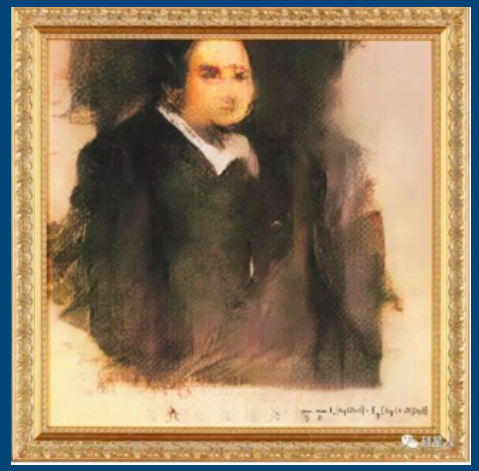


AI Portrait Sold for $432,000

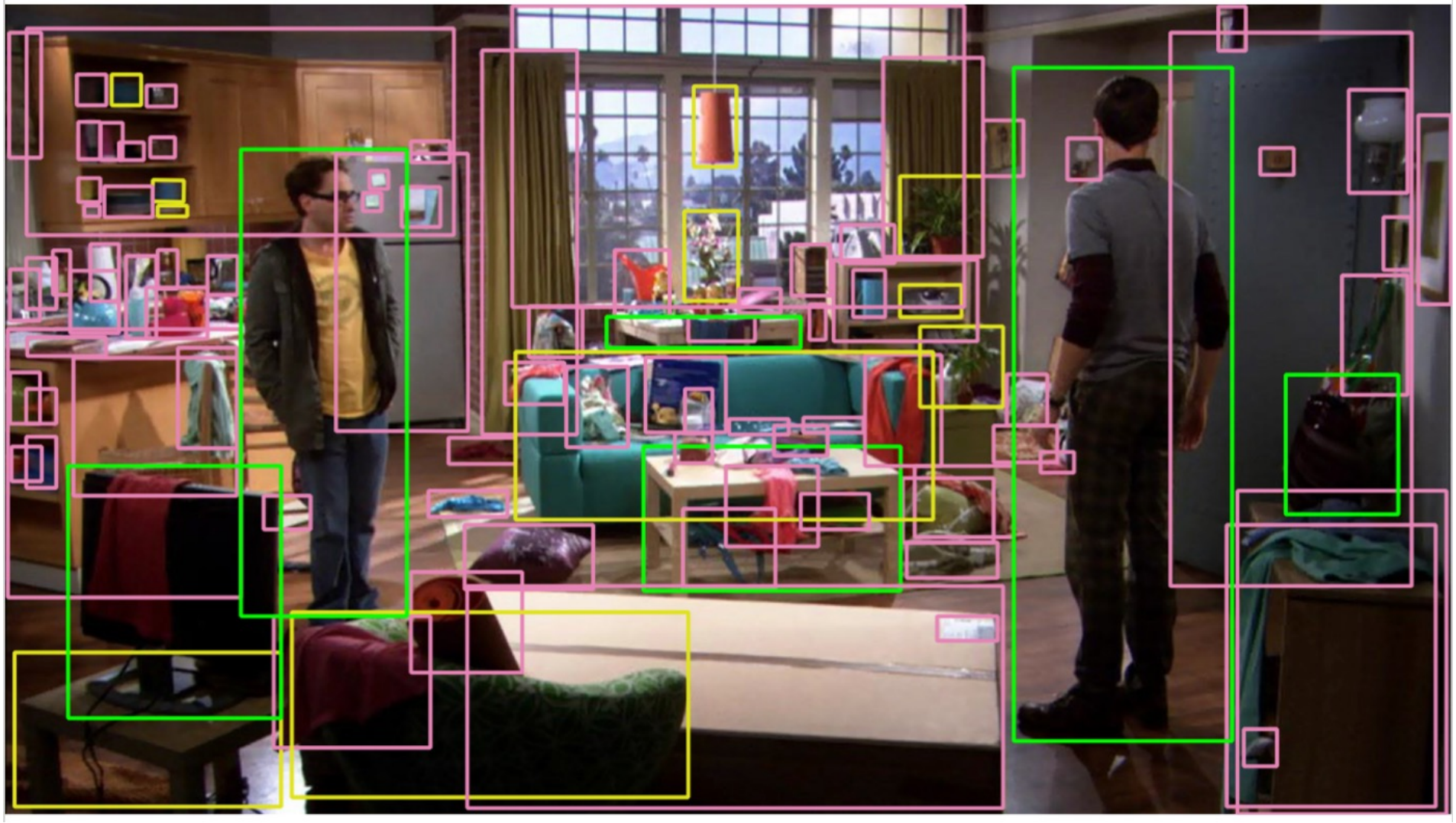
Deepfakes Generation and Detection

Guest Lecturer: Dr. Zuxuan Wu

School of Computer Science, Fudan University

Fall, 2022

深度学习在计算机视觉领域取得了显著的进展



深度学习在计算机视觉领域取得了显著的进展

媒体内容生成技术突飞猛进，各大公司纷纷交出惊艳答卷

DaLL•E2

OpenAI

从文本生成图像、图像编辑等

Imagen

Google

从文本生成图像和视频

Stable Diffusion

Stability AI

从文本生成图像、图像编辑等

深度学习在计算机视觉领域取得了显著的进展

- 图像视频自动生成技术近年来取得极大发展

2019

2021

人脸图像生成的分辨率、逼真度不断提升

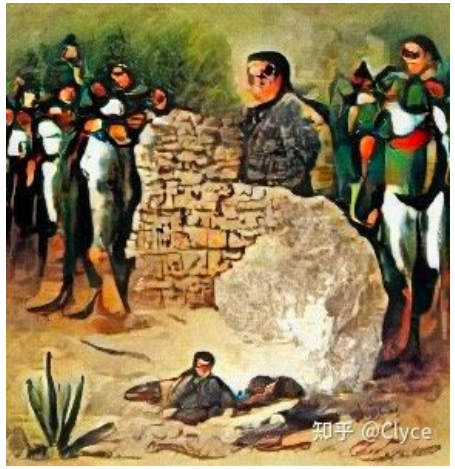
AI创作的肖像画在佳士得拍卖会上拍出43.2万美元（2018）

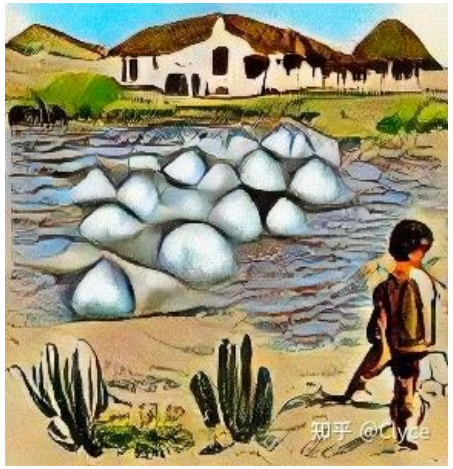
AI作画获得艺术比赛一等奖（2022）

深度学习在计算机视觉领域取得了显著的进展

- 扩散模型的快速发展，进一步展现出难以想象的跨模态生成能力

多年以后，奥雷连诺上校站在行刑队面前，准会想起父亲带他去参观冰块的那个遥远的下午。当时，马孔多是个二十户人家的村庄，一座座土房都盖在河岸上，河水清澈，沿着遍布石头的河床流去，河里的石头光滑、洁白，活象史前的巨蛋。这块天地还是新开辟的，许多东西都叫不出名字，不得不用手指指点点。每年三月，衣衫褴褛的吉卜赛人都要在村边搭起帐篷，在笛鼓的喧嚣声中，向马孔多的居民介绍科学家的最新发明。

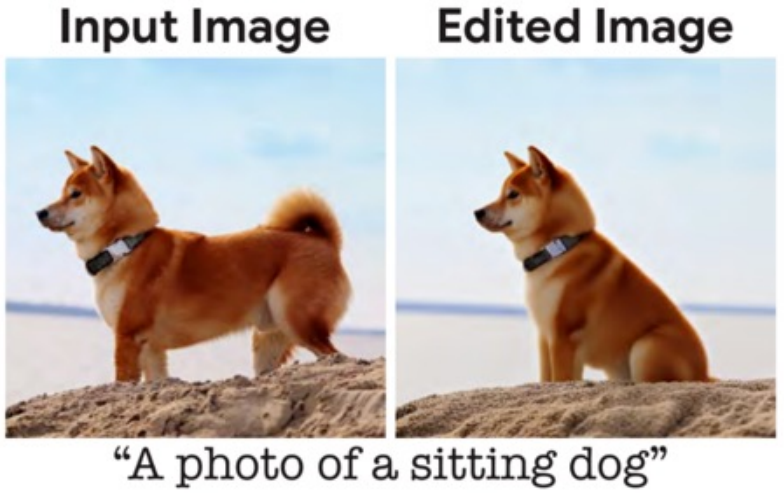




《百年孤独》第一段话生成（输入文本返回生成图像，2021）

DaLL•E2（2022）

基于“I have always wanted to be a cool panda riding a skateboard in Santa Monica.“生成



Imagic（2022）

用文字编辑照片

深度学习在计算机视觉领域取得了显著的进展

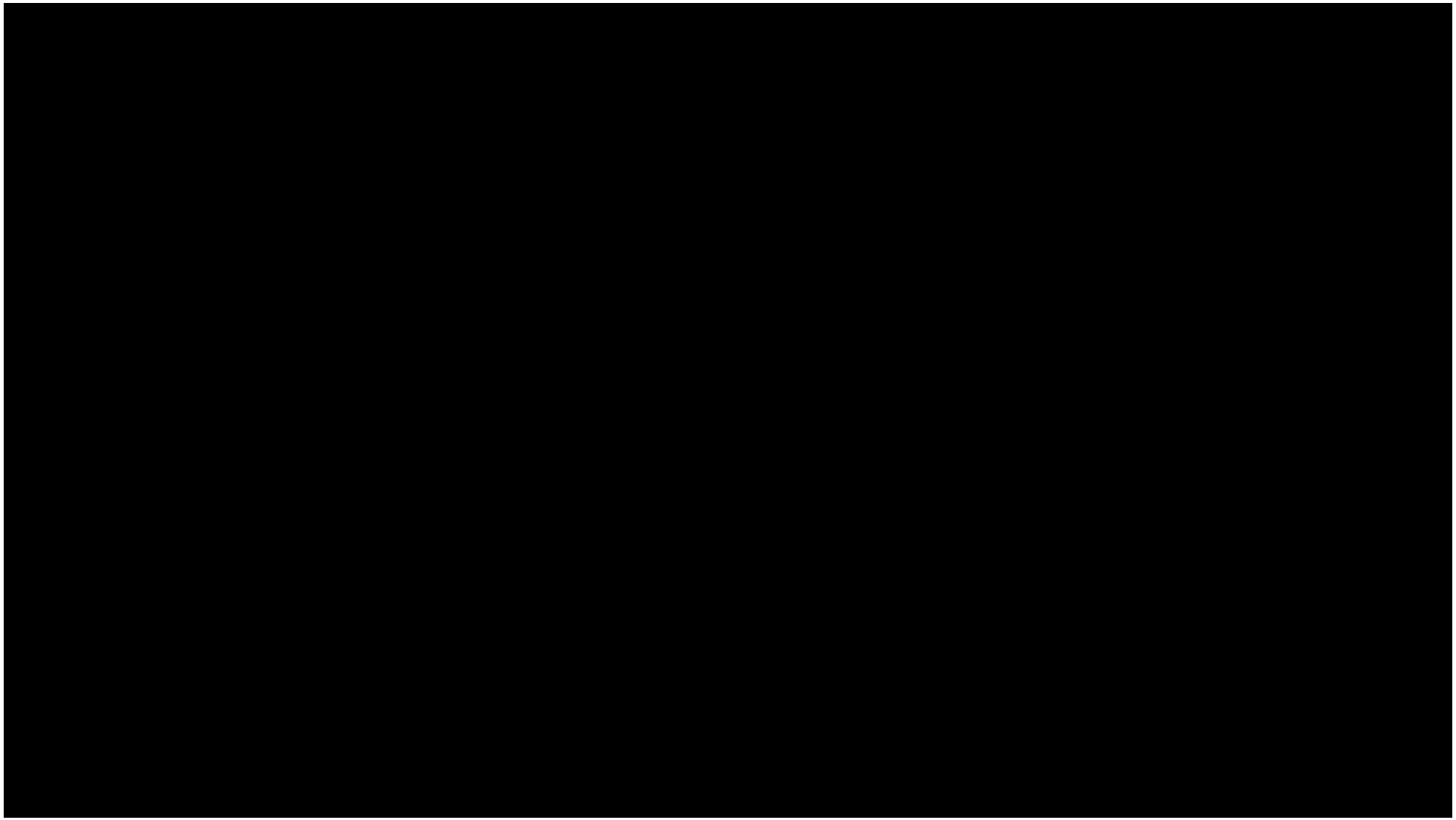
- 前沿生成技术带来技术变革，即将对艺术创作、美工、电影特效、游戏制作等带来冲击

Stable Diffusion

（2022）

深度学习在计算机视觉领域取得了显著的进展

- 前沿生成技术带来技术变革，即将对艺术创作、美工、电影特效、游戏制作等带来冲击



深度学习在计算机视觉领域取得了显著的进展

- 以换脸视频为代表的特定视频生成已足够逼真
- 在特定场景下媒体数据是否真实至关重要，迫切需要鉴别

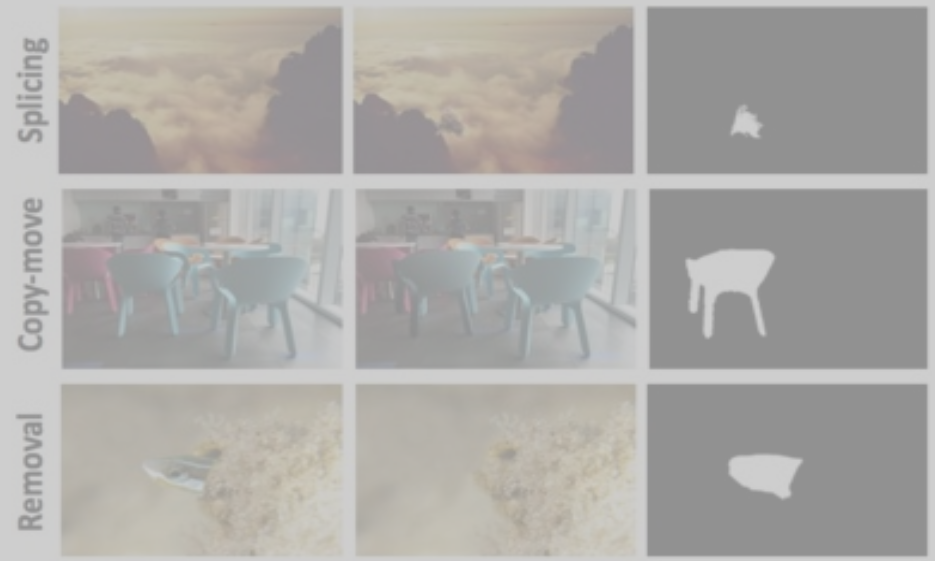
深度伪造

媒体数据篡改

深度学习在计算机视觉领域取得了显著的进展

- 以换脸视频为代表的特定视频生成已足够逼真
- 在特定场景下媒体数据是否真实至关重要，迫切需要鉴别

深度伪造

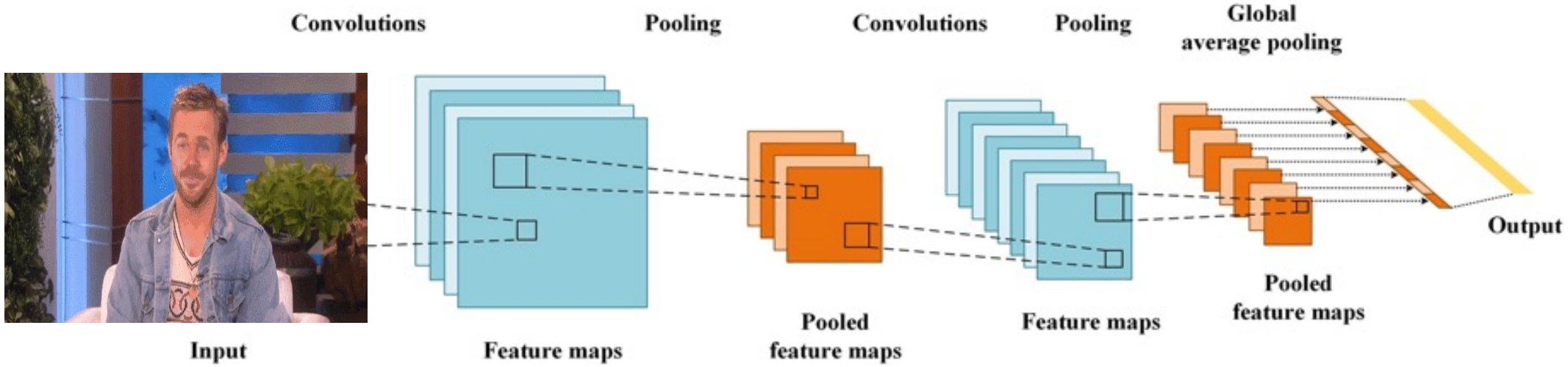


媒体数据篡改

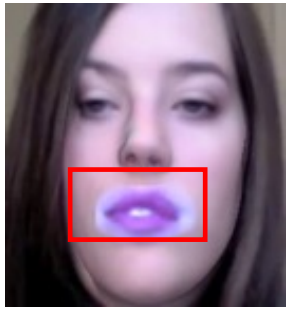
多尺度、多模态人脸伪造鉴别

Wang J et al. ICMR 2022.

多尺度特征捕捉： 受限于卷积操作固定的感受野，深度CNN网络只能捕捉单一尺度特征



深度伪造人脸的瑕疵大小不同



多尺度、多模态人脸伪造鉴别

Wang J et al. ICMR 2022.

