Cloud Gaming Quality Assessment

Saman Zadtootaghaj

Quality and Usability Lab, Technische Universität Berlin

Dec 14th, 2020

Standardization Activities

Standardization Activities

ITU-T Study Group 12

Ongoing ITU-T Activities

- ITU-T P.BBQCG: Parametric bitstream-based Quality Assessment of Cloud Gaming Services
- ITU-T G.OMMOG: Opinion Model for Mobile Online Gaming applications
- ITU-T P.CROWDG: Subjective evaluation of gaming quality with a crowdsourcing approach

Contributed to the following recommendations:

- ITU-T G.1032: Influence factors on gaming quality of experience
- ITU-T P.809: Subjective evaluation methods for gaming quality
- ITU-T G.1072: Opinion model predicting gaming quality of experience for cloud gaming services

Standardization Activities

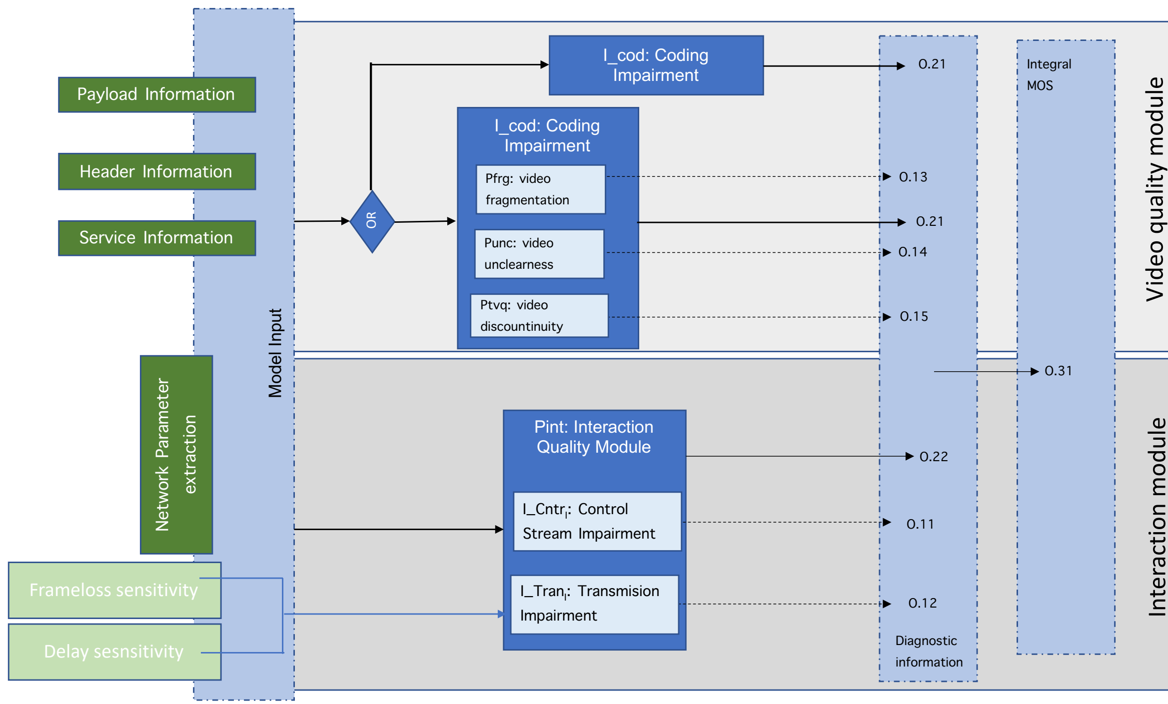
ITU-T P.BBQCG

Documents on progress

- Terms of Reference
- Plan for interactive tests
- Plan for passive viewing-listening tests

