# IMAGE COMPRESSION USING NEURAL NETWORK BASED PREDICTION: A REVIEW

Ayshath Afra TP
VLSI & Signal Processing, LBS College of Engineering, Kasaragod
ayshathafratp@gmail.com

**Abstract—The proposed intra predictor is a set of neural networks, called Prediction Neural Networks Set (PNNS), based on both fully connected and convolutional neural network for intra image prediction. The choice of neural network predicting a given image block depends on the block size, hence does not need to be signaled to the decoder. The size of the image block to be predicted determines which neural network of PNNS is used for prediction and defines the size of the causal neighborhood of the block, so-called context, fed into the neural network. A small image block is predicted via a fully-connected neural network of PNNS whereas a large one is predicted via a convolution neural network of PNNS. Thanks to the use of masks of random sizes during training, the neural networks of PNNS well adapt to the available context that may vary, depending on the position of the image block to be predicted. The use of neural networks for intra prediction within an image coding scheme addressed in this paper. when integrated into H.265 codec, the proposed neural networks are shown to give rate-distortion performance gains compared with the H.265 intra prediction. PNNS can model alargese to fcomplex textures.**

**Index Terms—Intra prediction, Neural network, H.265, Image compression, PNNS**

## I. INTRODUCTION

NTRAframe prediction exploits spatial redundancy, i.e., correlation among pixels within one frame, by calculating prediction values through extrapolation from already coded pixels for effective delta coding. The goal of intra predictionis to infer a block of pixels from the previously encoded anddecoded neighborhoods. It plays an important role in imagecoding. Currently, the state-of-the-art image coding solutionis High Efficiency Video Coding (HEVC) still picture profile,published by Joint Collaborative Team on Video Coding(JCTVC). the H.265 standard which selects according toa rate-distortion criterion one mode among 35 fixed andsimple prediction functions. The H.265 prediction functions consist in simply propagating the pixel values along specified directions [1]. A hierarchical generative model which scalesto realistic full-sized image. It fails in large areas usually containing more complex textures [2][3]. Two new spatial image prediction methods based on neighbor- embedding algorithms is used [4] it exploits the self-similarities within the image with more complex models defined as linear combinations of k-nearest patches in the neighborhood. However rescaling doesn't improve results [5]. In [6] analysis ofthe Intra Predictions in H.265/HEVC, it does not necessarily take into account the nature of the compressible material, which may result in inefficient coding of the transmitted data. The pixel-level annotations needed for semantic segmentation [7], hence they are expensive to acquire. Instead of using traditional methods that handle each component separately, in [8] it directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. But the network doesn't have different up scaling factors. A gated Markov Random Field (MRF) approach is suitable in predicting the specific textures [9]. Deep residual learning for image recognition overfits during training if no training data augmentation is used [10]. The authors in [11] demonstrate template matching Intra prediction.

Yet there are no self-similarities within the image. In [12]very deep convolutional networks for large-scale image recognition is considered, which is feasible due to the use of very small(3 × 3) convolution filters in all layers.

Currently, the state-of-the-art image coding solution is High Efficiency Video Coding (HEVC) [13]. When it is compared with its predecessor H.264/AVC, HEVC can achieve about 17% bitrates saving in image coding. Note that neural networks have already been considered for intra block prediction. The authors in [14] only take into consideration blocks of sizes 4*4, 8*8, 16*16, and 32*32 pixels and use fully-connected neural networks. Recently intra prediction of convolutional neural network (IPCNN) is considered, this network can learn the spatial correlation between current PU block and its nearest reference blocks[15].

In this paper, considered both fully-connected and convolutional neural networks. We show that, while fully-connected neural networks give good performance for small block sizes, convolutional neural networks are more appropriate, both in terms of prediction PSNR and PSNR-rate performance gains, for large block sizes. The choice of neural network is block size dependent, hence does not need to be signaled to the decoder. This set of neural networks, called Prediction Neural Networks Set (PNNS), has been integrated into a H.265 codec, showing PSNR-rate performance gains from 1.46 % to 5.20 %.