对图像压缩的鲁棒性： 检测方法在经过了压缩之后的图像上准确率锐减

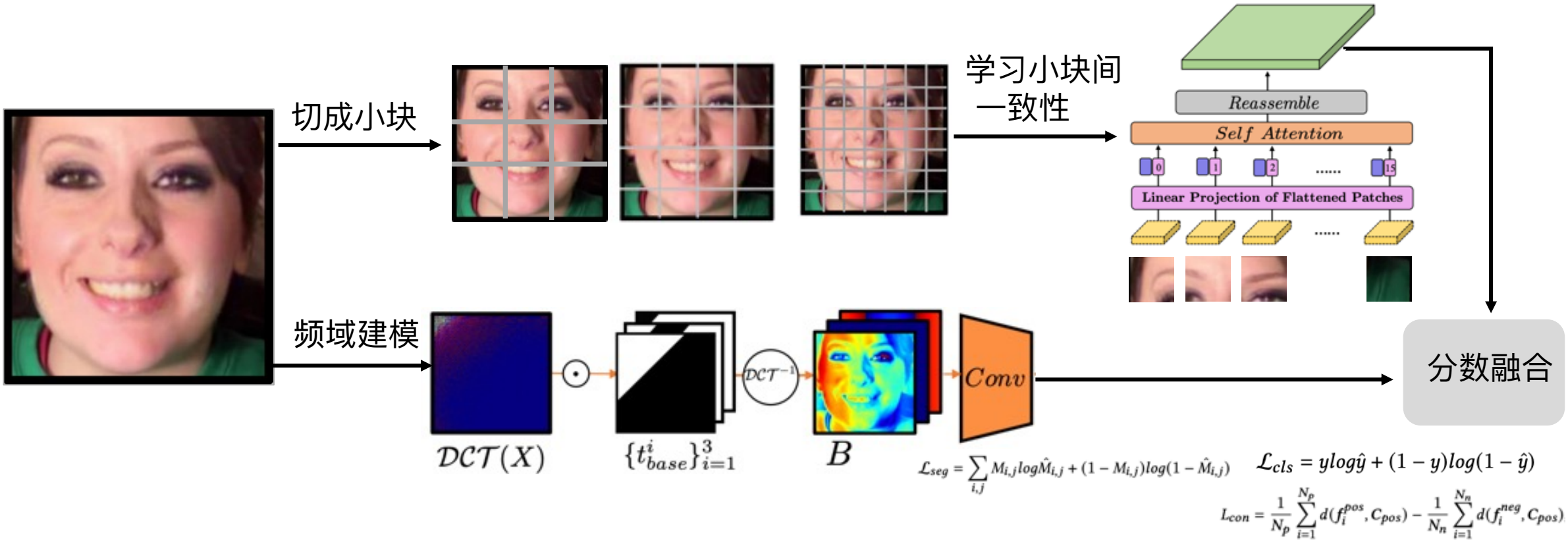
Methods	低质量图片（高度压缩） LQ		高质量图片（轻度压缩） HQ		原始图片（无压缩） RAW	
	ACC	AUC	ACC	AUC	ACC	AUC
Steg.Features [21]	55.98	-	70.97	-	97.63	-
LD-CNN [7]	58.69	-	78.45	-	98.57	-
MesoNet [1]	70.47	-	83.10	-	95.23	-
Face X-ray [32]	-	61.6	-	87.4	-	-
F^3-Net [44]	90.43	93.30	97.52	98.10	**99.95**	99.80
MaDD [66]	88.69	90.40	97.60	99.29	-	-

经过压缩后，现有鉴别方法性能大幅度下降

多尺度、多模态人脸伪造鉴别

Wang J et al. ICMR 2022.

基于Transformer的模型：



多尺度Transformer: 用于捕捉不同尺度的伪造特征

多模态特征: 用频域信息作为对RGB的有效补充

多尺度、多模态人脸伪造鉴别

Wang J et al. ICMR 2022.

实验结果：

Methods	LQ		HQ		RAW	
	ACC	AUC	ACC	AUC	ACC	AUC
Steg.Features [21]	55.98	-	70.97	-	97.63	-
LD-CNN [7]	58.69	-	78.45	-	98.57	-
MesoNet [1]	70.47	-	83.10	-	95.23	-
Face X-ray [32]	-	61.6	-	87.4	-	-
F^3-Net [44]	90.43	93.30	97.52	98.10	**99.95**	99.80
MaDD [66]	88.69	90.40	97.60	99.29	-	-
Ours	**92.35**	**94.22**	**98.23**	**99.48**	99.21	**99.91**

FF++数据集上，LQ、HQ和RAW三个版本都实现了SOTA

在LQ版本上比之前的方法有明显的性能提升。

深度伪造人脸鉴别的难点

不同伪造数据的泛化能力：某个数据集上训练得到的模型在其他数据集上性能较低

训练集　　测试集　　不同模型方法

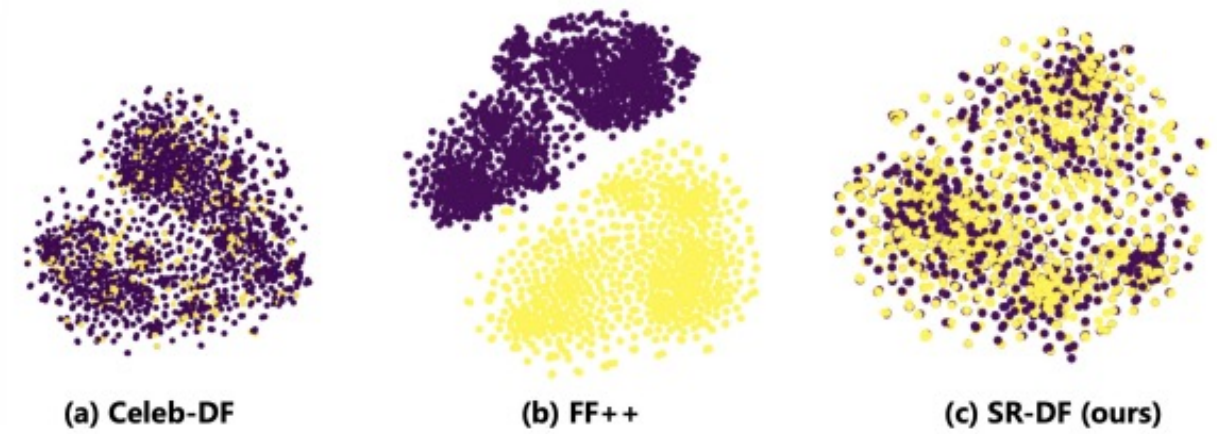
Training Set	Testing Set	Xception [48]	Multi-task [40]	Capsule [41]	DSW-FPA [33]
FF++	FF++	99.7	76.3	96.6	93.0
	Celeb-DF	48.2	54.3	57.5	64.6
	SR-DF	37.9	38.7	41.3	44.0
SR-DF	SR-DF	88.2	85.7	81.5	86.6
	FF++	63.2	58.9	60.6	69.1
	Celeb-DF	59.4	51.7	52.1	62.9

深度伪造人脸鉴别

Wang J et al. arxiv 2021.

现有数据集：普遍存在着视觉瑕疵明显、多样性不足等问题

(a) FF++ [48] (b) DFD [9] (c) DFDC [13] (d) Celeb-DF [35]



(a) Celeb-DF (b) FF++ (c) SR-DF (ours)

检测方法在现有数据集上训练后测试，能比较容易地实现较高的检测准确率，但是在高质量的伪造数据上性能很差。

深度伪造人脸鉴别

Wang J et al. ICMR 2022.

高质量的深度伪造数据集 SR-DF Dataset：

1. 利用SOTA的人脸操作方法生成伪造图像

   (1) First-order-motion [Aliaksandr Siarohin, NIPS 2019] **(2) IcFace** [Soumya Tripathy, WACV 2020] **(3) FSGAN** [Yuval Nirkin, ICCV 2019] **(4). FaceShifter** [Lingzhi Li, Arxiv].

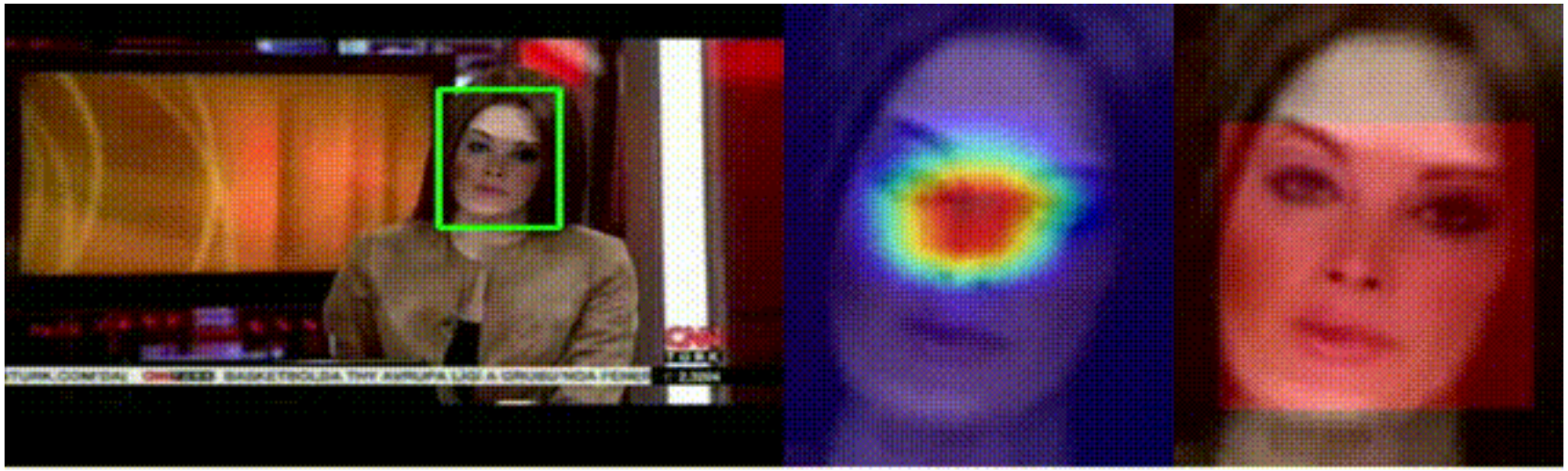
2. 利用图像和谐化方法DoveNet [Cong Wenyan et al. CVPR 2020.] 进行后处理。

深度伪造人脸鉴别

Wang J et al. ICMR 2022.

开源工具集合：

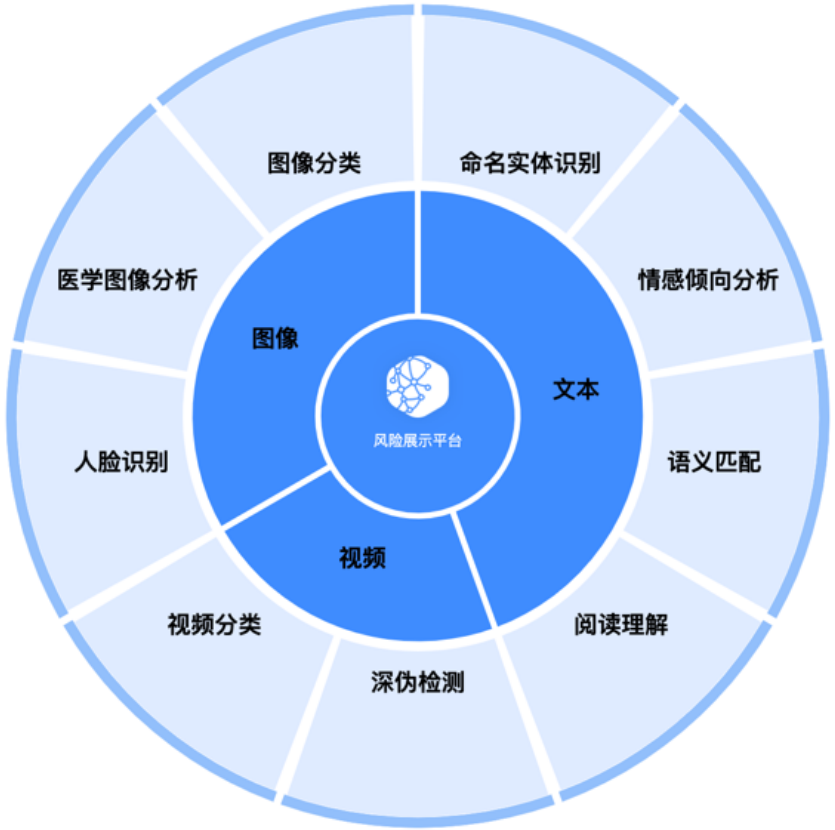
PyDeepFakeDet



深度伪造人脸鉴别

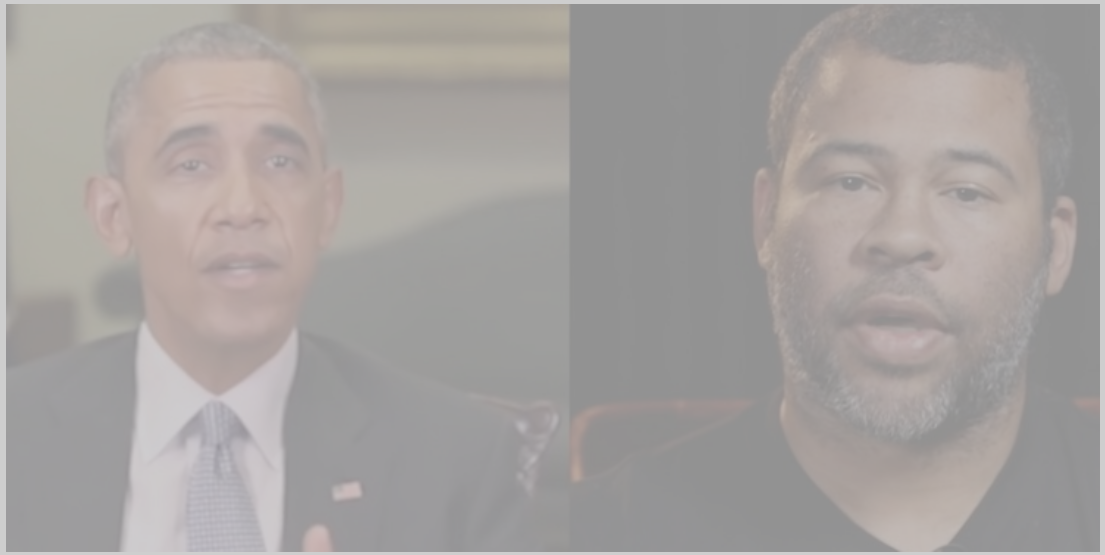
评测工具：

http://www.openeglab.org.cn/#/decision-support

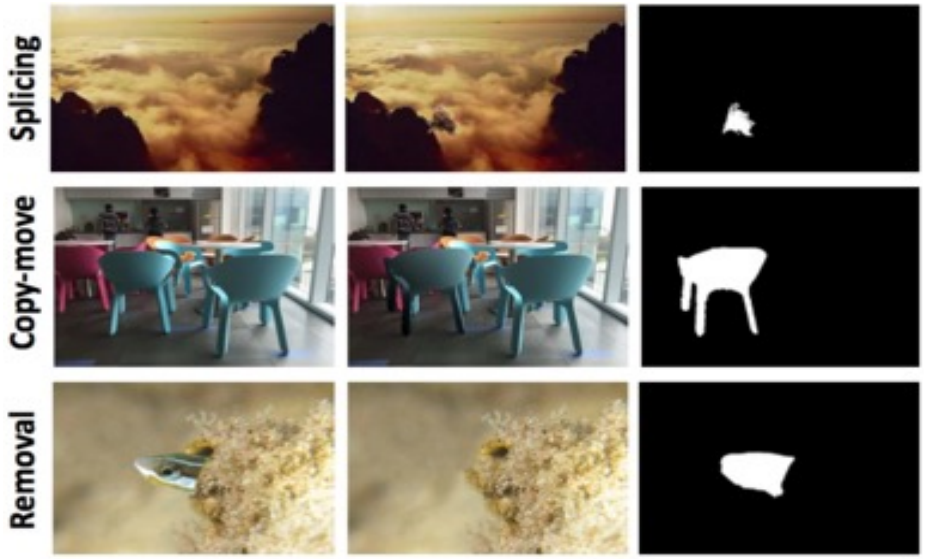


图像篡改检测

- 以换脸视频为代表的特定视频生成已足够逼真
- 在特定场景下媒体数据是否真实至关重要，迫切需要鉴别



深度伪造



媒体数据篡改

图像篡改检测

Wang J et al. CVPR 2022.

基于Transformer的模型：

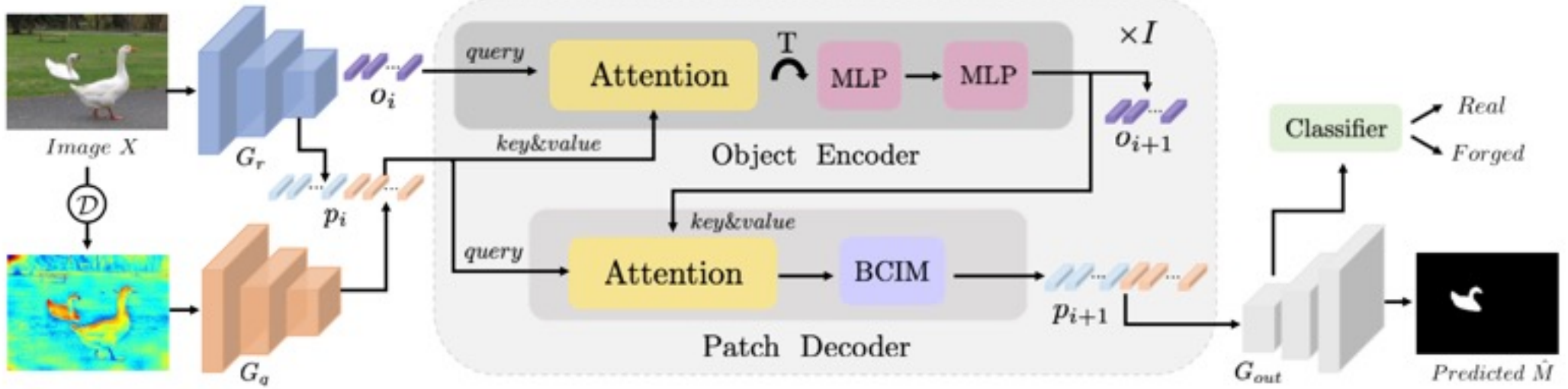


Figure 2. Overview of the proposed ObjectFormer. The input is a suspicious image ($H \times W \times 3$), and the output includes a tampering localization result and a predicted mask ($H \times W \times 1$), which localizes the manipulation regions.

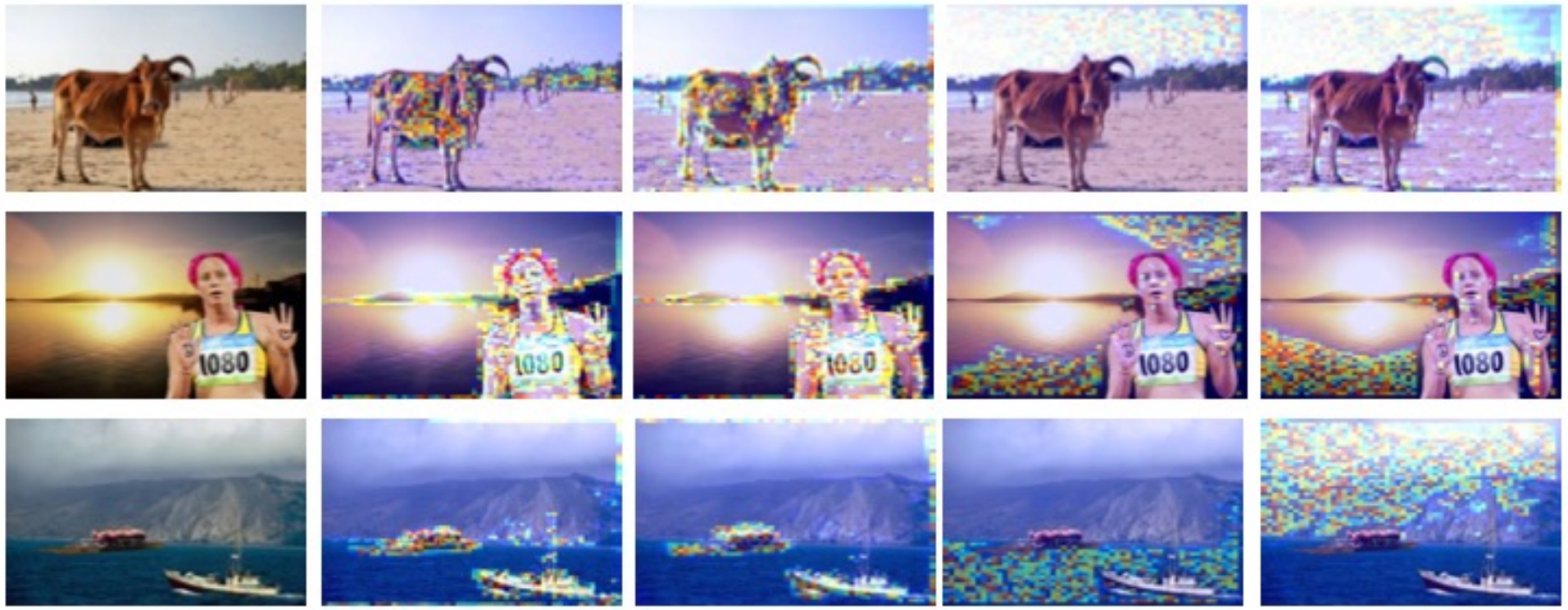
物体编码器: 学习物体表征和图像之间的交互。

图像块编码器: 学习关注到视觉异常的图像块。

图像篡改检测

Wang J et al. CVPR 2022.