Challenges

- Selection of cloud gaming service
- Integration of interactive & passive tests



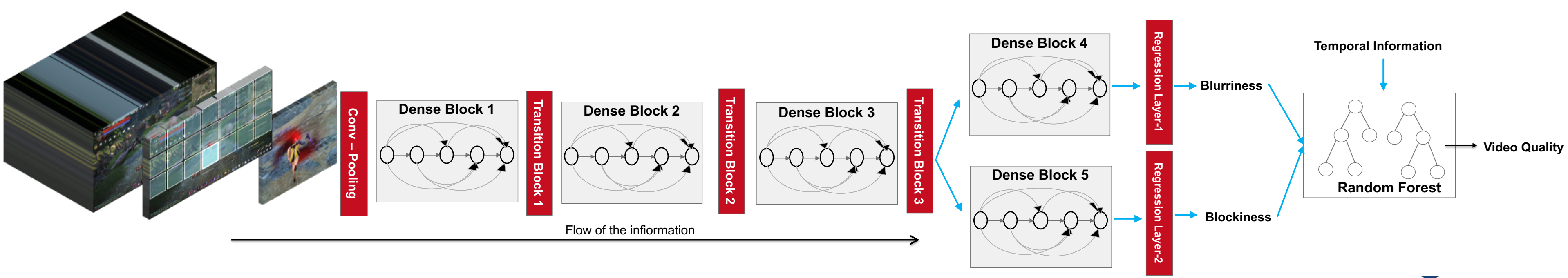
Quality & Usability Lab

Gaming Video Quality Models

Gaming Video Quality Predictions

Research

- Deep learning-based models
  - NDNetGaming, DEMI [1-2]
- Light-weight Quality metrics
  - NR-GVQM, BEG-VQ [3]
- Parametric and Bitstream Based models
  - GamingPara, ITU-T P.1204.3 on gaming [4-5]



DEMI model framework, based on DenseNet-121

Planning and Bitstream based Models

GamingPara

- **GamingPara**: Two gaming video quality models developed based on the multidimensional approach [4]:

    1. Planning model: developed based on the video quality dimensions:
        - Fragmentation: represent blockiness
        - Unclearness: represent blurriness
        - Discontinuity: represent jerkiness
    2. Bitstream model: the model developed based on the packet header information in a multidimensional approach

GamingPara

Planning Model

- Impairment of Video Discontinuity (I_{VD})

$$I_{VD} = exp(\frac{d1}{framerate}) + d2$$

- Impairment of Video Fragmentation (I_{VF})

$$bitperpixel = . \frac{bitrate}{framerate * hight * width}$$

$$I_{VF} = e1 + e2 * log(bitperpixel * bitrate) + e3 * bitrate$$

- Impairment of Video Unclearness (I_{VU})

$$scaleratio = \frac{hight_{encoding} * width_{encoding}}{hight_{display} * width_{display}}$$

$$I_{VU} = f1 + f2 * log(bitperpixel * bitrate) + f3 * \log(scaleratio)$$

GamingPara Bitstream Model

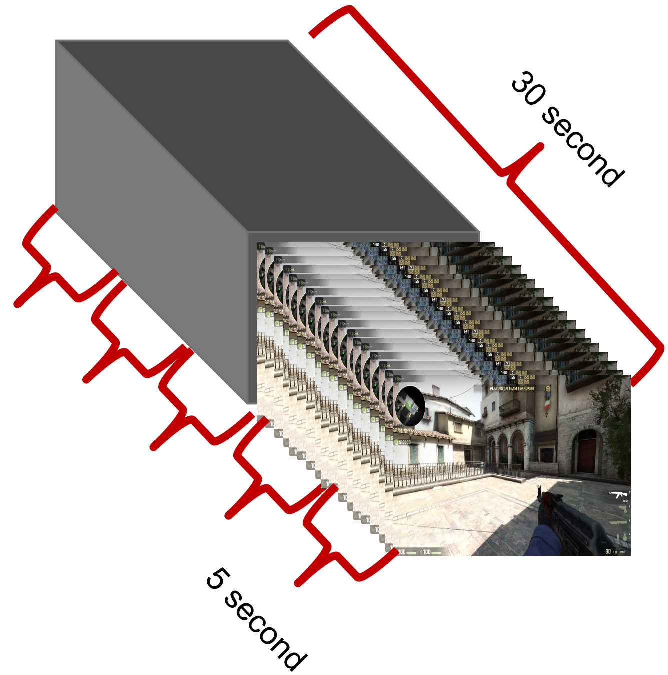
Short Interval

Weight of each frame in overall video quality:

- $w_i = \frac{NDG_i}{NDG_{final}}$ where NDG_i is the measured quality of frame i based on NDNetGaming. NDG_{final} is the final predicted video quality of a certain video based on NDNetGaming.

