Quality Enhancement of Gaming Content using Generative Adversarial Networks

Nasim Jamshidi[¥], **Saman Zadtootaghaj**[¥], Nabajeet Barman[ǂ], Steven Schmidt[¥], Maria Martini[ǂ], Sebastian Moller[¥]

[¥] Quality and Usability Lab, Berlin Institute of Technology, Germany

[ǂ] Kingston University, London, UK

VQEG Meeting, 9th March 2020

Image Enhancement Techniques

- Deblocking-oriented
  - E.g. Applying BM3D (block-matching and 3D filtering) [1]
  - Using Wavelets-based [2]
- Deep learning-based (Super Resolution Task)
  - CNN based methods
  - GANs-based methods

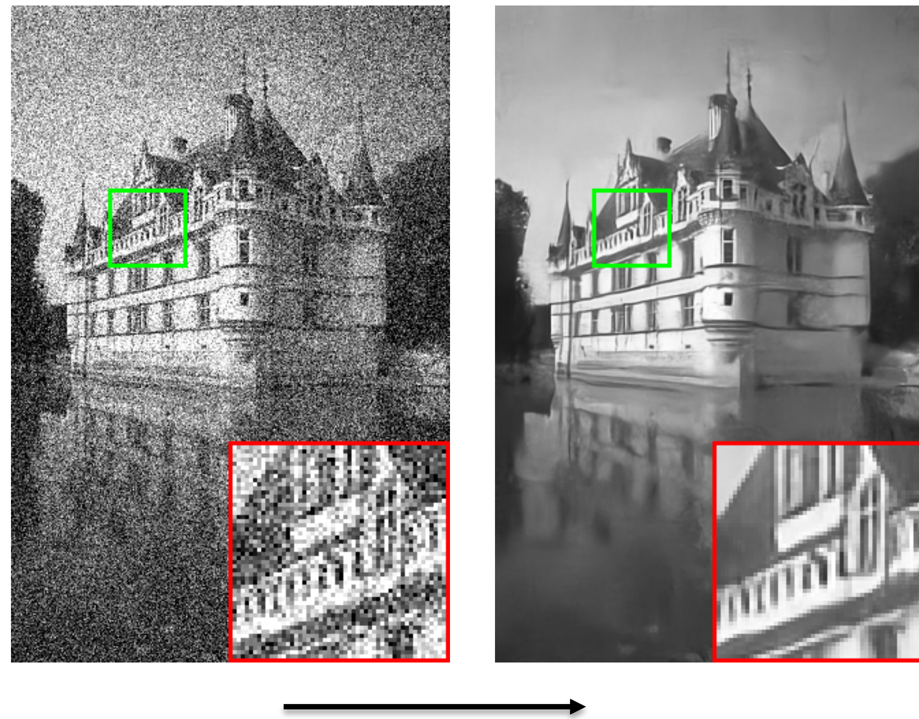
[1]. Self-learning-based post-processing for image/video deblocking via sparse representation

[2]. A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations

Super Resolution Tasks

Before Enhancement

After Enhancement



DnCNNs

CNN based methods

- DnCNNs [1] → for deblocking of Gaussian denoising with unknown noise level and Super resolution task

GAN based methods

- SRGAN [2] → Capable of inferring photo-realistic natural images for 4× upscaling factors

[1]. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising

[2]. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network

Quality & Usability Lab

Super Resolution Tasks

Bicubic
(21.59dB/0.6423)

SRGAN
(21.15dB/0.6868)

Original

Figure - Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network

Loss Function

Pixel-based vs Perceptual-based

- Perceptual quality or reconstruction accuracy?
- FR metric or NR metric for evaluation?

SRResNet
(23.53dB/0.7832)

SRGAN
(21.15dB/0.6868)

Original

Figure - Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network

Gaming Content

Special Temporal and Spatial Information

- Game is a **rule-based** system that has special characteristics.
- A game is usually constructed from a **pool of predesigned objects** which result in different level of details.
- A game has a **certain level of abstraction**, and that does not vary much during the gameplay.
- Many games have **specific motion pattern**, e.g. racing game or side scrolling games.

