Новые стандарты сжатия и устойчивость JPEG AI к состязательным атакам

Dmitriy Vatolin
MSU Institute for Artificial Intelligence
ISP RAS Research Center for Trusted AI
CS MSU Graphics&Media Lab

Collaborations

- **90% of our projects** are sponsored by companies
- We have experience of **long-term collaboration** with Intel, Samsung, Huawei and other
- All our research is aimed to be **extremely practical** for industry

and many others...

Our key results

#1 Video Codecs Analysis
compression.ru/video/codec_comparison

#1 3D Video Quality Estimation metrics
videoprocessing.ai/stereo_quality

#1 globalcompetition.compression.ru (www.gdcc.tech) —
world's biggest contest in lossless compression methods (50K & 200K EUR)

Top Video Processing Benchmarks Collection
videoprocessing.ai/benchmarks (18 benchmarks, 11 in top on paperswithcode.com)

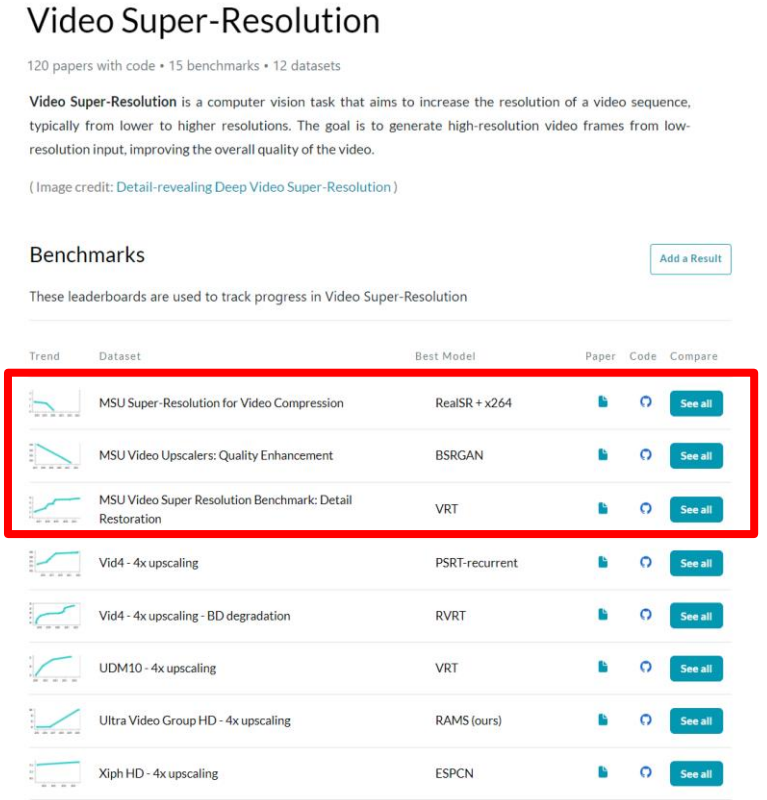
The most well known student — Karen Simonyan (author of VGG, Business Insiders's Top 100 AI persons 2023)

Our worldwide leading results

We declare that we are **in the top in benchmarks on Paperswithcode.com (11 benchmarks)**:

- the best in Russia,
- in the top of laboratories in the world

All our benchmarks are **dedicated to video processing**



Video Super-Resolution

120 papers with code • 15 benchmarks • 12 datasets

Video Super-Resolution is a computer vision task that aims to increase the resolution of a video sequence, typically from lower to higher resolutions. The goal is to generate high-resolution video frames from low-resolution input, improving the overall quality of the video.

(Image credit: Detail-revealing Deep Video Super-Resolution)

Benchmarks

Add a Result

These leaderboards are used to track progress in Video Super-Resolution

Trend	Dataset	Best Model	Paper	Code	Compare
	MSU Super-Resolution for Video Compression	RealSR + x264			See all
	MSU Video Upscalers: Quality Enhancement	BSRGAN			See all
	MSU Video Super Resolution Benchmark: Detail Restoration	VRT			See all
	Vid4 - 4x upscaling	PSRT-recurrent			See all
	Vid4 - 4x upscaling - BD degradation	RVRT			See all
	UDM10 - 4x upscaling	VRT			See all
	Ultra Video Group HD - 4x upscaling	RAMS (ours)			See all
	Xiph HD - 4x upscaling	ESPCN			See all

Вопрос в зал

Обучаете ли вы какие-нибудь сеточки для сжатия или обработки видео?

Вопрос в зал

Тестируете ли вы кодеки или настройки кодеков?

Вопрос в зал

Тестируете ли вы методы обработки видео (денойзеры, Super-Resolution и т.д.)?

Вопрос в зал

Как вы сравниваете результат?

Вопрос в зал

Используете ли вы субъективные сравнения?

