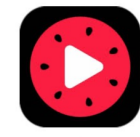
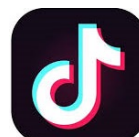


# JVET-X0081

## EE1-RELATED: CNN-BASED SUPER RESOLUTION FOR VIDEO CODING USING SEPARATE NETWORKS FOR CHROMA COMPONENTS

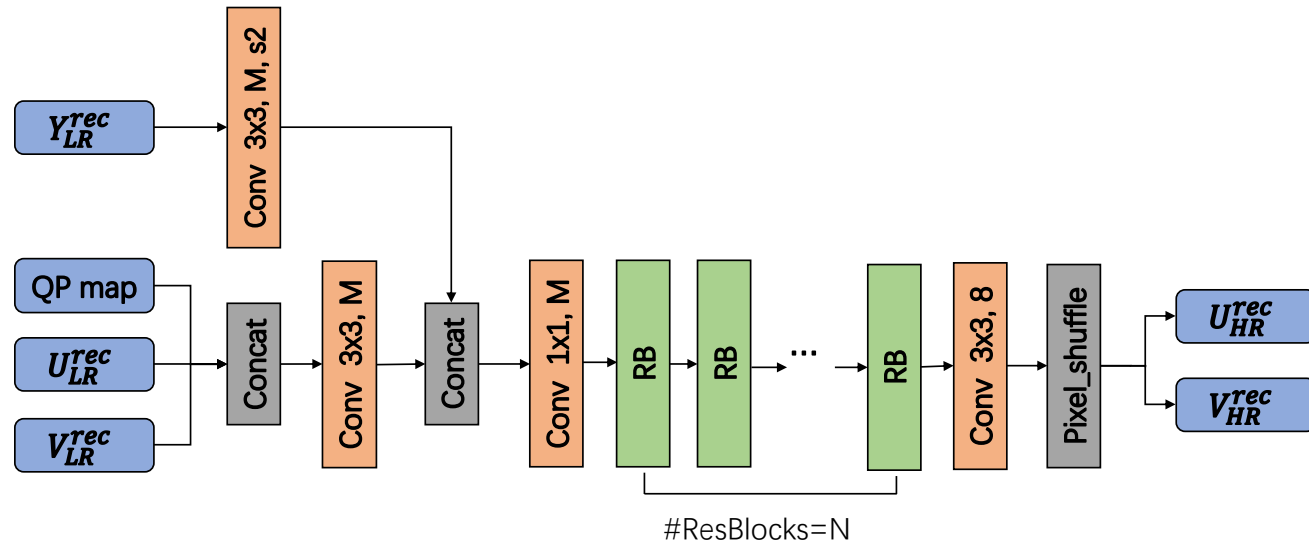
Chaoyi Lin, Yue Li, Kai Zhang, Li Zhang



# Introduction

## ■ Background

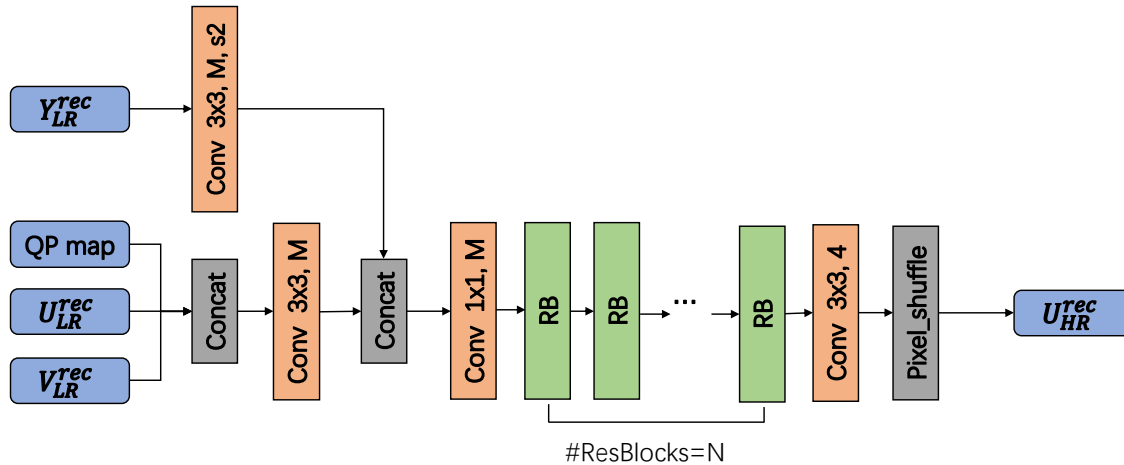
- In EE1-2.2, chroma up-sampling only requires a single model
- To achieve higher coding gains for chroma, this contribution utilizes two models to up-sample the U and V components, respectively.



