TinyVIRAT: Low-Resolution Video Action Recognition

CENTER FOR RESEARCH IN COMPUTER VISION

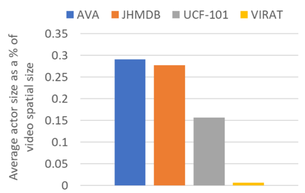
Ugur Demir
ugur@knights.ucf.edu

Yogesh S Rawat
yogesh@crcv.ucf.edu

Mubarak Shah
shah@crcv.ucf.edu

Motivation

- Action recognition in low resolution videos
  - Realistic low-resolution videos
  - **TinyVIRAT** benchmark
- Most of the existing action recognition datasets contain high resolution actor centric videos
- Downsampling does not reflect real-world scenario



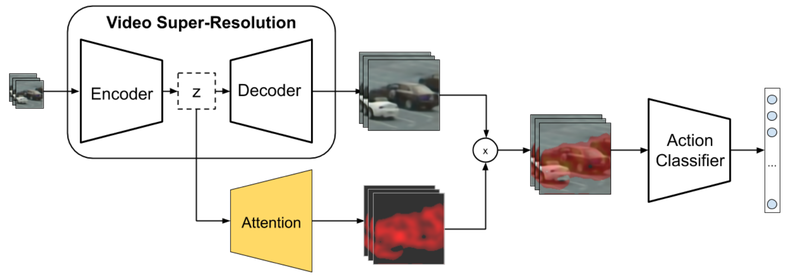
TinyVIRAT

- Collected from VIRAT dataset for real-life tiny action recognition problems
- Natural low-resolution actions
- Large variety of different actor sizes and multi-label actions

Size	Actions	Frames
20x20	standing	
28x28	carrying walking	
34x34	using tool walking	
38x38	carrying walking	
40x40	talking standing	
42x42	talking phone walking	
44x44	carrying	
48x48	walking	
52x52	moving riding	
58x58	pull walking	
76x76	vehicle moving	

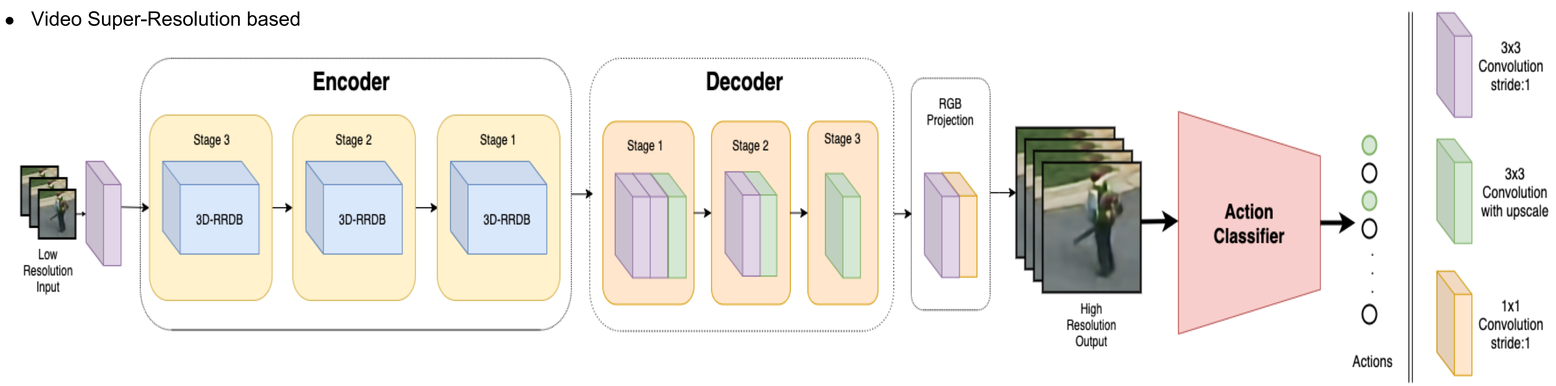
Weakly-Supervised Foreground Attention

- Video Super-Resolution based

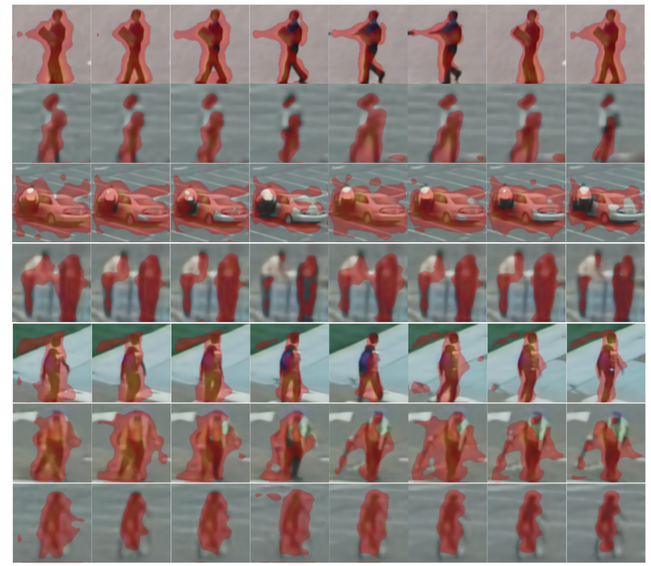


Progressive Video SR for Action Recogniton

- Video Super-Resolution based



Attention Visualization



Results on TinyVIRAT

Method	F1-Score
I3D	28.73
I3D + Prog. DVSR	32.55
I3D + Prog. DVSR + Att.	34.49
ResNet-50	29.08
ResNet-50 + Prog. DVSR	29.81
ResNet-50 + Prog. DVSR + Att.	30.80
WideResNet	32.66
WideResNet + Prog. DVSR	34.05
WideResNet + Prog. DVSR + Attn.	35.07

Ablation

Method	F1-Score
w/o DVSR	28.73
DVSR	30.45
Progressive DVSR	32.55
Progressive DVSR + Attention	34.49