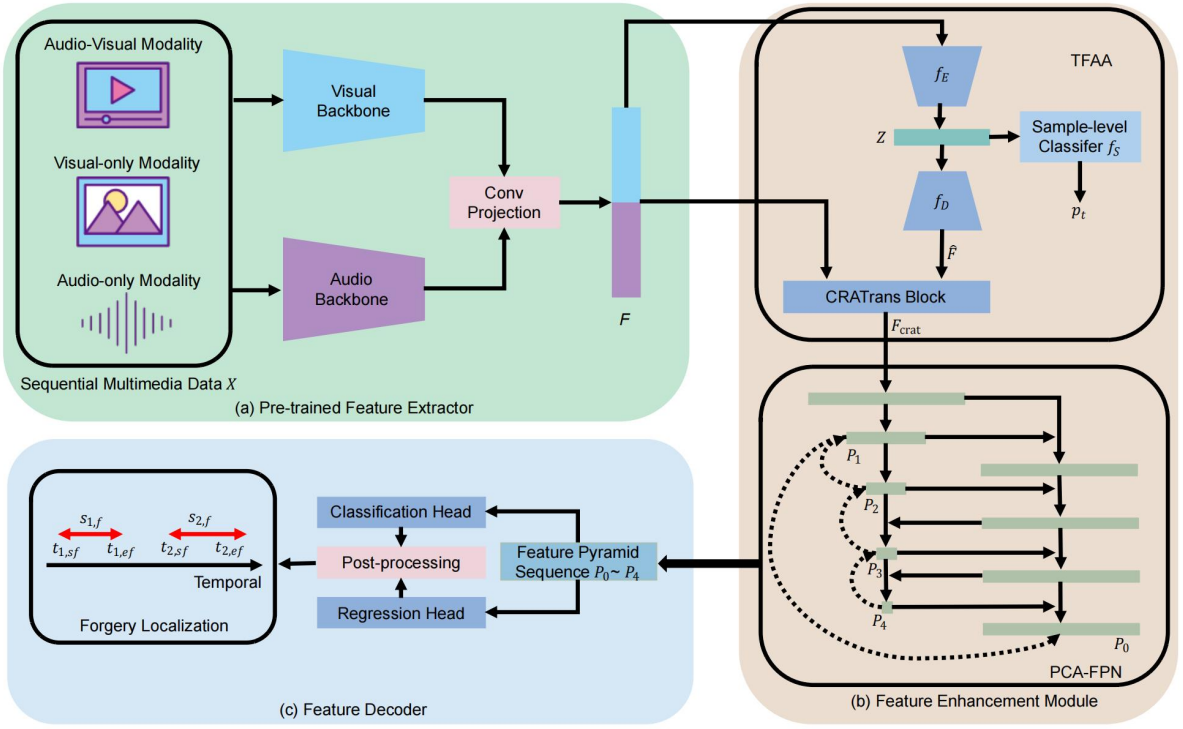
UMMAFormer: A Universal Multimodal-adaptive Transformer Framework For Temporal Forgery Localization

**23年 ACM MM（开源）**

1. 目标：多模态自适应，伪造检测及时间定位
2. 数据集：TVIL（v，自建，开源），LAV-DF（v+a），Psynd（a）
3. 方法：UMMAFormer：预训练模型提取特征+特征增强+解码器预测

   ①特征提取：时序分割网络（TSN）和自监督学习音频模型（BYOL - A）作为处理视觉和音频数据的骨干网络；

   ②两个模块：时间伪造注意力分析（TFAA）模块和并行交叉注意力特征金字塔网络（PCA - FPN）




| Methods | Feature | LAV-DF Full Set | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | AP@0.5 | AP@0.75 | AP@0.95 | AR@10 | AR@20 | AR@50 | AR@100 |
| MDS [11] | Visual | 12.78 | 1.62 | 0.00 | 37.88 | 36.71 | 34.39 | 32.15 |
| AGT [40] | Visual | 17.85 | 9.42 | 0.11 | 43.15 | 34.23 | 24.59 | 16.71 |
| BMN [36] | Visual | 24.01 | 7.61 | 0.07 | 53.26 | 41.24 | 31.60 | 26.93 |
| BMN (I3D) [36] | Visual | 10.56 | 1.66 | 0.00 | 48.49 | 44.39 | 37.13 | 31.55 |
| AVFusion [3] | Visual+Audio | 65.38 | 23.89 | 0.11 | 62.98 | 59.26 | 54.80 | 52.11 |
| BA-TFD [6] | Visual | 58.55 | 28.60 | 0.16 | 62.49 | 58.77 | 53.86 | 50.29 |
| | Visual+Audio | 76.90 | 38.50 | 0.25 | 66.90 | 64.08 | 60.77 | 58.42 |
| ActionFormer [59] | Visual | 95.34 | 90.20 | 23.73 | 88.41 | 89.63 | 90.33 | 90.41 |
| Ours | Visual | 97.30 | 92.96 | 25.68 | 90.19 | 90.85 | 91.14 | 91.18 |
| | Visual+Audio | **98.83** | **95.54** | **37.61** | **92.10** | **92.42** | **92.47** | **92.48** |

