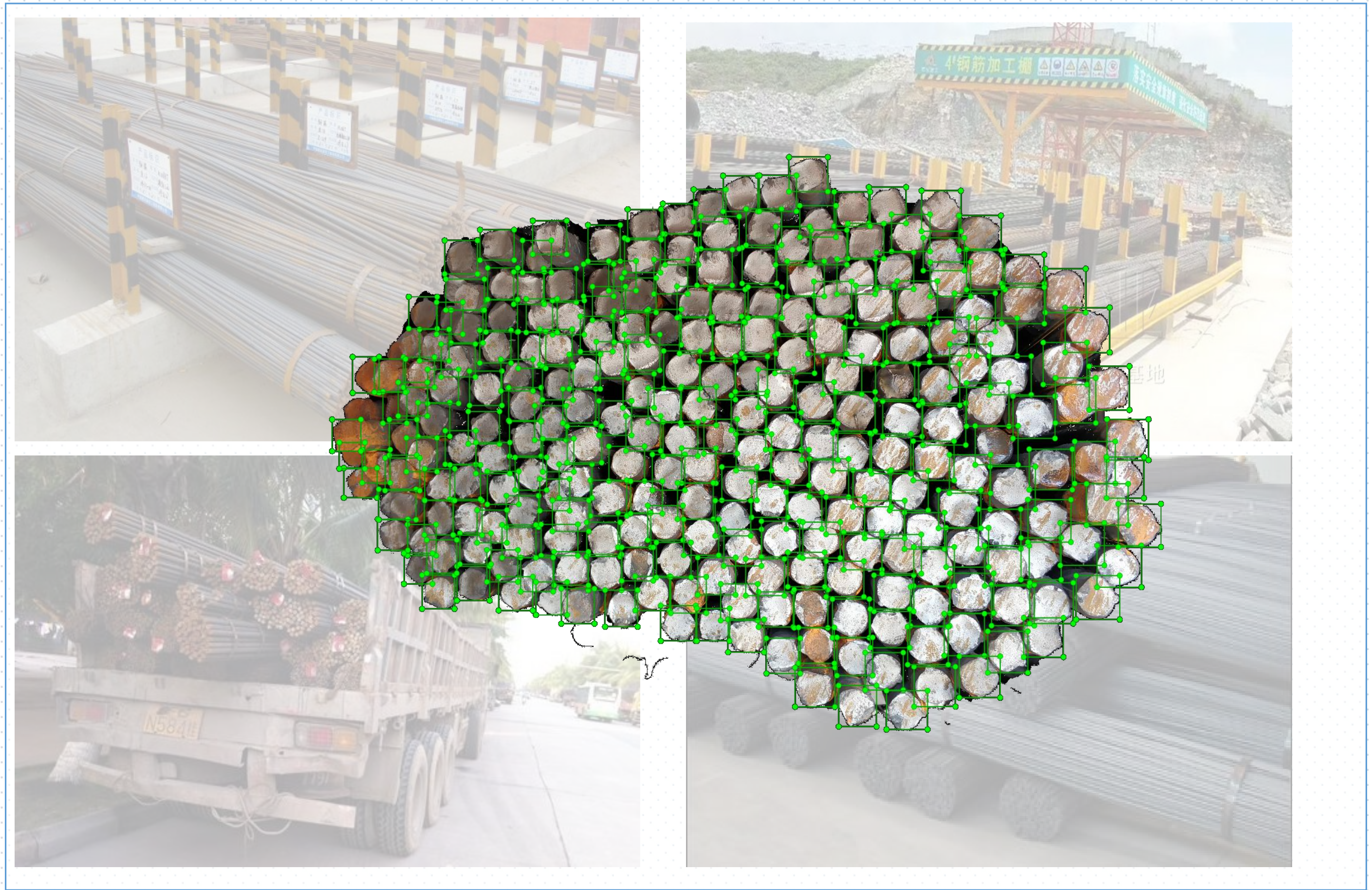
Long-distance and large-scale ship detection, tracking and recognition using multimodal data (video + radar + location)

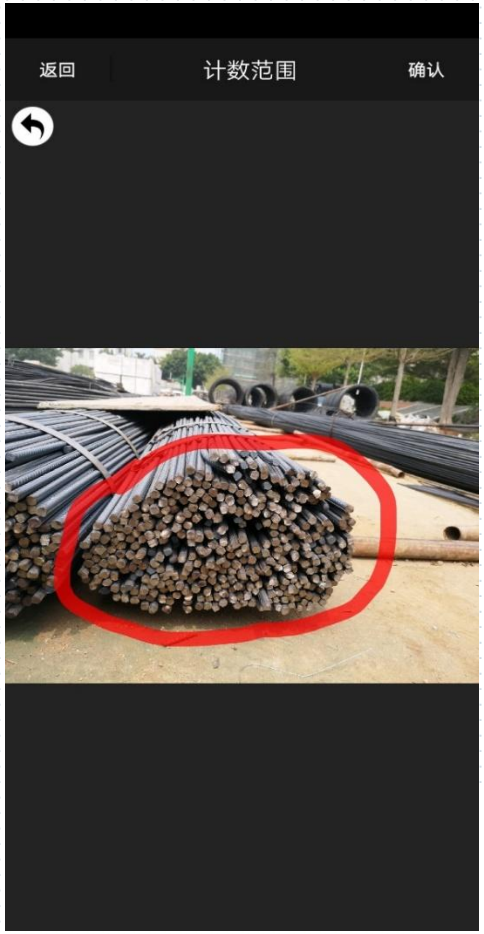
Risk prediction using multimodal data (analyzed ship information+ Customs clearance data)