- MOS of an interval duration from frame i to frame j of a video sequence is calculated as:

$$\widehat{MOS}_{i,j} = MOS_{video} * \sum_{k=i}^{j} \frac{w_k}{j-i}$$



Quality & Usability Lab

GamingPara

Bitstream Model

- Features extracted from each 5 second interval:
  - fr: Frame rate
  - sr: $scaleratio = \frac{hight_{encoding} * width_{encoding}}{hight_{display} * width_{display}}$
  - br_{avg}: Average bitrate of five second sequence
  - num_{I_frame}: Number of I-frames
  - $br_{avg\_I\_Frame}$: Average of I-frames bitrate
  - $stat_{\mathrm{P}\_Frame}$: Statistics of P-frames (e.g. average, standard deviation)
  - $stat_{I\_Frame}$: Standard deviation of P-frames bitrate
  - cp_{video}: $\mathrm{CP}_{\mathrm{video}}$: Video complexity based as follows $\frac{\overline{I_{bitrate}}}{\overline{P_{bitrate}}}$

Proposed Monitoring Model
Short Interval

- I_{VD} → fr, $std_{P\text{-}frame}$, CP_{video}
- I_{VF} → fr, $std_{P\text{-}frame}$, CP_{video} , sr, sr_{video}, $stat_{P\text{-}frame}$, br_{avg}
- I_{VU} → fr, $std_{P\text{-}frame}$, CP_{video} , sr, sr_{video}, $stat_{P\text{-}frame}$

- Core Perceptual Video Quality Model

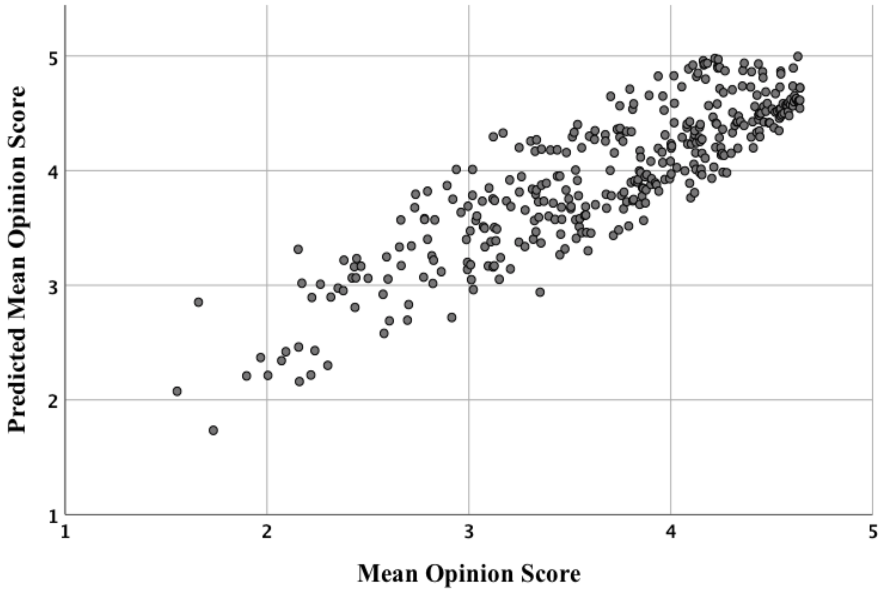
$$R_{QoE} = R_{max} - a \cdot I_{VD} - b \cdot I_{VF} - c \cdot I_{VU}$$

- Core Model

$$R_{QoE} = R_{max} - 0.259 \cdot I_{VD} - 0.554 \cdot I_{VF} - 0.341 \cdot I_{VU}$$

GamingPara

Core Model Results



Scatter plot of predicted MOS and actual MOS for all sequences in the test set of Bitstream model.

