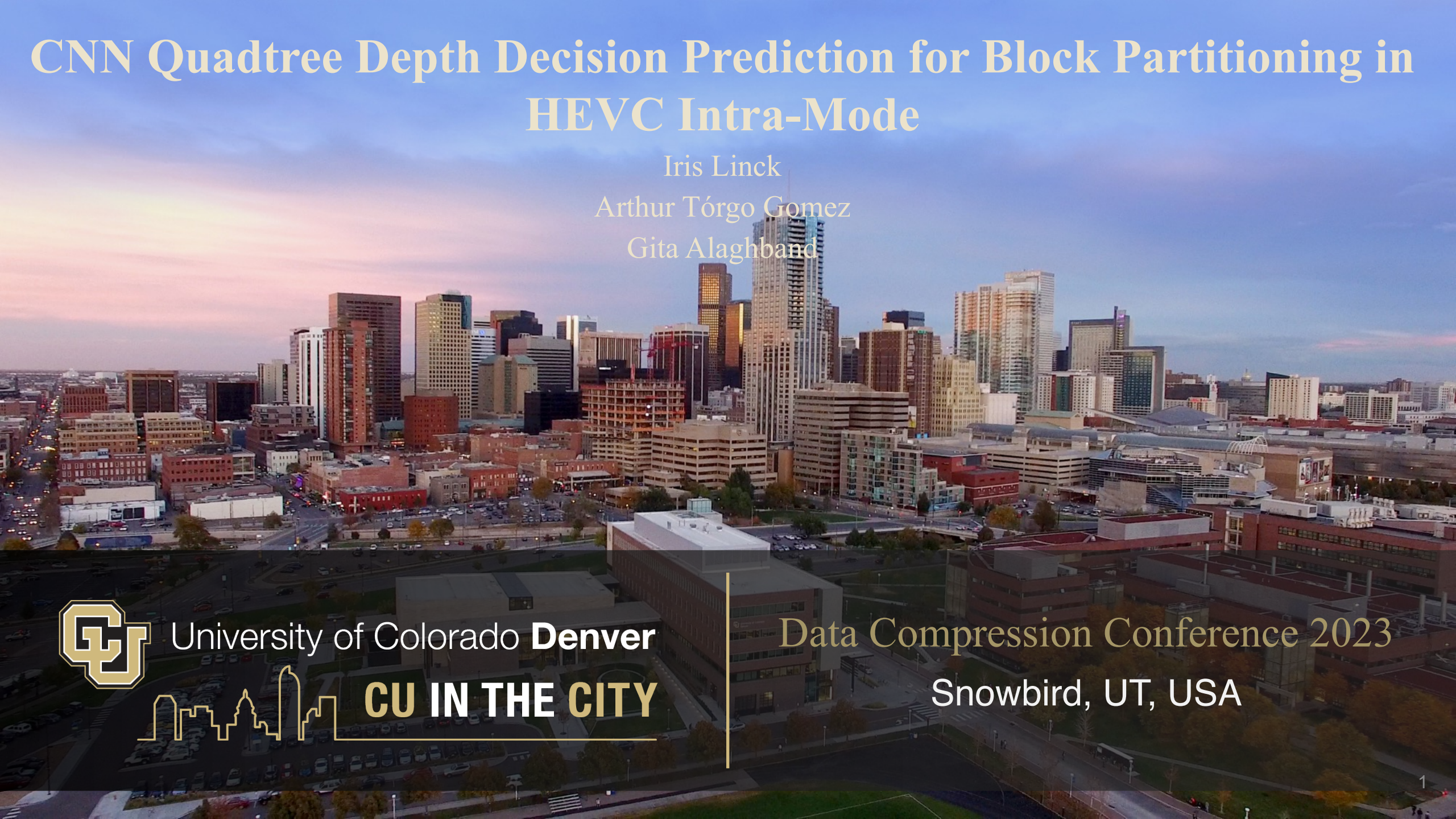
CNN Quadtree Depth Decision Prediction for Block Partitioning in HEVC Intra-Mode

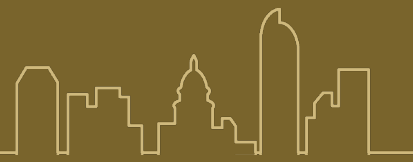
Iris Linck
Arthur Tórgo Gomez
Gita Alaghband



