

Supplementary Material for Sliced Wasserstein Generative Models

Jiqing Wu^{*1} Zhiwu Huang^{*1} Dinesh Acharya¹ Wen Li¹
 Janine Thoma¹ Danda Pani Paudel¹ Luc Van Gool^{1,2}

¹ Computer Vision Lab, ETH Zurich, Switzerland ² VISICS, KU Leuven, Belgium

{jwu, zhiwu.huang, liwen, jthoma, paudel, vangoool}@vision.ee.ethz.ch

^{*}Equal contribution

A Proof of Theorem 1

In order to prove Theorem 1, we need to apply the DvoretzkyKieferWolfowitz inequality,

Theorem. (DKW-inequality) Given $b \in \mathbb{N}$, let Z_1, Z_2, \dots, Z_b be real-valued i.i.d. random variables with a continuous CDF F_Z . Then we define the associated EDF $F_{Z,b}(t) = \frac{1}{b} \sum_{i=1}^b \mathbf{1}_{\{Z_i \leq t\}}$, then for all $\varepsilon \geq \sqrt{\frac{1}{2b} \ln 2}$ it holds

$$\Pr(\|F_{Z,b}(t) - F_Z(t)\|_\infty > \varepsilon) \leq e^{-2b\varepsilon^2}. \quad (1)$$

Based on the DKW-inequality we have the error estimation for Alg. 1:

Theorem 1. Given $b \in \mathbb{N}$, let Z_1, Z_2, \dots, Z_b be real-valued i.i.d. random variables with a continuous CDF F_Z^{-1} with domain $[0, 1]$. Then we define the associated EDF $F_{Z,b}^{-1}(t) = \frac{1}{b} \sum_{i=1}^b \mathbf{1}_{\{Z_i \leq t\}}$. Assume \tilde{F}_Y, F_Y are CDFs satisfying $\|\tilde{F}_Y - F_Y\|_\infty \leq \gamma$, then there exists a $\delta > 0$ such that for all $\varepsilon - \delta\gamma \geq \sqrt{\frac{1}{2b} \ln 2}$ it holds

$$\Pr(\|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}F_Y(t)\|_\infty > \varepsilon) \leq e^{-2b(\varepsilon - \delta\gamma)^2}. \quad (2)$$

Proof Let $\tilde{Z}_i = \tilde{F}_Y^{-1}Z_i$, it is not hard to see that $\tilde{Z}_1, \tilde{Z}_2, \dots, \tilde{Z}_b$ are i.i.d random variables with CDF $F_Z^{-1}\tilde{F}_Y$, since

$$F_Z^{-1}\tilde{F}_Y(t) = \Pr(Z_i < \tilde{F}_Y(t)) = \Pr(\tilde{F}_Y^{-1}Z_i < t). \quad (3)$$

Accordingly, we have the \tilde{Z}_i associated EDF

$$\tilde{F}_{Z,b}^{-1}(t) = \frac{1}{b} \sum_{i=1}^b \mathbf{1}_{\{\tilde{Z}_i \leq t\}} = \frac{1}{b} \sum_{i=1}^b \mathbf{1}_{\{Z_i \leq \tilde{F}_Y(t)\}} = F_{Z,b}^{-1}\tilde{F}_Y(t). \quad (4)$$

By applying the DKW-inequality it holds for all $\varepsilon \geq \sqrt{\frac{1}{2b} \ln 2}$

$$\Pr(\|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}\tilde{F}_Y(t)\|_\infty > \varepsilon) \leq e^{-2b\varepsilon^2}. \quad (5)$$

Since F_Z^{-1} is continuous with a compact convex domain $[0, 1]$, F_Z^{-1} satisfies Lipschitz continuity with a Lipschitz constant δ and it holds

$$\begin{aligned} & \|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}F_Y(t)\|_\infty \\ & \leq \|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}\tilde{F}_Y(t)\|_\infty + \|F_Z^{-1}\tilde{F}_Y(t) - F_Z^{-1}F_Y(t)\|_\infty \\ & \leq \|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}\tilde{F}_Y(t)\|_\infty + \delta\gamma, \end{aligned} \quad (6)$$

then for $\omega \in \{\|F_{Z,b}^{-1}\tilde{F}_Y - F_Z^{-1}F_Y\|_\infty > \varepsilon\}$ we have

$$\omega \in \{\|F_{Z,b}^{-1}\tilde{F}_Y - F_Z^{-1}\tilde{F}_Y\|_\infty > \varepsilon - \delta\gamma\}. \quad (7)$$

