SVD-InST

苏展 吴舒文 运嘉盛

2023年12月25日

风格迁移

Content	InST	CycleGAN	StyTR-2	fast-style-transfer	ours

风格迁移

Content | InST | CycleGAN | StyTR-2 | fast-style-transfer | ours

1 Related Works

2 Methods

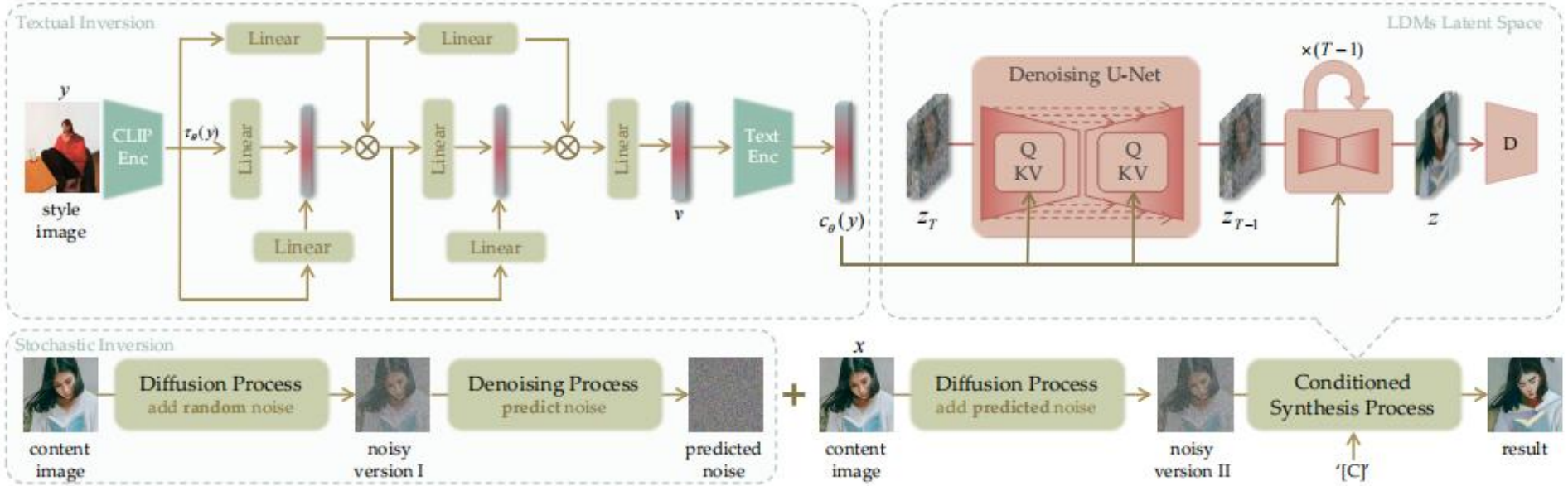
3 Experiments

4 Results

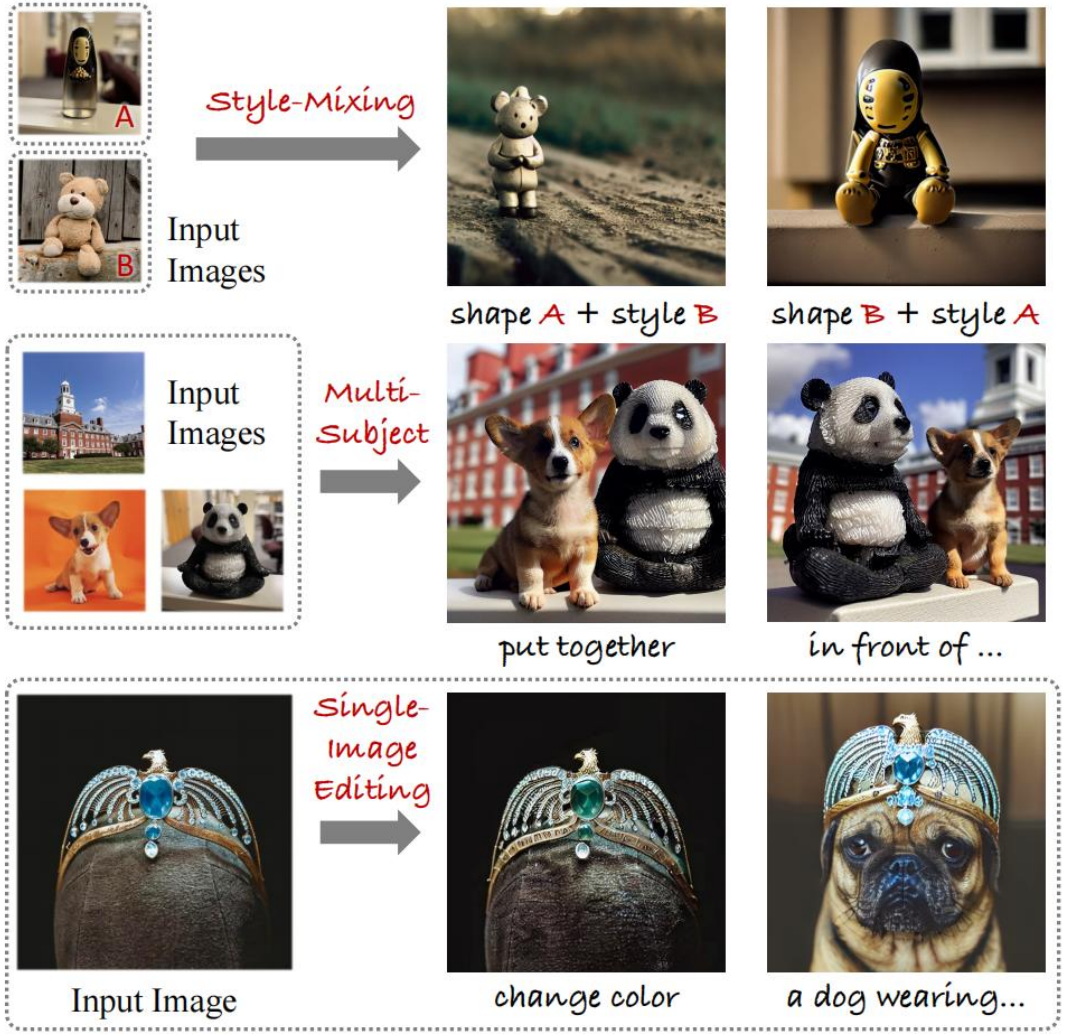
InST(CVPR2023)