| Methods | TVIL AP@0.5 | AP@0.75 | AP@0.95 | AR@10 | AR@20 | AR@50 | AR@100 |
|---|---|---|---|---|---|---|---|
| TAGS [39] | 18.40 | 12.68 | 0.09 | 24.41 | 25.05 | 25.56 | 25.56 |
| DCAN [8] | 82.75 | 75.00 | 3.22 | 64.73 | 66.02 | 68.82 | 69.97 |
| ActionFormer [59] | 86.27 | 83.03 | 28.17 | 84.82 | 85.77 | 88.10 | 88.49 |
| Ours | **88.68** | **84.70** | **62.43** | **87.09** | **88.21** | **90.43** | **91.16** |

| Methods | test set | special test set | landline | cellular |
|---|---|---|---|---|
| LFSS [58] | 98.58 | **99.35** | 80.29 | 80.94 |
| Ours | **98.70** | 98.24 | **92.04** | **86.57** |

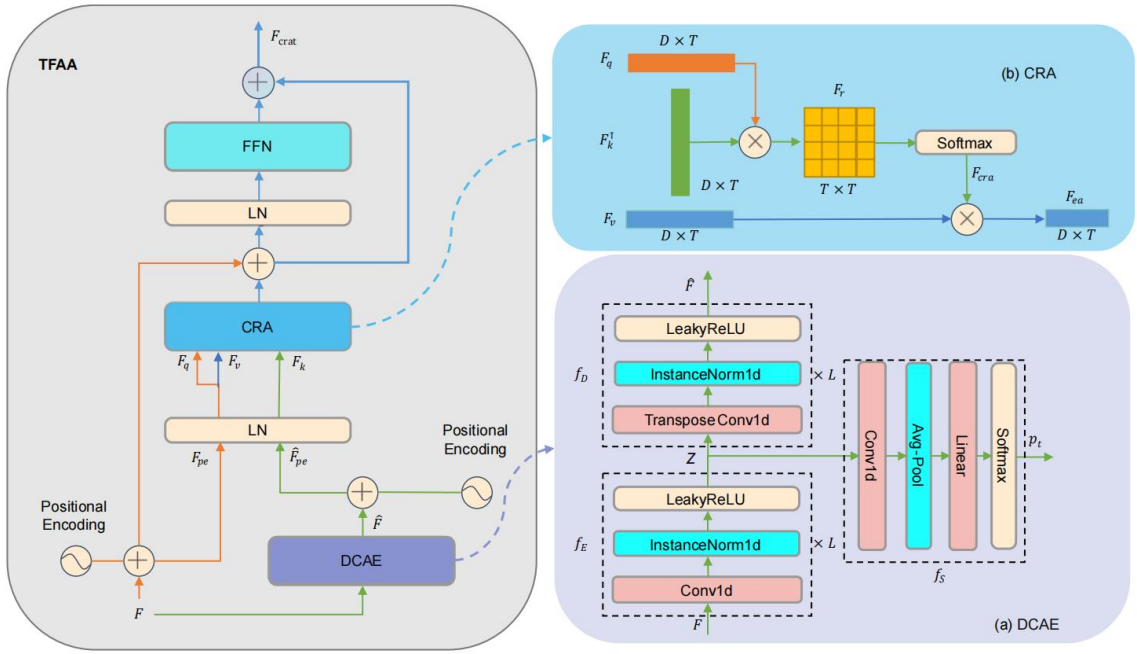
UMMAFormer: A Universal Multimodal-adaptive Transformer Framework For Temporal Forgery Localization