### A. CONDITIONS FOR EFFICIENT NEURAL NETWORK BASED INTRA PREDICTION

Intra-prediction algorithms depend on previously encoded portions of an image to form an estimate. While thisprocess could theoretically involve all previously decoded pixels and blocks, it is often restricted to the bounding neighbours to reduce complexity. optimal prediction, i.e., conditional expectation [12] requires knowing the conditional distribution of the image block to be predicted given causal and distorted data samples. The authors demonstrate that the optimal prediction, i.e. the prediction that minimizes the mean squared prediction error, is the conditional expectation E[X—B]. Yet, no existing model of natural images gives a reliable E[X—B][20]. Classical approaches in predictive coding consist in proposing a set of predefined functions and choosing the best of them in a rate-distortion sense. Thus, the number of possible functions is limited. However, neural networks have proved capable of learning a reliable model of the probability of image pixels for prediction. On the other hand, neural networks can approximate many functions, in particular complex predictive functions such as the generation of future video frames given an input sequence of frames[14].

## II. RELATEDWORKS

### A. The Intra Prediction in HEVC

The intra prediction coding in HEVC follows a block- wise approach. Specifically, HEVC partitions each frame hierarchically into non-overlapping coding units (CU) and further into prediction units (PU). Then each PU will be predicted according to the nearest reference pixels and the best prediction mode within the same frame. For each PU to be coded, prediction blocks for the luminance and chrominance are created by the reconstructed pixels surrounding thecurrent PU block. The quality of the prediction will have a direct influence on the bitrate saving and PSNR of video sequences. After prediction coding, the residual signal will be achieved which is the difference between the original block and its prediction values. Transform and quantization process will be implemented continuously. Ultimately, the quantized coefficients will be fed to the entropy coding process directly. In HEVC, a set of 35 modes including 33 angular modes and a DC and a PLANAR mode is available. In intra prediction, the set of bounding pixels are used to generate the block's estimate.

### B. Convolutional Neural Network based Video Coding

Recently, The Convolutional neutral network (CNN) is used universally in many domains. Many trials has been done by using CNN for super-resolution (SR) , image classification and object detection . The CNN is well known to perform well for image processing and classification. However, the CNN is rarely used in video coding. CNN is used for Intra CU mode decision. In these papers, the input of the network is a 8x8 block and two classes are extracted at the last layer. In fact, it is a classification problem. The result of the network is used for decision that if this input

block will be split or not. CNN is also used in in-loop filtering for coding efficiency improvement. Besides, proposed a CNN-based approach for post-processing.
In HEVC intra coding. It is from AR-CNN which is used for compression artifacts reduction. However, these methods mentioned above are not used for intra prediction directly.

C. Deep Learning for Image/VideoCoding
Recently, deep learning has achieved remarkable progress in many research areas. At the same time, some investigations on image/video coding have been carried out. These work scan be classified into two categories. One category is out of the traditional block-based coding frameworks, and usually takes an end-to-end scheme, like the auto-encoder. In contrast, the other category is still based on traditional frameworks. It usually improves a sub-part, such as the prediction and post processing modules.

D. Overview of the technical features and characteristics of the HEVC
High Efficiency Video Coding (HEVC) is currently being prepared as the newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The main goal of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standards—in the range of 50 bit-rate reduction for equal perceptual video quality. The video coding layer of HEVC employs the same hybrid approach (inter-/intra picture prediction and 2-D transform coding) used in all video compression standards since H.261.HEVC represents a number of advances in video coding technology[24]. Its video coding layer design is based on conventional block-based motion compensated hybrid video coding concepts, but with some important differences relative to prior standards.

E. context-based pixel prediction
Unsupervised visual feature learning algorithm driven by context-based pixel prediction. By analogy with auto- encoders, we propose Context Encoders – a convolutional neural network trained to generate the contents of an arbitrary image region conditioned on its surroundings. In order to succeed at this task, context encoders need to both understand the content of the entire image, as well as produce a plausible hypothesis for the missing part(s). During training context encoders, experimented with both a standard pixel-wise reconstruction loss, as well as a reconstruction plus an adversarial loss. The latter produces much sharper results because it can better handle multiple modes in the output. context encoders trained to generate images conditioned on context advance the state of the art in semantic in painting, at the same time learn feature representations that are competitive with other models trained with auxiliary supervision.

F. Intra prediction using video compression standard H.265/HEVC
Generally, HEVC dataflow, briefing of the intra prediction mechanism and detailed illustration of service information transmission part[17]. In this paper its present statistics of using each intra prediction mode and statistics of modes becoming part of the most probable mode array obtained inthe experiments.. In HEVC this feature is often violated in all experiment's elements MPM [1] and MPM [2] were rarer than the Outside MPM situation. A more flexible approach to the selection of the array of the most probable modes by, perhaps, taking into account the nature of video material, could be the subject of further development of this area of video encoding.