GRAPHICS & MEDIA LAB
VIDEO GROUP

Бенчмаркинг
метрик качества видео

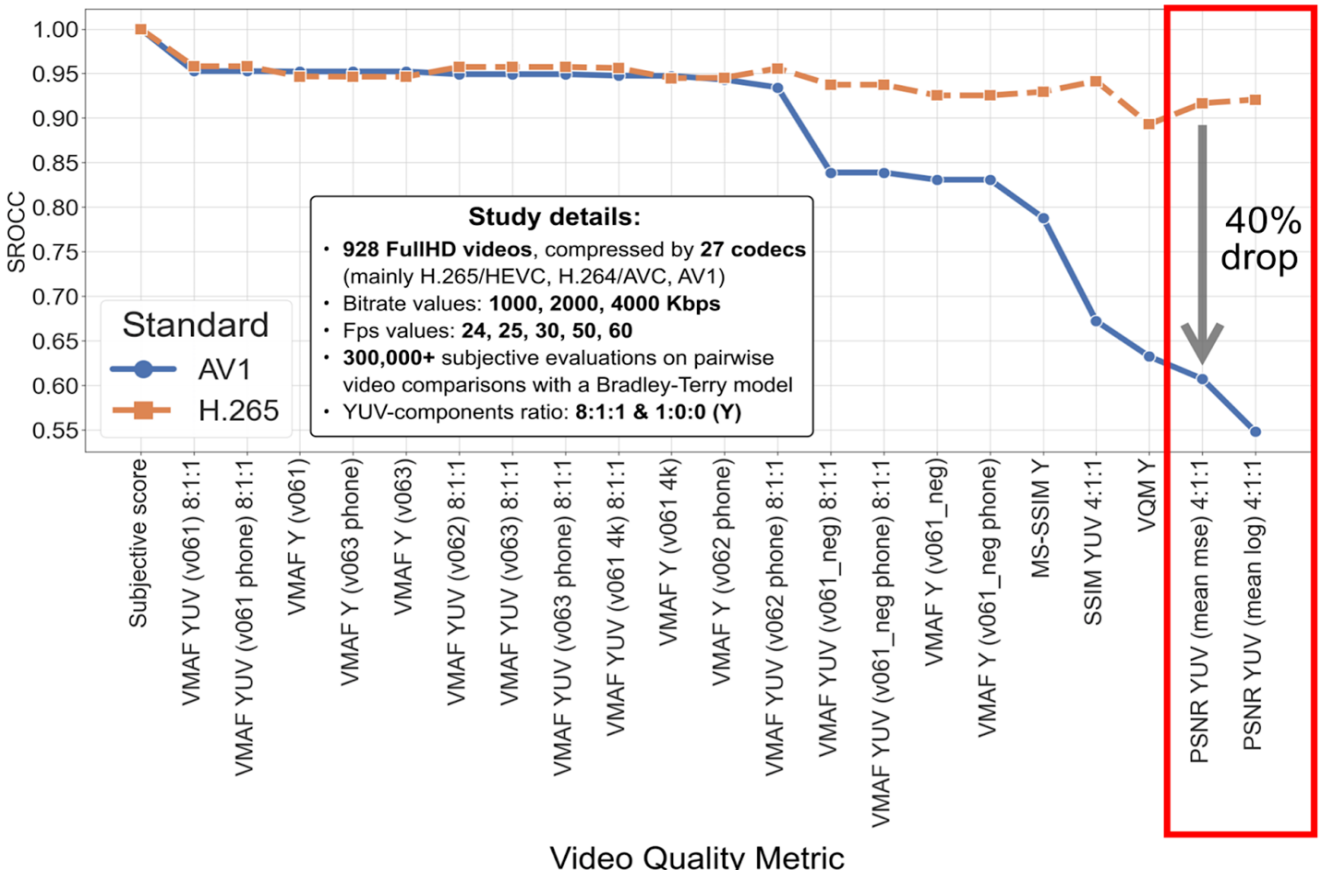
Comparison of quality metrics

MSU Video Quality Metrics Benchmark Dataset

Dataset	Original videos	Average duration (s)	Distorted videos	Distortion	Subjective framework	Subjects	Answers
MCL-JCV (2016) [H.Wang et al.]	30	5	1,560	Compression	In-lab	150	78K
VideoSet (2017) [H. Wang et al.]	220	5	45,760	Compression	In-lab	800	-
UGC-VIDEO (2020) [Y. Li et al.]	50	>10	550	Compression	In-lab	30	16.5K
CVD-2014 [M. Nuutinen el al.]	5	10-25	234	In-capture	In-lab	210	-
LIVE-Qualcomm [D. Ghadiyaram et al.]	54	15	208	In-capture	In-lab	39	8.1K
GamingVideoSET [N. Barman et al.]	24	30	576	Compression	In-lab	25	-
KUGVD (2019) [N. Barman et al.]	6	30	144	Compression	In-lab	17	-
KoNViD-1k (2017) [V. Hosu et al.]	1,200	8	1,200	In-the-wild	Crowdsource	642	205K
LIVE-VQC (2018) [Z. Sinno et al.]	585	10	585	In-the-wild	Crowdsource	4,776	205K
YouTube-UGC (2019) [Y. Wang et al.]	1,500	20	1,500	In-the-wild	Crowdsource	>8,000	600K
LSVQ (2020) [Z. Ying et al.]	39,075	5-12	39,075	In-the-wild	Crowdsource	6,284	5M
MSU VQM Benchmark Dataset (2022)	36	10, 15	2,486	Compression (83 codecs)	Crowdsource	10,800	766K

Background

PSNR degradation on new compression standards



Good question

GRAPHICS & MEDIA LAB
VIDEO GROUP

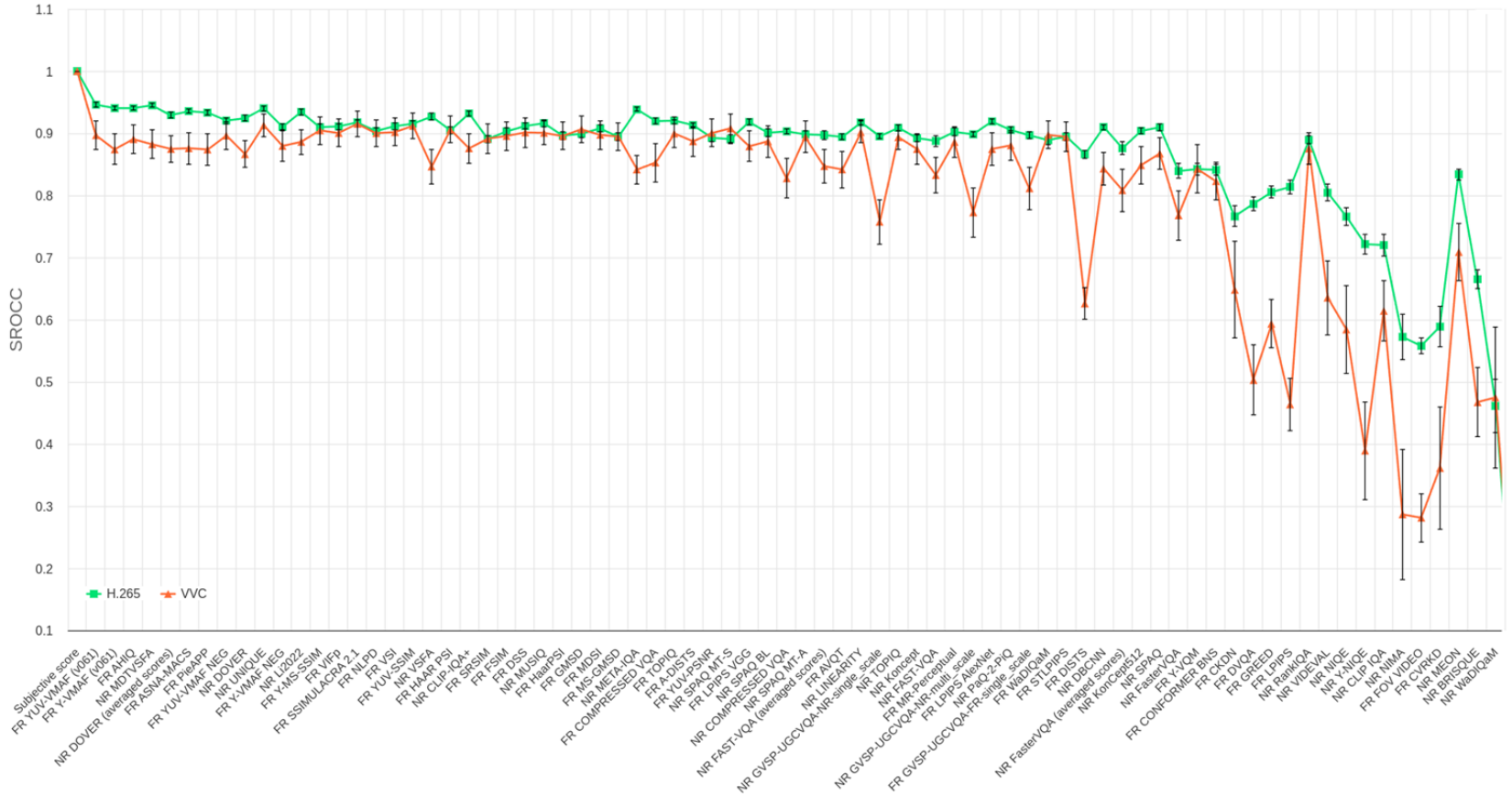
Still using PSNR?

(for AV1/VVC)

Good luck!

