Stochastic Digital Twin for Copy Detection Patterns

Y. Belousov, O. Taran, V. Kinakh and S. Voloshynovskiy

Department of Computer Science, University of Geneva, Switzerland

UNIVERSITÉ DE GENÈVE

IEEE Signal Processing Society

IEEE WIFS 2023 December 6, 2023

Copy detection pattern (CDP) integration: a practical example

Figure 1: CDP can be naturally integrated into many types of packaging[1]

[1] Justin Picard, Paul Landry, and Michael Bolay. "Counterfeit Detection with QR Codes". In: *Proceedings of the 21st ACM Symposium on Document Engineering*. DocEng '21. Limerick, Ireland: Association for Computing Machinery, 2021.

Copy detection patterns' applications

- A product authentication service
- Securing identification documents
- Packaging authenticating (pharmaceutical, etc).

CDP lifecycle

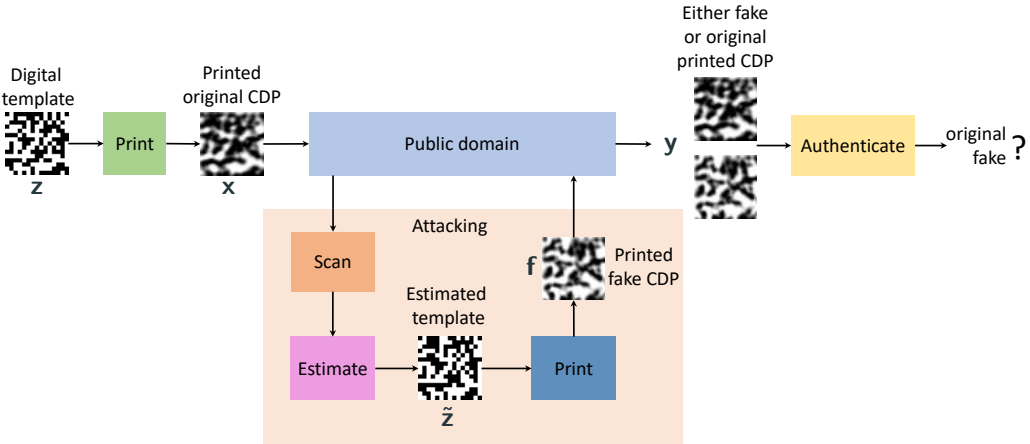


Figure 2: CDP: the general pipeline

Indigo 1x1 base and variability datasets

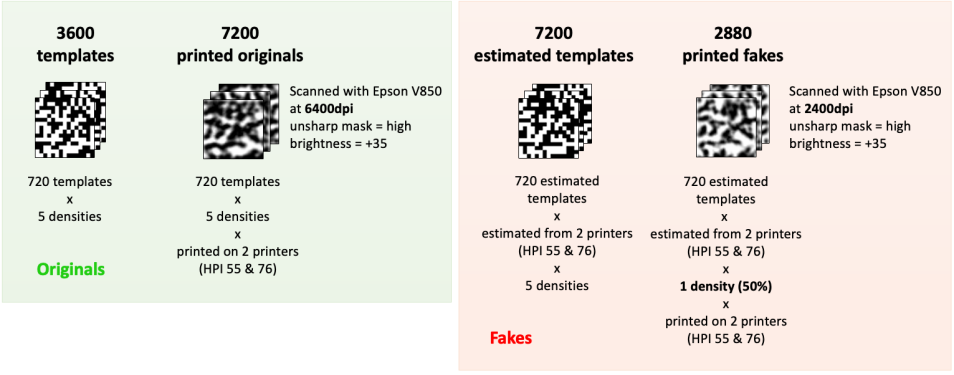


Figure 3: Indigo 1x1 base dataset[2,3]

[2] Roman Chaban et al. "Machine learning attack on copy detection patterns: are 1x1 patterns cloneable?" In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Montpellier, France, Dec. 2021.