物体表征–图像注意力可视化



不同的物体表征可以关注到图像的前景和背景中不同物体对应的区域。

图像篡改检测

Wang J et al. CVPR 2022.

量化结果对比：

Method	Type	Coverage		CASIA		NIST16	
		AUC	F1	AUC	F1	AUC	F1
ELA	U	58.3	22.2	61.3	21.4	42.9	23.6
NOI1	U	58.7	26.9	61.2	26.3	48.7	28.5
CFA1	U	48.5	19.0	52.2	20.7	50.	0.28
J-LSTM	F	61.4	-	-	-	76.4	-
H-LSTM	F	71.2	-	-	-	79.4	-
RGB-N	F	81.7	43.7	79.5	40.8	93.7	72.2
SPAN	F	93.7	55.8	83.8	38.2	96.1	58.2
PSCC-Net	F	94.1	72.3	87.5	55.4	99.6	81.9
Ours	F	**95.7**	**75.8**	**88.2**	**57.9**	**99.6**	**82.4**

篡改定位

Methods	AUC	F1
Mantra-Net	59.94	56.69
SPAN	67.33	63.48
PSCC-Net	99.65	97.12
Ours	**99.70**	**97.34**

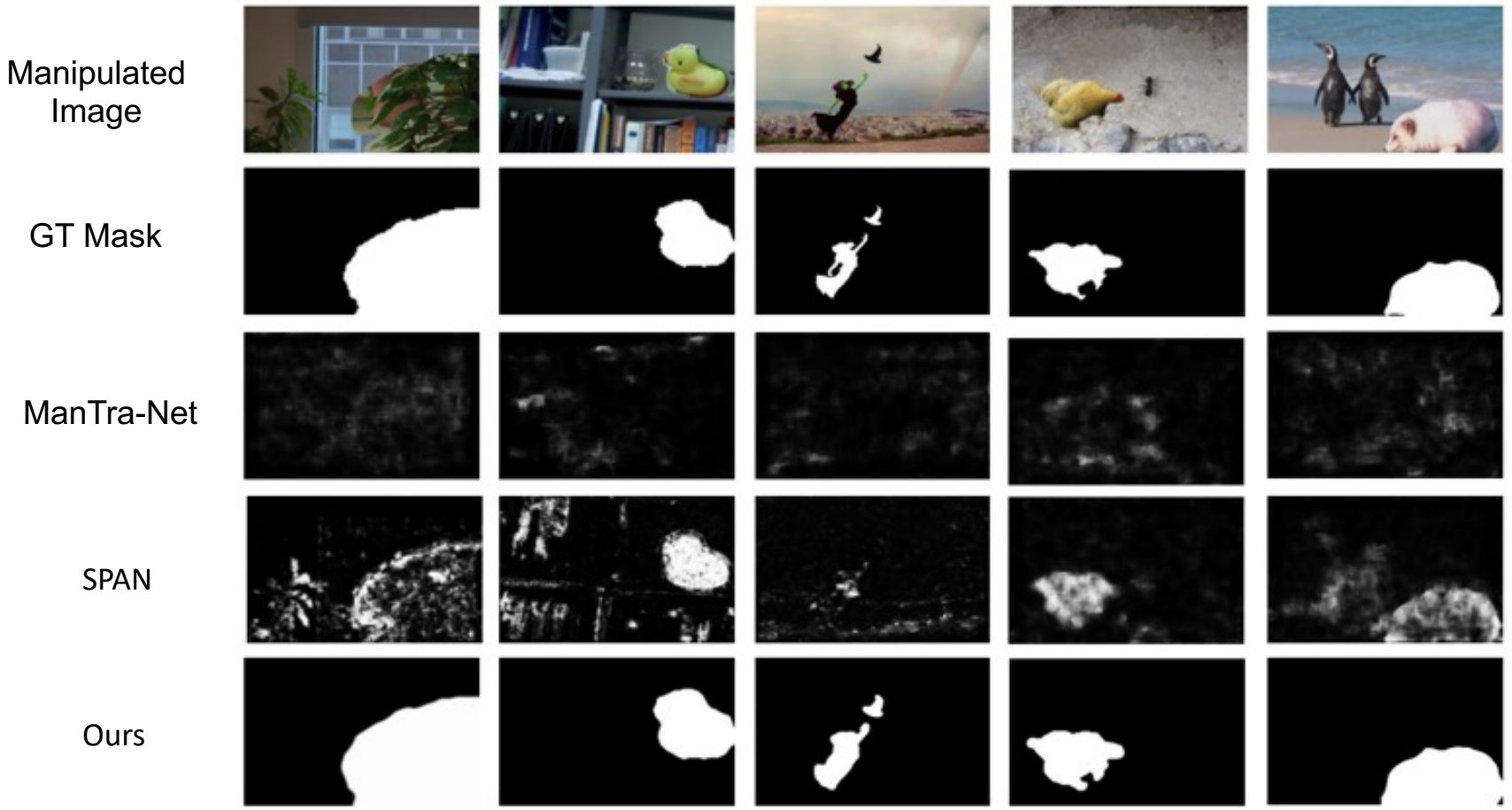
篡改检测

在图像篡改定位和鉴别上，我们的方法都取得了最好的结果。

图像篡改检测

Wang J et al. CVPR 2022.