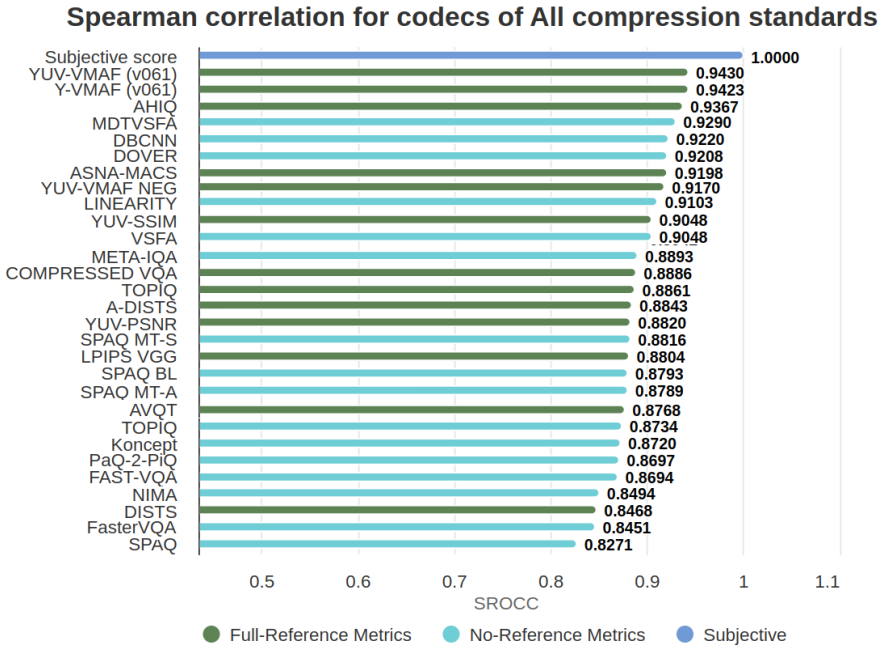
CVQA benchmark

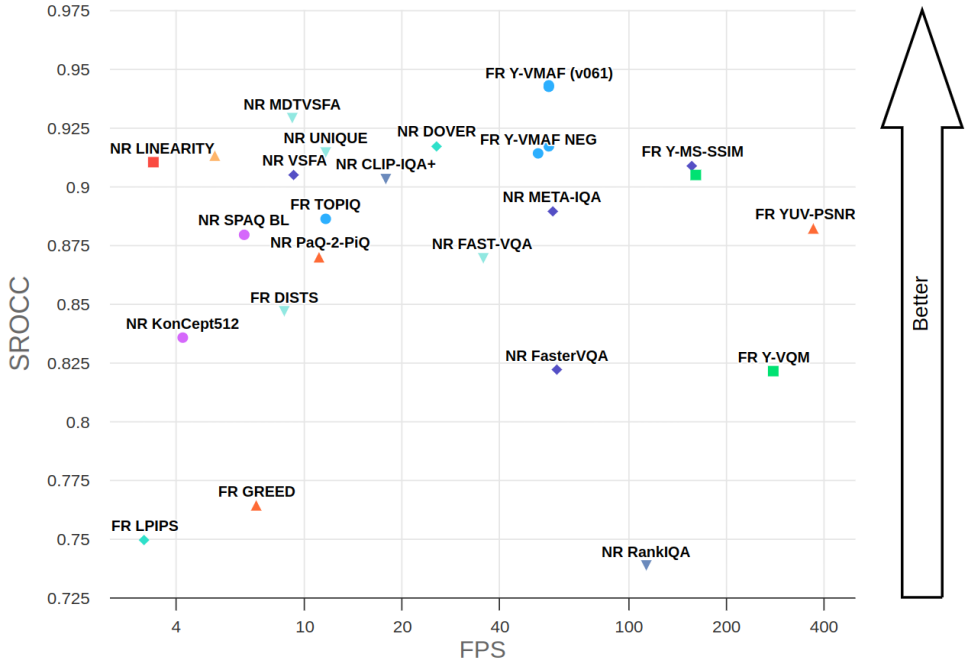
New compression standards are more complex



CVQA benchmark

Current results for selected metrics





MSU Institute for AI Video Lab

https://videoprocessing.ai/

Biggest problem of modern metrics

The results of ABSOLUTELY ALL video quality metrics that outperformed VMAF on open datasets were not reproduced on our dataset!!!

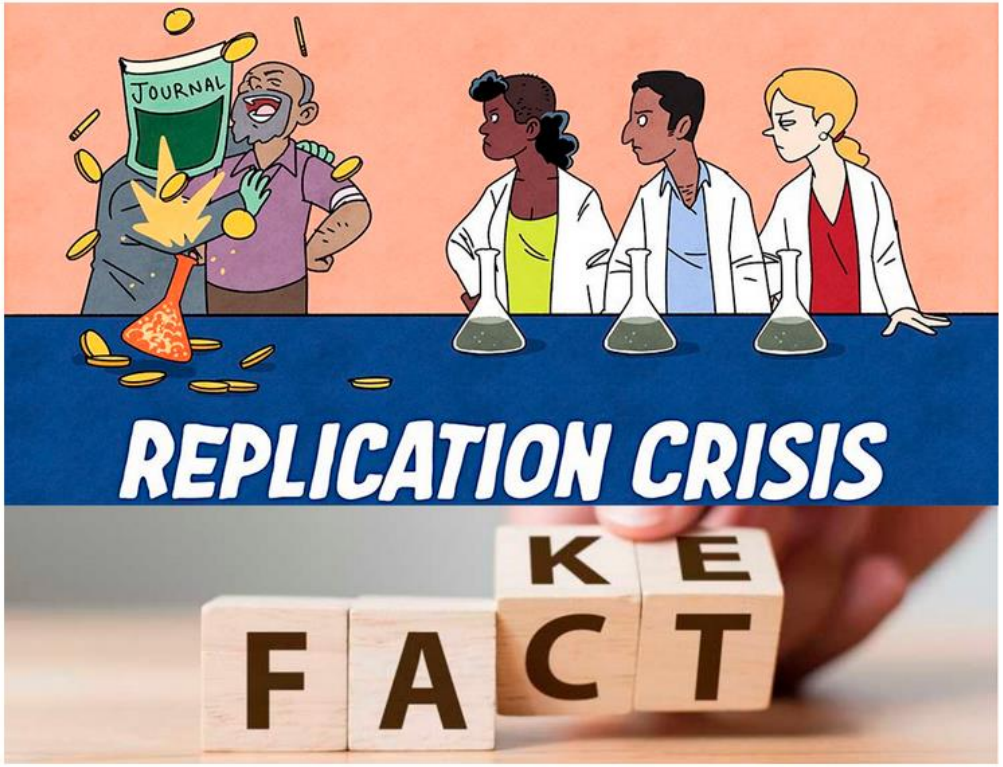
3Dvideo 17 дек 2019 в 10:11

Deep Fake Science, кризис воспроизводимости и откуда берутся пустые репозитории

13 мин 60K

Open source*, Big Data*, Машинное обучение*, Научно-популярное, Искусственный интеллект

Технотекст 2020



https://habr.com/ru/articles/480348/

GRAPHICS & MEDIA LAB
VIDEO GROUP

Поиск устойчивой метрики

Reviewer's insight

This paper is sound, interesting, but in my opinion does not innovate enough to be published in high profile journal like IJCV. The paper can be easily compressed into a conference paper. As a side note, I'd say that the ethics of such research is questionable in that it fosters fraud in the evaluation of results, but does not offer a solution. The only deduction one can make from such papers is that NR metrics should be banned from benchmarks and challenges, or that they could no longer be public, so that nobody can train on them. But, perhaps, this deduction is too much hurried up and there might be ways to make any NR metrics robust to such attacks. That would be for sure a valuable contribution.

Review of our paper about adversarial attacks on NR-metrics

This conclusion is applicable for ALL FR- RR-metrics as well

Are you ready to ban all NN-metrics from all benchmarks, challenges and papers?

