

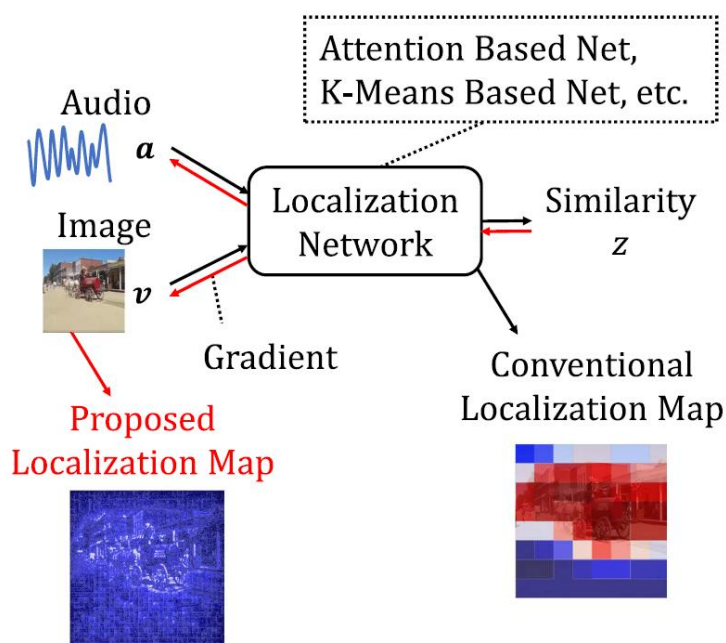
(PS T1.3, 1783) Unsupervised Sound Source Localization From Audio-Image Pairs Using Input Gradient Map

Tomohiro Tanaka, Takahiro Shinozaki
Tokyo Institute of Technology



- We proposed an input gradient-based unsupervised sound source localization method
- It has a **low computational cost** and **high resolution**, and **can be combined with any existing methods**
- We obtained **consistent improvement** by **combining** it with conventional methods

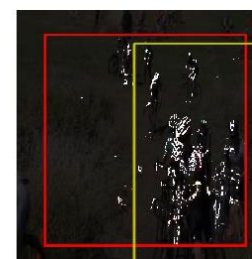
| Model | AUC | p-value |
|-------------------|--------------|----------|
| Baseline†(no box) | (56.8%) | N/A |
| Baseline(no box) | 56.0% | N/A |
| Baseline(box) | 57.4% | N/A |
| Proposed(BP) | 55.4% | .062 |
| Proposed(GBP) | 56.4% | .349 |
| Combined(BP) | 59.7% | <.001*** |
| Combined(GBP) | 59.9% | <.001*** |



Original



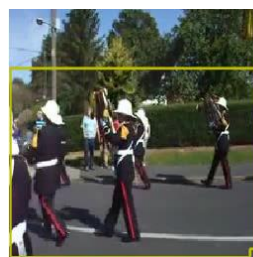
Baseline



Proposed



Combined



Original



Baseline



Proposed



Combined