篡改定位的可视化结果：相比于现有的SOTA方法有明显改善

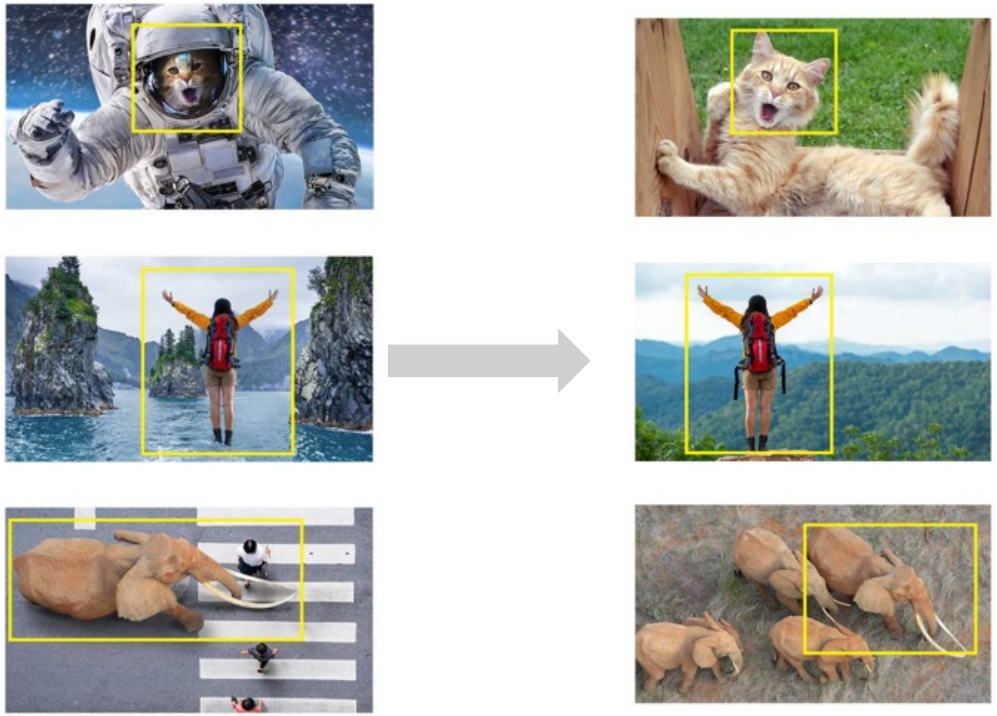


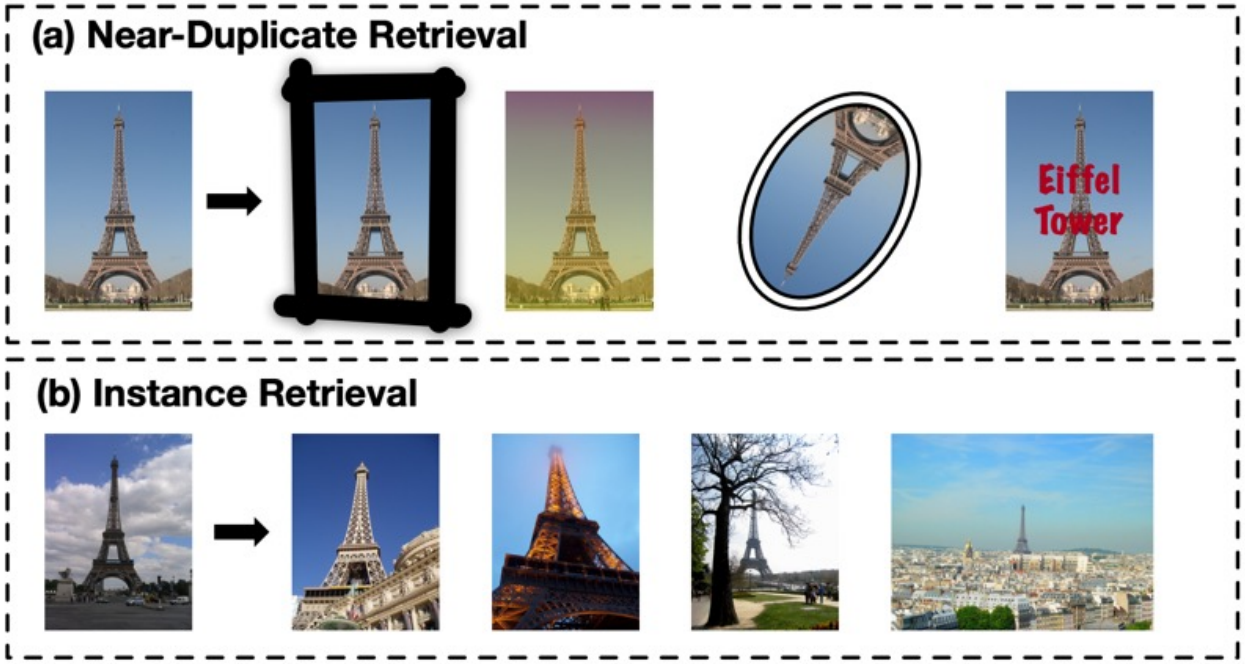
物体拼接检测与定位

Chen B et al. In CVPR 2021

拼接检测：

找出用于拼接的原始图像/视频帧



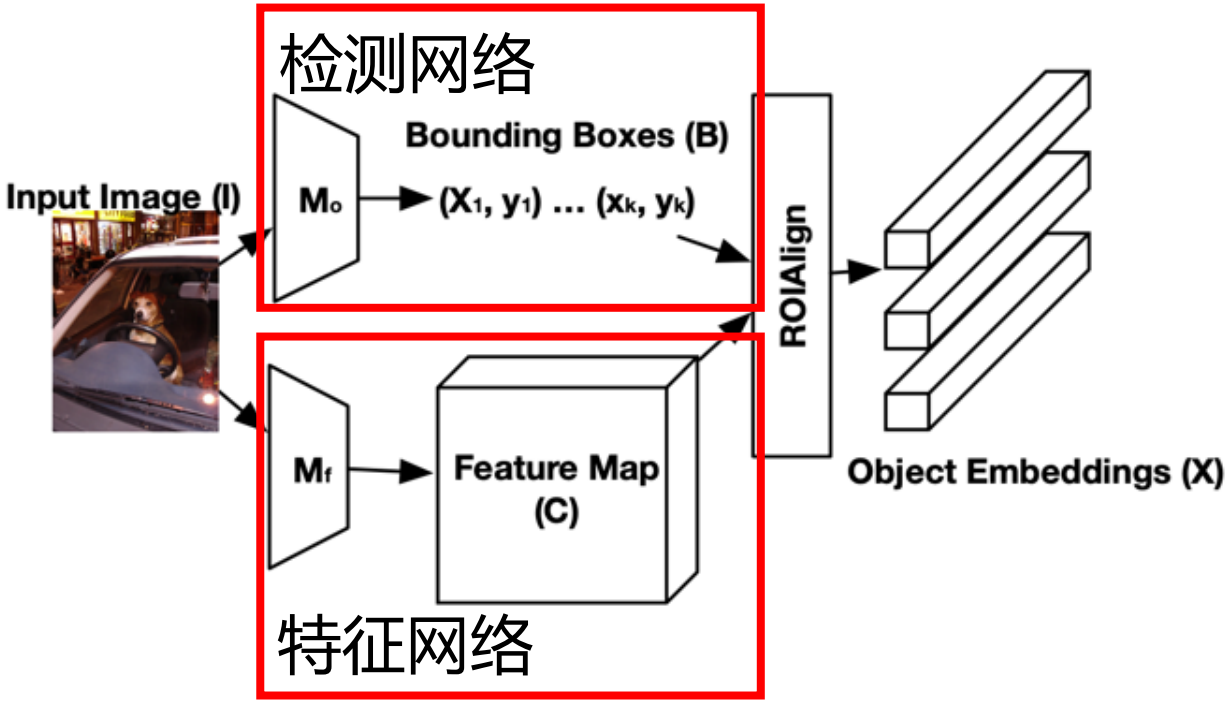


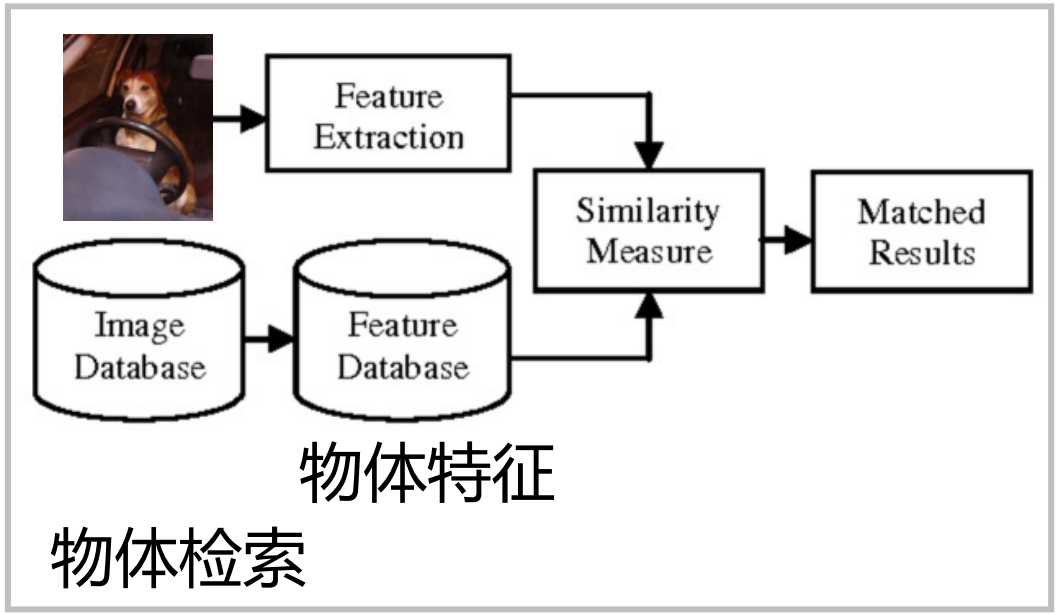
传统检测基于全局的信息，缺乏物体级别的表征能力

物体拼接检测与定位

Chen B et al. In CVPR 2021

使用同一个网络对特征输入进行特征提取和定位：



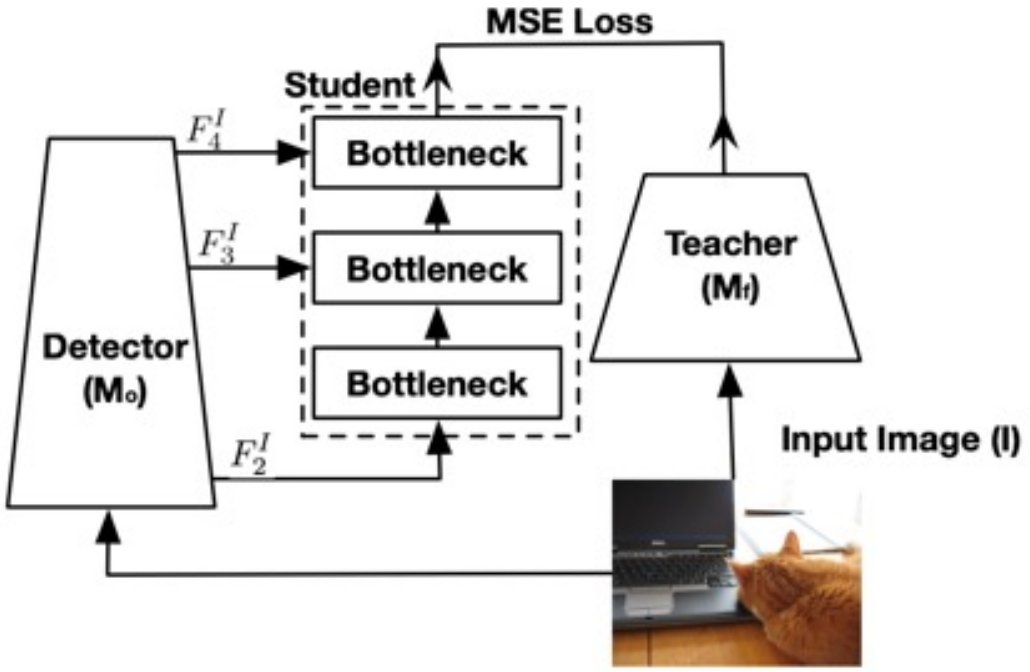


❖ 使用分开的网络有助于提升检索性能、特征网络可以随时更换

物体拼接检测与定位

Chen B et al. In CVPR 2021

知识蒸馏：



$$y_i = y_i + F_{i+2}^I, \quad \text{if } 1 \leq i \leq 2$$

$$\min_{\theta_s} \sum_I ||f(F_2^I, F_3^I, F_4^I; \theta_s) - M_f(I)||_2$$

❖ 使用知识蒸馏加速网络推理

物体拼接检测与定位

Chen B et al. In CVPR 2021

在多个数据集上取得最好的效果：

COCO-Fake

- 58张拼接图像
- 10000张真实图像

PIR

- 70,389张拼接图像
- 10,592张真实图像

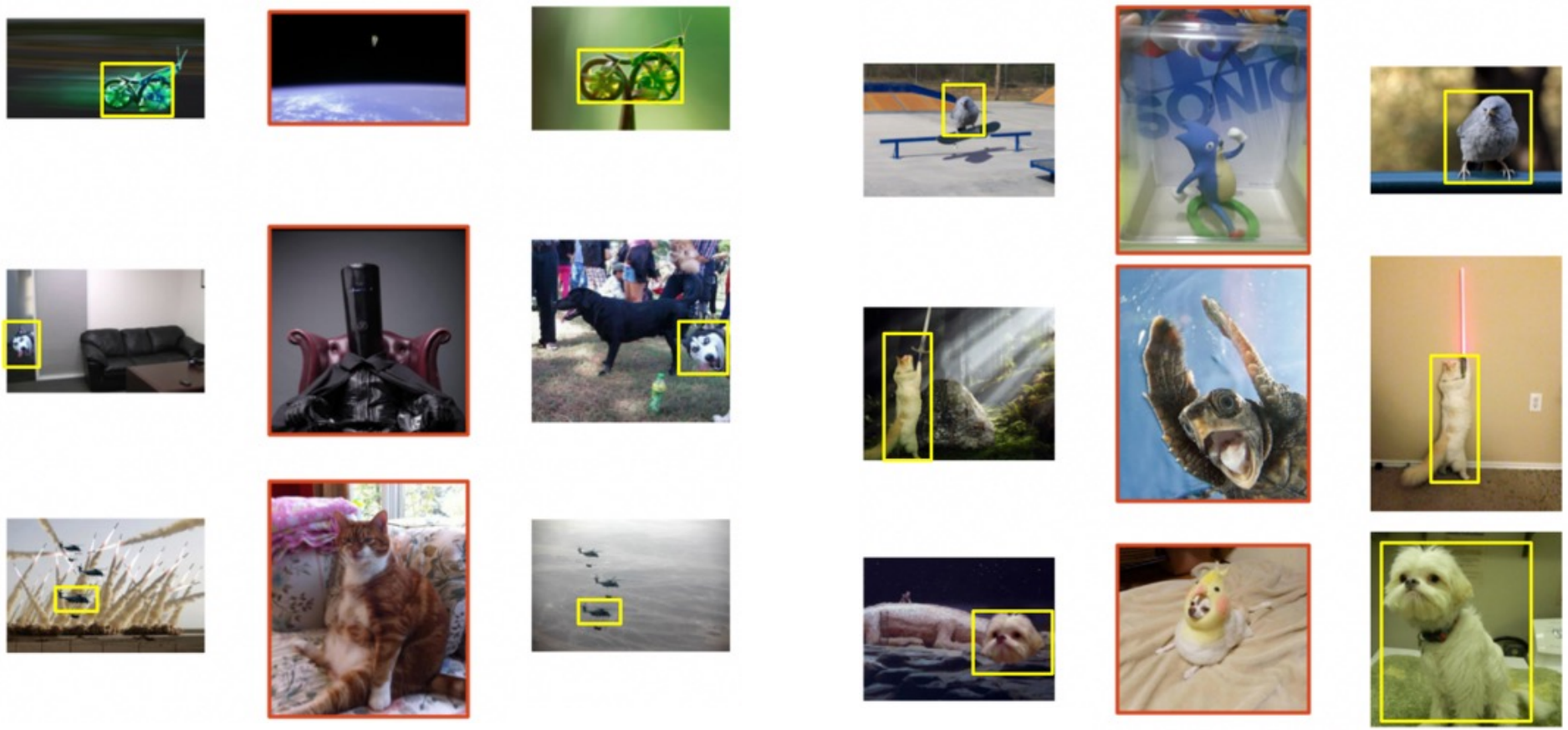
Method	COCO-Fake		PIR	
	R@1	R@10	R@1	R@10
SPoC [4]	29.3	34.3	43.2	46.6
MAC [51]	29.3	34.8	52.6	59.9
R-MAC [64]	37.9	42.5	51.6	58.5
GeM [50]	37.9	43.7	48.2	54.2
OE-HoG [13]	43.1	48.3	49.8	53.6
OE-FasterRCNN [53]	39.7	55.1	48.7	54.8
OE-SIR (Ours)	**70.7**	**84.5**	**58.6**	**67.7**

物体拼接检测与定位

Chen B et al. In CVPR 2021

在多个数据集上取得最好的效果：

图像检索方法　我们的算法　　图像检索方法　我们的算法

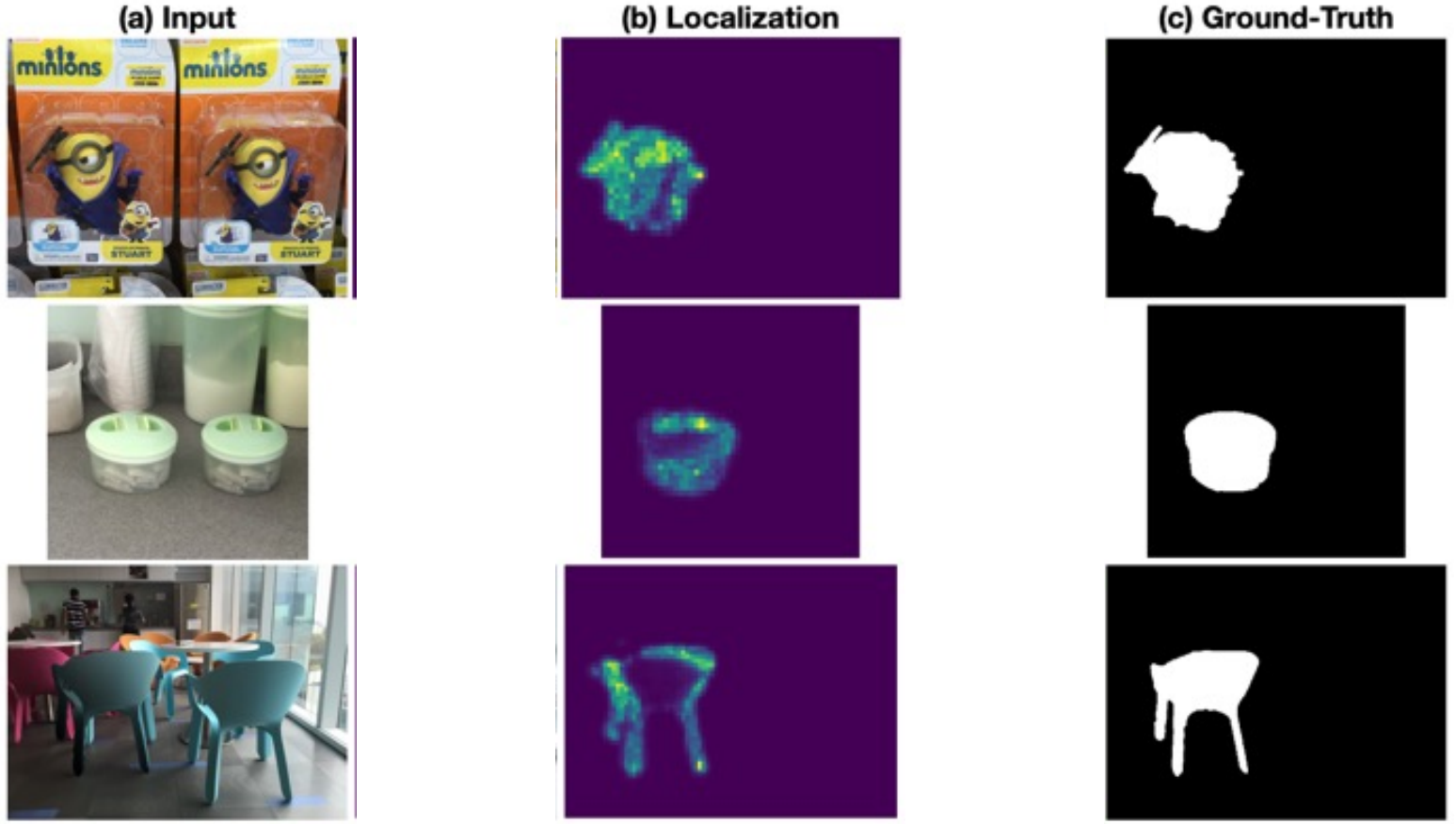


$$D(I_q, I_d) = \min_{i,j} ||x_i^q - x_j^d||_2^2.$$

物体拼接检测与定位

Chen B et al. In CVPR 2021

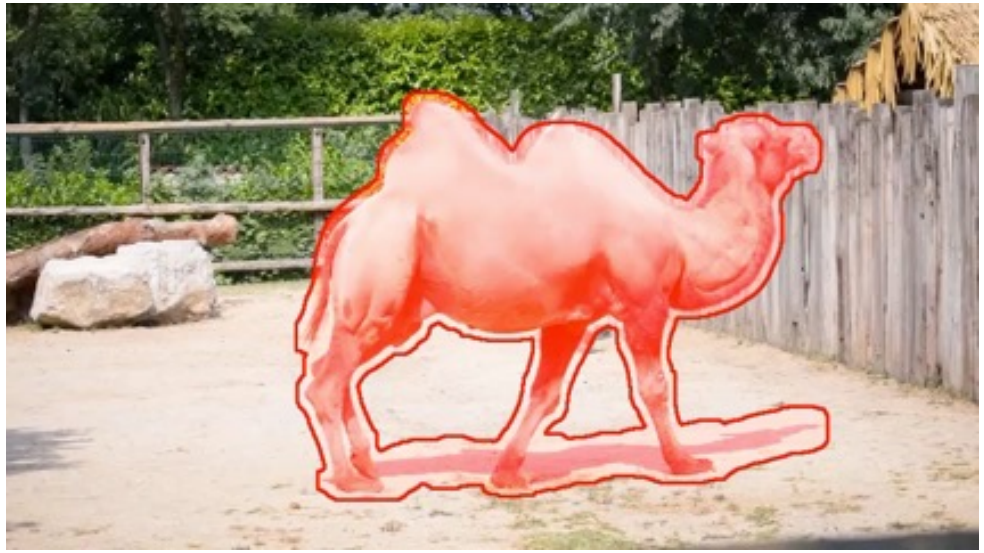
物体的标注框可以进一步帮助定位找到更改区域：

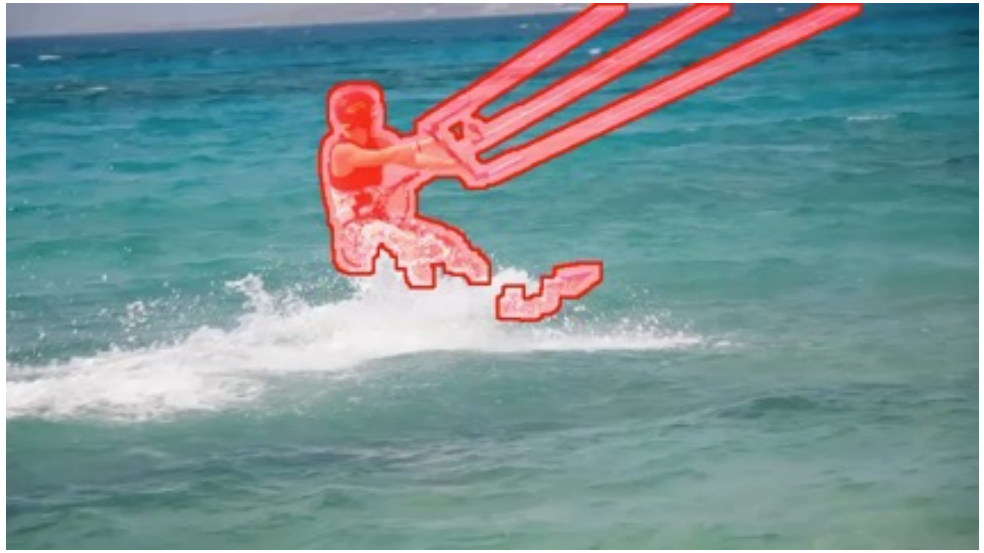


$$(i,j) = \arg\min_{i,j} ||x_i^q - x_j^d||_2^2 \qquad Q = ||M_f(T(I_q)) - M_f(I_d^*)||_2^2$$

视频篡改检测

Zhou P et al. BMVC 2021





视频篡改检测

Zhou P et al. BMVC 2021

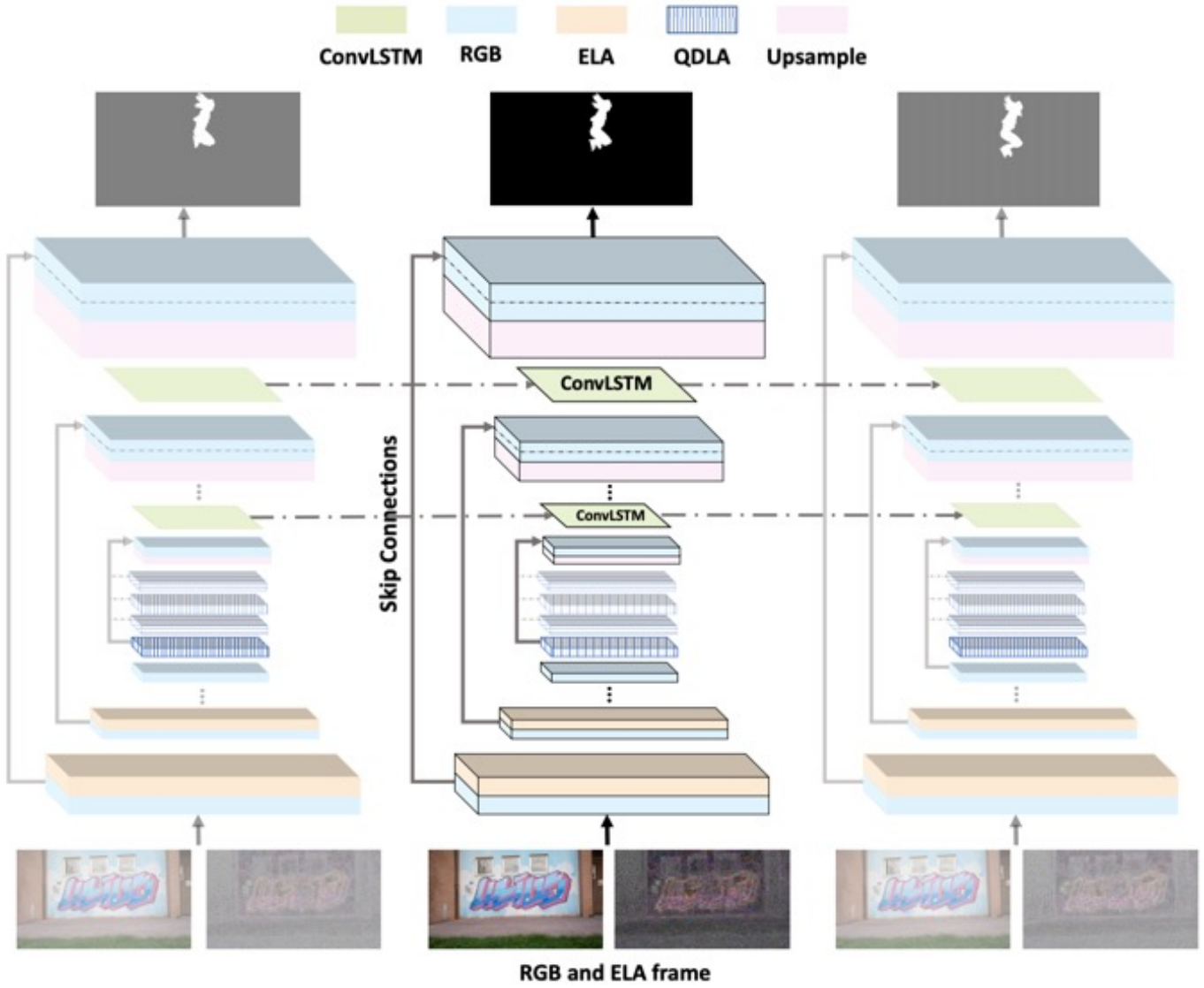
Error Level Analysis (ELA) 图像

视频篡改检测

Zhou P et al. BMVC 2021

多模态的框架：

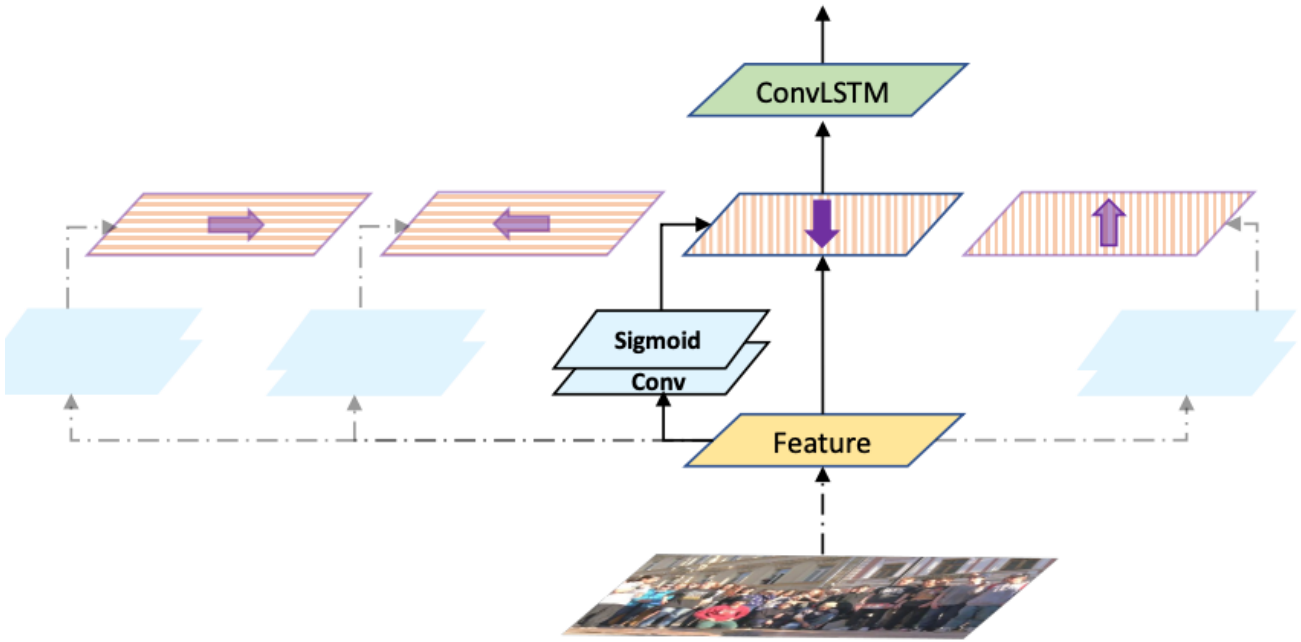
- 采用RGB和ELA特征同时建模
- 四个方向的注意力对相似度建模
- ConvLSTM时序建模