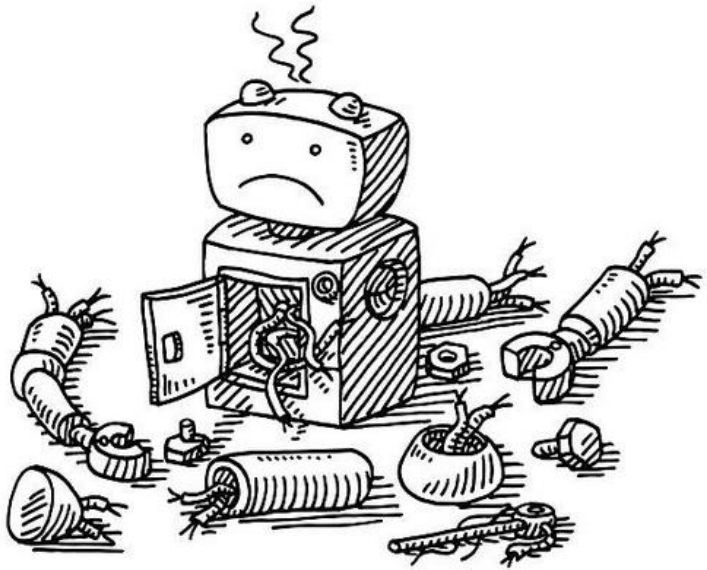
Хабр | КАК СТАТЬ АВТОРОМ | Карьерный трек

Моя лента • Все потоки Разработка Администрирование Дизайн Менеджмент Маркетинг Научпоп

3Dvideo 22 ноября в 11:02

Хакинг метрик качества видео или как с приходом ИИ все становится намного сложнее

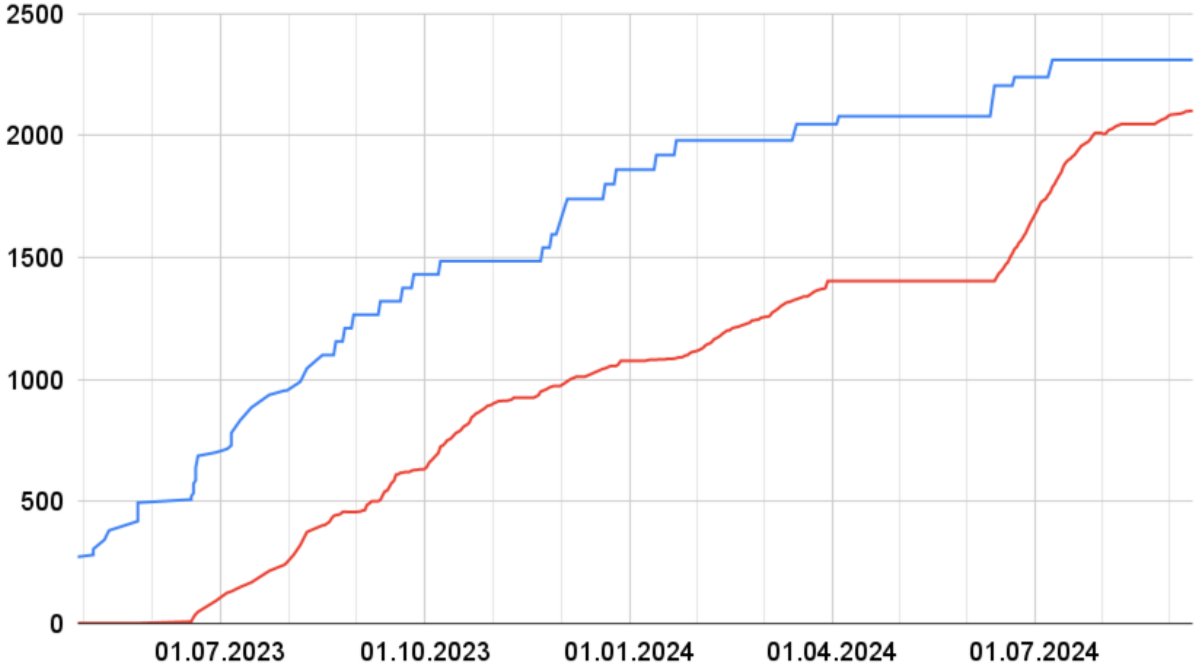
Программирование*, Сжатие данных*, Машинное обучение*, Научно-популярное, Искусственный интеллект



Сейчас модно писать, что ML пришел туда и все стало отлично, DL пришел сюда и все стало замечательно. А к кому-то пришел сам AI, и там все стало просто сказочно! Возможна ли

https://habr.com/ru/articles/700726/

Scope of computations



Implemented attack-on-metrics

Computed attack-on-metrics

We are planning to evaluate:

- 66 metrics
- 35 attacks

Main practical impact

- **Robustness metrics can be used <u>in loss function</u>**
  - Simple metrics were used for JPEG AI training
  - PSNR is still the most popular metric for SR (!)
- **Robust metrics should be used <u>in benchmarking</u>**
  - A lot of scientific benchmarks are useless now
  - In-house comparison of best methods
- **Attacks detection is necessary for results analysis**
  - For example in MSU Annual Codecs Comparison
- **Protection from attacks will be necessary**
  - For JPEG AI practical implementation and not only

Ключевой момент

На данный момент не удалось найти НИ ОДНОЙ новой метрики, которая бы не была подвержена атакам

(пока идет накопление уникального опыта атак)

GRAPHICS & MEDIA LAB
VIDEO GROUP

Бенчмарк защит метрик

Вопрос в зал

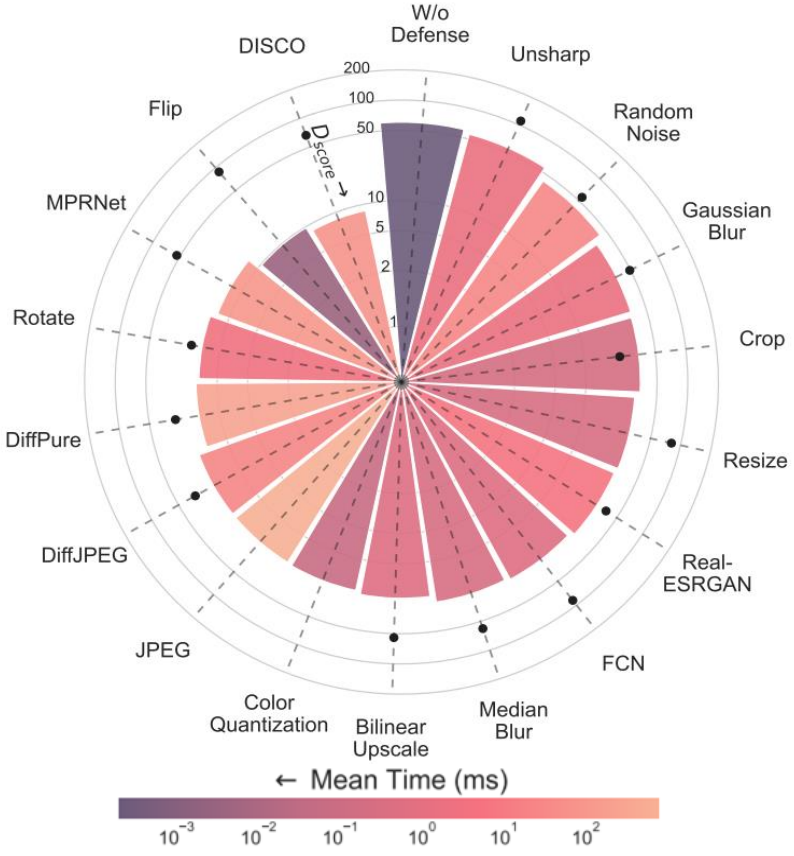
Есть отчет компании/внутреннее сравнение: кто готов по графикам VMAF делать выводы по отчету?