[3] Roman Chaban et al. "Printing variability of copy detection patterns". In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Shanghai, China, Dec. 2022.

Turbo digital twin motivation

- The production of datasets of real CDP is a costly and time-consuming process
- The lack of an accurate mathematical model leads to a need to collect a huge amount of data
- Existing mathematical models of printing-imaging channel are non-differentiable

(a) Direct path

(b) Reverse path

Figure 4: *Turbo digital twin* system: direct and reverse paths[4]

[4] Yury Belousov et al. “Digital twins of physical printing-imaging channel”. In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Shanghai, China, Dec. 2022.

Stochastic Digital Twin motivation

- Turbo model is deterministic by nature, i.e. for one input there is one and only one output
  1. *"provided Gaussian noise z as an input to the generator, in addition to x ... the generator simply learned to ignore the noise"* – pix2pix paper[5]
  2. *"The current model does not take z as input. In both pix2pix and CycleGAN, we tried to add z to the generator but often found that z got ignored. So we decided to only take real_A as input"* – one of authors of CycleGAN[6]
- But the printing and acquisition processes are stochastic due to printer condition, type of paper, dot gain, lighting, ...

[5] Phillip Isola et al. "Image-to-Image Translation with Conditional Adversarial Networks". In: *CVPR* (2017).
[6] Jun-Yan Zhu. *Random noise Z augmentation with cgan · issue #152 · JUNYANZ/Pytorch-Cyclegan-and-pix2pix*.

DDPM (Palette) scheme

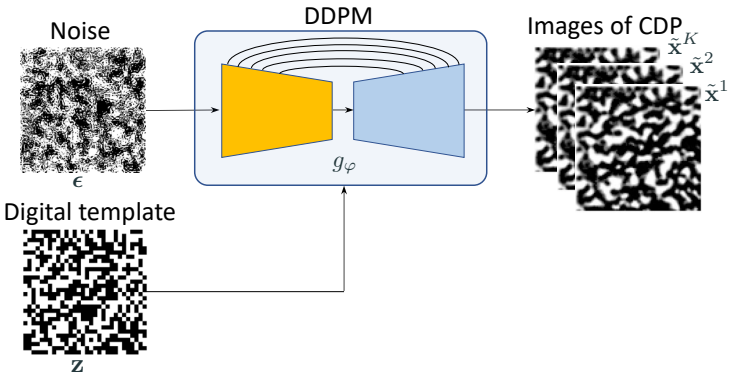


Figure 5: Schematic block-diagram of DDPM[7] generative model g_φ based on Palette[8] architecture

[7] Jonathan Ho, Ajay Jain, and Pieter Abbeel. "Denoising diffusion probabilistic models". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 6840–6851.

[8] Chitwan Saharia et al. "Palette: Image-to-image diffusion models". In: *ACM SIGGRAPH 2022 Conference Proceedings*. 2022, pp. 1–10.

Losses

- Turbo:

$$\begin{aligned}\mathcal{L}^{\text{Turbo}}(\phi,\theta) = \mathcal{L}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}}) + \mathcal{D}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}}) \\ + \lambda_D \mathcal{L}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}}) + \lambda_D \mathcal{D}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}}) \\ + \lambda_T \mathcal{L}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}}) + \lambda_T \mathcal{D}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}}) \\ + \lambda_T \lambda_R \mathcal{L}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}}) + \lambda_T \lambda_R \mathcal{D}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}})\end{aligned}$$

1. $\mathbf{x}$ — the image of CDP, $\mathbf{z}$ — the digital template
2. $\tilde{\mathbf{x}}$, $\tilde{\mathbf{z}}$ — the generated estimation images, $\hat{\mathbf{x}}$, $\hat{\mathbf{z}}$ — the reconstructions
3. $\mathcal{L}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}})$, $\mathcal{L}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}})$, $\mathcal{L}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}})$ and $\mathcal{L}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}})$ — ℓ_1-norm pair-wise losses
4. $\mathcal{D}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}})$, $\mathcal{D}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}})$, $\mathcal{D}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}})$ and $\mathcal{D}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}})$ — adversarial losses between the corresponding distributions
5. λ_T, λ_D and λ_R — trade-off parameters.