Bitstream Model						
Dataset	Training Set			Test Set		
Metric	RSME	SRCC	PLCC	RMSE	SRCC	PLCC
Score	0.21	0.93	0.94	0.34	0.89	0.90

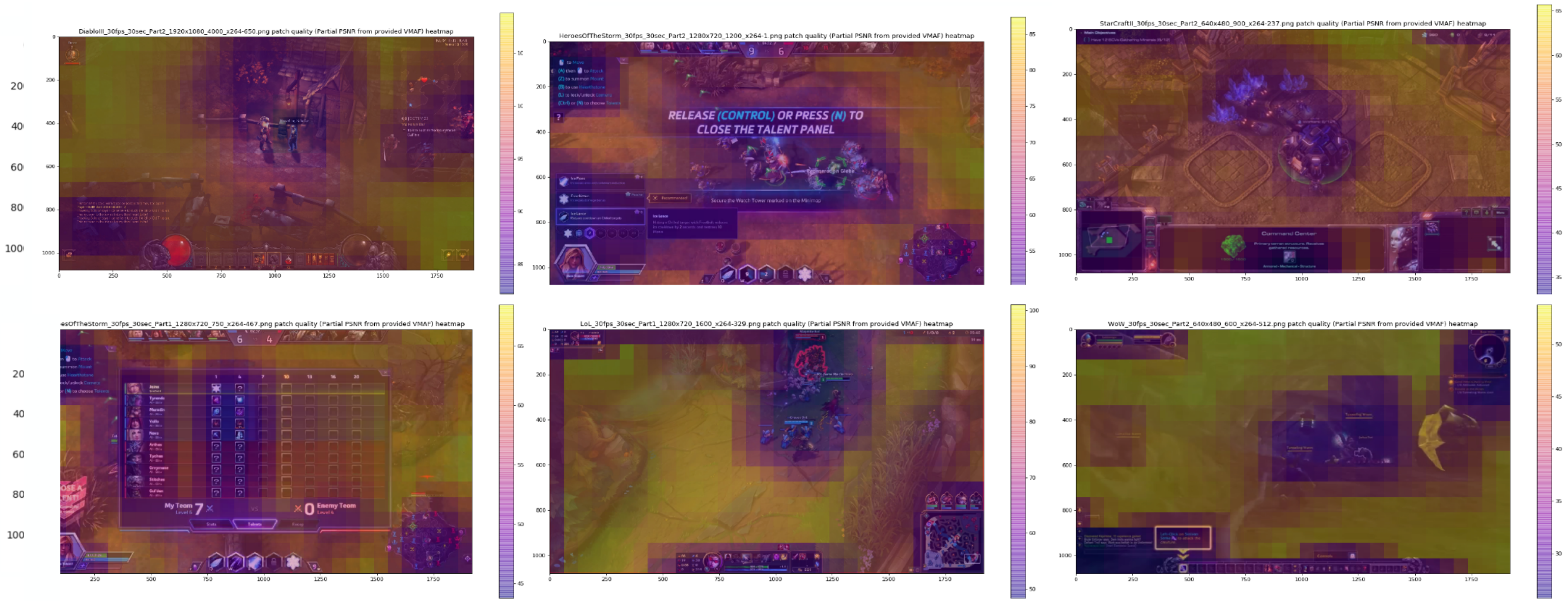
Deep Learning based Quality Models

DEMI

- Three phases followed to develop a deep learning based gaming video quality model
  1. Train the network based on an Objective metric on the image level
     - Allows the CNN learn image artifacts (e.g. blurriness, blockiness)
  2. Fine-tune the model based on a small set of subjective test, using transfer learning
  3. Pool the frame level prediction to the video level based on the temporal complexity
- NDNetGaming: uses DenseNet-121 and lightweight temporal pooling [2]
- DEMI: similar to NDNetGaming but using multidimensional approach [1]