Defenses benchmark

Key numbers and results

- 7 IQA metrics
- 14 attacks
- 17 evaluated defenses
- Adaptive and non-adaptive cases

Adversarial defenses efficiency for IQA metrics in terms of metrics gain. Bars and dots are for non-adaptive and adaptive attacks, respectively



Практический вопрос

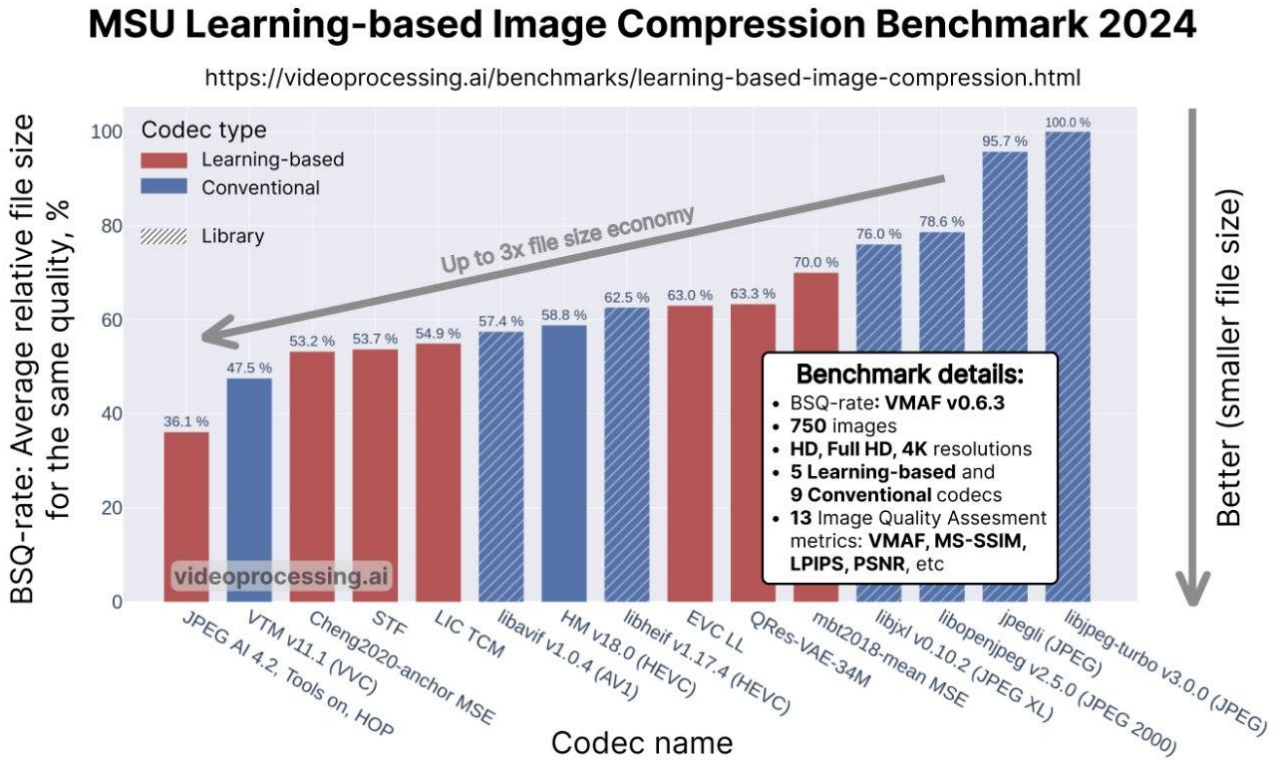
В КАЖДОМ уважающем себя кодеке сегодня есть атаки на метрики (тюнинг под метрики качества).

Вопрос: есть кодек-«черный ящик», включен ли в нем тюнинг? (и какой)

GRAPHICS & MEDIA LAB
VIDEO GROUP

JPEG AI и его устойчивость

Learning-based Image Compression Benchmark
Current leaderboard



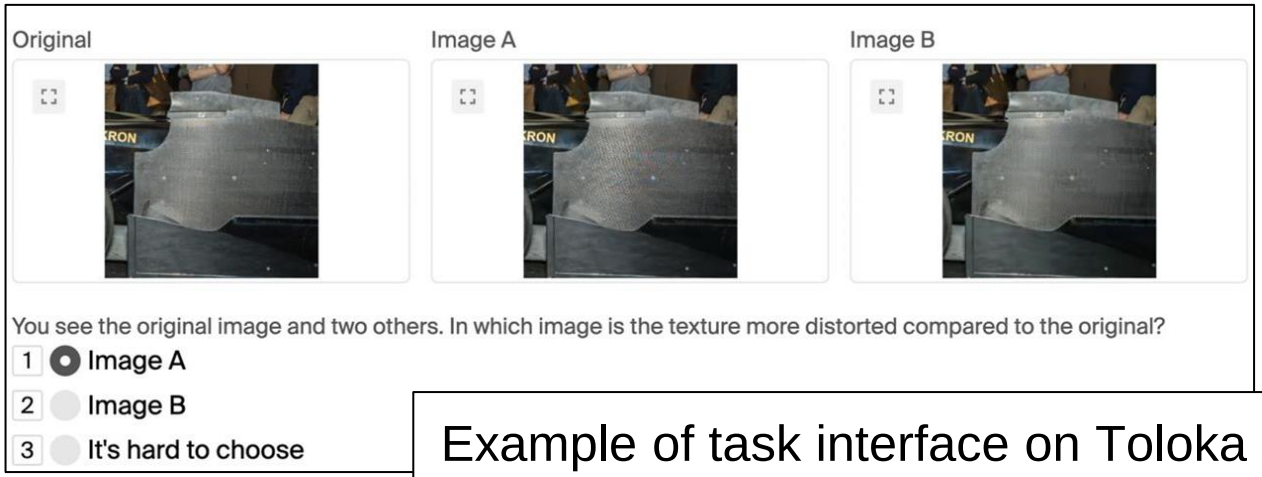
MSU Institute for AI Video Lab
https://videoprocessing.ai/

Not published yet!

Subjective Evaluation

To confirm artifacts in the neural compressed images, subjective comparisons were conducted on the **Toloka AI** platform:

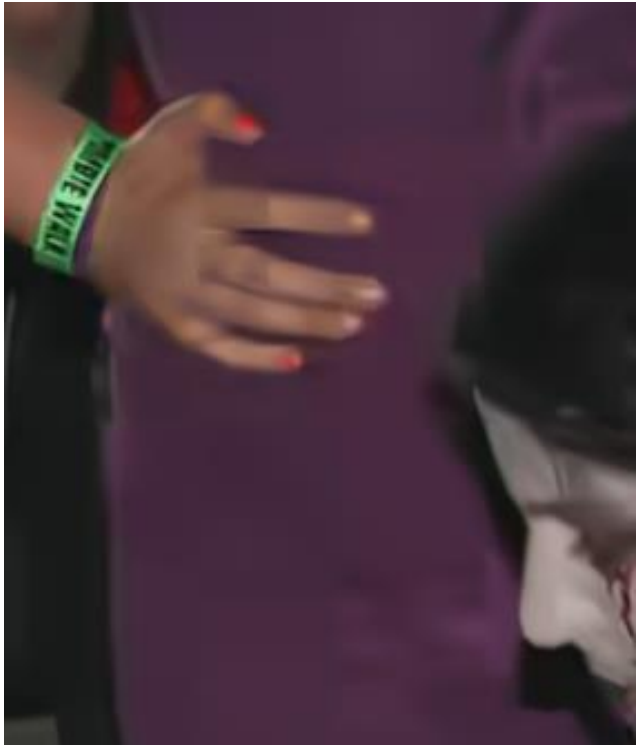
- A task and instructions were created for each artifact type
- Participants were asked to choose the most significant distortion



