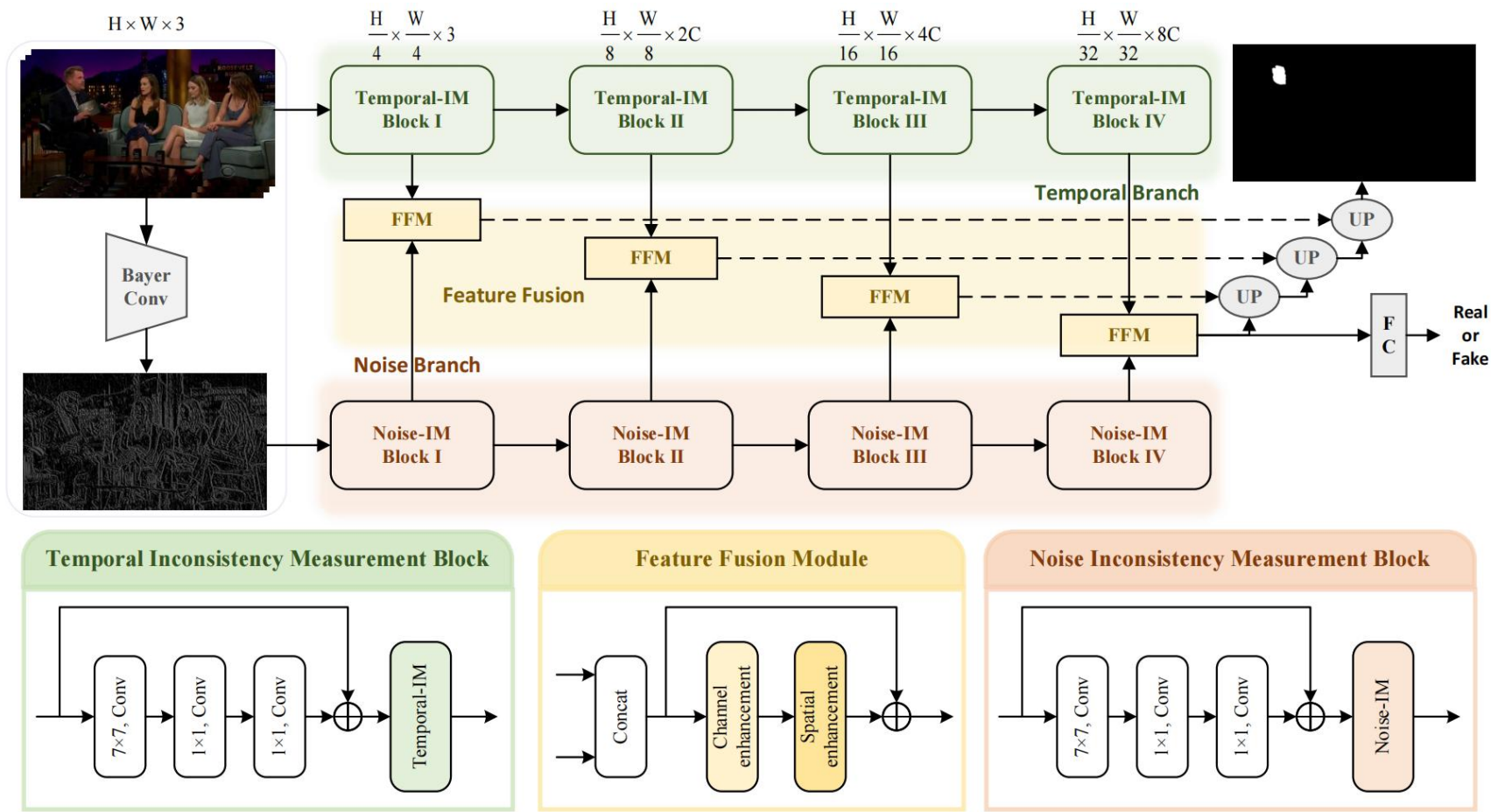


Deepfake Detection and Localization Using Multi-View Inconsistency Measurement

24年 TDSC (未开源)