Inversion-Based Style Transfer with Diffusion Models

Yuxin Zhang[1,2] Nisha Huang[1,2] Fan Tang[3] Haibin Huang[4]
Chongyang Ma[4] Weiming Dong[1,2*] Changsheng Xu[1,2]
[1]MAIS, Institute of Automation, Chinese Academy of Sciences [2]School of AI, UCAS
[3]Institute of Computing Technology, Chinese Academy of Sciences [4]Kuaishou Technology



SVDiff(ICCV2023)



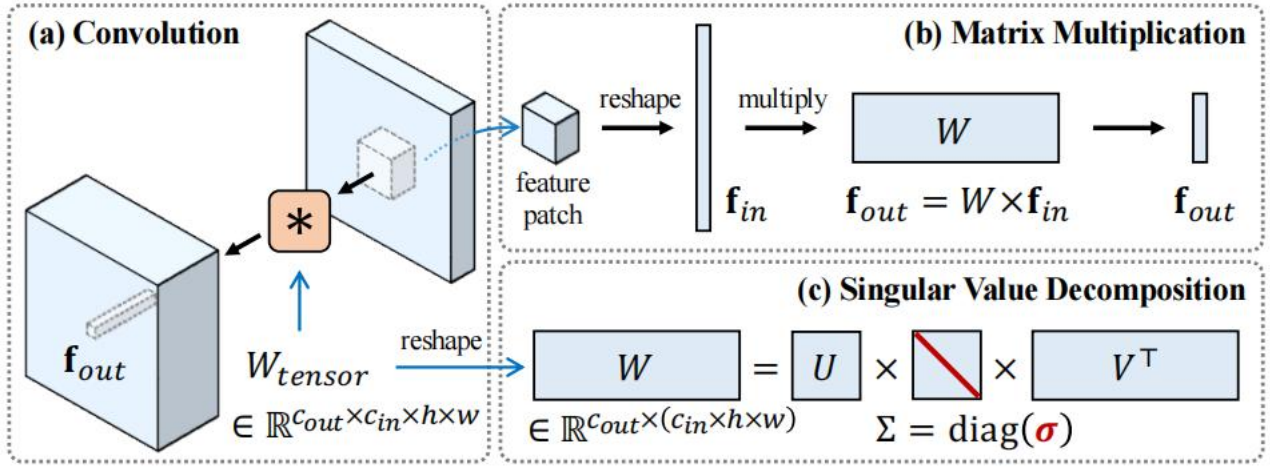


Figure 2. Performing singular value decomposition (SVD) on weight matrices. In an intermediate layer of the model, **(a)** the convolutional weights W_{tensor} **(b)** serve as an associative memory [8]. **(c)** SVD is performed on the reshaped 2-D matrix W.