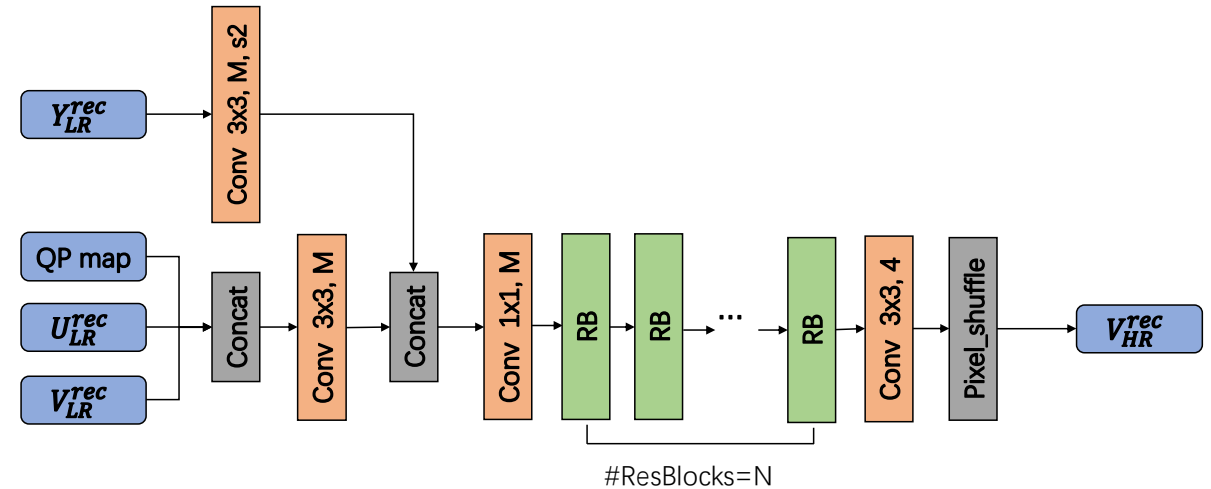
The chroma up-sampling model in EE1-2.2

# Proposed solution

- Using separate models for up-sampling the U and V components, respectively
  - One model for up-sampling U component (M=64, N=16)
  - One model for up-sampling V component (M=64, N=16)
  - The two models are used for both I-slice and B-slice



The U up-sampling network



The V up-sampling network

# Training details

- Training dataset: DIV2K, BVI-DVC
- Training data generation
  - QPs: {22, 27, 32, 37, 42}
- Number of models
  - 2 luma up-sampling models
  - 1 U up-sampling model
  - 1 V up-sampling model

| Network Information in Training Stage |  |  |
|---------------------------------------|--|--|
| Mandatory                             | GPU Type   | GPU: Tesla-V100-SXM2-32GB                    |
|                                       | Framework:   | PyTorch v1.8                                 |
|                                       | Number of GPUs per Task  | 1  |
|                                       | Epoch:   | 120  |
|                                       | Batch size:  | 16   |
|                                       | Training time:   | 64h/model                                    |
|                                       | Training data information:   | DIV2K, BVI DVC                               |
|                                       | Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0 + new MCTF, QP {22, 27, 32, 37, 42} |
|                                       | Loss function:   | L1   |
| Optional                              | Number of iterations   |  |
|                                       | Patch size   | 128×128                                      |
|                                       | Learning rate:   | 1e-4   |
|                                       | Optimizer:   | ADAM   |
|                                       | Preprocessing:   |  |
|                                       | Other information:   |  |