University of Colorado **Denver**
CU IN THE CITY

Data Compression Conference 2023
Snowbird, UT, USA

Agenda

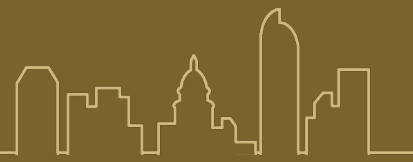


- Introduction
- Motivation
- CNN architecture based on VGGNet
- Results
- Conclusion

CNN Quadtree Depth Decision Prediction for Block Partitioning in HEVC Intra-Mode

Introduction

Introduction



- High Efficient Video coding (HEVC/H265) [1] improved compression rate around 40-50% bit rate reduction in comparison to its antecessor H.264/AVC.
- The high compression improvement of HEVC is mostly due to the new quad-tree partitioning of coding tree unit (CTU) at the expense of increasing coding complexity.
- We propose three CNNs based on VGGNet to predict quadtree levels for CTUs in Intra-Mode to reduce code complexity in HEVC.

CNN Quadtree Depth Decision Prediction for Block Partitioning in HEVC Intra-Mode

MOTIVATION

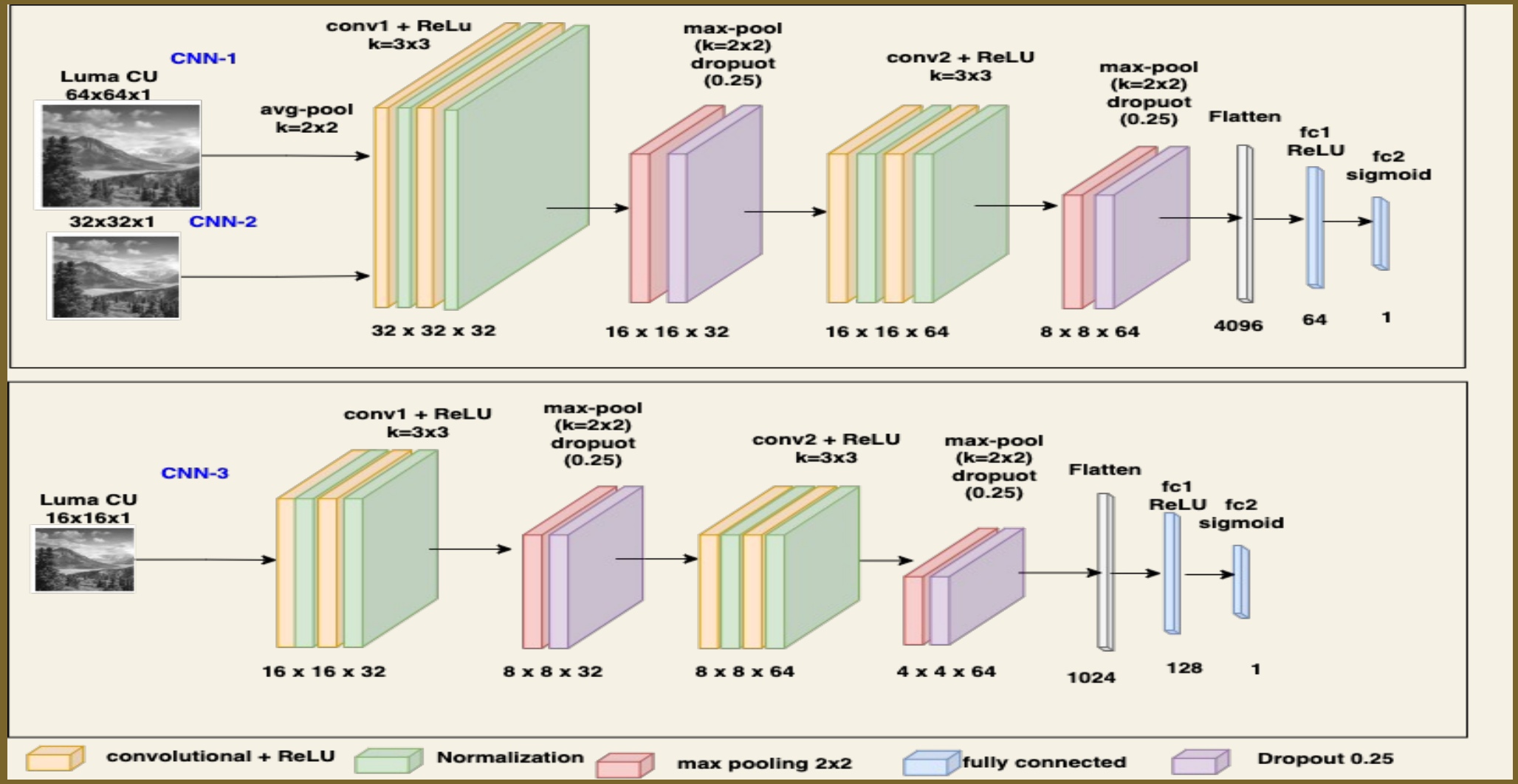
MOTIVATION

- Quad-tree seems to be a tendency in the next generations of video coding, and it was introduced for the first time in HEVC.
- An example of this tendency is showed in the new standard and a successor to HEVC, called Versatile Video Coding (VVC) [2] finalized in 2020, which inherits the Quadtree (QT) partitioning structure from HEVC.