视频篡改检测

Zhou P et al. BMVC 2021

多模态的框架：



$$f_{5\rightarrow}[k] = (1 - A_{\rightarrow}[k]) f_5[k] + A_{\rightarrow}[k] f_5[k-1]$$

$$A_{\rightarrow} = \sigma(F_{\rightarrow}(f_5; W_{\rightarrow})),$$

视频篡改检测

Zhou P et al. BMVC 2021

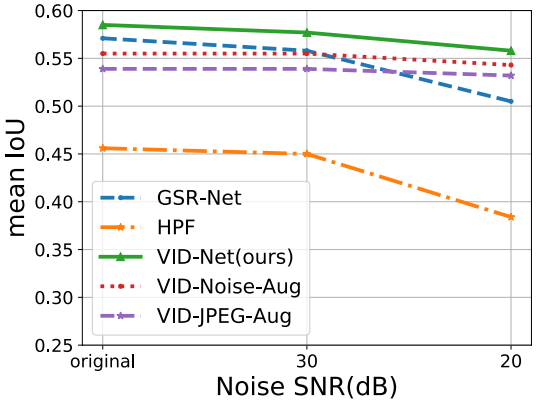
泛化性能较好：

Methods	VI* IoU/F1	OP* IoU/F1	CP IoU/F1	VI IoU/F1	OP* IoU/F1	CP* IoU/F1	VI* IoU/F1	OP IoU/F1	CP* IoU/F1
NOI [23]	0.08/0.14	0.09/0.14	0.07/ 0.13	0.08/0.14	0.09/0.14	0.07/0.13	0.08/0.14	0.09/0.14	0.07/ 0.13
CFA [7]	0.10/0.14	0.08/0.14	0.08/0.12	0.10/0.14	0.08/0.14	0.08/0.12	0.10/0.14	0.08/0.14	0.08/0.12
COSNet [22]	0.40/0.48	0.31/0.38	0.36/0.45	0.28/0.37	0.27/0.35	0.38/0.46	0.46/0.55	0.14/0.26	0.44/0.53
HPF [18]	0.46/0.57	0.49/0.62	0.46/0.58	0.34/0.44	0.41 /0.51	0.68/0.77	0.55/0.67	0.19/ 0.29	0.69/0.80
GSR-Net [42]	0.57/0.69	0.50/0.63	0.51/0.63	0.30 /0.43	0.74/0.82	0.80/0.85	0.59 /0.70	0.22/0.33	0.70/0.77
Ours RGB (baseline)	0.55/0.67	0.46/0.58	0.49/0.63	0.31/0.42	0.71 /0.77	0.78/0.86	0.58/0.69	0.20/0.31	0.70/0.82
VIDNet-BN (ours)	**0.62/0.73**	**0.75/ 0.83**	**0.67/0.78**	0.30/0.42	**0.80/0.86**	**0.84/0.92**	0.58 /0.70	0.23/0.32	0.75/0.85
VIDNet-IN (ours)	0.59/0.70	0.59/ 0.71	0.57/0.69	**0.39 /0.49**	0.74/0.82	0.81/0.87	**0.59/ 0.71**	**0.25/0.34**	**0.76/0.85**