# Inference details

- The number of parameters of the largest model: 2.7M
- Total parameter number: 6.95M
- MAC: 854 kMAC/pixel

| <u>Network Information in Inference Stage</u> |                                   |  |
|---|-----------------------------------|--|
| Mandatory                                     | HW environment:                   | Intel(R) Xeon(R) Platinum 8260 CPU @ 2.40GHz   |
|   | Framework:                        | LibTorch v1.8.1  |
|   | Number of GPUs per Task           | 0  |
|   | Number of Parameters (Each Model) | luma up-sampling model for I slice: 2.7 M<br>luma up-sampling model for B slice: 1.37M<br>U up-sampling model: 1.44M<br>V up-sampling model: 1.44M |
|   | Total Parameter Number            | 6.95 M   |
|   | Parameter Precision (Bits)        | 32 (F)   |
|   | Memory Parameter (MB)             | 4 models in total: 27.80 MB  |
|   | Multiply Accumulate (MAC)         | 854 kMAC/pixel   |
|   |                                   |  |
| Optional                                      | Total Conv. Layers                | 34 for up-sampling the luma, 36 for up-sampling the chroma   |
|   | Total FC Layers                   | 0  |
|   | Batch size:                       | 1  |
|   | Patch size                        | Whole frame  |

# Simulation results

- Anchor: VTM-11.0 + new MCTF
- QP for anchor and test: 22, 27, 32, 37, 42

|                |               | Random Access Main10                             |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 22, 27, 32, 37, 42) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -9.37%   | -15.03% | -14.73% | 128% | 29%  |
|                | FoodMarket4   | -8.02%   | -5.71%  | -7.54%  |      |      |
|                | Campfire      | -18.95%  | 33.33%  | -15.31% |      |      |
| Class A2<br>4K | CatRobot1     | -0.21%   | -12.98% | -5.34%  | 117% | 28%  |
|                | DaylightRoad2 | 8.89%  | -18.54% | -2.46%  |      |      |
|                | ParkRunning3  | -9.23%   | 34.11%  | 7.41%   |      |      |
| Average on A1  |               | -12.11%  | 4.20%   | -12.53% | 128% | 29%  |
| Average on A2  |               | -0.18%   | 0.86%   | -0.13%  | 117% | 28%  |
| overall        |               | -6.15%   | 2.53%   | -6.33%  | 123% | 29%  |

|                |               | All Intra Main10                                 |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 22, 27, 32, 37, 42) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -11.14%  | -14.71% | -15.68% | 252% | 31%  |
|                | FoodMarket4   | -5.95%   | -3.21%  | -4.28%  |      |      |
|                | Campfire      | -16.20%  | 132.42% | -11.85% |      |      |
| Class A2<br>4K | CatRobot1     | -9.64%   | -14.87% | -11.70% | 150% | 32%  |
|                | DaylightRoad2 | -4.39%   | -20.96% | -9.60%  |      |      |
|                | ParkRunning3  | -12.94%  | 10.72%  | -3.12%  |      |      |
| Average on A1  |               | -11.10%  | 38.17%  | -10.60% | 252% | 31%  |
| Average on A2  |               | -8.99%   | -8.37%  | -8.14%  | 150% | 32%  |
| overall        |               | -10.04%  | 14.90%  | -9.37%  | 201% | 32%  |

## JVET-X0064, EE1-2.2

|                |               | Random Access Main10                             |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 22, 27, 32, 37, 42) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -9.37%   | -22.03% | -13.07% | 110% | 29%  |
|                | FoodMarket4   | -8.02%   | -3.68%  | -4.67%  |      |      |
|                | Campfire      | -18.95%  | 40.95%  | -11.12% |      |      |
| Class A2<br>4K | CatRobot1     | -0.21%   | -14.75% | -3.56%  | 100% | 29%  |
|                | DaylightRoad2 | 8.89%  | -17.76% | -1.43%  |      |      |
|                | ParkRunning3  | -9.23%   | 50.77%  | 19.96%  |      |      |
| Average on A1  |               | -12.11%  | 5.08%   | -9.62%  | 110% | 29%  |
| Average on A2  |               | -0.18%   | 6.09%   | 4.99%   | 100% | 29%  |
| overall        |               | -6.15%   | 5.58%   | -2.31%  | 105% | 29%  |