Challenges in Deep Learning based QE Models

Diverse patch quality distribution



Quality & Usability Lab

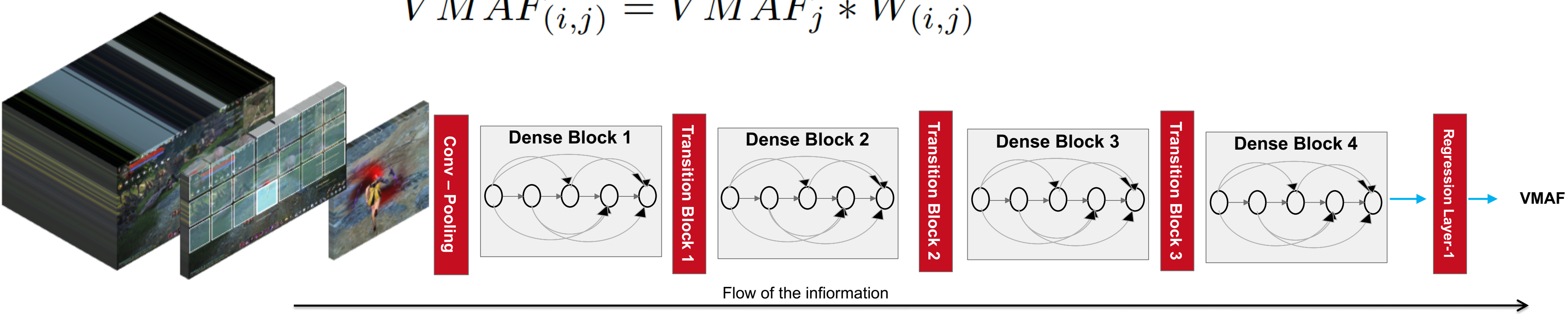
DEMI- Phase 1: Training VMAF

Structure of the model

- DenseNET-121 architecture with patch size of *299 x 299*
- Patch quality is weighted based on Partial PSNR weight