Quality & Usability Lab

GAN for Enhancement of Gaming Content

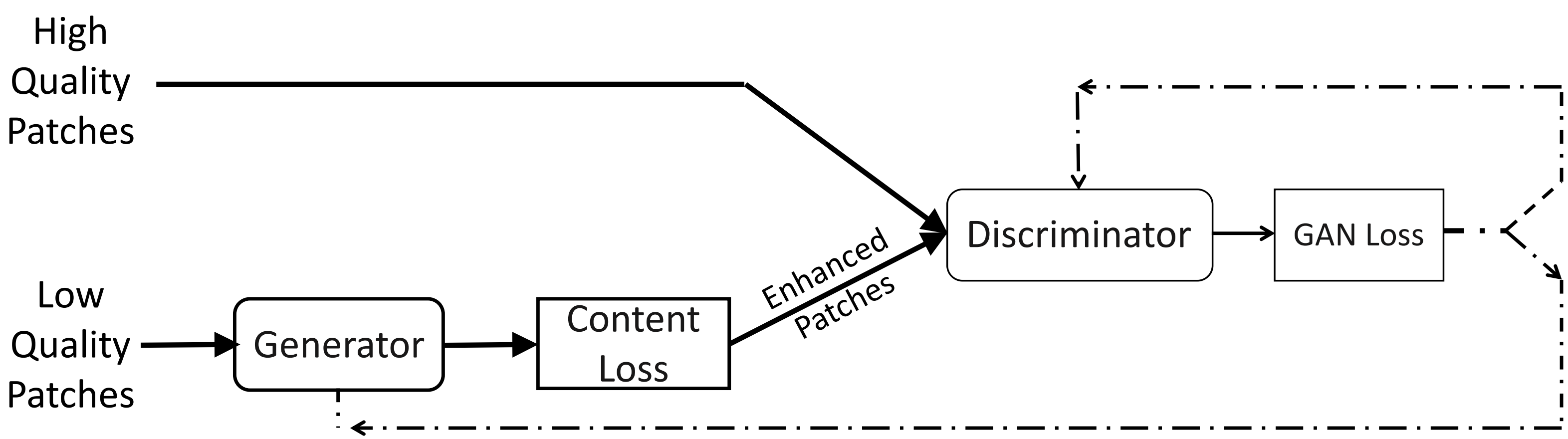
Research Questions

- Investigate the performance of GANs for quality enhancement of gaming content
  - **Content Diversity**: Can we have better enhancement for a specific game?
  - **Enhancement Power**: How much enhancement we may gain?
  - **Blurriness vs. Blockiness**: Is there any difference in enhancement of different image artifact?