Example of task interface on Toloka AI

Text distortion

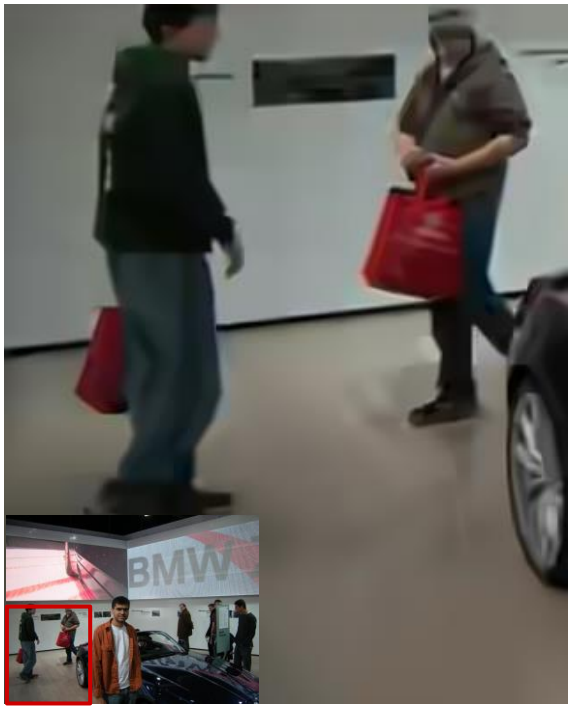
Original



VTM 20.0,
311.2 times

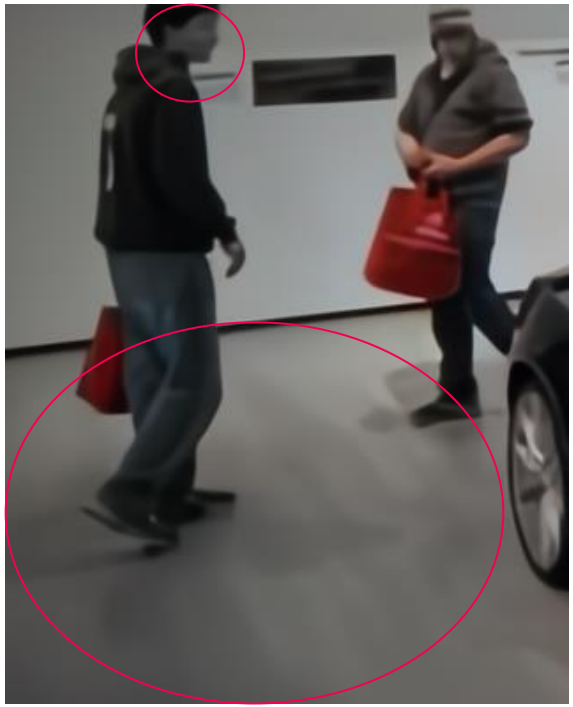
JPEG AI 4.6 tools on, high,
310.3 times

Color distortion



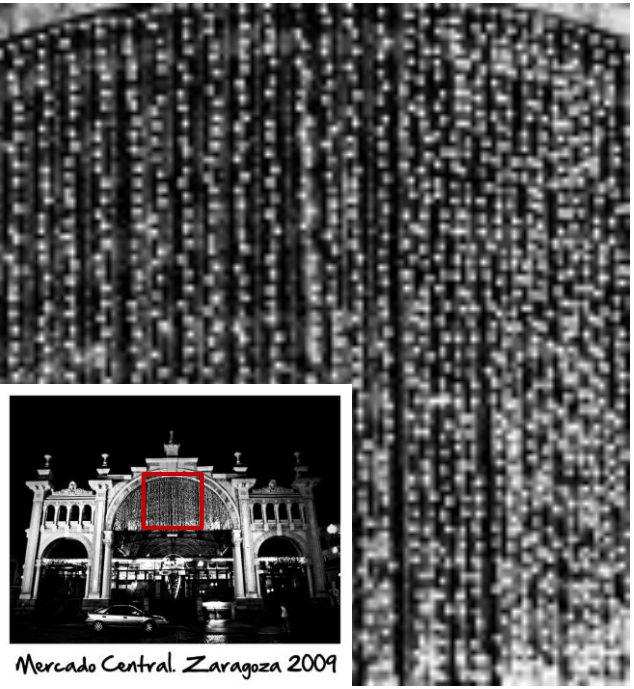
Original

VTM 20.0,
208.7 times

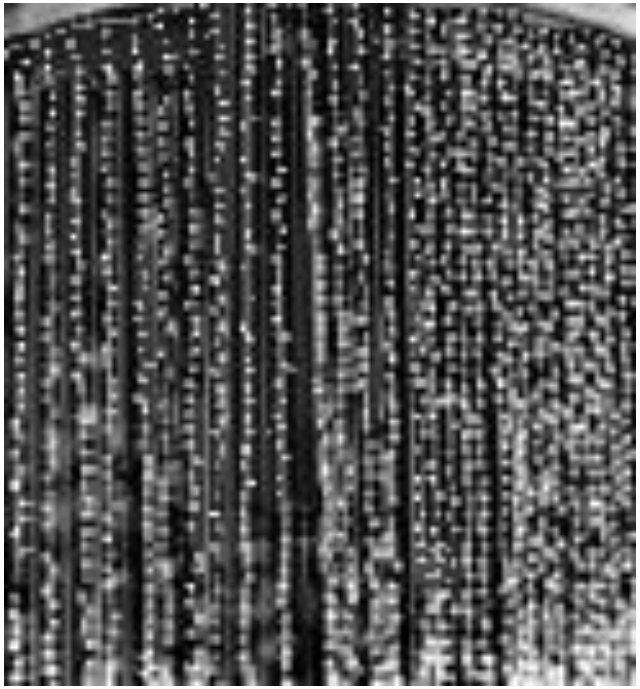


JPEG AI 4.6 tools on, high,
204.1 times

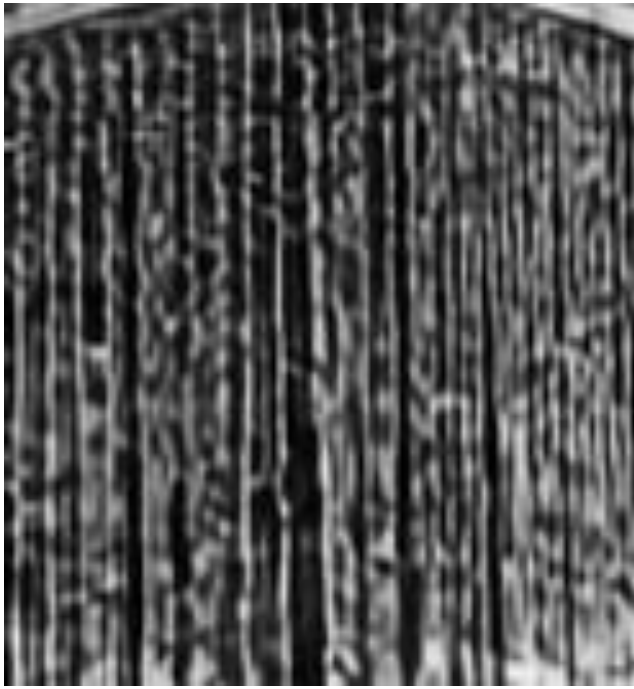
Texture distortion



Original



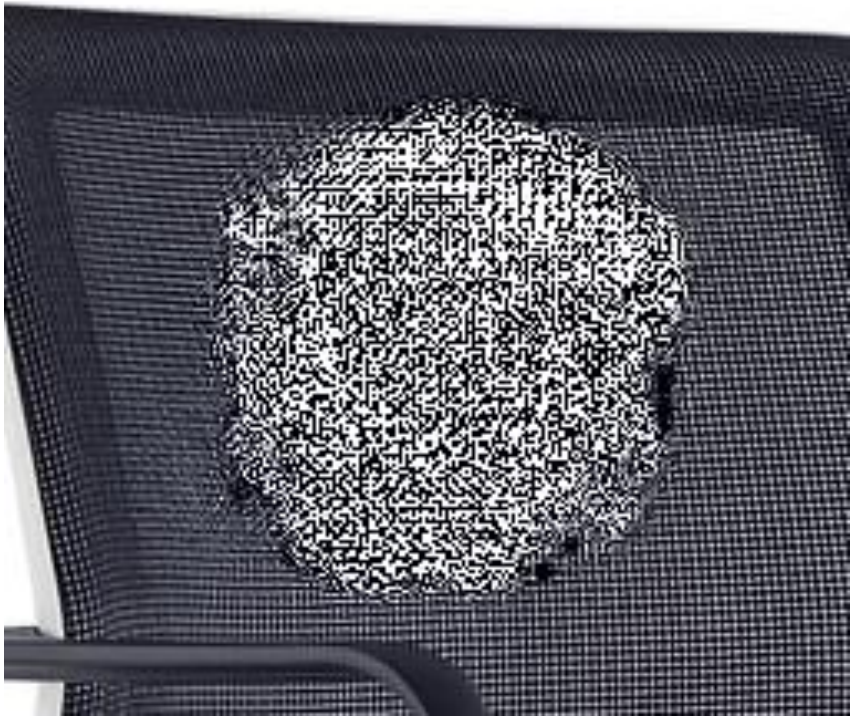
VTM 20.0,
208.7 times



JPEG AI 4.6 tools on, high,
204.1 times

Image corrupted by JPEG AI

Original



JPEG AI 4.6 tools on, high,
24.1 times

Tested attacks

- 6 types of existing adversarial attacks on NNs + a random noise
- 6 different losses to guide attacks: 5 losses to attack quality of the decoc image, and 1 to increase bitrate of encoded image