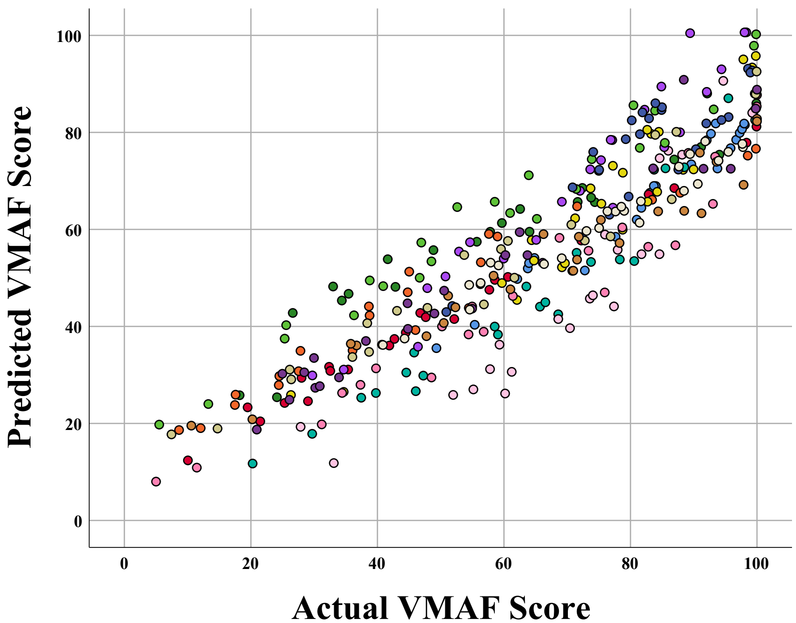
$$W_{(i,j)} = PPSNR_{(i,j)}/PSNR_j$$

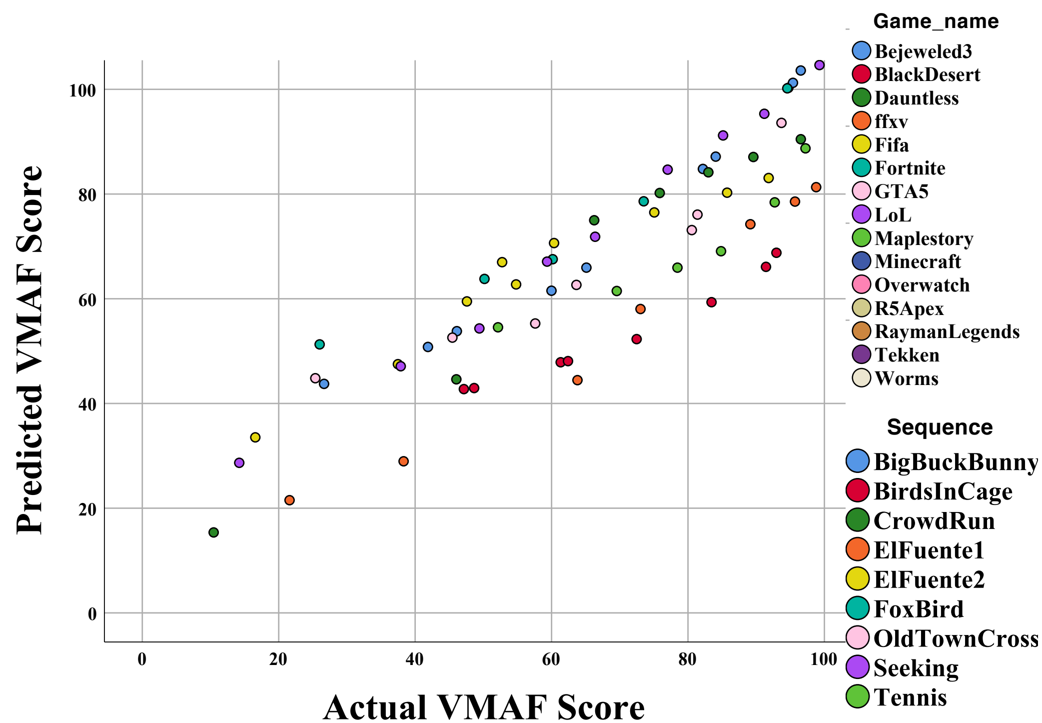
$$VMAF_{(i,j)} = VMAF_j * W_{(i,j)}$$



DEMI- Phase 1: Training VMAF

Performance of Training





CGVDS dataset

NFLX-PD dataset

Predicted VMAF vs. Actual VMAF scores

Quality & Usability Lab

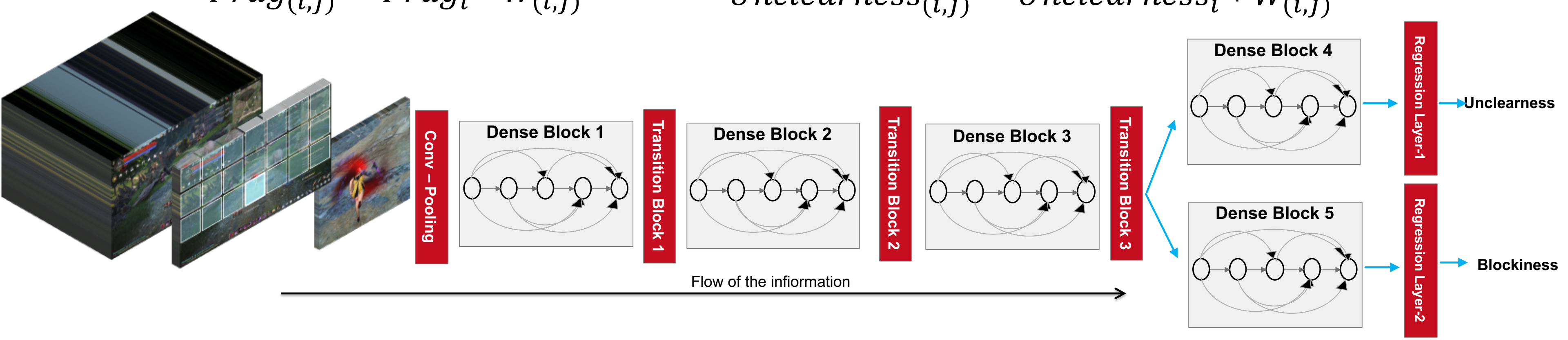
DEMI- Phase 2: Training Unclearness and Fragmentation