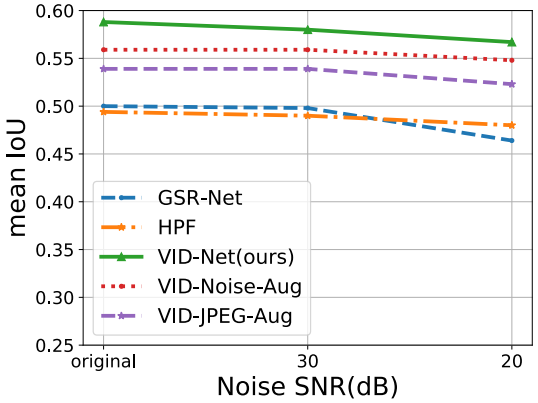
视频篡改检测

Zhou P et al. BMVC 2021

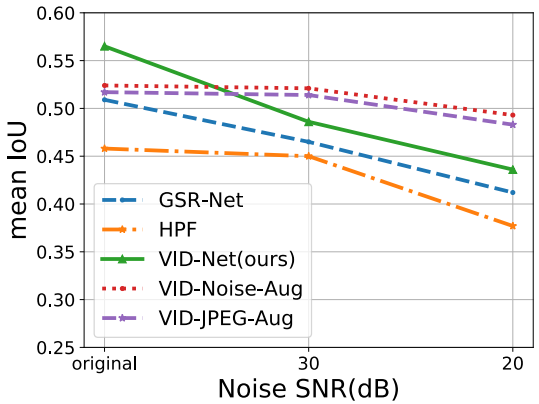
对噪声很鲁棒：



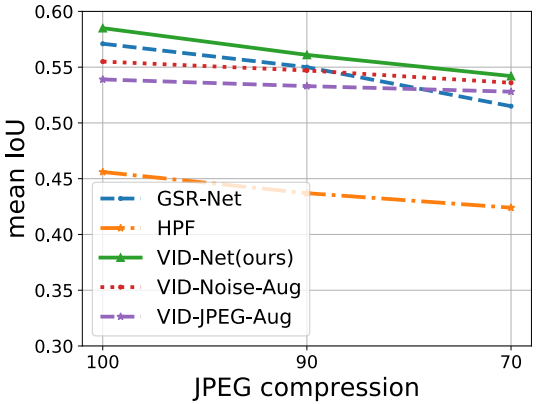
VI*



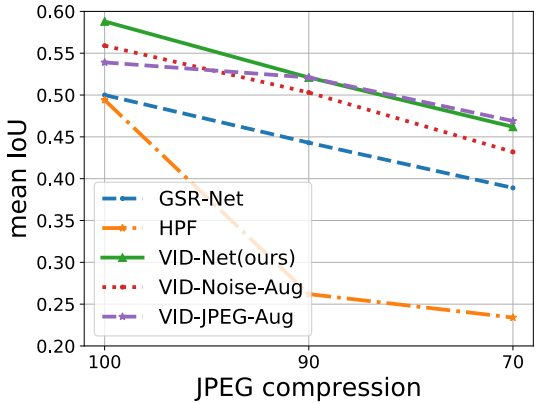
OP*



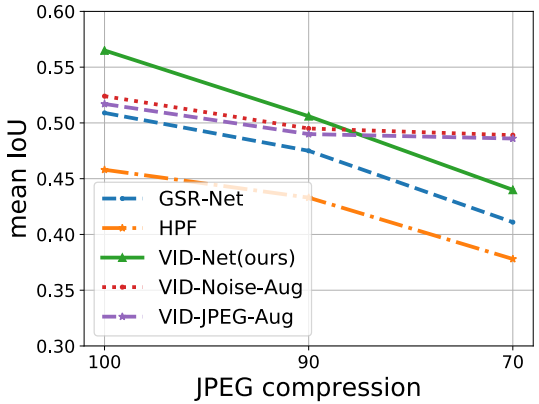
CP



VI*



OP*



CP

可视化结果

Original video with mask

Inpainted video

Original video with mask

Inpainted video

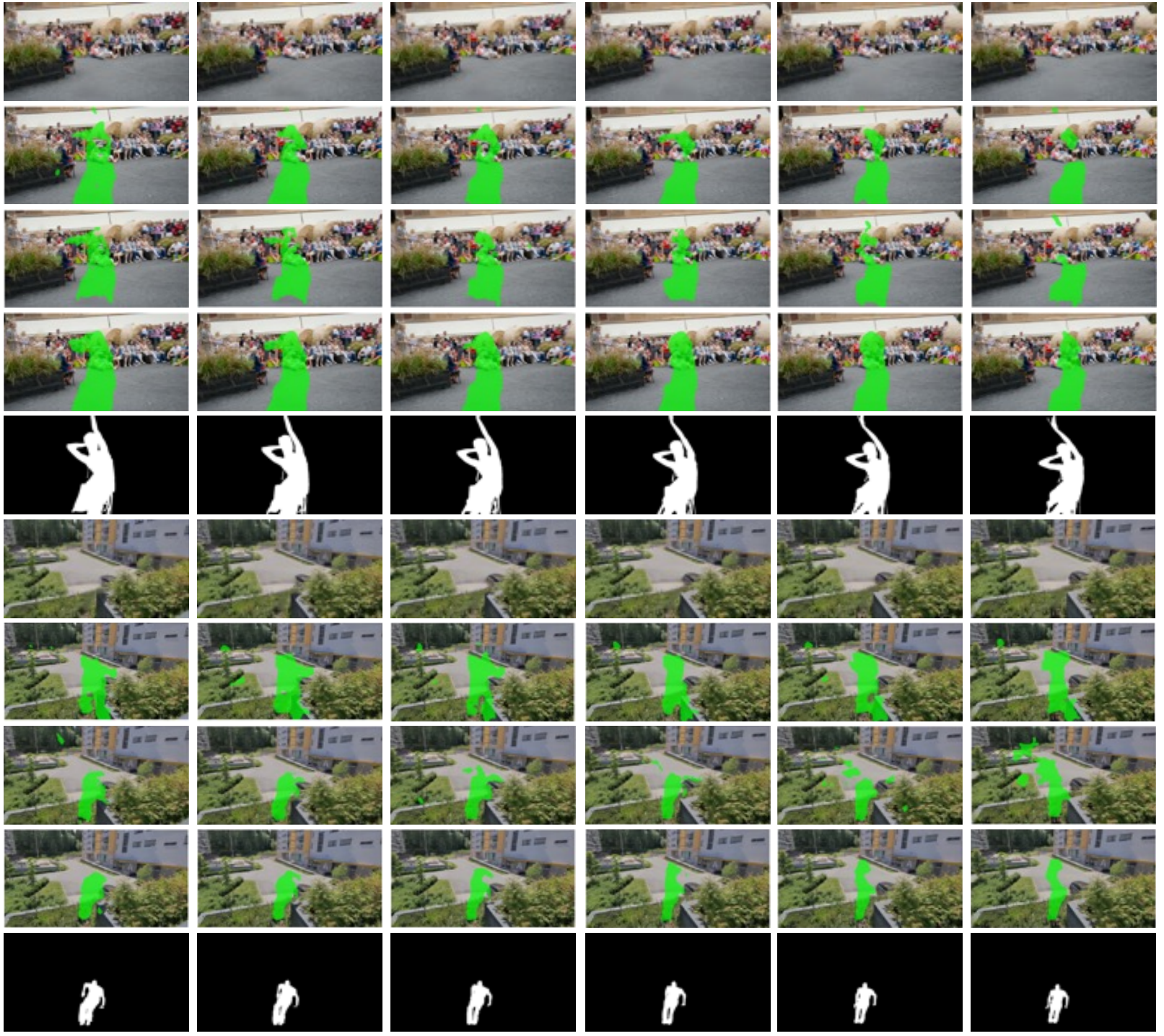
Input frame

HPF

GSR-Net

Ours

Ground truth



Input frame

HPF

GSR-Net

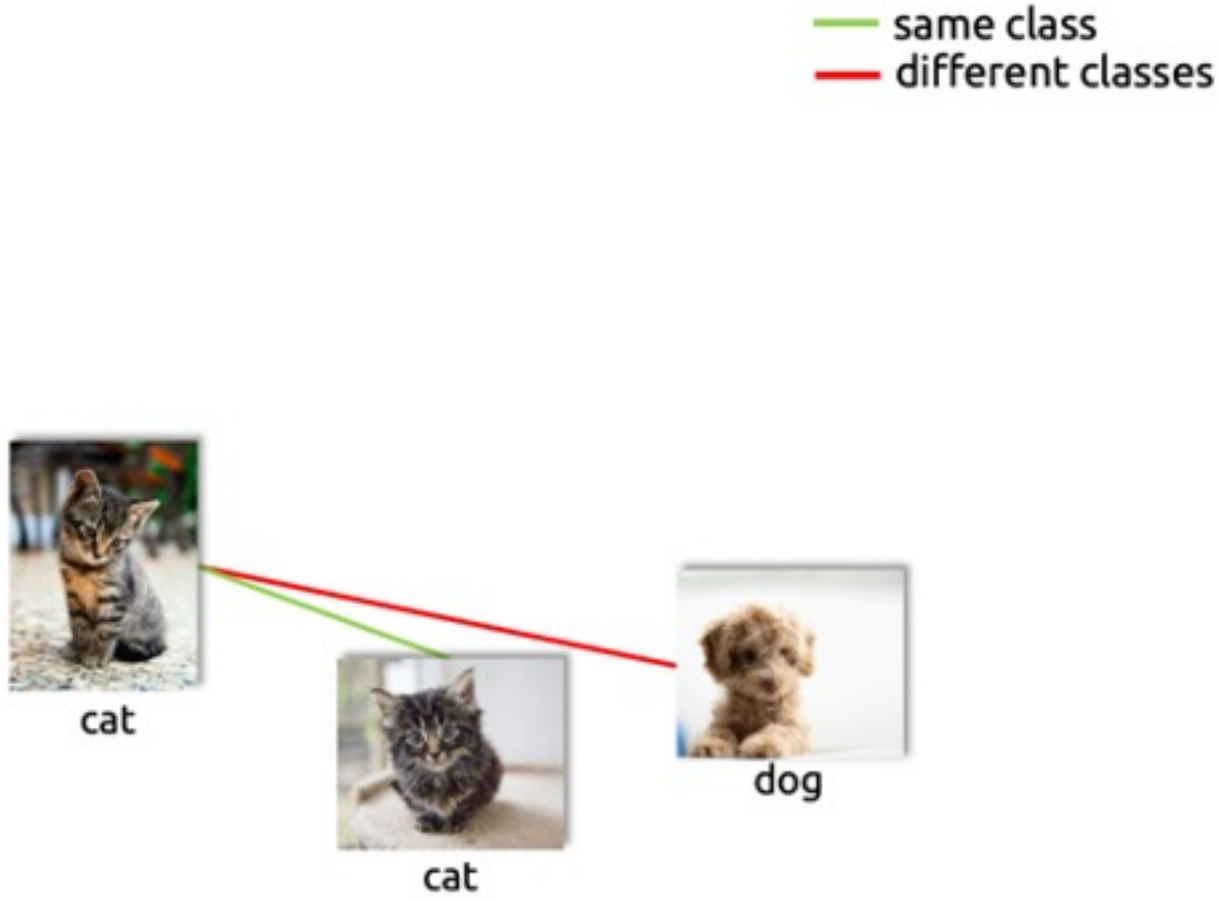
Ours

Ground truth

图像意图理解

Jia M et al. CVPR 2021

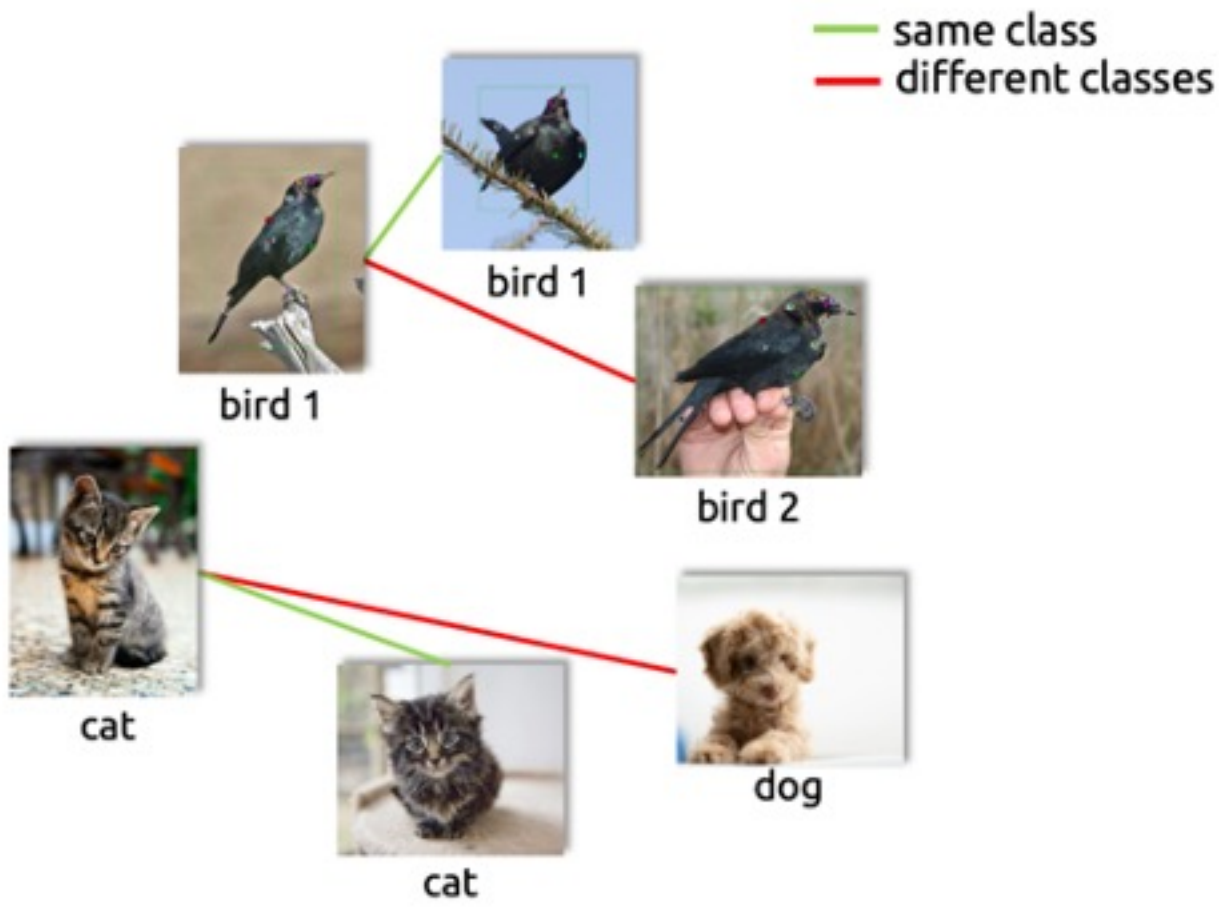
理解社交媒体意图：



图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：

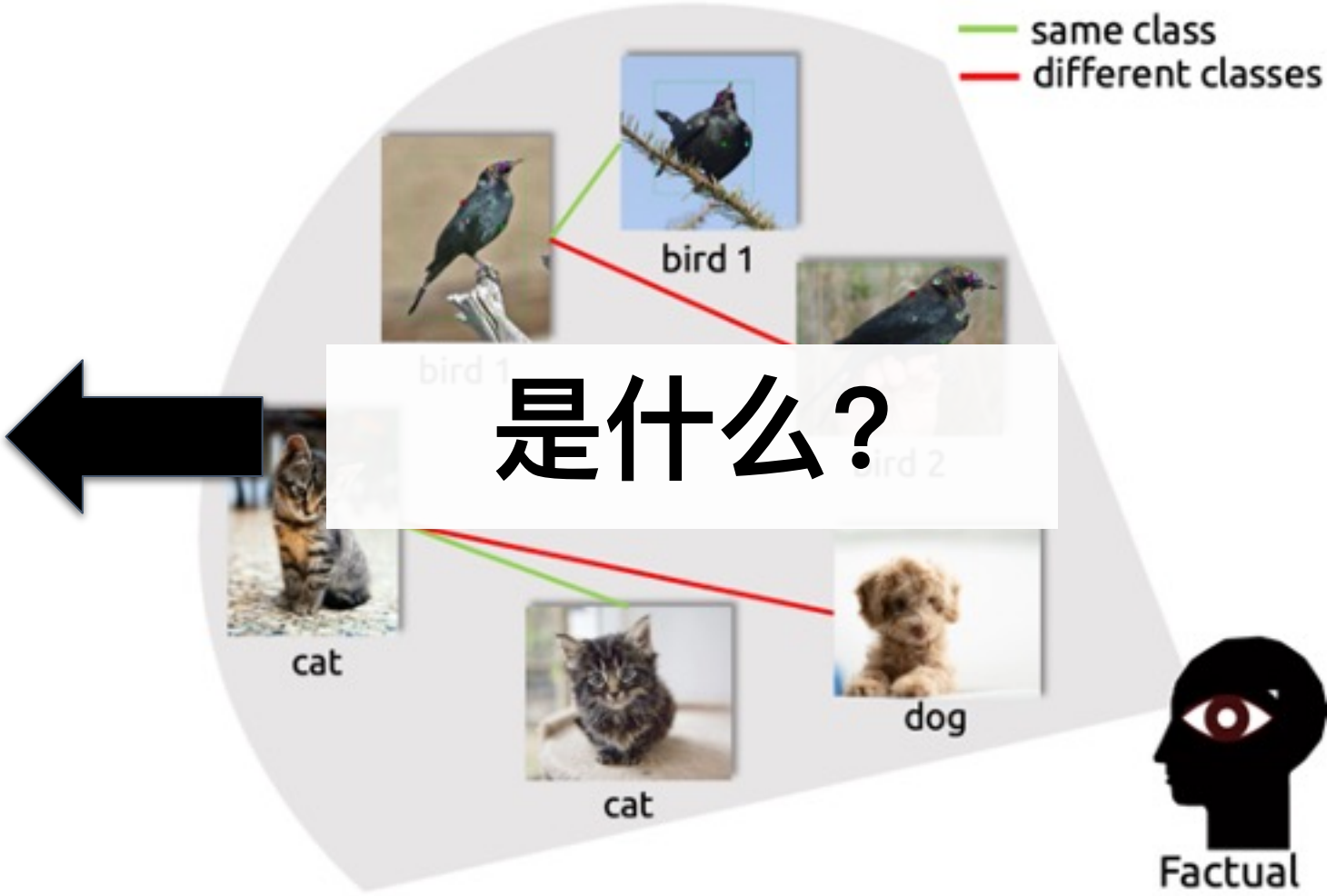


图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：

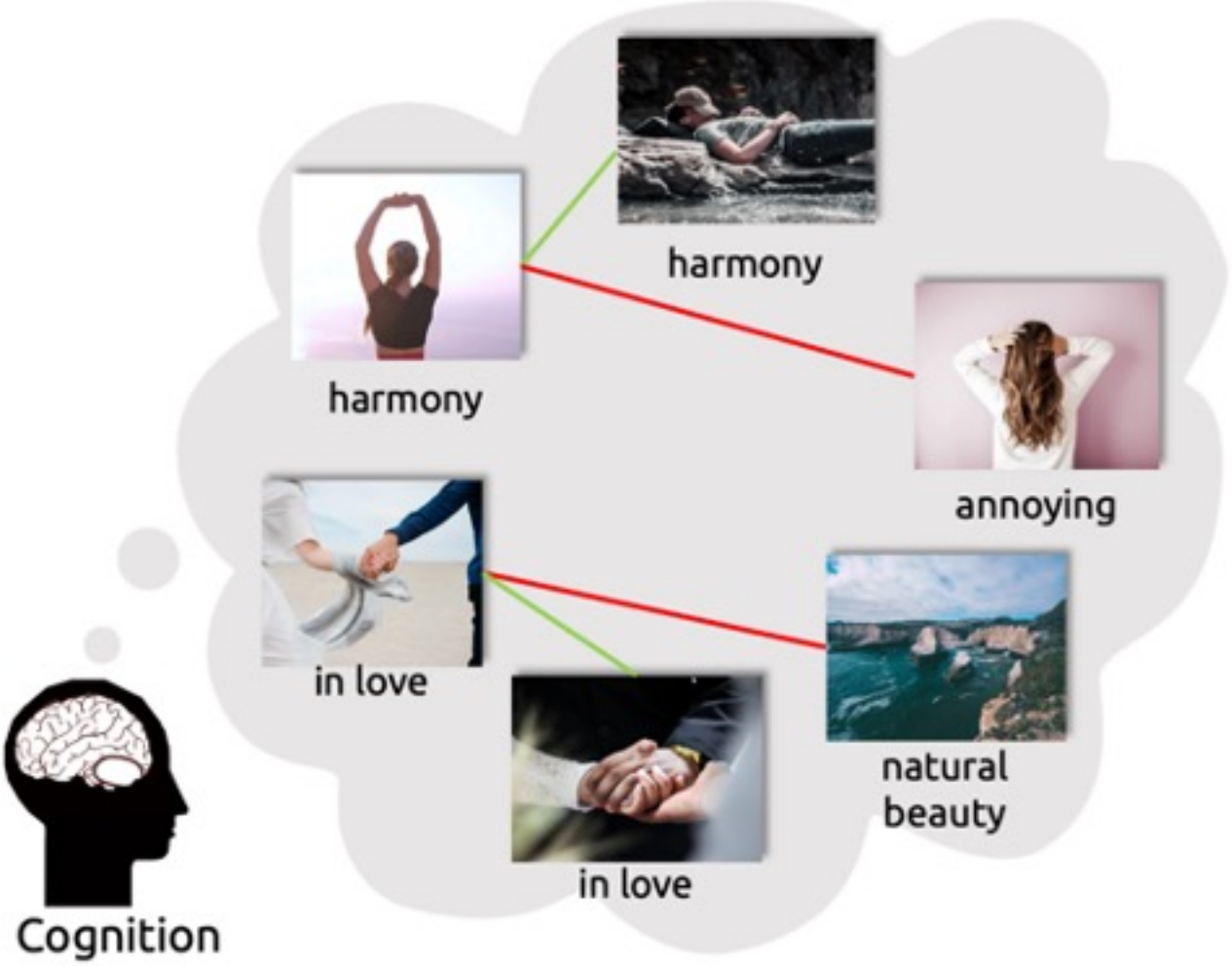
为什么？



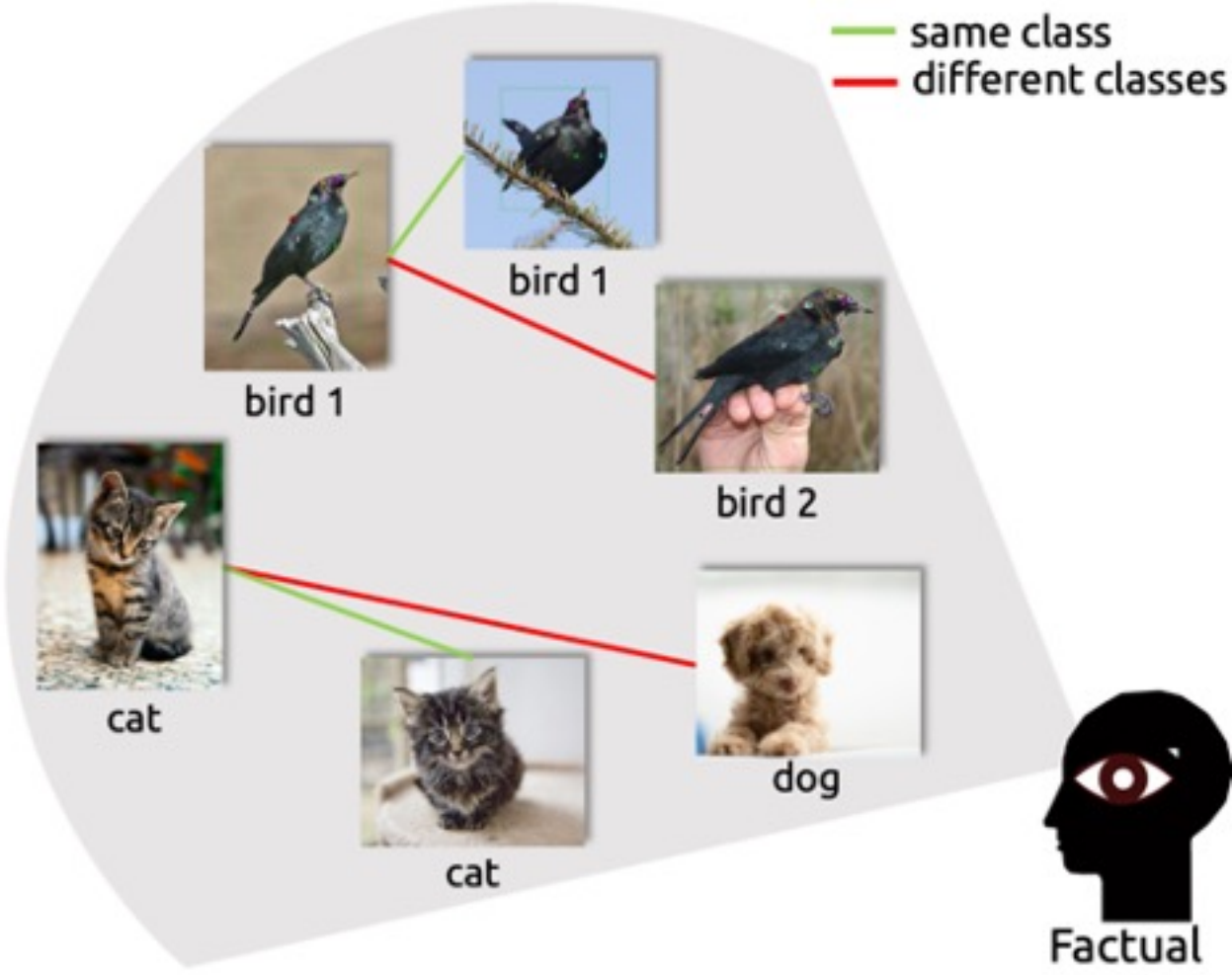
图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：



Why does my friend post this photo?

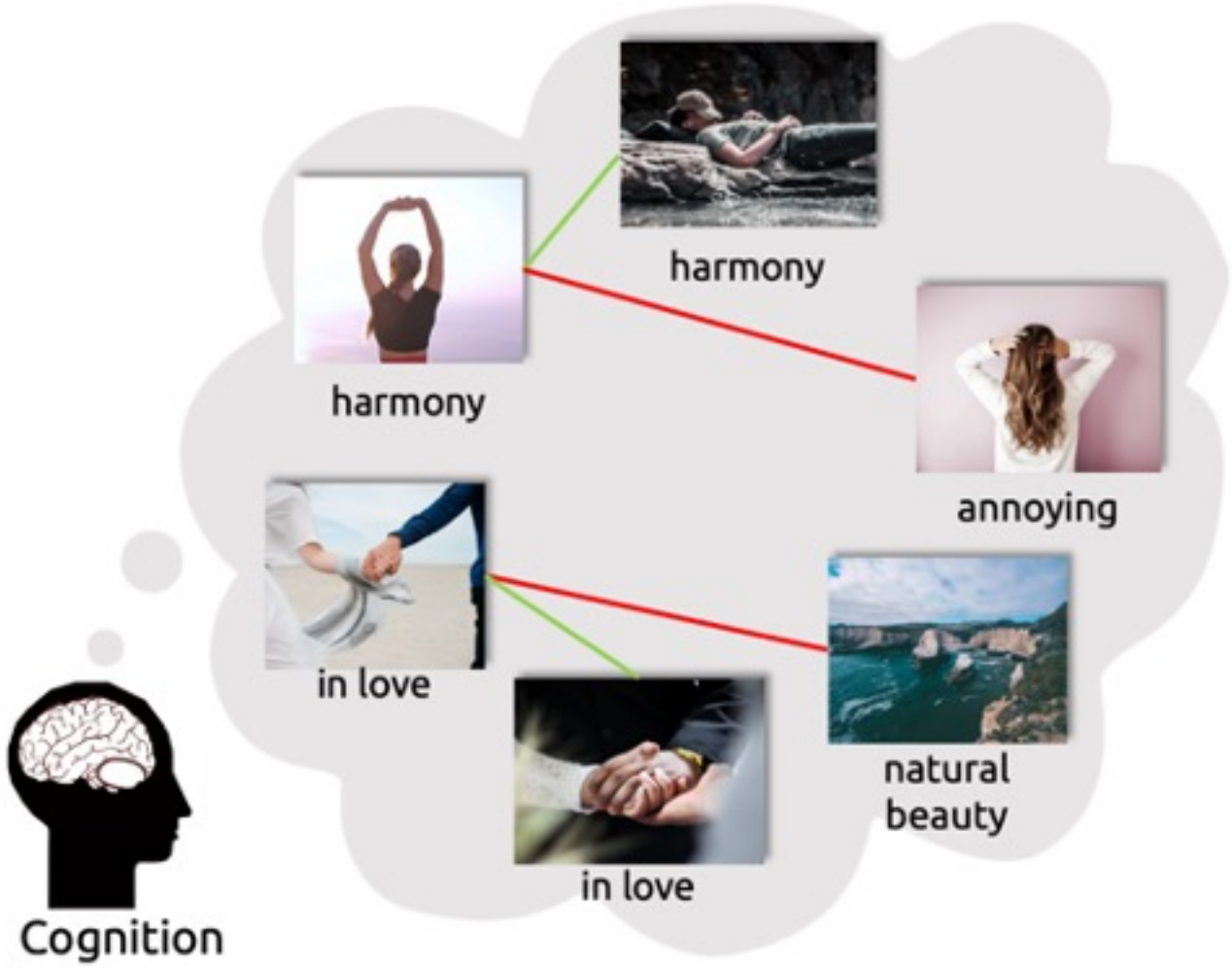


What's inside of the image?

图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：

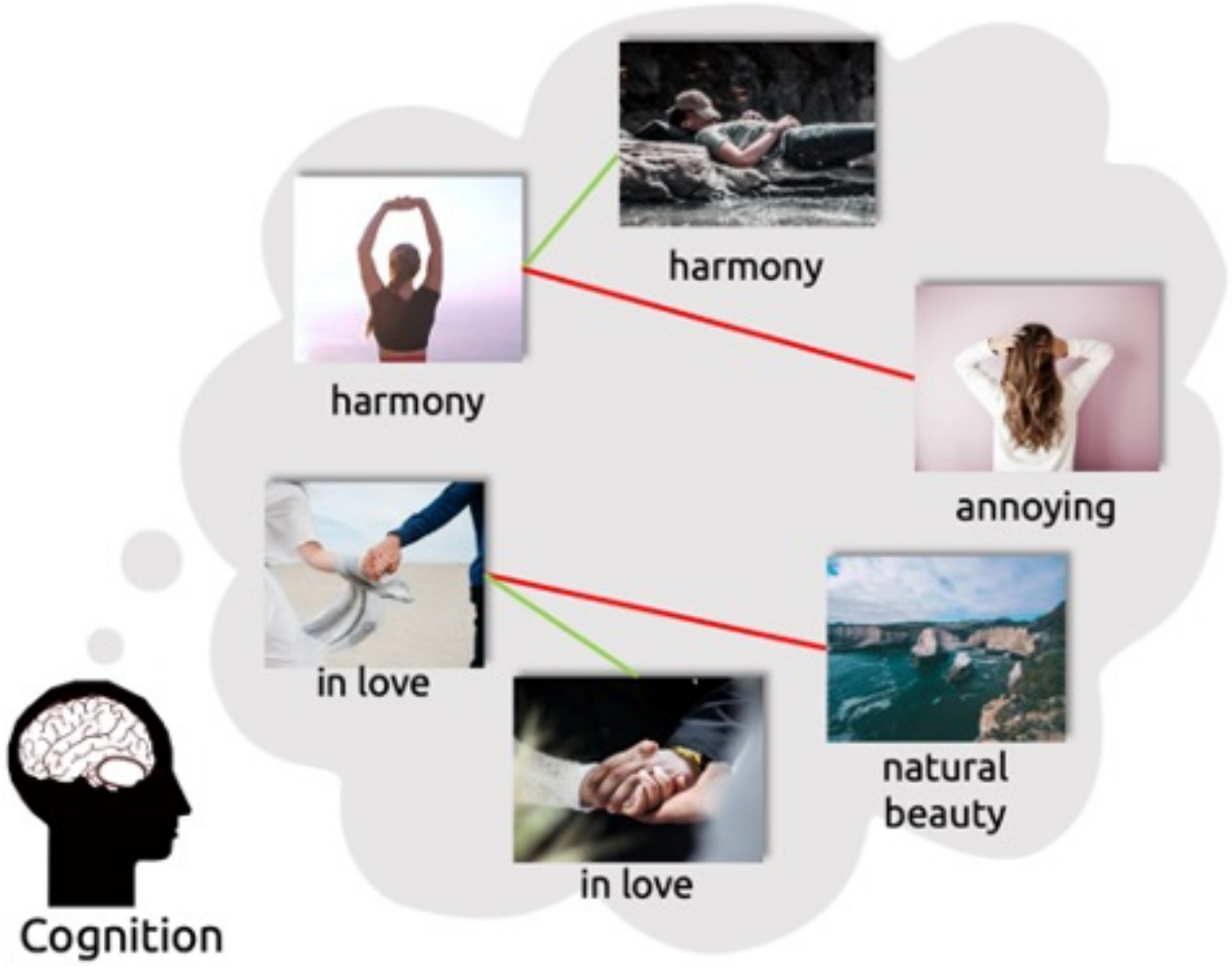


物体及上下文是否足够理解意图?