- 1. 目标：视频，伪造检测及空间定位
- 2. 方法：MVIM：噪声不一致性测量（Noise-IM）模块 + 时间不一致性测量（Temporal-IM）模块 + 特征融合模块（FFM）

- ①Noise-IM 通过掩码注意力机制利用噪声不一致性信息来指导网络定位篡改区域；
- ②Temporal-IM 使用自注意力和双向卷积，捕捉帧间不一致，性提高定位和分类性能



Methods	Localization		Classification	
	F1	IoU	ACC	AUC
FFD	71.71	67.02	86.62	86.53
D&L	51.46	46.67	90.32	90.34
M2TR	65.32	61.19	96.79	96.80
SFFs	92.18	90.65	94.74	94.75
SDIML	94.17	92.68	93.65	93.61
LVNet	68.57	61.52	93.86	93.86
MVIM(Ours)	96.38	94.83	99.63	99.62

在 FFIW 数据集上与近期检测+空间定位方法比较

Methods	ACC	AUC
Xception	93.52	93.51
PCL	90.92	91.93
GFFD	98.62	98.61
LDAM	90.07	90.05
DDLm	94.10	94.09
MINet	89.70	89.67
MVIM(Ours)	99.63	99.62

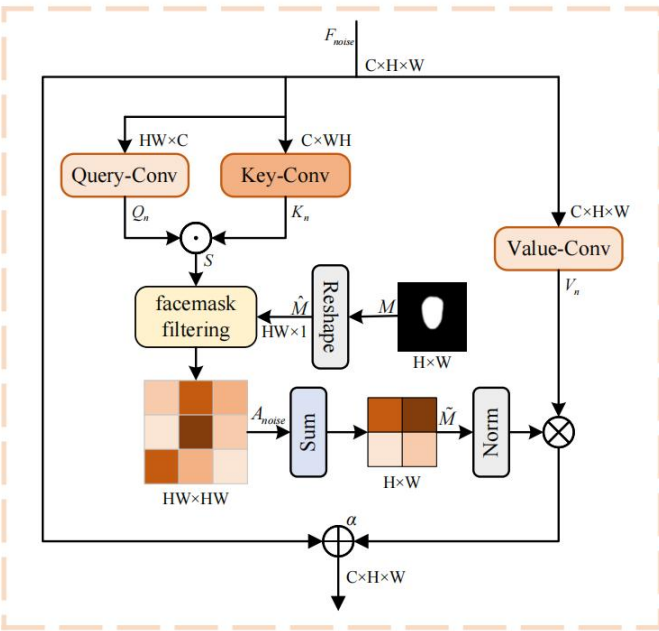
在 FFIW 数据集上与近期检测方法比较

Deepfake Detection and Localization Using Multi-View Inconsistency Measurement

24年 TDSC (未开源)

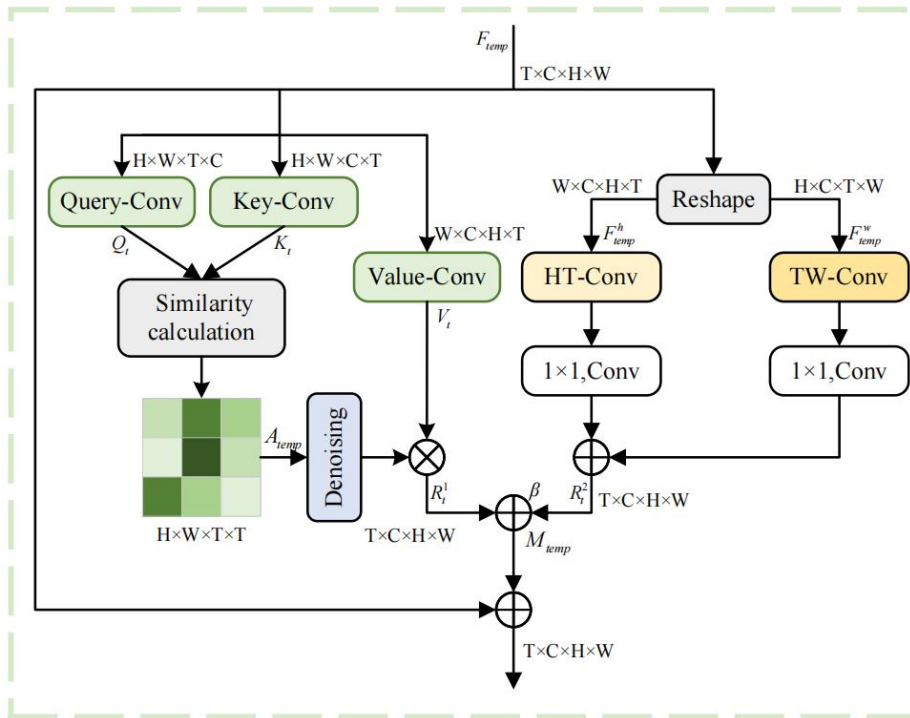
3. Noise-IM 噪声不一致性测量模块:

- ①将输入噪声特征映射到Q、K、V, 然后通过计算Q和K的负点积相似性来获得注意力图, 以测量不同位置之间的不一致性 $S = softmax(\frac{-Q_n K_n^T}{\sqrt{d}})$,
- ②在数据预处理阶段获得的人脸掩码M, 匹配S大小后添加到负点积中 $A_{noise} = softmax(\frac{\hat{M}' - Q_n K_n^T}{\sqrt{d}})$.
- ③按求和并将归一化 A_{noise} , 然后将其重塑为HxM, 更新噪声特征 $F_{noise} = \alpha \tilde{M} \otimes V_n + F_{noise}$,
- ④注意力图 \tilde{M} 在真实篡改掩码 M_{gt} 的监督下 $\mathcal{L}_{noise} = \mathcal{L}_{bce}(\tilde{M}, \bar{M}_{gt})$.



4. Temporal-IM 时间不一致性测量模块:

- ①对于要检测的帧 I_t , 我们利用其相邻帧 I_{t-1} 和 I_{t+1} 来提取时间不一致性信息
- ②将时间特征映射到Q、K, 计算负余弦相似性得分, 学习相邻帧之间的一致程度, 经过去噪后与V相乘, 得到 $R_t^1 = f_{denoise}(m_{cos}(Q_t, K_t))V_t$,
- ③为从不同角度提取帧间差异, 将时间特征重塑为两个坐标方向的表示和, 然后沿高度和宽度方向进行卷积, 得到 $R_t^2 = f_{1 \times 1}(f_{ht}(F_{temp}^h)) + f_{1 \times 1}(f_{tw}(F_{temp}^w))$,
- ④更新时间特征 $F_{temp} = (\beta R_t^1 + (1 - \beta)R_t^2) + F_{temp}$,
- ⑤从不一致性特征 F_{temp} 中提取中间帧 M_{temp} , 表示帧间的不一致关系, 用真实掩码进行监督: $\mathcal{L}_{temp} = \mathcal{L}_{bce}(M_{temp}, M'_{gt})$.



24年 TDSC (未开源)

