**1. Motivation**
- Frame reconstruction is critical in applications like retrieving missing frames in surveillance videos, anomaly detection, data compression, video editing, video post-processing, animation, spoofing and so on.
- When multiple frames are missing and adjacent frames within the camera are far apart, realistic coherent frames can still be reconstructed using corresponding frames from other overlapping cameras.

**2. Contributions**
- We tackle a novel problem of frame reconstruction in multi-camera scenario using an adversarial approach.
- We perform extensive experiments on a challenging multi-camera video dataset to show the effectiveness of our method and on a single-camera video dataset to provide qualitative comparison with the state-of-the-art.

**3. Solution Overview**
- We learn the representations of the missing frame conditioned on the preceding and following frames within the camera and on the corresponding frames in other overlapping cameras using CGAN.
- These representations are merged together using a weighted average where the weights are chosen by maximizing the average PSNR on a smaller validation set.

**4. Network Architecture**
- “U-Net”-based architecture of the generator with skip connections which directly connect encoder layers to decoder layers.
- The discriminator tries to differentiate at patch-level and runs convolutionally across the image to generate an averaged output.

**5. Model Training Approach**
- We use a combination of J1 loss and adversarial loss in the objective function.
- We alternate between a gradient descent step upon D and one upon G and the training maximizes $\log D(x, G(x, z))$.
- To optimize the network, we use a minibatch stochastic gradient descent with an adaptive subgradient method (Adam) and a learning rate of 0.0002.

**6. Datasets and Experimental Results**
- **KTH Human Action Dataset:** Single-view dataset with 6 types of human activities
  - Proposed Method: 3.05 3.03 0.93
  - Table 1: Single-view Reconstruction Performance Comparisons for KTH Human Action Dataset
- **Office Lobby Dataset:** Multi-view dataset with 3 video clips captured by 3 cameras
  - Table 2: Multi-view Reconstruction Performance for Office Lobby Dataset

**7. Acknowledgements**
This work was partially supported by NSF grant 1544969 from the Cyber-Physical Systems program.