Losses

- Turbo:

$$\begin{aligned}\mathcal{L}^{\text{Turbo}}(\phi,\theta) = \mathcal{L}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}}) + \mathcal{D}_{\tilde{\mathbf{z}}}(\mathbf{z},\tilde{\mathbf{z}}) \\ + \lambda_D \mathcal{L}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}}) + \lambda_D \mathcal{D}_{\hat{\mathbf{x}}}(\mathbf{x},\hat{\mathbf{x}}) \\ + \lambda_T \mathcal{L}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}}) + \lambda_T \mathcal{D}_{\tilde{\mathbf{x}}}(\mathbf{x},\tilde{\mathbf{x}}) \\ + \lambda_T \lambda_R \mathcal{L}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}}) + \lambda_T \lambda_R \mathcal{D}_{\hat{\mathbf{z}}}(\mathbf{z},\hat{\mathbf{z}})\end{aligned}$$

- DDPM (Palette):

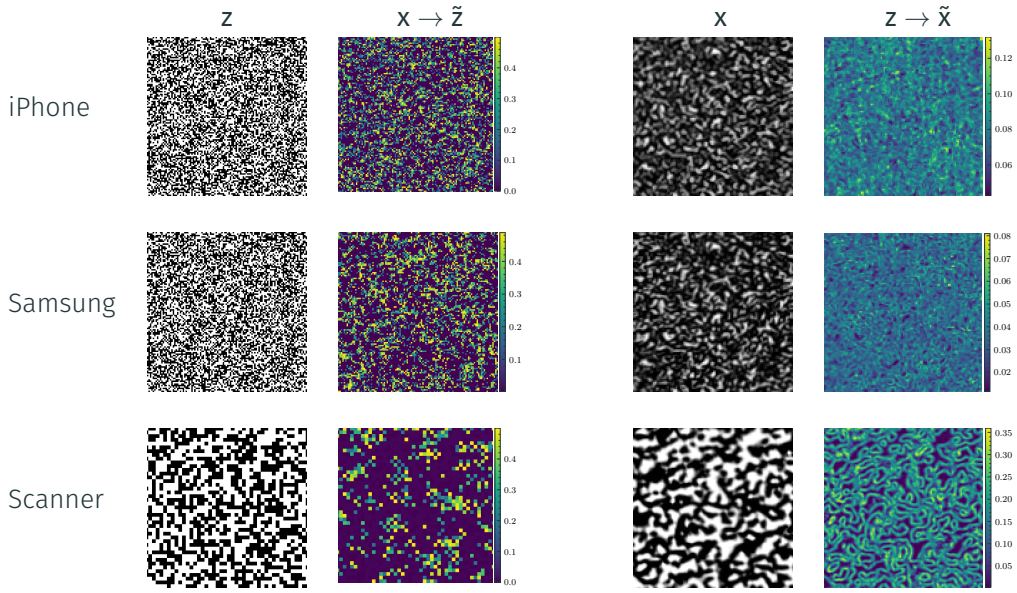
$$\mathcal{L}^{DDPM}(\varphi) = \mathbb{E}_{t,\mathbf{z},\mathbf{x},\boldsymbol{\epsilon}}\left[\left\|\boldsymbol{\epsilon} - g_\varphi\left(\sqrt{\bar{\alpha}_t}\mathbf{x} + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, \mathbf{z}, t\right)\right\|^2\right].$$

1. $\mathbf{x}$ — the image of CDP, $\mathbf{z}$ — the digital template
2. $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ — the noise added at step t
3. g_φ — the parametrized denoiser model
4. $\bar{\alpha}_t$ — the noise scale parameter