	Attack	Paper
1	FTDA	T.Chen, 2023
2	I-FGSM	A.Kurakin, 2018
3	MADC	Z.Wang, 2008
4	PGD	A.Madry, 2019
5	SSAH	C.Luo, 2022
6	CAdv	A.Bhattad, 2019
7	Random noise	

	Optimisation target	Formula
1	FTDA default	$\|\|f(x) - f(x')\|\|_2$
2	Added-noises	$\|\|f(x') - f(x) - (x' - x)\|\|_2$
3	Reconstruction	$\|\|f(x') - x'\|\|_2$
4	FTDA-msssim	$\mathrm{SSIM}(x', f(x'))$
5	Reconstruction_msssim	$\text{MS-SSIM}(x', f(x'))$
6	BPP increase	$\mathrm{bpp}(f(x'))$

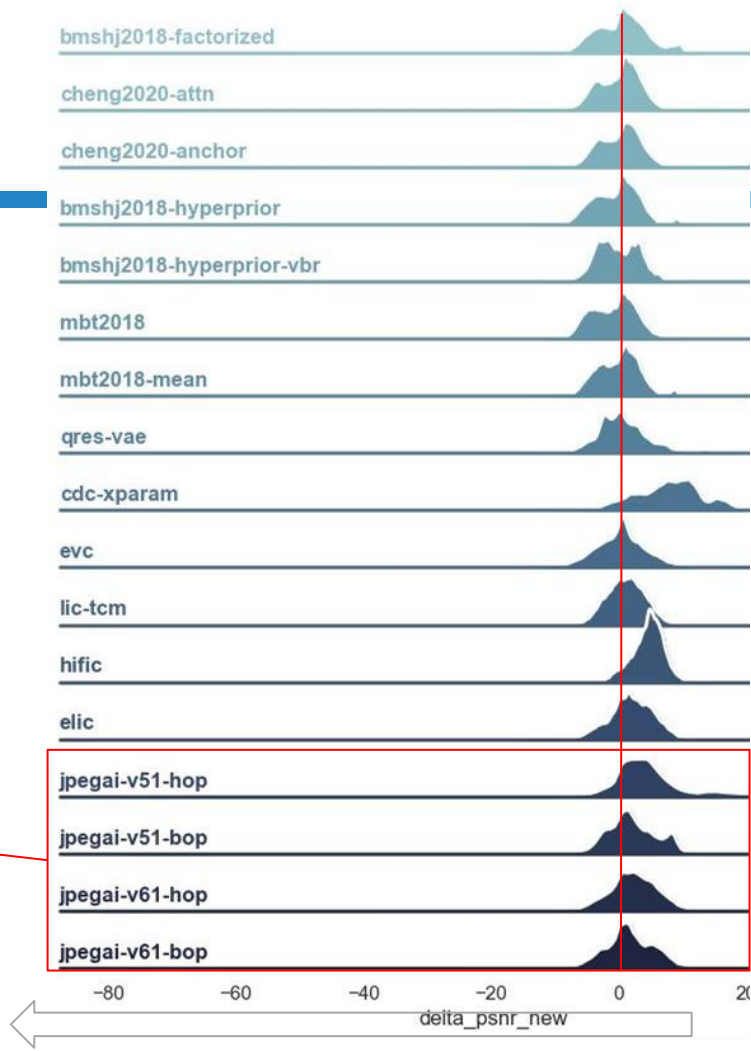
Preliminary results

Codecs' robustness

Worst robustness: cdc-xparam (diffusion-based), Hific

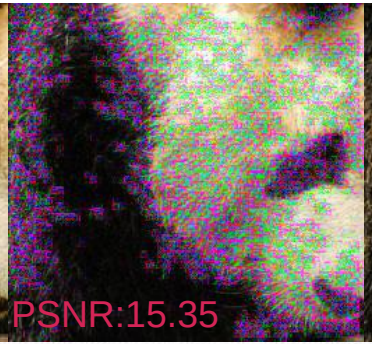
Best robustness: Qres-VAE

JPEG AI shows average robustness, HOP (high operation point) mode is more vulnerable to attacks