- In this scenario, algorithms to improve quadtree partition structure will benefit the next generations of video coding.

CNN Quadtree Depth Decision Prediction for Block Partitioning in HEVC Intra-Mode

CNN architecture based on VGGN

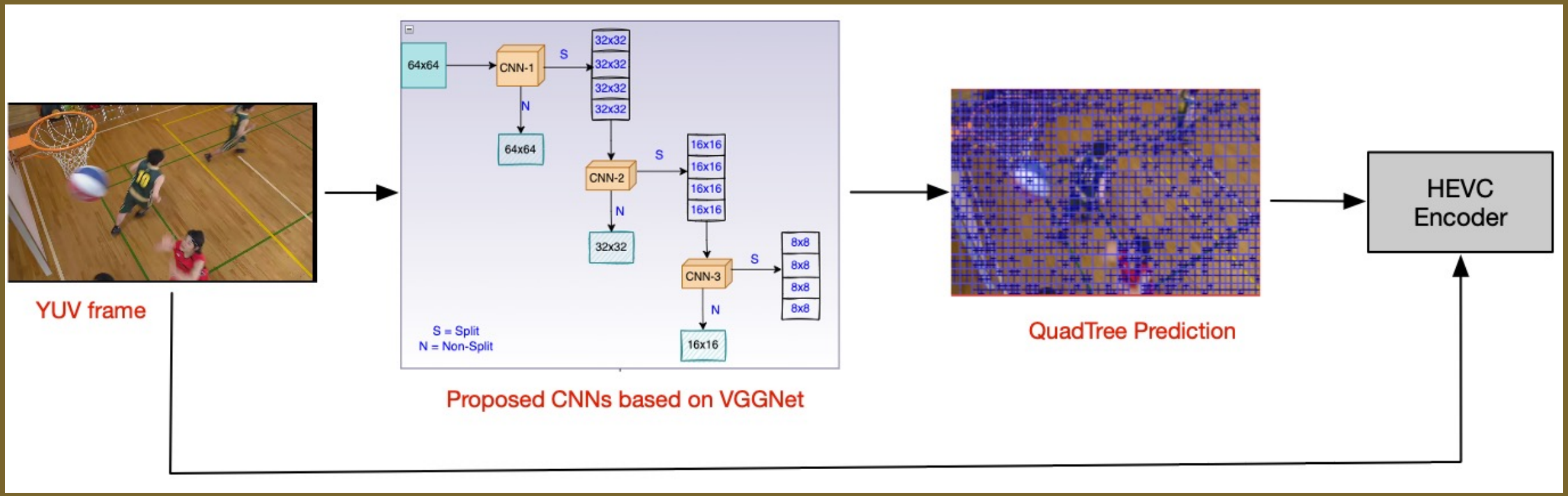
CNN Architecture based on VGGNet



Architecture

- Dataset:
  - Each CNN has its own CU dataset with a specific size encoded with a specific QP.
  - Each dataset is composed by CU blocks from frames of videos with different resolutions.
- CNN Model:
  - CNN-1 applies avg pool to the input CU_{64x64}
  - The models use binary cross entropy with Adam optimizer.
  - The conv layers use ReLU
  - The output layer is activated with sigmoid function

CNN prediction integrated to HEVC



Configuration:

- HM-16.5, which uses all-intra main configuration with QP value set to 32 and applied to videos of classes B, C and E.
- The complete quadtree prediction for each frame is generated and sent to HEVC as a text file.
- A threshold of 0.5 is applied to decide if a block may or may not split.

