

Contents

1	Introduction	1
2	Video Coding and Neural Networks	7
2.1	Fundamentals of Video Coding	7
2.1.1	Representation and Properties of Video Sequences	8
2.1.2	The Hybrid Coding Approach	9
2.1.3	Picture Partitioning	11
2.1.4	Predictive Coding	14
2.1.5	Further Essential Coding Tools	15
2.1.6	Encoder Optimization and Evaluation Methods	24
2.2	Intra Prediction Methods	30
2.2.1	Classic and Wide-Angle Prediction	31
2.2.2	Matrix-based Intra Prediction	37
2.2.3	Further VVC Intra Improvements	40
2.2.4	Chroma Intra Prediction	43
2.2.5	Intra Mode Signaling	45
2.3	Neural Networks	50
2.3.1	Basic Machine Learning Terminology	50
2.3.2	Common Architecture Elements	52
2.3.3	Loss Functions and Evaluation Metrics	57
2.3.4	Learning Methods	61
2.3.5	Pruning	65
2.3.6	Dedicated Network Types	66
2.4	State of the Art in NN-based Intra Prediction	76
2.4.1	A History of Intra NNs	76
2.4.2	NN-based Predecessors and Training Methods of MIP	79
2.4.3	GAN-based Intra Approaches	82
2.4.4	U-Nets and Autoencoders	84
2.4.5	NN-based Approaches for Related Tasks	86
3	Non-adaptive NN-based Intra Prediction	89
3.1	Prediction Problem Assessment and Requirements	89
3.1.1	Intra Prediction Optimization Aim	89
3.1.2	Probabilistic Modeling	91
3.1.3	Codec Integration Requirements	93
3.2	Training Considerations	96
3.2.1	Datasets	96
3.2.2	Preprocessing	111
3.2.3	Training Methods and Hyper Parameters	114
3.2.4	Loss Function Choice	115

3.3	Basic Network Architectures	121
3.3.1	Single Output Luma Prediction	121
3.3.2	Chroma Intra Prediction	124
3.4	Architecture Optimization and Network Pruning	127
3.4.1	Number of reference lines	128
3.4.2	Architectural Hyperparameters	129
3.4.3	Adding Deconvolutional Layers	132
3.4.4	Pruning	133
3.4.5	Transference to other Network Versions	135
3.4.6	Combined Results	136
3.4.7	Visual Evaluation	137
3.5	Integration and Signaling	138
3.5.1	Integration into HEVC	139
3.5.2	Integration into VVC	157
3.6	Comparison to Other Works	162
4	Adaptive Prediction Networks	167
4.1	Problem Analysis	167
4.2	Multiple Output Networks	169
4.2.1	Multiple Hypotheses Training	170
4.2.2	HM Integration	171
4.2.3	Results	172
4.3	Autoencoder Architectures	176
4.3.1	Conditional Autoencoder Model	176
4.3.2	Architecture Adjustments	178
4.3.3	Results Analysis	181
4.4	General Variational Autoencoder Models	186
4.4.1	Model Description	187
4.4.2	Results and Analysis	188
4.5	Variational Autoencoders with Joint Vector Quantization	191
4.5.1	Model Setup	191
4.5.2	Training Methods and Loss Adaptation	192
4.5.3	Result Evaluation	202
4.6	VVC Integration and Latent Space Coding	204
4.6.1	Embedding Binarization	204
4.6.2	Further Integration Considerations	208
4.6.3	Final Analysis	209
5	Intra Mode Prediction	213
5.1	Preliminary Considerations and Problem Analysis	214
5.1.1	Problem Assessment	214
5.1.2	Data Analysis	217
5.2	Occurring Problems and Solution Ideas	220
5.2.1	General Approach Setup	220
5.2.2	The Data Imbalance Problem	223
5.2.3	Further General Challenges	227

5.2.4	Joint Mode and Block Prediction	230
6	Conclusion and Outlook	235
6.1	Summary	235
6.2	Outlook	237
A	Self-assembled Dataset Details	241
B	Training Runs for Architecture Optimization	249
	Bibliography	257