Model | 2014以前 ~ 17 | 18 ~ 21 | 22 ~ 24

GAN系 [発表年度]

AR系 [発表年度]

その他系 [発表年度]

評価系 [発表年度]

VAE系 [発表年度]

Flow系 [発表年度]

Neural Style Transfer [CVPR2016]
AdaIN [ICCV2017]
スタイル変換

SAGAN [ICML2019]
BigGAN [ICLR2019]
省データで高効率な学習
StyleGAN2-ADA [NeurIPS2020]
ProjectedGAN [NeurIPS2021]
FastGAN [ICLR2021]
生成画像の高解像度化

GigaGAN [CVPR2023]
StyleGAN-XL [SIGGRAPH2022]
近年のText-to-Image手法との融合
StyleGAN-T [ICML2023]

PGGAN [ICLR2018]
StyleGAN [CVPR2019]
StyleGAN2 [CVPR2020]
StyleGAN3 [NeurIPS2021]
高解像度顔画像生成の代表的手法: StyleGAN ファミリー

GAN Inversionのコンセプト
Invertible cGAN [NeurIPS2016]

LSGAN [ICCV2017]
WGAN [arXiv2017]
WGAN-GP [NeurIPS2017]
Spectral Norm [ICLR2018]
学習安定化手法の確立

Image2StyleGAN [ICCV2019]
pSp [CVPR2021]
StyleSpace [CVPR2021]
StyleCLIP [ICCV2021]
PTI [TOG2022]
StyleGANを利用したInversion/画像編集が多数登場

SAN [ICLR 2024]

GAN [NeurIPS2014]
DCGAN [ICLR2016]
cGAN [arXiv2014]
ACGAN [ICML2016]
条件付き生成手法の確立

FOMM [NeurIPS2019]
StyleRig [CVPR2020]
PIRenderer [ICCV2021]
顔のコントロール手法は年々高解像度化し、Conditionとして様々な情報が利用されるように
Face Vid2Vid [CVPR2021]
StyleHEAT [ECCV2022]
EMOPortraits [CVPR2024]
MegaPortraits [ACM2022]
MetaPortrait [CVPR2023]

CycleGAN [ICCV2017]
Pix2Pix [CVPR2017]
Pix2PixHD [ICCV2017]
UNIT [NeurIPS2017]
Image-to-Image Translation (I2I)の隆盛
X2Face [ECCV2018]
GANimation [ECCV2018]
Vid2Vid [NeurIPS2018]
StarGAN [CVPR2018]

可変解像度画像生成
SinGAN [ICCV2019]
MSG-GAN [CVPR2020]
InfinityGAN [ICLR2021]
Anyres-GAN [ECCV2022]

CIPS [CVPR2021]
Implicit Neural Representationを利用した生成手法も登場
CRESP [CVPR2023]

I2Iの超解像への利用
SRGAN [CVPR2017]
ESRGAN [ECCV2018]

SPADE [CVPR2019]
MaskGAN [CVPR2020]
EditGAN [NeurIPS2021]
インタラクティブ性のある画像生成
DragGAN [SIGGRAPH2023]

Text-to-Image (T2I)生成
StackGAN [ICCV2017]
AttnGAN [CVPR2018]
DM-GAN [CVPR2019]
CPGAN [ECCV2020]
CLIPやBERTなどの導入が始まる
XMC-GAN [CVPR2021]
DF-GAN [CVPR2022]
GALIP [CVPR2023]

VGAN [NeurIPS2016]
TGAN [ICCV2017]
MoCoGAN [CVPR2018]
GANによるVideo Generation

GANによるデータ拡張
DatasetGAN [CVPR2021]

Denoising Diffusion GANs [ICLR2022]
Diffusion Modelの利用
Diffusion-GAN [ICLR2023]

Inception Score [NeurIPS2016]
FID [NeurIPS2017]
KID [ICLR2018]
LPIPS [CVPR2018]
FVD [ICLR2019]
$\overline{\mathrm{IS}}_\infty$ / $\overline{\mathrm{FID}}_\infty$ [CVPR2020]
生成画像の評価指標 (実質的にはFIDがデファクト)
Clean-FID [CVPR2022]

PixelRNN [ICML2016]
PixelCNN [NeurIPS2016]
PixelSNAIL [ICLR2018]
ImageGPT [ICML2020]
VQGAN [CVPR2021]
DALL-E1 [ICML2021]
ViT-VQGAN [ICML2022]
MaskGIT [CVPR2022]
Parti [arXiv2022]
Muse [arXiv2022]