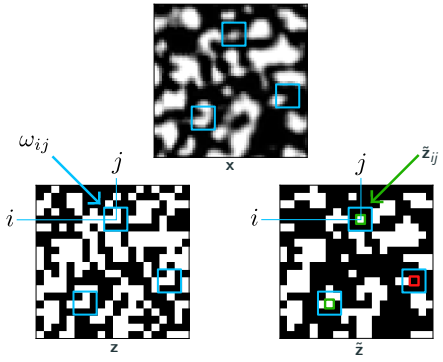
Stochasticity of the DDPM model

Figure 6: The stochasticity in the DDPM Model. Original $\mathbf{z}$ and $\mathbf{x}$ are in the first column, followed by the stochastic estimations $\{\tilde{\mathbf{x}}^k\}_{k=1}^5$ and $\{\tilde{\mathbf{z}}^k\}_{k=1}^5$.

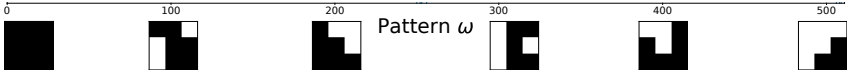
2D variability for the same template



Patterns

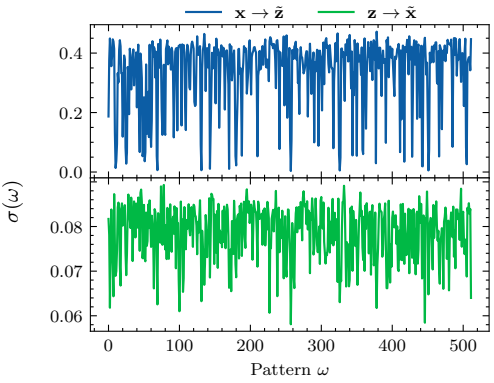


(a) Illustration of pattern

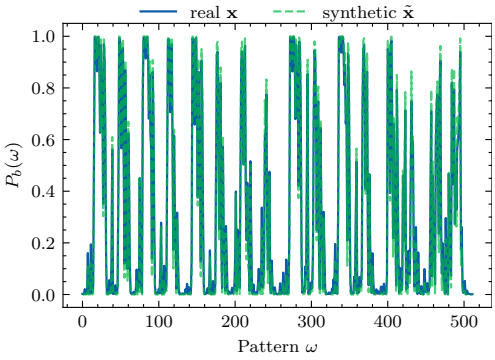


(b) Examples of patterns

Stochasticity per pattern



(a) Standard deviation of the central pixel



(b) Probability of bit-flipping for the central pixel

Figure 8: Properties of central pixel of pattern ω for the 512 different possible patterns ω ordered by their flattened binary representations.

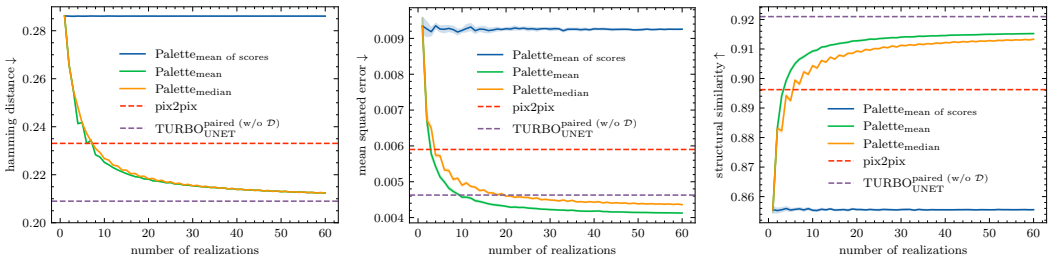
Metrics

- Hamming distance $d_H(\mathbf{z}, \mathbf{binary}(\tilde{\mathbf{z}}))$, where $\mathbf{binary}(.)$ is a binarization function
- Mean square error (MSE) distance $d_2(\mathbf{x}, \tilde{\mathbf{x}})$
- Structural similarity index (SSIM) $d_{SSIM}(\mathbf{x}, \tilde{\mathbf{x}})$[9]
- Fréchet Inception Distance (FID): $\text{FID}_{\,z \to \tilde{x}}$ and $\text{FID}_{\,x \to \tilde{z}}$ proposed in[10]

[9] Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.

