**VIDEO SUPER RESOLUTION WITH CONVOLUTION NEURAL NETWORK (CNN)**

**Introduction**
Propose a CNN that is trained on both the spatial and the temporal dimensions of videos to enhance their spatial resolution. Consecutive frames are motion compensated and used as input to a CNN that provides super-resolved video frames as output.

**Method**
- **Weight Transfer from pretraining:**
  - Ensure Reference and video SR networks conditions to be equal
  - Pretrained filters values from the reference model
- **Motion Compensation:**
  - Predict an image frame in a video, given the previous and future image frames by accounting for motion of the object in the video.

**Results**
- **SRCNN**
- **Bicubic**

---

**VIDEO FRAME INTERPOLATION (VFI)**

**What is VFI?**
Video frame interpolation (VFI) converts low frame rate (LFR) contents to high frame rate (HFR) videos by synthesizing one or more intermediate frames between given two consecutive frames.

**Applications**
- Slow motion generation
- Frame rate upconversion
- Frame recovery in video streaming
  - The videos with a high frame rate can avoid common artifacts, and therefore are visually more appealing to the viewers.

**Results**
- Where it fails – fast moving object interpolated
- Where it works – rigid object with small relative movement interpolated

---

**BLIND SUPER RESOLUTION**

**Introduction**
The blind super-resolution aims to solve the problem of super-resolution images which unknown degradations. The state-of-the-art super-resolution networks are trained using Bicubically downsampled i.e images that are downsampled using bicubic interpolation. This creates model bias and the networks get overfit towards this type of super-resolution. And when we try to deploy these trained networks in a real-life scenario, they fail because of this false assumption that all low-resolution images are downsampled using bicubic interpolation.

**Method**
There are several ways to solve this problem depending on different scenarios. The overall research is summarized in the taxonomy diagram shown on the right side. The most promising method out of these is degradation modeling-based methods. These methods can estimate degradation from any image and that is why these methods are the most promising.

**Model**
We present the state-of-the-art model for Blind super-resolution called KOALAnet. This model overcomes several disadvantages over the previous methods. It can handle complex degradations as well as to adapt to new and unseen degradations very well because of the dynamic kernels. The idea is to generate the actual degradation kernels instead of just the degradation representation unlike the previous methods. These dynamic kernels are used in both, degradation estimation network and upsampling network.

Reference: https://github.com/hjSim/KOALAnet

---

**References**
- [Video Super Resolution with ESTRNN](https://github.com/zzh-tech/ESTRNN)
- [KOALAnet](https://github.com/hjSim/KOALAnet)