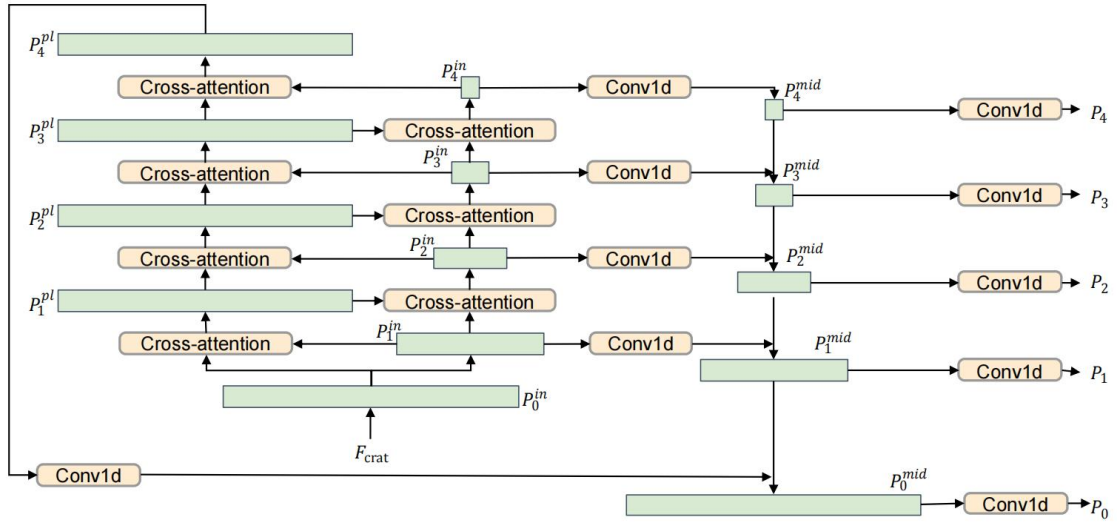
**23年 ACM MM（开源）**

4. 两个模块：时间伪造注意力分析（TFAA）模块和并行交叉注意力特征金字塔网络（PCA - FPN）

①TFAA：深度卷积自编码器 (DCAE)【重建损失+分类损失】，交叉重建注意力Transformer（CRA） 增强对时间差异的检测

②PCA - FPN：下采样（融合交叉注意力机制）+上采样，增强微妙特征




(c) PCA-FPN

**TFAA**

| Dataset | Methods | AP@0.5 | AP@0.75 | AP@0.95 | AR@10 | AR@20 | AR@50 | AR@100 |
|---|---|---|---|---|---|---|---|---|
| Lav-DF Full Set | Baseline | 97.58 | 93.75 | 40.38 | 92.23 | 92.71 | 92.87 | 92.90 |
| | Baseline+TFAA | 97.57 | 93.74 | **40.53** | **92.31** | **92.80** | **92.98** | **92.99** |
| | Baseline+PCA-FPN+TFAA (ours) | **98.83** | **95.54** | 37.61 | 92.10 | 92.42 | 92.47 | 92.48 |
| TVIL | Baseline | 86.10 | 82.86 | 28.11 | 84.68 | 85.71 | 88.04 | 88.43 |
| | Baseline+TFAA | 85.82 | 83.23 | 51.71 | 86.32 | 87.48 | 89.31 | 89.55 |
| | Baseline+PCA-FPN+TFAA (ours) | **88.68** | **84.70** | **62.43** | **87.09** | **88.21** | **90.43** | **91.16** |
| Psynd-Test | Baseline | **100.00** | **100.00** | 71.08 | 95.95 | 95.95 | 95.95 | 95.95 |
| | Baseline+TFAA | **100.00** | 98.41 | 76.23 | 97.09 | 97.09 | 97.09 | 97.09 |
| | Baseline+PCA-FPN+TFAA (ours) | **100.00** | **100.00** | **79.87** | **97.60** | **97.60** | **97.60** | **97.60** |

**FPN对比PCA-FPN**

| Dataset | Methods | AP@0.5 | AP@0.75 | AP@0.95 | AR@10 | AR@20 | AR@50 | AR@100 |
|---|---|---|---|---|---|---|---|---|
| Lav-DF Full Set | Baseline | 97.58 | 93.75 | 40.38 | 92.23 | 92.71 | 92.87 | 92.90 |
| | Baseline+FPN | **98.84** | **95.61** | 38.63 | 92.30 | 92.59 | **92.65** | **92.66** |
| | Baseline+PCA-FPN | 98.72 | 95.52 | **39.00** | **92.31** | **92.60** | **92.65** | **92.66** |
| TVIL | Baseline | 86.10 | 82.86 | 28.11 | 84.68 | 85.71 | 88.04 | 88.43 |
| | Baseline+FPN | 88.50 | 84.35 | 38.95 | **85.91** | 87.26 | **89.63** | **90.09** |
| | Baseline+PCA-FPN | **88.57** | **84.82** | **40.37** | 85.56 | **87.44** | 89.53 | 89.78 |
| Psynd-Test | Baseline | **100.00** | **100.00** | 71.08 | 95.95 | 95.95 | 95.95 | 95.95 |
| | Baseline+FPN | 43.28 | 5.13 | 0.11 | 47.22 | 48.48 | 48.86 | 48.86 |
| | Baseline+PCA-FPN | **100.00** | 98.54 | **77.72** | **97.34** | **97.34** | **97.34** | **97.34** |