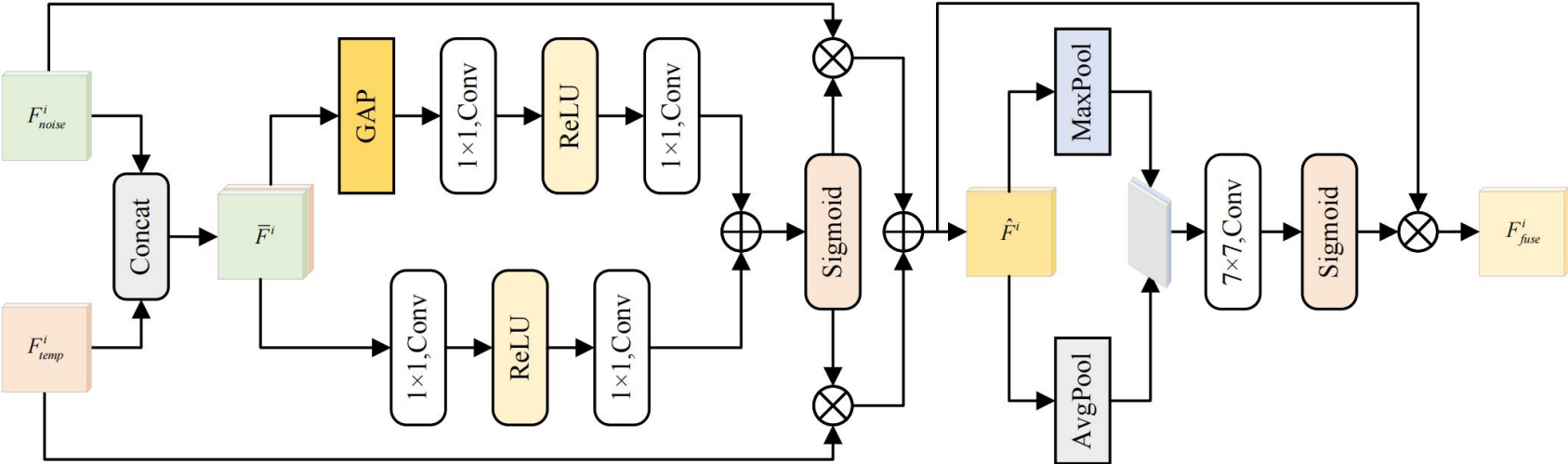
5. 特征融合模块 (FFM) :

①GAP表示全局平均池化;

②从全局和局部角度应用通道增强:

$$W = \sigma(f_{act}(GAP(\bar{F}^i) + f_{act}(\bar{F}^i))),$$
$$\hat{F}^i = W \otimes F_{noise}^i + (1 - W) \otimes F_{temp}^i,$$

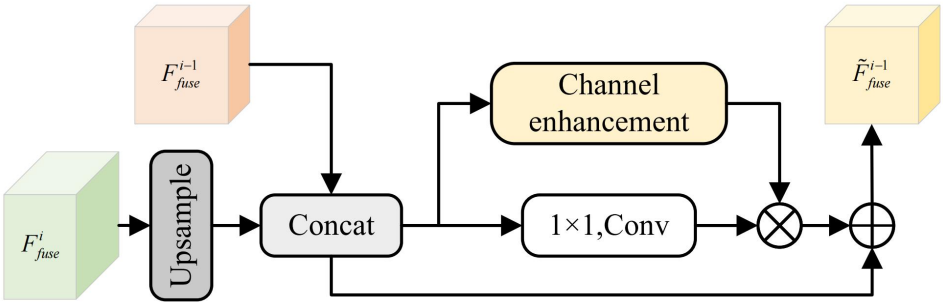
③得到 $F_{fuse}^i = \sigma(f_{7 \times 7}([f_{avg}(\hat{F}^i); f_{max}(\hat{F}^i)])) \otimes \hat{F}^i$,



6. 分类和定位, 引入分类损失和定位损失, 定位的细节暂略 (右下图) , 总损失函数为 $\mathcal{L} = \lambda_1 \mathcal{L}_{noise} + \lambda_2 \mathcal{L}_{temp} + \lambda_3 \mathcal{L}_{cls} + \lambda_4 \mathcal{L}_{loc}$,

7. 消融实验结果 (在FFIW数据集上)

Module settings			Localization		Classification	
Noise-IM	Temporal-IM	FFM	F1	IoU	ACC	AUC
✓	✓	✓	92.64	91.07	95.62	95.61
			94.47	93.04	96.82	96.81
			94.61	93.23	96.99	96.98
✓	✓	✓	93.69	92.26	96.20	96.19
			95.02	93.49	98.55	98.54
			95.16	93.71	98.99	98.98
✓	✓	✓	95.47	93.97	98.50	98.49
			96.38	94.83	99.63	99.62



Module settings	Localization		Classification	
	F1	IoU	ACC	AUC
Noise-IM w/o facemask filtering	95.47	93.97	99.26	99.26
Temporal-IM w/o denoising	96.11	94.30	99.54	99.53
Temporal-IM w/o bi-direction conv	95.93	94.20	99.38	99.38
MVIM(Ours)	96.38	94.83	99.63	99.62