GAN for Enhancement of Gaming Content

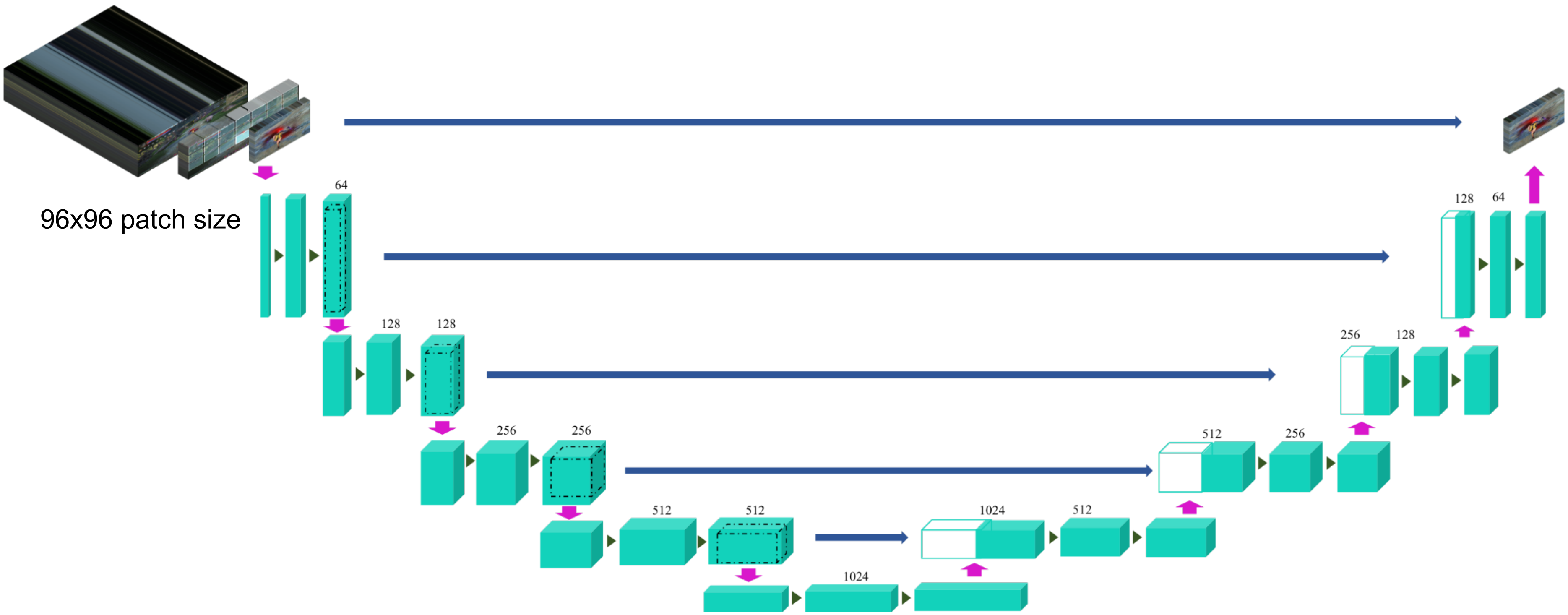
Updating the model

- We used SRGAN architecture
  - Updated the loss function
  - Used U-NET for generator block
  - Used deeper CNN for Discriminator



GAN for Enhancement of Gaming Content

Generator



96x96 patch size

GAN for Enhancement of Gaming Content

Loss Function

$$l^{SR} = \underbrace{\underbrace{l^{SR}_{\mathrm{X}}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}}_{\text{perceptual loss (for VGG based content losses)}}$$

- Content Loss: Euclidean distance between the feature representations of a generated image and the reference image
  - Feature map of the last convolution before max-pooling (VGG19)
- Adversarial Loss: from discriminator network
- **Quality Loss**: Pretrained VGG19 for quality task using VMAF.

GAN for Enhancement of Gaming Content

Datasets

- **Part-1**: The dataset is created based on 100k image patches which are extracted from a single game, League of Legends (LoL).
- **Part-2**: The dataset is created based on 100k image patches which are extracted from 12 different video games.
- **Part-3**: The dataset is created based on 100k image patches which are extracted from LoL, but consists of two sub-parts
  - **Blur**: consisting of patches extracted from 480p and 720p videos up-scaled to 1080p videos using bicubic method.
  - **Blockiness**: consisting of patches extracted from frames from 1080p videos encoded at various bitrate levels.

GAN for Enhancement of Gaming Content

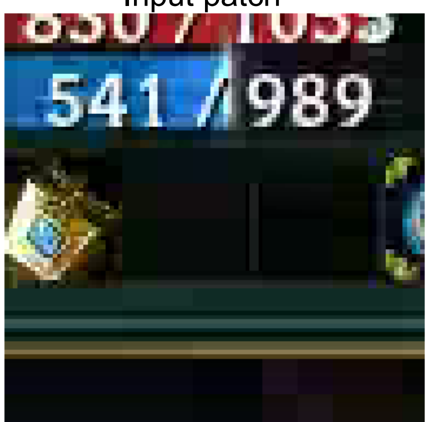
Patch Selection

Results

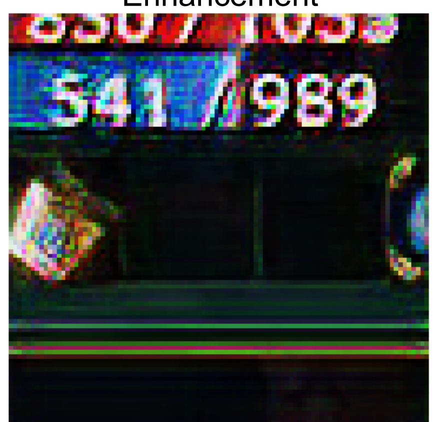
Observation

- Perceptual Quality vs. Image reconstruction
  - Using pixel-wise metrics in the loss function allows the model to predict the text in the image well
- Add-up Distortion:
  - If the quality is already high, we observed additional distortion

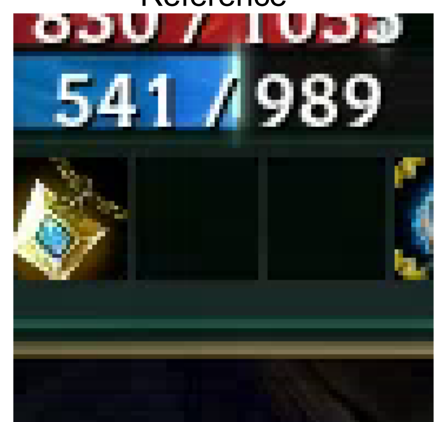
Input patch



Enhancement



Reference



GAN for Enhancement of Gaming Content

Enhancement Power

- **Dataset Part-1:**
  - Class 1 and 2 consisting of frames with Blockiness artifact with VMAF values range between 20 - 40 and 40-60 respectively
  - Class 3 and 4 consisting of frames with Blurriness artifact with VMAF values range between 20 - 40 and 40-60 respectively.