Application 3: Intelligent Software for Steel Bar Counting

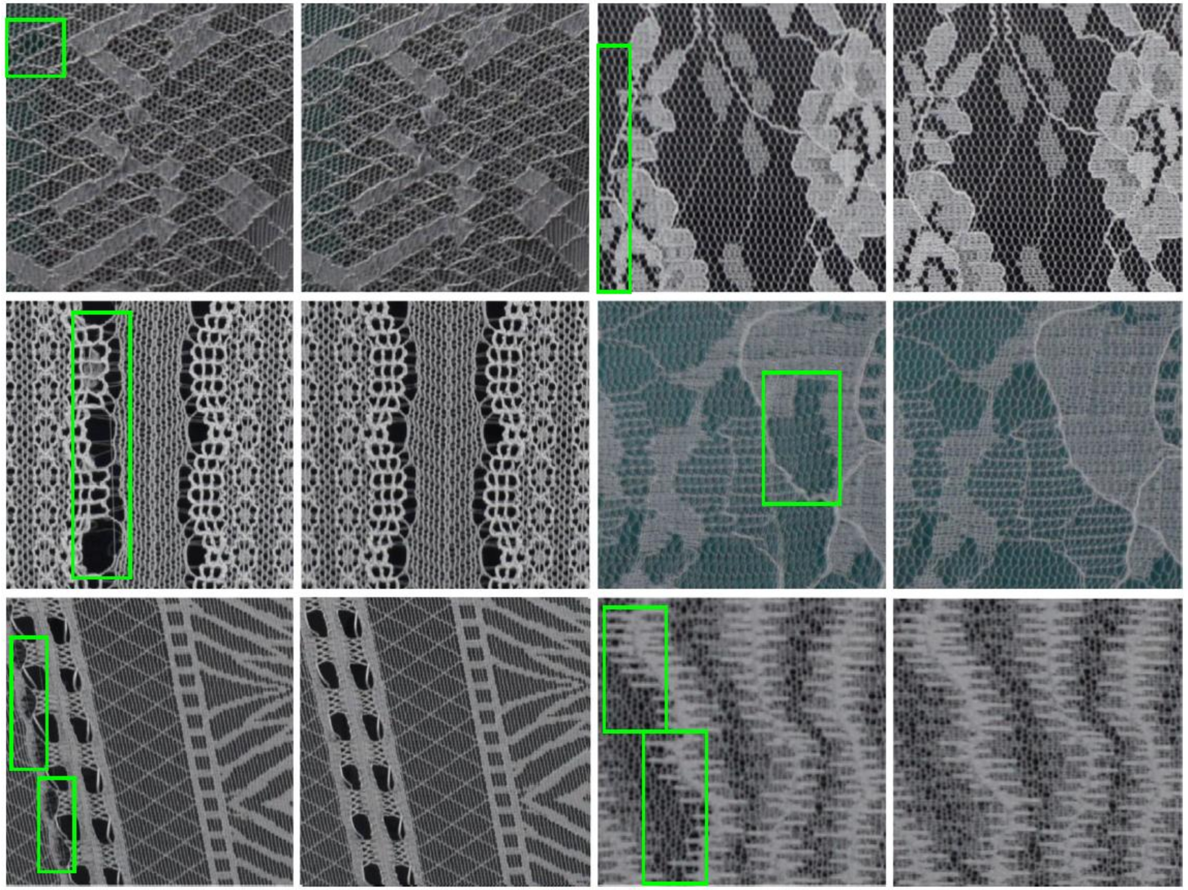
- Cooperate with Shanghai Ganglian E-Commerce Co., Ltd. (largest steel trading platform in China)
- Our vision technology can obtain the recognition accuracy over 99.5%.
- Integrated into Ganglian trading platform to provide the service for more than 100,000 users.





Application 4: Intelligent System for Lace Defect Detection

- Cooperate with Jiayou Weaving Co., Ltd." (annual export volume for lace ranks second in China)
- 32 types of defects can be successfully detected



(a) Defective images (b) Reference of (a) (c) Defective images (d) Reference of (c)

Application 5: Intelligent Robot Arm for Defective Rubber Removal

- Cooperate with Beijing Yanshan Petrochemical Co.,Ltd. (largest petrochemical factory in the China)
- The recognition accuracy for the defective rubber can obtain over 99%
- Deployed in the Beijing Yanshan Petrochemical

Thank You !!