Therefore, for all $\varepsilon - \delta\gamma \geq \sqrt{\frac{1}{2b} \ln 2}$ we have

$$\Pr(\|F_{Z,b}^{-1}\tilde{F}_Y(t) - F_Z^{-1}F_Y(t)\|_\infty > \varepsilon) \leq e^{-2b(\varepsilon - \delta\gamma)^2}. \quad (8)$$

B Complementary study of k, k'

k	k'	0	0.1	1
10^{-4}		14.8	15.2	14.9
10^{-3}		13.0	15.0	15.5
0.1		20.9	17.9	16.7
1		22.0	22.2	21.5
10		22.5	19.3	22.7

Figure 1. Complementary FID scores to Fig. 3 (Left h).

C More Hyperparameter Studies

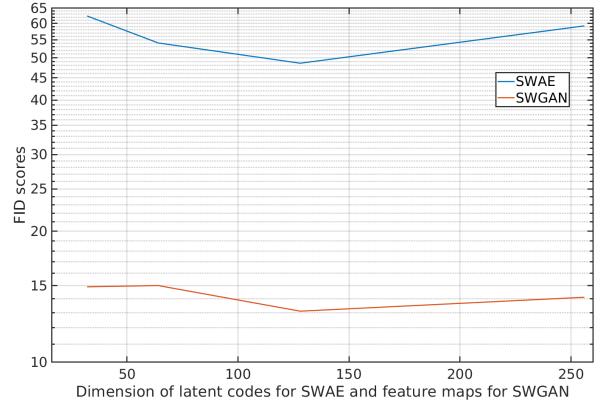
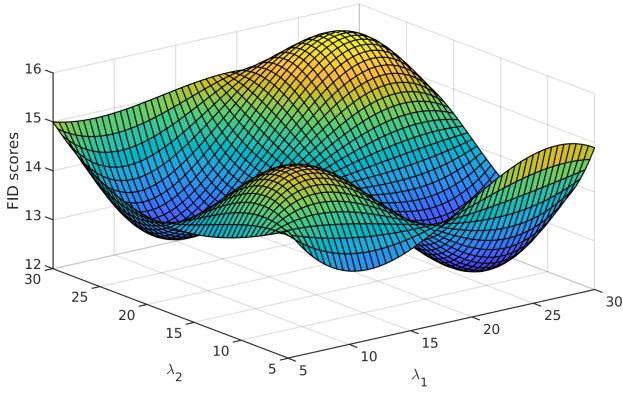


Figure 2. The FID scores of various λ_1, λ_2 (left) and dimensions of latent codes for SWAE and feature maps for SWGAN (right). We can see that the FID scores are not very sensitive to the changes of λ_1, λ_2 , while the optimal dimension for SWAE and SWGAN is 128.

D Model Stability Study

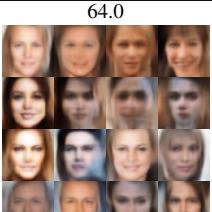
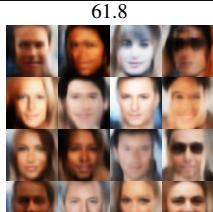
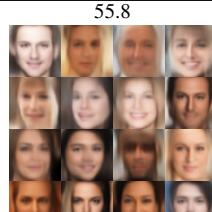
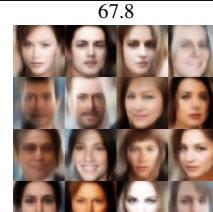
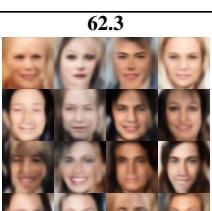
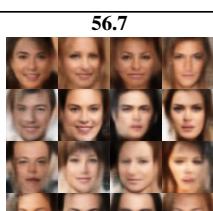
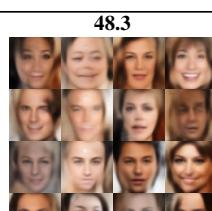
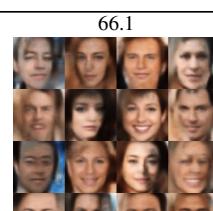
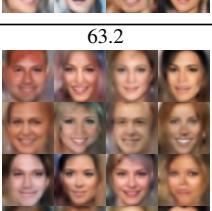
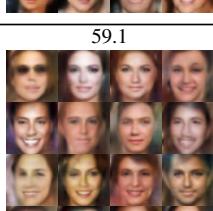
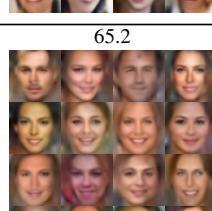
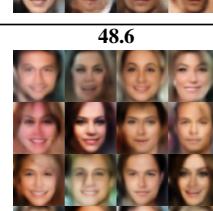
	ResNet (w/ norm)	ResNet (w/o norm)	ConvNet (w/ norm)	ConvNet (w/o norm)
CTGAN	16.0	16.5	19.5	19.7
				