JPEG AI v5.1 (hop) examples

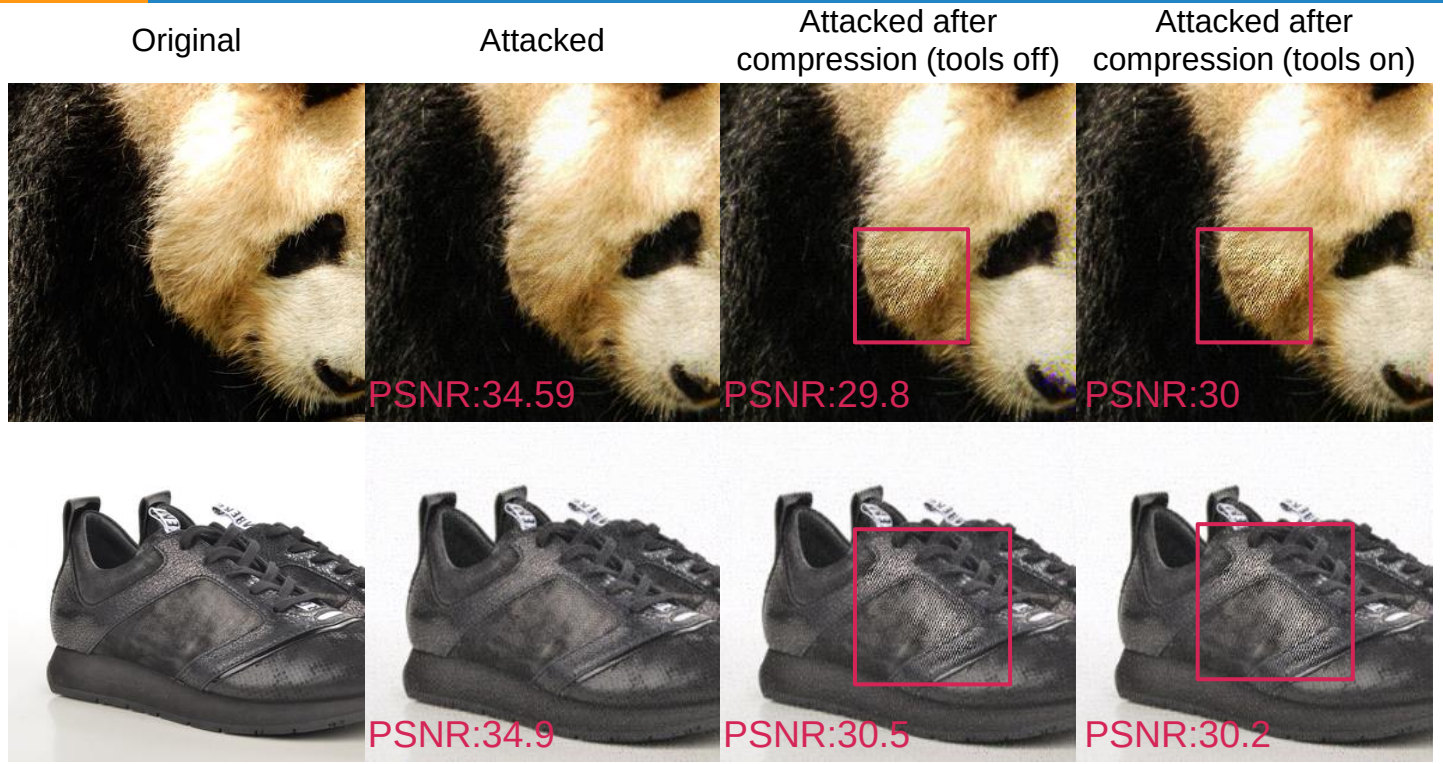
MADC
Added-noises

Original	Attacked	Attacked after compression (tools off)	Attacked after compression (tools on)
	 PSNR:34.6	 PSNR:15.35	 PSNR:32
	 PSNR:34.9	 PSNR:22.2	 PSNR:22.9

MSU Institute for AI Video Lab
https://videoprocessing.ai/

JPEG AI v6.1 (hop) examples

MADC
Added-noises



MSU Institute for AI Video Lab
https://videoprocessing.ai/

Cheng2020 anchor examples

MADC
Added-noises

Original	Attacked	Attacked after compression
	 PSNR:35.68	 PSNR:19.5
 PSNR:35.8		 PSNR:14.8

JPEG AI v5.1 (bop) examples

MADC
Added-noises-Y

Original

Attacked

MSU Institute for AI Video Lab
https://videoprocessing.ai/

JPEG AI v6.1 (bop) examples

MADC
Added-noises-Y

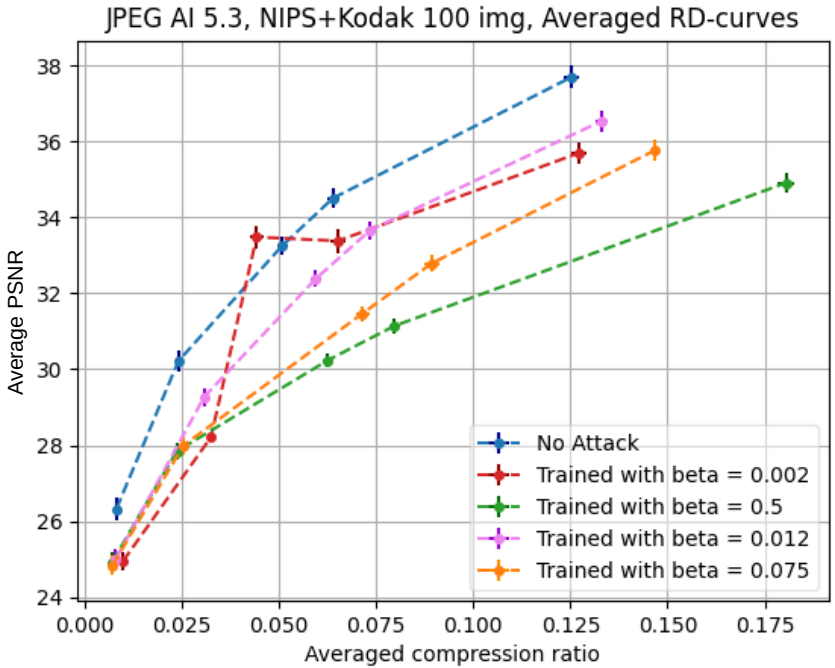
Original

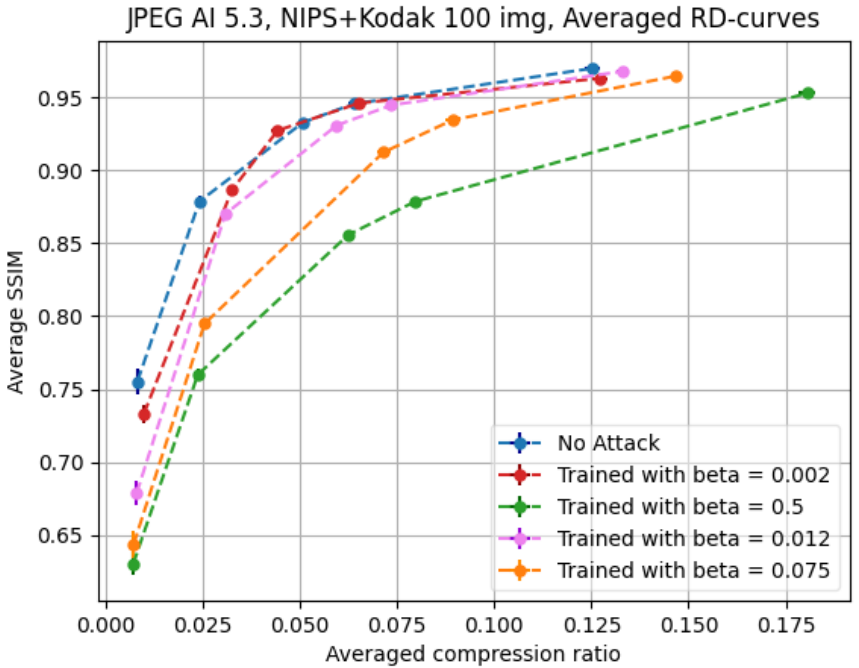
Attacked

JPEG AI (hop) v5.1 vs v6.1 examples

MSU Institute for AI Video Lab
https://videoprocessing.ai/

Attacks that increase bitrate





Итого

- JPEG AI будет жать в 3-3,5 раза лучше JPEG
- Возможны необычные артефакты
- Возможны атаки на создание артефактов
- Возможны атаки на увеличение размера
- Возможны защиты, в том числе встроенные в контур сжатия

Качество библиотек JPEG AI будет заметно отличаться