图像意图理解

Jia M et al. CVPR 2021

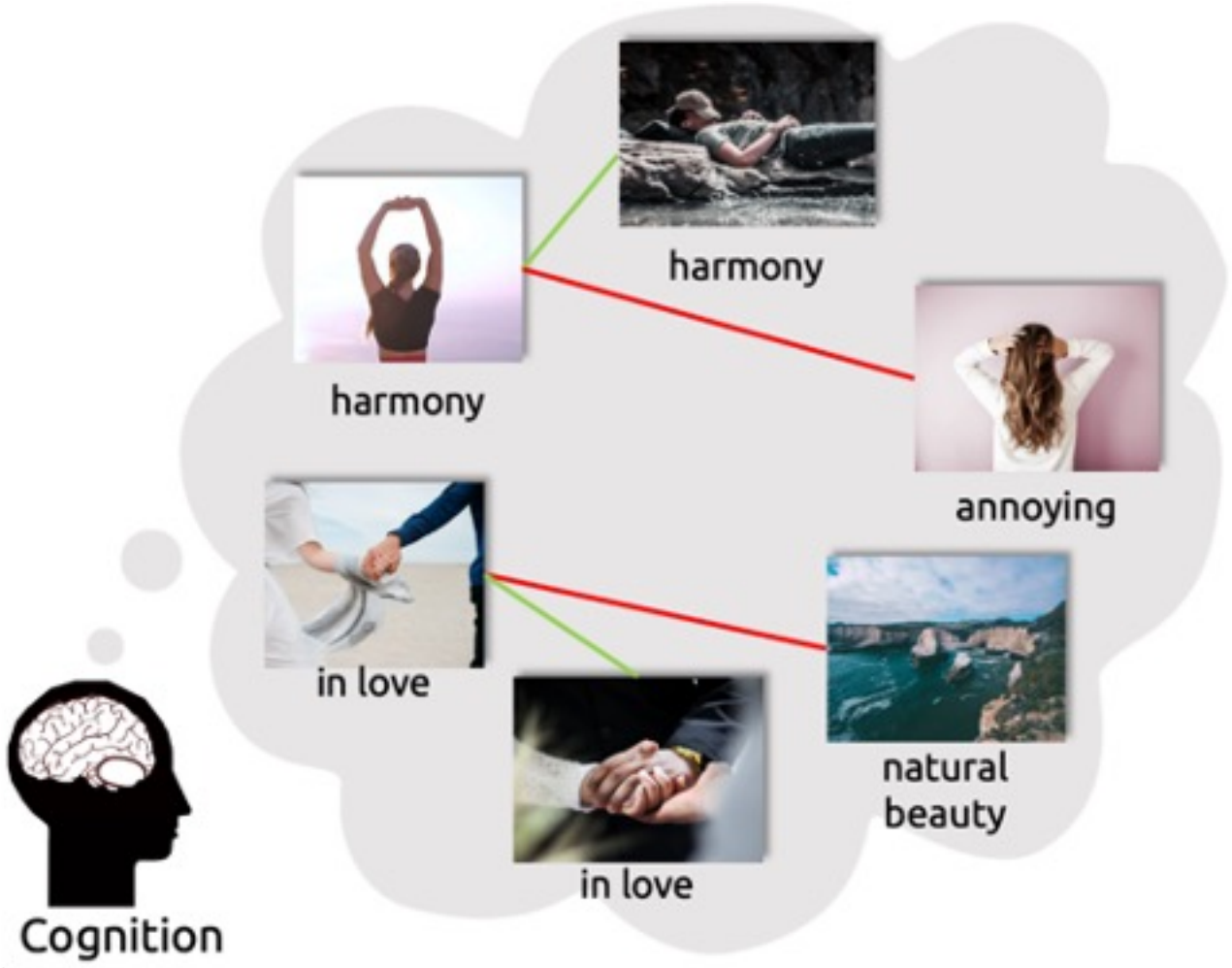
理解社交媒体意图：



图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：



Intentonomy

- 数据集
- 物体及上下文
- 弱监督的框架

图像意图理解

Jia M et al. CVPR 2021

理解社交媒体意图：

- 14,455图像 （Upsplash ）
- 28个根据心理学定义的意图类别
- 多标签分类

图像意图理解

Jia M et al. CVPR 2021

从视觉内容到意图识别：

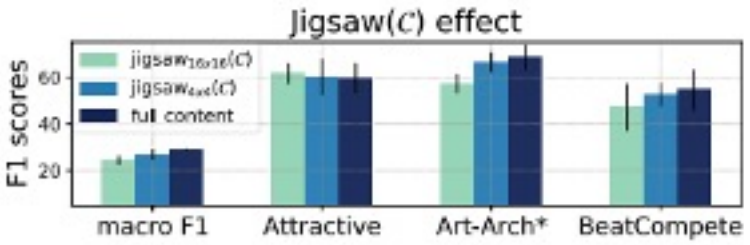
识别精度作为一个函数:

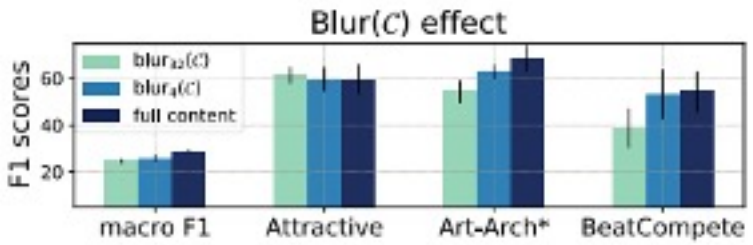
- object/contex信息的多少;
- object/context的性质，如空间、分辨率、纹理

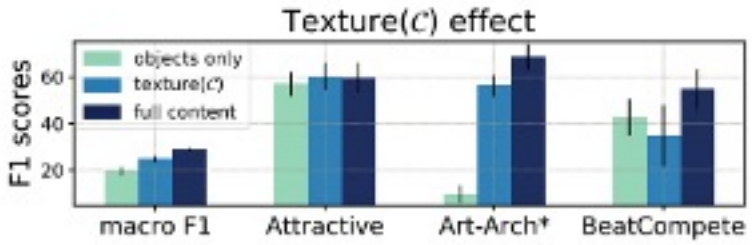
图像意图理解

Jia M et al. CVPR 2021

从视觉内容到意图识别：







上下文调整

(b) Content geometry.

(c) Content resolution.

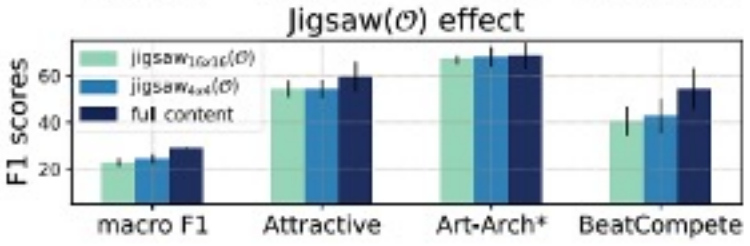
(d) Content texture.

图像意图理解

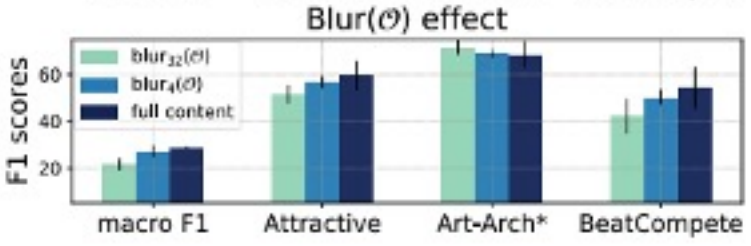
Jia M et al. CVPR 2021

从视觉内容到意图识别：

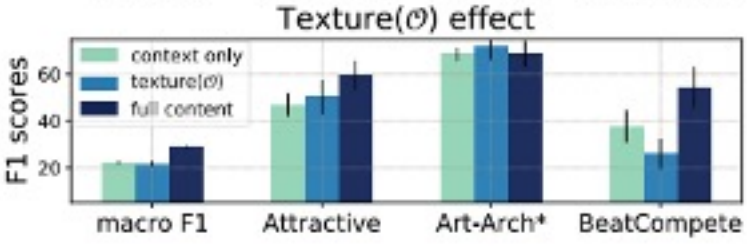
物体调整



(b) Content geometry.



(c) Content resolution.



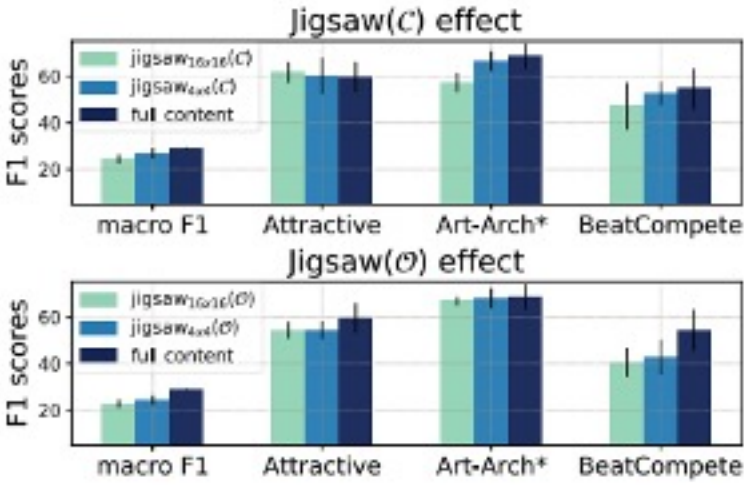
(d) Content texture.

图像意图理解

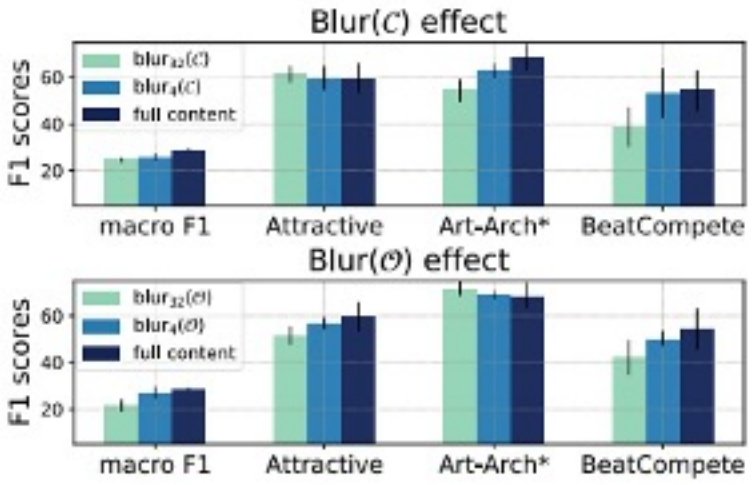
Jia M et al. CVPR 2021

从视觉内容到意图识别：

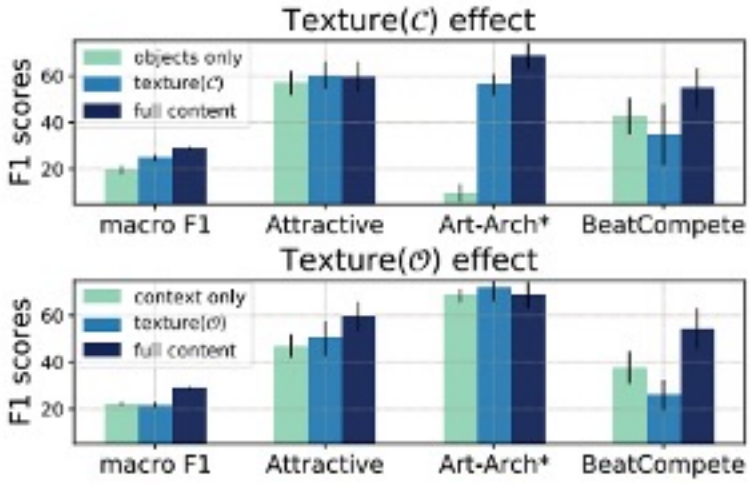
不同意图所依赖的信息不同



(b) Content geometry.



(c) Content resolution.

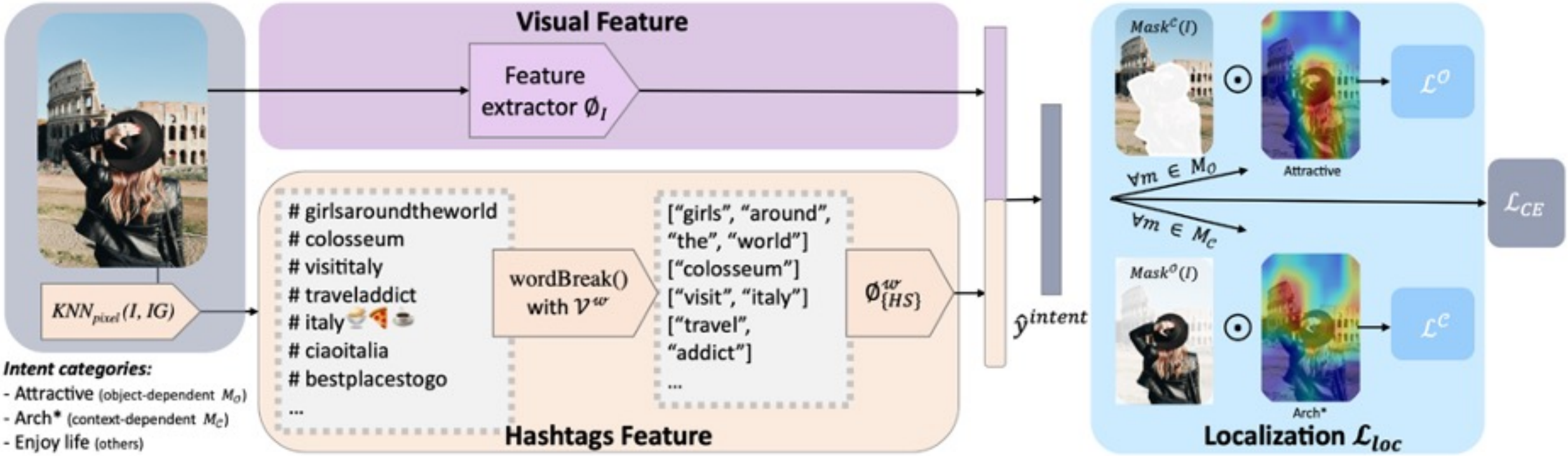


(d) Content texture.

图像意图理解

Jia M et al. CVPR 2021

从视觉内容到意图识别：



1. 不同类别关注不同的区域
2. 语义信息是视觉信息的有效补充.

图像意图理解

Jia M et al. CVPR 2021

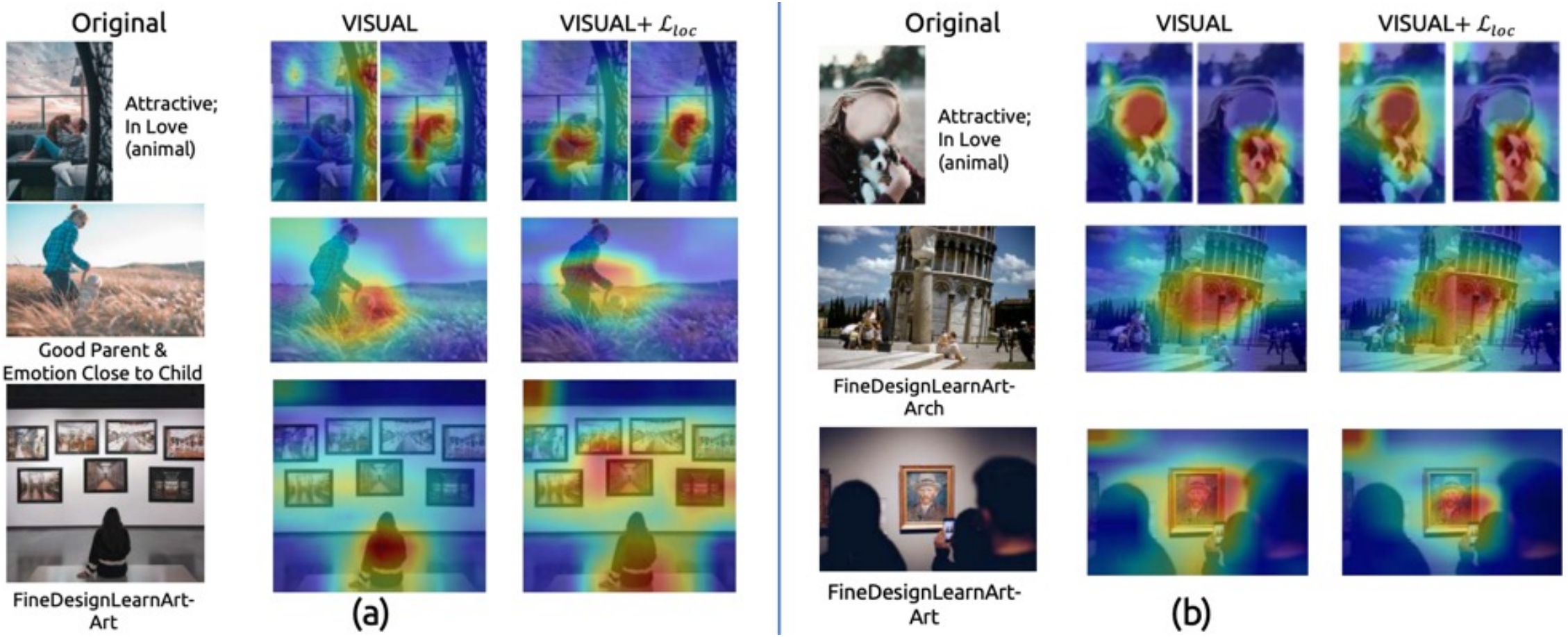
实验结果：

Method	Content			Difficulty		
	$\mathcal{O}$-classes	$\mathcal{C}$-classes	*Others*	Easy	Medium	Hard
RANDOM	7.75 ± 5.47	12.53 ± 5.96	6.05 ± 5.23	19.86 ± 1.28	7.11 ± 3.40	2.81 ± 1.80
VISUAL	34.92 ± 3.63	41.27 ± 3.53	25.34 ± 1.13	61.84 ± 4.90	33.71 ± 2.24	11.73 ± 1.74
HT	26.96 ± 0.80	35.15 ± 4.18	15.43 ± 0.87	63.58 ± 1.79	19.68 ± 1.70	6.63 ± 1.43
VISUAL + $\mathcal{L}_{loc}$	38.82 ± 1.95 (+3.9)	**43.14** ± 3.00 (+1.87)	25.90 ± 1.35 (+0.56)	63.67 ± 1.47 (+0.09)	**34.72** ± 1.26 (+1.01)	13.83 ± 1.13 (+2.10)
VISUAL + HT	37.71 ± 2.70 (+2.79)	42.17 ± 3.62 (+0.90)	26.36 ± 1.17 (+1.02)	**66.67** ± 2.12 (+4.83)	32.93 ± 1.57 (-0.78)	15.52 ± 0.98 (+3.79)
VISUAL + $\mathcal{L}_{loc}$ + HT	**39.82** ± 1.56 (+4.90)	42.09 ± 2.57 (+0.82)	**26.77** ± 1.13 (+1.43)	66.18 ± 4.56 (+4.34)	33.86 ± 1.08 (+0.15)	**16.50** ± 1.80 (+4.77)

图像意图理解

Jia M et al. CVPR 2021

实验结果：

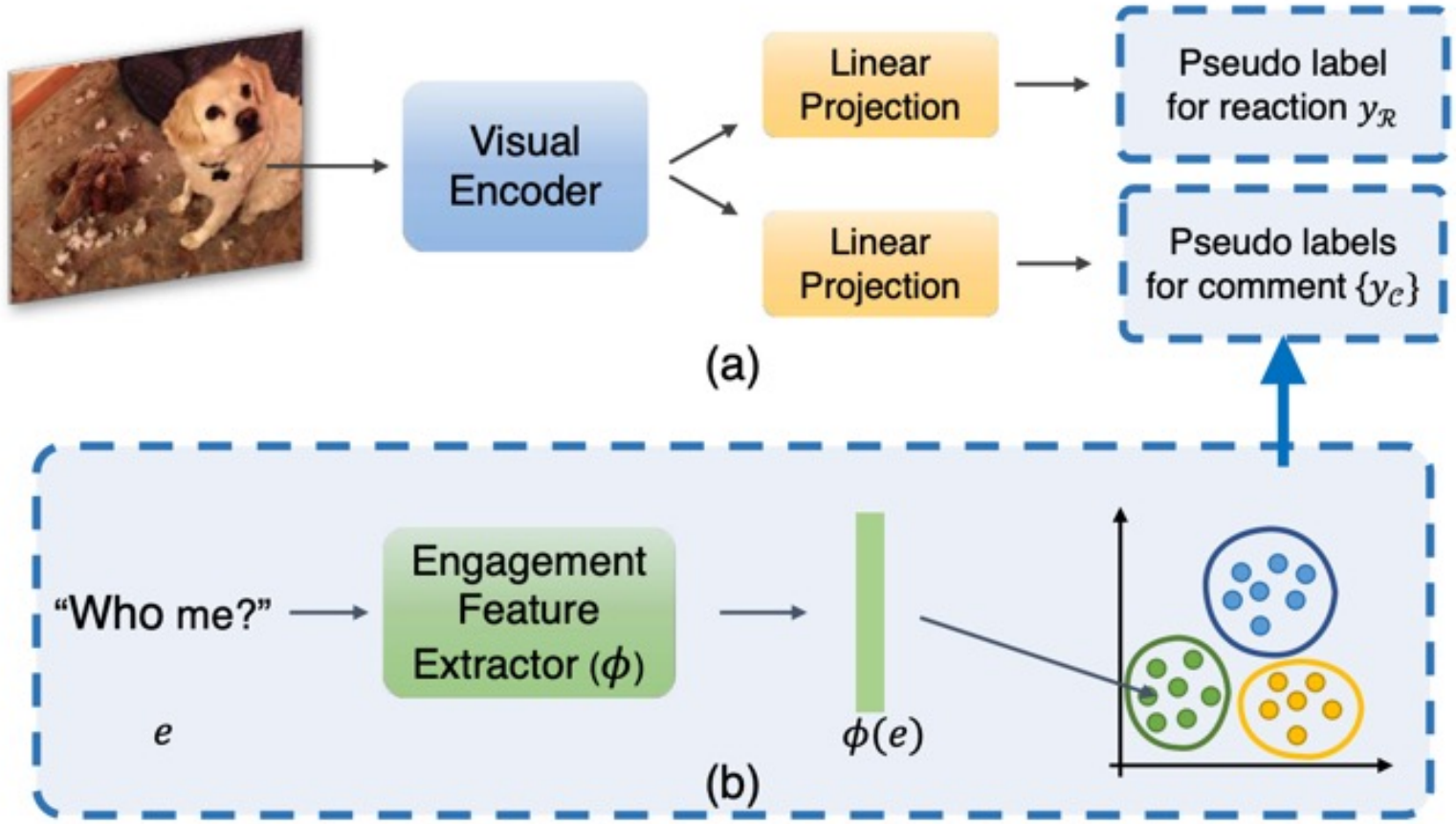


面向主观识别的特征学习

Jia M et al. ICCV 2021

面向主观识别的特征学习

Jia M et al. ICCV 2021



$$\arg\min_{W} \mathbb{E}_{(x_i,\{y_e\}_i)\sim\mathcal{D}} \mathcal{L}_{\mathcal{C}}(f(x_i),\{y_{\mathcal{C}}\}_i;W) + \mathcal{L}_{\mathcal{R}}(f(x_i),\{y_{\mathcal{R}}\}_i;W),$$