CNN Quadtree Depth Decision Prediction for Block Partitioning in HEVC Intra-Mode

Results

Results

Test Sequence		Ref. [7]		Ref.[8]		Ref. [5]		Proposed	
		BD-rate (%)	TS (%)	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)	BD-rate (%)	TS (%)
Class B 1920 x 1080	Kimono	1.73	78.40	1.93	69.20	1.49	59.53	0.95	84.67
	ParkScene	1.70	64.80	2.06	58.90	1.47	63.25	0.09	80.08
	Cactus	1.90	63.80	2.18	59.40	2.07	61.56	-1.22	77.69
	BQTerrace	1.90	57.20	1.48	63.30	1.09	67.97	1.21	72.42
	BaskballDrive	2.82	68.70	3.58	65.10	2.26	60.18	1.04	81.41
Class C 832x480	BaskballDrill	2.41	53.30	2.35	59.60	2.80	62.11	1.35	66.65
	BQMall	1.65	54.00	1.15	55.00	2.09	54.50	0.54	66.44
	RaceHorses	1.52	56.20	1.43	61.00	1.97	48.65	1.54	63.66
	PartyScene	0.49	41.40	1.07	57.20	1.01	49.00	-0.38	50.75
Class E 1280 x 720	FourPeople	2.71	65.00	3.51	65.70	1.83	67.89	0.89	78.49
	Johnny	3.16	72.20	3.42	66.30	1.69	71.34	4.71	83.56
	Kristen&Sara	2.68	71.4	2.91	69.5	1.55	72.19	2.49	78.52
Classes B, C, E Average		**2.06**	**62.20**	**2.26**	**62.52**	**1.78**	**61.51**	**1.10**	**73.69**

Conclusion

Conclusion

- Our method reduces the encoding time by 73.69% on average at the cost of 1.10% BD-rate increase and no significant loss in PSNR compared with HM16.5 reference software.
- On average, our proposed method outperformed the results presented by current machine learning approaches [3], [4] and [5], in terms of BD-BR and encoding time.

References

- [1] G. J. Sullivan, J. Ohm, H. Woo-Jin, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 22, no. 12, pp. 1649-1668, 2012, doi: 10.1109/TCSVT.2012.2221191.
- [2] M. Corrêa *et al.*, “AV1 and VVC Video Codecs: Overview on Complexity Reduction and Hardware Design,” *IEEE Open J. Circuits Syst.*, vol. 2, pp. 564–576, 2021, doi: 10.1109/OJCAS.2021.3107254.
- [3] A. Feng, C. Gao, L. Li, D. Liu, and F. Wu, "Cnn-Based Depth Map Prediction for Fast Block Partitioning in HEVC Intra Coding," in 2021 IEEE International Conference on Multimedia and Expo (ICME), 5-9 July 2021, pp. 1-6, doi: 10.1109/ICME51207.2021.9428069.

References

- [4] Y. Zhang, G. Wang, R. Tian, M. Xu, and C. C. J. Kuo, "Texture-Classification Accelerated CNN Scheme for Fast Intra CU Partition in HEVC," in 2019 Data Compression Conference (DCC), 26-29 March 2019, pp. 241-249, doi: 10.1109/DCC.2019.00032.
- [5] M. Xu, T. Li, Z. Wang, X. Deng, R. Yang, and Z. Guan, "Reducing Complexity of HEVC: A Deep Learning Approach," IEEE Transactions on Image Processing, vol. 27, no. 10, pp. 5044-5059, 2018, doi: 10.1109/TIP.2018.2847035.


University of Colorado Denver
CU IN THE CITY
Thanks