	NIQE		PIQE	
	Distorted	Enhanced	Distorted	Enhanced
Class-1	5.5	2.79	70.67	24.1
Class-2	3.69	2.73	56.1	18.06
Class-3	5.8	2.69	74.8	27.21
Class-4	3.94	2.62	61.36	20.1

GAN for Enhancement of Gaming Content

Content Diversity

- **Comparing Dataset Part-1 and Part-2:**
  - Class 1 and 2 consisting of frames with Blockiness artifact with VMAF values range between 20 - 40 and 40-60 respectively
  - Class 3 and 4 consisting of frames with Blurriness artifact with VMAF values range between 20 - 40 and 40-60 respectively.

	NIQE		
	Distorted	Enhanced Part-1	Enhanced Part-2
Class 1	5.5	2.79	3.21
Class 2	3.69	2.73	3.32
Class 3	5.8	2.69	3.11
Class 4	3.94	2.62	3.91

Quality & Usability Lab

GAN for Enhancement of Gaming Content

Blurriness vs. Blockiness

- **Dataset Part-3:**
  - Selected 40 pairs of LoL frames with similar quality level in terms of VMAF, one part with blockiness and the other with blurriness artifacts.
  - Class-1 with VMAF value ranges from 20 to 40.
  - Class-2 with VMAF value ranges from 40 to 60.

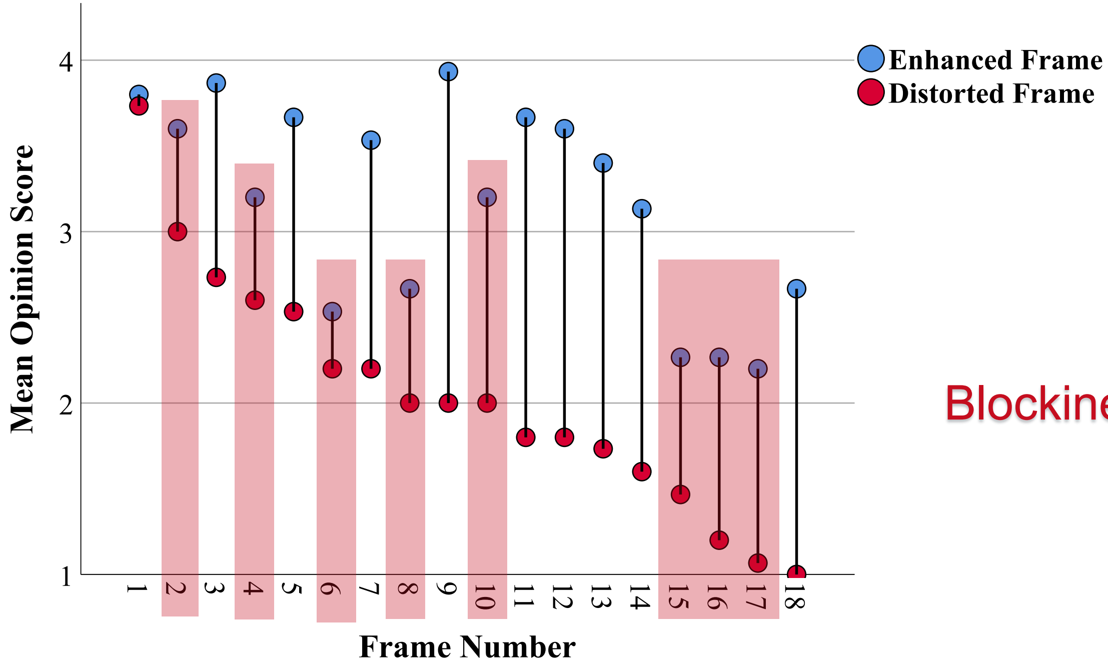
	NIQE Improvement		PIQE Improvement	
	Bluriness	Blockiness	Bluriness	Blockiness
Class-1	2.68	2.21	45.15	37.36
Class-2	1.54	1.14	34.35	28.16

Quality & Usability Lab

GAN for Enhancement of Gaming Content

Subjective Quality Assessment

- **Dataset Part-3:**
  - Selected 2 reference frames from the game LoL
  - 9 distorted frames are selected from each reference frame



Blockiness


Before enhancement
After enhancement
Before enhancement
After enhancement
Single Gaming Video Frame

Discussion and Conclusion

- Copy-right issue
- Video quality enhancement
- There is a lot to do on loss function
- Mixture of deblocking and deep learning methods might be a good option for image quality enhancement of blockiness artifact
- Gaming content benefits from similarity of content
- Research on lighter version of deep learning models for enhancement technique

Thank you for your Attention!!

Any Question?

Saman Zadtootaghaj

saman.zadtootaghaj@qu.tu-berlin.de

Visit www.qu.tu-berlin.de for more information.