[10] Martin Heusel et al. "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017.

Impact of the number of realizations



(a) Hamming distance (b) MSE (c) SSIM

Figure 9: Impact of the number of realizations on different metrics for the iPhone dataset.

Models performance on the iPhone data

Model	FID x→z̃	Hamm. dist.	FID z→x̃	MSE	SSIM
W/O processing	*289.68*	*0.300*	*289.68*	*0.254*	*0.249*
pix2pix	11.82	0.232	11.64	0.005	0.910
CycleGAN	20.69	0.268	12.59	0.014	0.782
$\text{TURBO}^{\text{paired (w } \mathcal{D})}_{\text{CNN-RESNET-CNN}}$	6.56	0.239	10.20	0.005	0.915
$\text{TURBO}^{\text{paired (w/o } \mathcal{D})}_{\text{UNET}}$	35.35	**0.210**	12.22	**0.004**	**0.925**
$\text{Palette}_{\text{mean}_{21}}$	**4.64**	0.211	**9.00**	**0.004**	0.915

Models performance on the Samsung data

Model	FID x→z̃	Hamm. dist.	FID z→x̃	MSE	SSIM
W/O processing	*381.44*	*0.314*	*381.44*	*0.278*	*0.193*
pix2pix	8.53	0.241	20.18	**0.004**	0.908
CycleGAN	8.85	0.283	22.85	0.015	0.694
$\text{TURBO}^{\text{paired (w } \mathcal{D})}_{\text{CNN-RESNET-CNN}}$	7.01	0.247	17.34	**0.004**	0.914
$\text{TURBO}^{\text{paired (w/o } \mathcal{D})}_{\text{UNET}}$	54.80	**0.211**	28.88	**0.004**	**0.922**
$\text{Palette}_{\text{mean}_{21}}$	**4.46**	0.215	**10.72**	**0.004**	0.908

Models performance on the Scanner data

Model	FID x→z̃	Hamm. dist.	FID z→x̃	MSE	SSIM
W/O processing	*304.13*	*0.238*	*304.01*	*0.181*	*0.480*
pix2pix	3.37	0.111	8.57	0.045	0.758
CycleGAN	3.87	0.155	4.45	0.049	0.732
$\text{TURBO}^{\text{paired (w } \mathcal{D})}_{\text{CNN-RESNET-CNN}}$	3.16	0.086	6.60	0.040	**0.779**
$\text{TURBO}^{\text{paired (w/o } \mathcal{D})}_{\text{UNET}}$	6.21	0.100	28.110	**0.036**	0.778
$\text{Palette}_{\text{mean}_{21}}$	**2.90**	**0.081**	**4.36**	0.038	0.769

Conclusions

- Extended the digital twin system for the simulation of the physical printing-imaging channel to support *stochasticity*

Conclusions

- Extended the digital twin system for the simulation of the physical printing-imaging channel to support *stochasticity*
- **Showed that the synthetic CDP images produced by DDPM reflect the natural randomness of the printing process**

Conclusions

- Extended the digital twin system for the simulation of the physical printing-imaging channel to support *stochasticity*
- Showed that the synthetic CDP images produced by DDPM reflect the natural randomness of the printing process
- **Compared the proposed system with several state-of-the-art methods used for image-to-image translation applications.**

Future work

- The investigation of more advanced sampling techniques at the training and inference stages for the improvement of DDPM complexity (30 minutes for 1 realization for DDPM compared to just 15 seconds for Turbo)