G. Intra prediction using fully connected network
A deep learning method for intra prediction. Different from traditional methods utilizing some fixed rules, we propose using a fully connected network to learn an end-to-end mapping from neighbouring reconstructed pixels to the current block. Here, the network is fed by multiple reference lines. Firstly, introduce the architecture of the proposed intra prediction using fully connected network(IPFCN)[14].
The experimental results demonstrate its effectiveness. When compared with HEVC, an average of 3.4% bitrate saving is achieved. For 4K sequences, achieved 4.5% bitrate saving on average, and up to 7.4%. The proposed method can handle various bitrate settings without retraining the model.

H. Intra prediction convolutional neural network(IPCNN)

An intra prediction convolutional neural network (IPCNN) is proposed for intra prediction, which exploits the rich context of the current block and therefore is capable of improving the accuracy of predicting the current block. This is the paper that directly applies CNNs to intra prediction for HEVC.Experimental results validate the effectiveness of applying CNNs to intra prediction and achieved significant performance improvement compared to traditional intra predictionmethods[15].

## III. EXPERIMENTAL RESULTS ANDINFERENCES

In IPFCN experimental results demonstrate its effectiveness. When compared with HEVC[24], an average of 3.4% bitrate saving is achieved[14].When integrating PNNS into a H.265 codec, PSNRrate performance gains going from 1.46% to 5.20% are obtained. These gains are on average 0.99% larger than those of prior neural network based methods. Unlike the H.265 intra prediction modes, which are each specialized in predicting a specific texture, the proposed PNNS can model a large set of complex textures[16].

## IV. COMPARISON

| Techniques | Method | Drawbacks |
|---|---|---|
| Intra Predictions in H.265/HEVC | 1)intra prediction using video compression standard H.265/HEVC | 1)Result in In efficient coding of the transmitted data. |
| An Optimized Tem- plate Matching Ap- proach to Intra Cod- ing in Video/Image Compression | Template matching Intra prediction | Self-similarities within the image. |
| Overview of the High Efficiency Video Coding (HEVC)Standard | High Efficiency Video Coding (HEVC). | HEVC, which are each specialized in predicting a specific texture. |
| Fully-connected network based intra prediction for image coding. | IPFCN | 1)Only take into consideration blocks of sizes 4*4, 8*8,16*16, and 32*32 pixels. 2)Computational complexity is relative high. |

## V. RESEARCH GAPS OR FUTURESCOPE

The results of this study open new avenues for future research and may serve as source of hypotheses for further quantitative research Future work in the area includes, extending and tuning the tools for Multiview/scalable coding, higher dynamic range operation, and 4:4:4 sampling formats [1].In paper[14], the two chroma components reuse the network model of the luma component. Actually, the chroma and luma components have different signal characteristics. Training separate models will probably boost the performance, and this will be investigated in the future. Deep leaning based methods may be practicable for low complexity. CNN to intra prediction in HEVC is one of the earliest attempt[15]. This paper can provide more inspiration for further researches.But, the use of neural networks for intra prediction within an image coding scheme raises several questions. Thus, the number of possible functions is limited. On the other hand, neural networks can approximate many functions[23], in particular complex predictive functions such as the generation off uture video frames given an input sequence of frames.

## VI. CONCLUSION

In this paper, a comprehensive analysis of various intra prediction techniques that are used for image compression with a comparison study of them is presented. Compared to the traditional method this paper presented a set of neural network architectures, including both fully-connected neural networks and convolutional neural networks, for intra prediction [16]. It is shown that fully-connected neural networks are well adapted to the prediction of image blocks of small sizes where as convolutional one sprovide better predictions in large blocks[25]. When integrated in to a H.265 codec, the proposed neural networks are shown to give rate-distortion performance gains compared with the H.265 intra prediction. Moreover, it is shown that these neural networks can cope with the quantization noise[21] present in the prediction context, i.e. they can be trained on undistorted contexts, and then generalize well on distorted contexts in a coding scheme.

## REFERENCES

[1] J. Lainema, F. Bossen, W. J. H an, J. Min, and K. Ugur, "Intra coding of the HEVC standard," IEEE Transactions on Circuits and Systems for video Technology, vol.22,no.12,pp.17921801,December2012.