Structure of the model

- DenseNET-121 architecture with patch size of *299 x 299*
- Patch quality is weighted based on Partial PSNR weight

$$W_{(i,j)} = PPSNR_{(i,j)}/PSNR_j$$

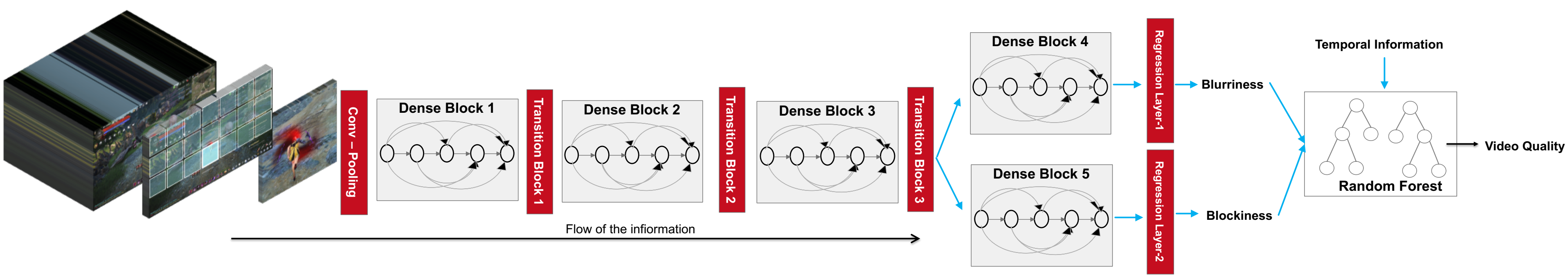
$$Frag_{(i,j)} = Frag_i * W_{(i,j)} \qquad Unclearness_{(i,j)} = Unclearness_i * W_{(i,j)}$$



DEMI- Phase 3: Video Quality

Structure of the model

- Train and Random Forest model based on the temporal Information and Frame level prediction.
  - Block Motion (BM) estimation with a block size of 8x8 is calculated based on Sci-kit video library
  - Temporal Information Index between two consequent frames



Performance of VQA

Signal/Bitstream Based Models

		Full Reference Metrics			No Reference Metrics						
Dataset	Metric	PSNR	SSIM	VMAF	BRISQUE	PIQE	NIQE	NR-GVQM	NDG	DEMI	GamingPara
CGVDS	PLCC	0.66	0.64	0.87	-0.48	-0.41	-0.53	0.80	0.92	0.92	0.89*
	SRCC	0.67	0.76	0.87	-0.46	-0.41	-0.55	0.79	0.93	0.92	0.90*

Saman Zadtootaghaj
Researcher: Quality and Usability Lab
Email: Saman.zadtootaghaj@qu.tu-berlin.de

References

[1] Utke M, Zadtootaghaj S, Schmidt S, Bosse S, Möller S. NDNetGaming-development of a no-reference deep CNN for gaming video quality prediction. Multimedia Tools and Applications. 2020 Jul 24:1-23.

[2] Zadtootaghaj S, el. al. DEMI: Deep Video Quality Estimation Model using Perceptual Video Quality Dimension. IEEE 22nd International Workshop on Multimedia Signal Processing. 2020.

[3] Zadtootaghaj S, Barman N, Schmidt S, Martini MG, Möller S. NR-GVQM: A no reference gaming video quality metric. In2018 IEEE International Symposium on Multimedia (ISM) 2018 Dec 10 (pp. 131-134).

[4] Zadtootaghaj S, Schmidt S, Sabet SS, Möller S, Griwodz C. Quality estimation models for gaming video streaming services using perceptual video quality dimensions. InProceedings of the 11th ACM Multimedia Systems Conference 2020 May 27 (pp. 213-224).

[5] Ramachandra Rao, R.R el al, A Large-scale Evaluation of the bitstream-based video-quality model ITU-T P.1204.3 on Gaming Content, IEEE 22nd International Workshop on Multimedia Signal Processing. 2020.