Future work

- The investigation of more advanced sampling techniques at the training and inference stages for the improvement of DDPM complexity (30 minutes for 1 realization for DDPM compared to just 15 seconds for Turbo)
- **Use the generated examples to build a classifier based on the augmented synthetic samples of both original and fake CDP**

Future work

- The investigation of more advanced sampling techniques at the training and inference stages for the improvement of DDPM complexity (30 minutes for 1 realization for DDPM compared to just 15 seconds for Turbo)
- Use the generated examples to build a classifier based on the augmented synthetic samples of both original and fake CDP

- **Content-style disentanglement for printer-specific codes**

Future work

- The investigation of more advanced sampling techniques at the training and inference stages for the improvement of DDPM complexity (30 minutes for 1 realization for DDPM compared to just 15 seconds for Turbo)
- Use the generated examples to build a classifier based on the augmented synthetic samples of both original and fake CDP
- Content-style disentanglement for printer-specific codes
- **Adversarial attacks in the physical world.**

Thank you for your attention!

Data:
http://sip.unige.ch/projects/snf-it-dis/datasets/

Code:
https://gitlab.unige.ch/sip-group/stochastic-digital-twin

Paper:
http://sip.unige.ch/articles/2023/Belousov_WIFS2023.pdf

Document identification: FIFA World Cup 2006

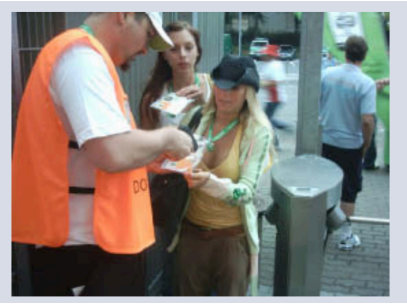
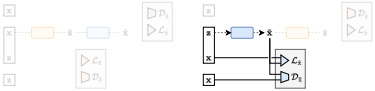


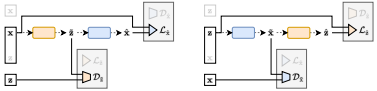
Figure 10: ID control at the turnstiles

Figure 11: ID card with access legitimation

Turbo as Generalisation of Many Models



(a) pix2pix



(b) CycleGAN

Figure 12: The pix2pix and CycleGAN architectures expressed in the TURBO framework.

References i

[1] Justin Picard, Paul Landry, and Michael Bolay. “Counterfeit Detection with QR Codes”. In: *Proceedings of the 21st ACM Symposium on Document Engineering*. DocEng ’21. Limerick, Ireland: Association for Computing Machinery, 2021.

[2] Roman Chaban et al. “Machine learning attack on copy detection patterns: are 1x1 patterns cloneable?” In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Montpellier, France, Dec. 2021.

[3] Roman Chaban et al. “Printing variability of copy detection patterns”. In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Shanghai, China, Dec. 2022.

[4] Yury Belousov et al. “Digital twins of physical printing-imaging channel”. In: *IEEE International Workshop on Information Forensics and Security (WIFS)*. Shanghai, China, Dec. 2022.

References ii

[5] Phillip Isola et al. “Image-to-Image Translation with Conditional Adversarial Networks”. In: *CVPR* (2017).

[6] Jun-Yan Zhu. *Random noise Z augmentation with cgan · issue #152 · JUNYANZ/Pytorch-Cyclegan-and-pix2pix*.

[7] Jonathan Ho, Ajay Jain, and Pieter Abbeel. “Denoising diffusion probabilistic models”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 6840–6851.

[8] Chitwan Saharia et al. “Palette: Image-to-image diffusion models”. In: *ACM SIGGRAPH 2022 Conference Proceedings*. 2022, pp. 1–10.

[9] Zhou Wang et al. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.

References iii

[10] Martin Heusel et al. “GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium”. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017.