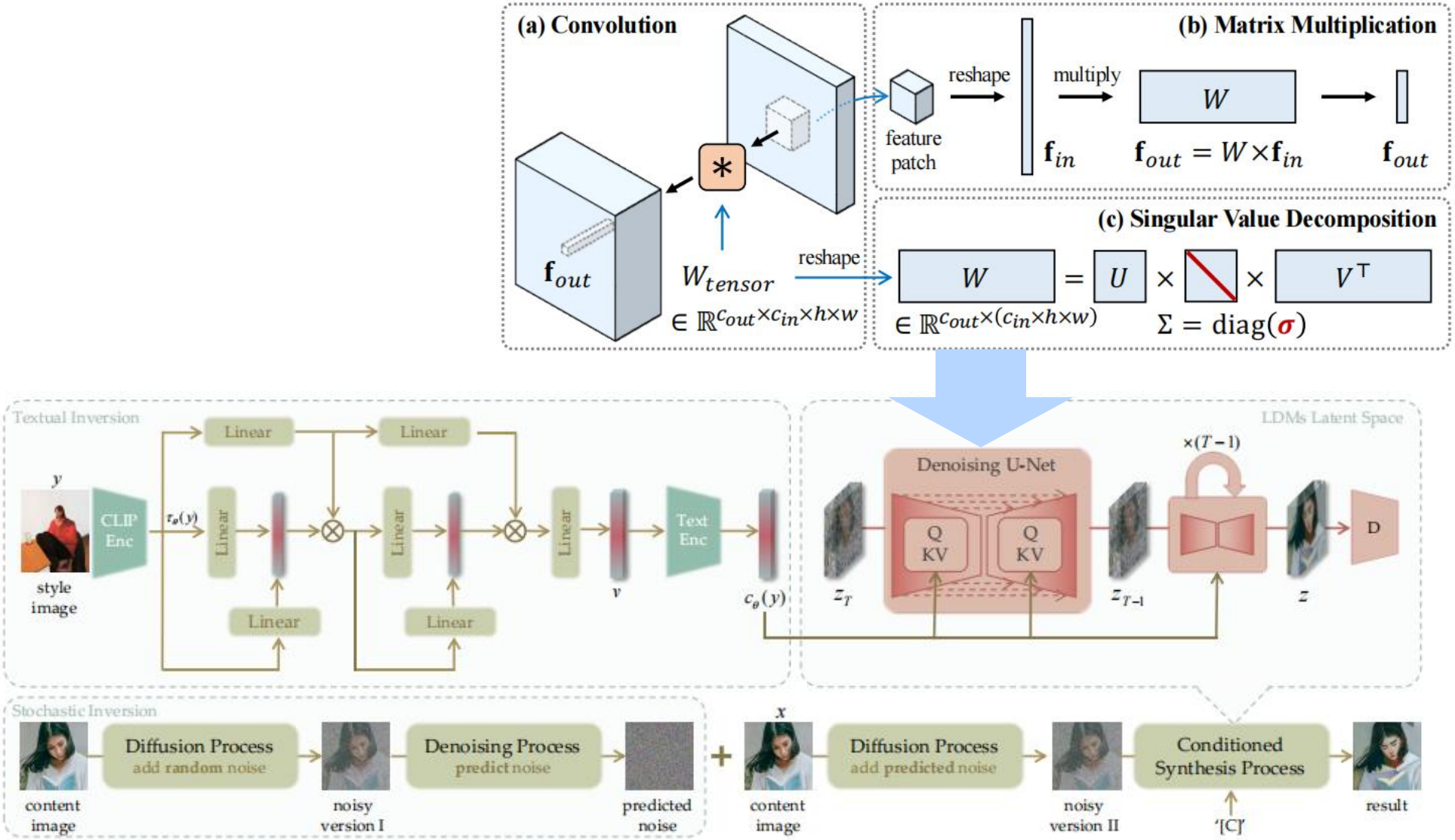
1 Related Works

2 Methods

3 Experiments

4 Results

SVD-InST


(a) Convolution
(b) Matrix Multiplication
(c) Singular Value Decomposition
feature patch
reshape
multiply
W
$\mathbf{f}_{in}$
$\mathbf{f}_{out} = W \times \mathbf{f}_{in}$
$\mathbf{f}_{out}$
W_{tensor}
$\in \mathbb{R}^{c_{out} \times c_{in} \times h \times w}$
$W = U \times \Sigma \times V^{\top}$
$\in \mathbb{R}^{c_{out} \times (c_{in} \times h \times w)}$
$\Sigma = \mathrm{diag}(\boldsymbol{\sigma})$
Textual Inversion
LDMs Latent Space
style image
CLIP Enc
Linear
Text Enc
Denoising U-Net
Q KV
$\times(T-1)$
D
Stochastic Inversion
content image
Diffusion Process
add random noise
noisy version I
Denoising Process
predict noise
predicted noise
Diffusion Process
add predicted noise
noisy version II
Conditioned Synthesis Process
'[C]'
result

1 Related Works

2 Methods

3 Experiments

4 Results

风格迁移

Content	InST	CycleGAN	StyTR-2	fast-style-transfer	ours

风格迁移

Content | InST | CycleGAN | StyTR-2 | fast-style-transfer | ours

Model	FID	LPIPS
InST	127.5	0.53
CycleGAN	178.3	\
StyTR-2	171.3	\
fast-style-transfer	172.7	\
ours	125.1	0.54