|                |               | All Intra Main10                                 |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 22, 27, 32, 37, 42) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -11.14%  | -19.08% | -14.26% | 215% | 31%  |
|                | FoodMarket4   | -5.95%   | -1.98%  | -2.11%  |      |      |
|                | Campfire      | -16.20%  | 159.49% | -5.58%  |      |      |
| Class A2<br>4K | CatRobot1     | -9.64%   | -15.79% | -10.61% | 131% | 32%  |
|                | DaylightRoad2 | -4.39%   | -20.32% | -9.21%  |      |      |
|                | ParkRunning3  | -12.94%  | 20.33%  | 4.84%   |      |      |
| Average on A1  |               | -11.10%  | 46.14%  | -7.32%  | 215% | 31%  |
| Average on A2  |               | -8.99%   | -5.26%  | -5.00%  | 131% | 32%  |
| overall        |               | -10.04%  | 20.44%  | -6.16%  | 173% | 32%  |

# Simulation results (cont.)

- Anchor: VTM-11.0 + new MCTF
- QP for anchor and test: 27, 32, 37, 42, 47
- For RA configuration, {-9.73%, -2.32%, -7.44%} coding gain is achieved for {Y, U, V} , respectively

|                |               | Random Access Main10                             |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 27, 32, 37, 42, 47) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -11.84%  | -18.16% | -17.34% | 221% | 32%  |
|                | FoodMarket4   | -10.96%  | -5.58%  | -8.77%  |      |      |
|                | Campfire      | -18.75%  | 15.61%  | -16.97% |      |      |
| Class A2<br>4K | CatRobot1     | -5.70%   | -14.81% | -7.52%  | 194% | 31%  |
|                | DaylightRoad2 | -0.78%   | -19.15% | -0.31%  |      |      |
|                | ParkRunning3  | -10.32%  | 28.18%  | 6.30%   |      |      |
| Average on A1  |               | -13.85%  | -2.71%  | -14.36% | 221% | 32%  |
| Average on A2  |               | -5.60%   | -1.93%  | -0.51%  | 194% | 31%  |
| overall        |               | -9.73%   | -2.32%  | -7.44%  | 208% | 32%  |

|                |               | All Intra Main10                                 |         |         |      |      |
|----------------|---------------|--|---------|---------|------|------|
|                |               | Over VTM-11.0 + new MCTF (QP 27, 32, 37, 42, 47) |         |         |      |      |
|                |               | Y  | U       | V       | EncT | DecT |
| Cass A1<br>4K  | Tango2        | -12.54%  | -17.21% | -16.79% | 494% | 37%  |
|                | FoodMarket4   | -6.85%   | -1.72%  | -4.28%  |      |      |
|                | Campfire      | -19.70%  | 94.43%  | -14.68% |      |      |
| Class A2<br>4K | CatRobot1     | -11.54%  | -16.03% | -11.89% | 288% | 40%  |
|                | DaylightRoad2 | -7.12%   | -20.48% | -7.35%  |      |      |
|                | ParkRunning3  | -14.16%  | 7.78%   | -2.70%  |      |      |
| Average on A1  |               | -13.03%  | 25.16%  | -11.92% | 494% | 31%  |
| Average on A2  |               | -10.94%  | -9.58%  | -7.31%  | 288% | 32%  |
| overall        |               | -11.99%  | 7.79%   | -9.61%  | 391% | 32%  |

# Conclusion

- This contribution presents a CNN-based super resolution method
  - Different networks are trained for up-sampling U and V components, respectively
  - It brings significant coding gains for U and V components
  - Specially, for anchor QPs {27, 32, 37, 42, 47}, the simulation results show {-9.73%, -2.32%, -7.44%} coding gains for {Y, U, V} under RA configuration, respectively
- Recommend to further study the proposed scheme