SWG	24.3	29.1	22.2	28.5
				
SWGAN	13.0	14.8	19.2	18.8
				
WAE-MMD	64.0	61.8	55.8	67.8
				
AAE (WAE-GAN)	62.3	56.7	48.3	66.1
				
SWAE	63.2	59.1	65.2	48.6
				

Table 1. FID scores of various architectures on CelebA. The optimal architectures are ConvNet for SWG, WAE-MMD, AAE (WAE-GAN), ResNet for CTGAN, SWGAN, and ConvNet without normalization (w/o norm) for SWAE.

E Architecture Details of Our Proposed Models

Encoder	Kernel size	Resampling	Output shape
NearestNeighbor	–	Down	$3 \times 16 \times 16$
Linear	–	–	128
3 Primal SWD blocks	–	–	128
Decoder			
Noise	–	–	128
Linear	–	–	$64 \times 8 \times 8$
2 (Conv, ELU) blocks	3×3	–	$64 \times 8 \times 8$
NearestNeighbor	–	Up	$64 \times 16 \times 16$
2 (Conv, ELU) blocks	3×3	–	$64 \times 16 \times 16$
NearestNeighbor	–	Up	$64 \times 32 \times 32$
2 (Conv, ELU) blocks	3×3	–	$64 \times 32 \times 32$
NearestNeighbor	–	Up	$64 \times 64 \times 64$
2 (Conv, ELU) blocks	3×3	–	$64 \times 64 \times 64$
Conv, tanh	3×3	–	$3 \times 64 \times 64$

Table 2. Network architecture for SWAE

Generator	Kernel size	Resampling	Output shape
Noise	–	–	128
Linear	–	–	$128 \times 4 \times 4$
2 (Res, ReLU) blocks	3×3	Up	$128 \times 8 \times 8$
2 (Res, ReLU) blocks	3×3	Up	$128 \times 16 \times 16$
2 (Res, ReLU) blocks	3×3	Up	$128 \times 32 \times 32$
2 (Res, ReLU) blocks	3×3	Up	$128 \times 64 \times 64$
Conv, tanh	3×3	–	$3 \times 64 \times 64$
Discriminator			
2 (Res, ReLU) blocks	3×3	Down	$128 \times 32 \times 32$
2 (Res, ReLU) blocks	3×3	Down	$128 \times 16 \times 16$
2 (Res, ReLU) blocks	3×3	Down	$128 \times 8 \times 8$
2 (Res, ReLU) blocks	3×3	–	$128 \times 8 \times 8$
Linear	–	–	128
4 (Dual, LeakyReLU) SWD blocks	–	–	128

Table 3. Network architecture for SWGAN

F More Visual Results for SWAE, SWGAN and Compared Methods

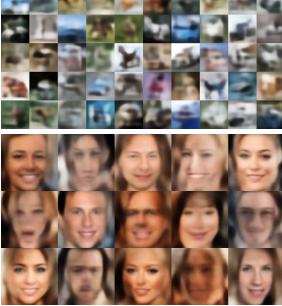
VAE	WAE-MMD	AAE (WAE-GAN)	SWAE
			
SWG	CTGAN	WGAN-GP	SWGAN
			
			

Table 4. Visual results of AE-based (top 2 rows) and GAN (bottom 3 rows) models on CIFAR-10, CelebA and LSUN.



Table 5. Interpolation results of the proposed SWAE (left) and SWGAN (right) models on CelebA.

G More Visual Results for PG-WGAN and PG-SWGAN



Figure 3. Visual results for PG-WGAN and PG-SWGAN on CelebA-HQ and LSUN.