面向主观识别的特征学习

Jia M et al. ICCV 2021

Hateful Memes

https://ai.facebook.com/blog/hateful-memes-challenge-and-data-set/

Unbiased Emotion

面向主观识别的特征学习

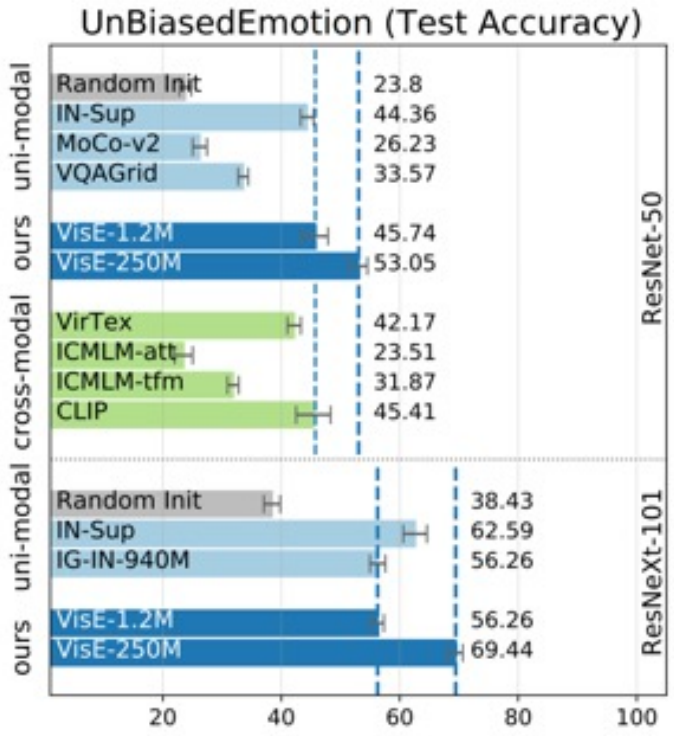
Jia M et al. ICCV 2021

实验设置：

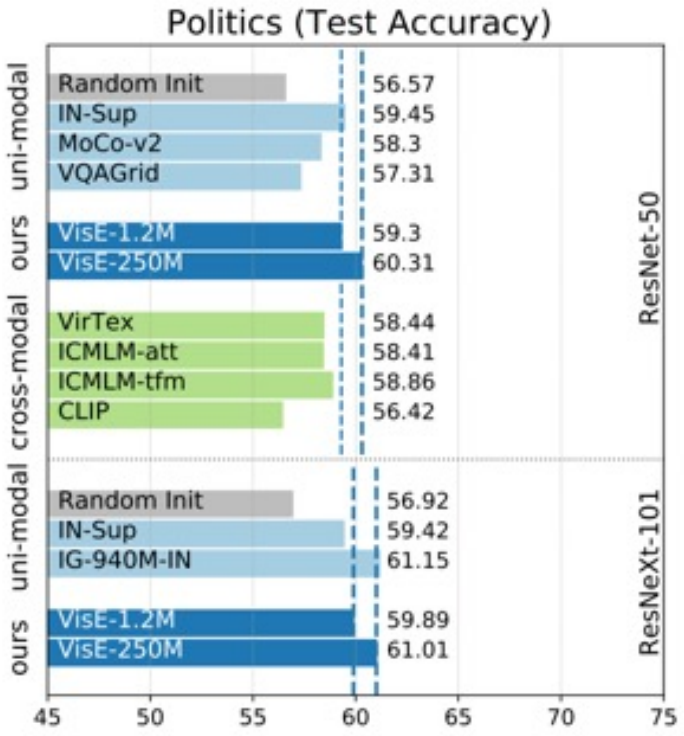
	Method	Input Type	Annotation Type	Noisy Labels	Pre-trained Data	Data Size
	Random Init	-	-	-	train from scratch	-
(1)	IN-Sup	images	object labels		ImageNet [11]	1.28M
	IG-940M-IN [55]	images	hashtags + labels	✓	IG [55] + ImageNet [11]	940M + 1.28M
	MoCo-v2 [7]	image pairs	-		ImageNet [11]	83.9B*
	VQAGrid [35]	images	object + attribute		VisualGenome [44]	103k
(2)	VirTex [12]	images	captions		COCO-caption [8]	118k
	ICMLM [2]	images + captions	masked token from captions		COCO-caption [8]	118k
	CLIP [66]	images + text	-		WebImageText [66]	13.1T*
	ours	images	pseudo labels	✓	VisE-1.2M	1.23M
					VisE-250M	250M

面向主观识别的特征学习

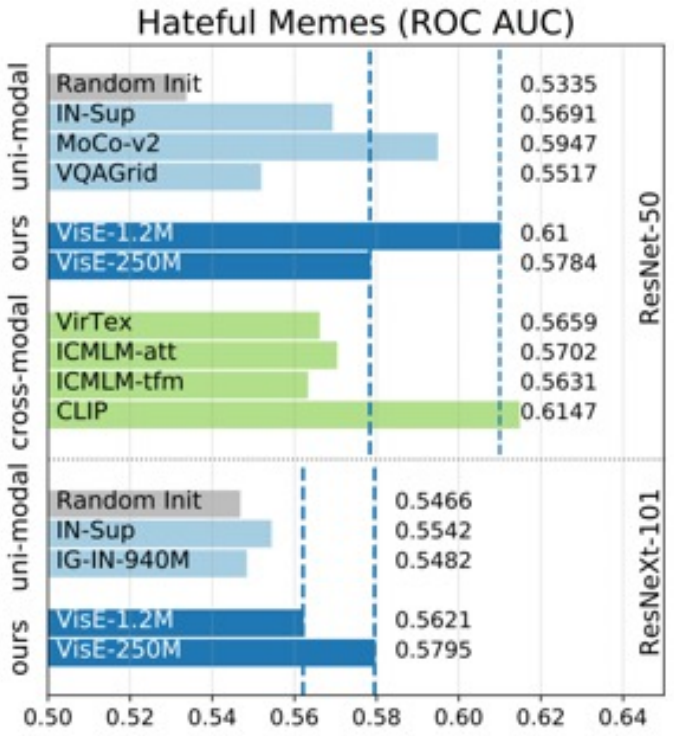
Jia M et al. ICCV 2021



(a) Linear evaluation: UnbiasedEmotion.



(b) Linear evaluation: Politics.



(c) Linear evaluation: Hateful Memes.

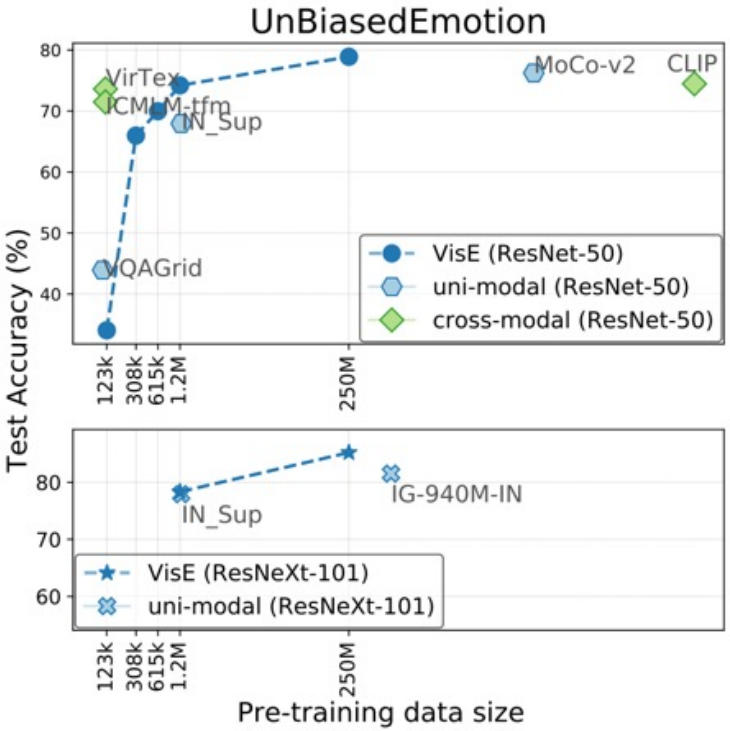
面向主观识别的特征学习

Jia M et al. ICCV 2021

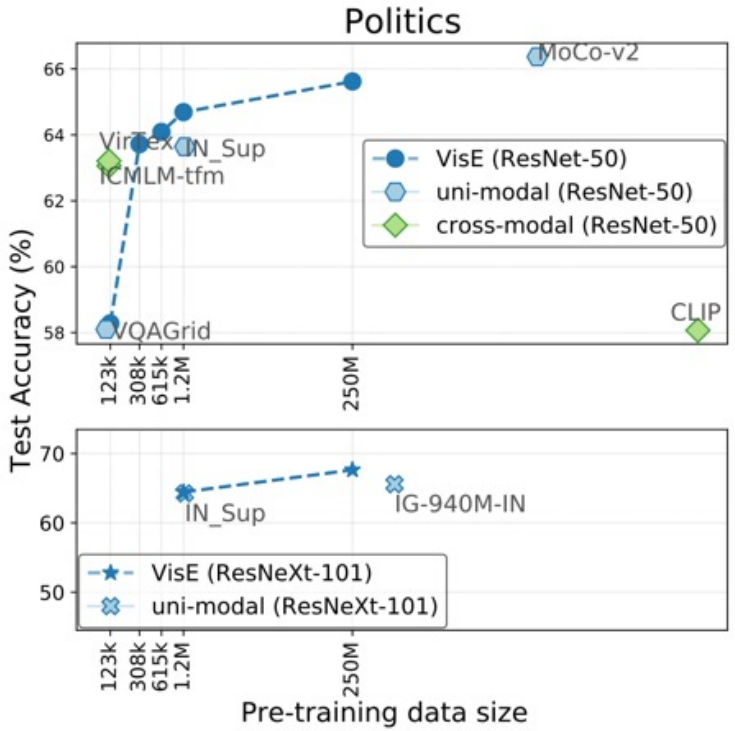
Backbone	Method	UnbiasedEmotion		Politics		Hateful Memes	
		Val Accuracy	Test Accuracy	Val Accuracy	Test Accuracy	ROC AUC	Accuracy
ResNet-50	Random Init	39.01 ± 0.99	37.25 ± 2.12	58.28	58.31	0.5833	51.84
	Uni-modality pre-training methods						
	IN-Sup	69.87 ± 3.27	67.94 ± 3.18	63.87	63.64	0.6005	54.32
	VQAGrid [35]	42.17 ± 3.07	43.93 ± 1.56	58.31	58.1	0.5906	53.24
	Cross-modalities pre-training methods						
	VirTex [12]	72.24 ± 2.13 ↑2.37	73.61 ± 1.94 ↑5.67	63.24	63.06	0.5898	53.84
	$ICMLM_{att\text{-}fc}$ [2]	71.65 ± 2.31 ↑1.78	70.98 ± 2.01 ↑3.05	63.3	63.2	0.5846	53.52
	$ICMLM_{tfm}$ [2]	70.92 ± 1.65 ↑1.05	71.48 ± 1.78 ↑3.54	63.43	63.21	0.5842	53.40
	Contrastive learning pre-training methods						
	MoCo-v2 [7]	77.63 ± 1.78 ↑7.76	76.23 ± 1.88 ↑8.29	**66.24** ↑2.37	**66.37** ↑2.73	0.5884	52.48
	CLIP [66]	73.68 ± 0.93 ↑3.81	74.46 ± 1.21 ↑6.53	58.08	58.07	0.5470	53.48
	Ours						
	VisE-1.2M	73.82 ± 1.07 ↑3.95	74.20 ± 1.93 ↑6.26	64.69 ↑0.82	64.69 ↑1.05	**0.6070** ↑0.0039	**55.88** ↑0.96
	VisE-250M	**79.74** ± 1.54 ↑9.87	**78.89** ± 2.23 ↑10.95	65.83 ↑1.96	65.62 ↑1.98	0.6060 ↑0.0055	55.00 ↑0.68
ResNet-101	Random Init	40.20 ± 2.76	39.08 ± 1.72	58.18	58.07	0.5868	53.48
	IN-Sup	71.84 ± 2.72	72.43 ± 2.24	58.28	58.42	0.5939	**54**
	VisE-1.2M (ours)	**73.82** ± 0.77 ↑1.97	**74.52** ± 1.18 ↑2.10	**63.92** ↑5.64	**63.85** ↑5.43	**0.5958** ↑0.0019	52.96 ↓1.04
ResNeXt-101 32 × 16d	Random Init	40.20 ± 2.65	38.59 ± 0.91	58.26	58.39	0.5959	**54.68**
	IN-Sup	79.00 ± 2.33	77.92 ± 2.38	64.22	64.25	0.5903	52.92
	IG-940M-IN [55]	83.24 ± 1.68 ↑4.24	81.52 ± 1.76 ↑3.60	65.90 ↑1.68	65.58 ↑1.33	0.5951 ↑0.0048	54.28 ↑1.36
	VisE-1.2M (ours)	77.57 ± 2.43 ↓1.43	78.33 ± 1.39 ↑0.41	64.61 ↑0.39	64.44 ↑0.19	0.5976 ↑0.0073	54.40 ↑1.48
	VisE-250M (ours)	**84.08** ± 1.87 ↑5.08	**85.21** ± 1.24 ↑7.29	**67.61** ↑3.39	**67.64** ↑3.39	**0.5957** ↑0.0054	**54.96** ↑2.04

面向主观识别的特征学习

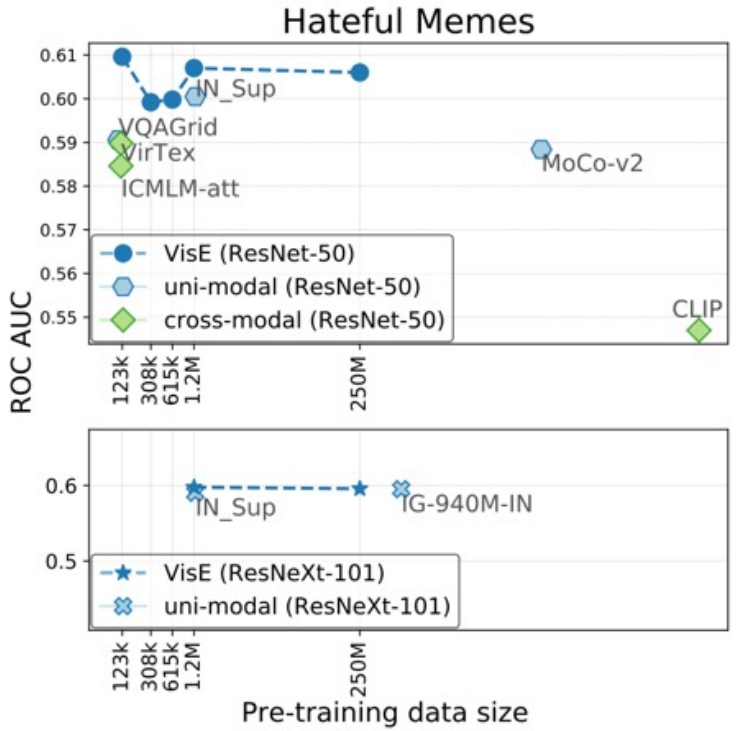
Jia M et al. ICCV 2021



(a) UnbiasedEmotion.



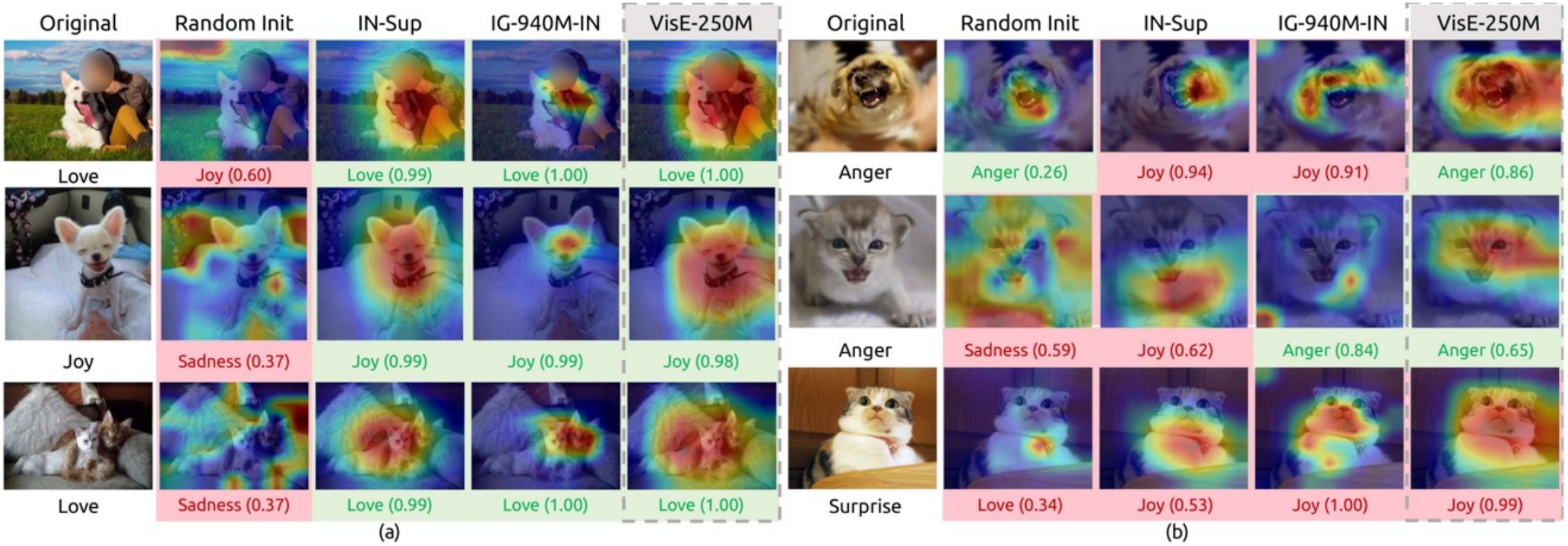
(b) Politics.



(c) Hateful Memes.

面向主观识别的特征学习

Jia M et al. ICCV 2021



谢谢!