[2] Visual Information Processing and Communication V, edited by Amir Said, Onu r G. Guleryuz, Robert L.Steve nson, Proc. of SPIE IST ElectronicImaging,SPIEVol.9029,902904©2014SPIEIST

[3] H. Lee, C. Ekanadham, and A. Y. Ng, "Sparse deep belief net model for visual area V2," in NIPS,2007.

[4] H.Lee, R. Grosse, R. Ranganath , an d A. Y.Ng,"Convolutionaldeep belief networks for scalable unsupervised learning of hierarchical representations," in ICML, 20092009.

[5] M. Turkan and C. Guillemot, "Image prediction based on neighbor embedding methods," IEEE Transaction on Image Proces sing , vol. 21,no.4,pp.18851898,April2012.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for largescaleimagerecognition,"inICLR,2015.

[7] D. Pathak, P. Kr¨ahenb¨uhl, J. Donahue, T. Darell, and A.A. Efros,"Context encoder s: f eature learning by inpainting," in CVPR, 2016.

[8] R. I. Chernyak, "Analysis of the intra predictions in H.265/HEVC,"Applied Mathematical Sciences, vol. 8, no.148, pp. 7389 7408,2014.

[9] P. Luc, N. Neverova, C. Couprie, J. Verbeek, and Y. LeCun, "Pre dict ingdeeperintothefutureofsemanticsegmentation, "inICCV,2017.

[10] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super resolution using deep convolutional networks," IEEE Transactions on P attern Analysis andMachineIntelligence,vol.38,no.2,pp.295307, February2016.

[11] M. Ranzato, V. Mnih, J. M. Susskind, and G.E. Hinton, "Modeling nat- ural images using gated MRFs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 9, pp. 2206 2222, September 2013.

[12] G. J. Sulliv an, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding(HEVC) Standard," IEEE Transactionson Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649 1667, December2012.

[13] B. Lim, S. Son, H. Kim, S. Nah, an d K. M. Lee, "Enhanced deep resid- ual networks for single image super resolution," in C VPR Workshops, 2017

[14] J. Li, B. Li, J. Xu, R. Xiong, and W. Gao, "Fully connected network- based intra prediction for image coding,"IEEE Transactions on Image Processing,vol.27,no.7,pp.32363247,July2018.

[15] Wenxue Cui , Tao Zhang Zhang, Shengping ZhangZhang, Feng Jiang Jiang, Wangmeng ZuoZuo, Debin Zhao Convolutional Neural Networks based Intra Prediction for HEVC," Computer Vision and Pattern Recog- nition (cs.CV). DCC2017

[16] Thierry Dumas, Aline Roumy, Christine Guillemot. Context-adaptive neural network based prediction for image compression. IEEE Trans- actions on Image Processing, Institute of Electrical and Electronics Engineers, 2019, 29, pp.679 - 693. 10.1109/TIP.2019.2934565. hal-01841034v3.

[17] M. Wien, "High efficiency video coding: coding tools and specifica- tion,"Springer, September2014.

[18] V. Dumoulin and F. Visin, "A guide to convolutional arithmetic for deep learning," arXiv:1603.07285v2, January2018

[19] S. Niklaus, L. Mai, and F. Liu, "Video frame interpolation via adaptive convolution," in CVPR,2017.

[20]A.vanderOord,N.Kalchbrenner,andK.Kavukcuoglu,“Pixelrecurrent neural networks,” in ICML,2016.
[21] R. M. Gray and D. L. Neuhoff, “Quantization,” IEEE Transactions on InformationTheory,,vol.44,no.6,pp.2325–2383,October1998.
[22] M. Ranzato, A. Szlam, J. Bruna, M. Mathieu, R. Collobert, and S. Chopra, “Video (language) modeling: a baseline for generative models ofnaturalvideos,”arXiv:1412.6604v4,April2015.
[23] Y.LeCun,L.Bottou,G.B.Orr,andK.-R.M¨uller,“Efficientback-prop,”Neural Networks: Tricks of the Trade, Springer, 1998.
[24] O. Bougacha, H. Kibeya, N. Belhadj, M. A. Ben Ayed, and N. Masmoudi, “Statistical analysis of intra prediction in HEVC video encoder,”in IPAS,2016.
[25] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, July2006.