VAE [ICLR2013]
AAE [arXiv2015]
DRAW [ICML2015]
VAE/GAN [ICML2016]
IAF-VAE [NeurIPS2016]
β-VAE [ICLR2017]
VQ-VAE [NeurIPS2017]
dVAE [NeurIPS2018]
VQ-VAE2 [NeurIPS2019]
NVAE [NeurIPS2020]
生成の多段化
Very Deep VAE [ICLR2021]
HQ-VAE [TMLR]
RQ-VAE [CVPR 2022]
SQ-VAE [ICML2022]
DQ-VAE [CVPR 2023]

NICE [ICLR2015]
Normalizing flow [ICML2015]
RealNVP [ICLR2016]
Glow [NeurIPS2018]
FFJORD [ICLR2019]
モデル構造の改良
ALAE [CVPR2020]
高解像度画像生成手法も増加
Flow Straight and Fast [ICLR23]
InstaFlow [ICLR24]
Diffusion Model の Rectified Flowへ応用

※分類は便宜上のものも含む

2014 ~ 17 | 18 ~ 21 | 22 ~ 24

Diffusion系 [発表年度]

モデル構造の発展

Diffusion models [ICML15]
DDPM [NeurIPS20]
Improved DDPM [ICLR21] / ADM [NeurIPS21]
GANの性能上回る!!
LDM (SD) [CVPR23]
潜在空間でモデル構築
UViT / DiT[ICCV2023]
Transformer ベースのモデルの登場
高精細な生成が可能に

Text2Image
GLIDE[ICML22] / DALLE2/ Imagen [Neurips22]
PixArt系/ DALLE3 / SD
テキストに対する忠実性向上

Personalization
DreamBooth / Custom Diffusion [CVPR23]
Textual Inversion[ICLR23] / LoRA [NeurIPS23]
特定の被写体の画像を用意しモデルをFinetuning
IP-Adaptor /Composer /DisenBooth [ICLR24]
EncoderベースのFinetuning不要なカスタマイズ
学習不要で軽量なカスタマイズへ

Higher Controllability
GLIGEN [CVPR23] / ControlNet [ICCV23] / T2I Adapter [AAAI24]
追加入力を受けつけ制御性が向上

生成高速化

SDE / ODE-based formulation
Score-based Generative Modeling / Gotta Go Fast / DDIM [ICLR21]
SDE/ODE ベースの手法の登場
PLMS [ICLR22] / EDS-DGM / DPM-Solver [NeurIPS22]
微分方程式の解法を利用したサンプラーの登場
Consistency Models [ICML23]
軌跡の終点を直接推定
Improved Consistency Model [ICLR24] / CTM [ICLR24] LCM [ICLR24]
Flow Straight and Fast [ICLR23]
直線の軌跡を持つODEを利用
InstaFlow [ICLR24]
学習済みT2Iを用いてReflowし、T2Iを高速化

Distillation
Progressive Distillation [ICLR22]
Guidance Distillation [CVPR23]
Adversarial Diffusion Distillation
SDSを高速化に利用した蒸留手法の登場
1Stepでの高速な生成も可能に

（3D Generation）
DreamFusion / Score Jacobian Chaining / Prolific Dreamer
Score-Distillation Samplingを利用した3D生成手法の登場

Img2Img (Image Editing)
SDEdit [ICLR22]
P2P [ICRL23] / SINE / Instruct P2P / Imagic / Plug-and-Play [CVPR23]
DiffEdit [ICCV23] / Null-text Inversion [CVPR23]
詳細なテキスト指示に基づいた画像編集法
DragDiffusion / DragonDiffusion / DragNoise / FreeDrag [CVPR24]
Drag操作で直感的な編集が可能に
Blended Diffusion / DiffusionCLIP/ Diffusion Auto Encoder [CVPR22]
簡単なtextによる編集 / attribute操作
Blended Latent Diffusion [SIGGRAPH23]
Paint-by-Example / ImagenEditor / SmartBrush [CVPR23]
局所的なエリアの編集法
詳細なテキストによる編集 / ポイントベースの直感的な編集が可能に

Video Generation
MM-Diffusion [CVPR23] Audioと動画の同時生成
Video Diffusion Model [NeurIPS23]
DiffVideoAE /Align your Latents / PVDM / VideoFusion [CVPR23]
Gen-1 [ICCV23] / Make-a-Video [ICLR23]
基盤動画生成モデルの登場
Disco / Animate Anyone / MagicAnimate [CVPR24]
ControlVideo / AnimateDiff [ICLR24] / Sora
時系列一貫性 & 制御性の向上
制御性も高く、高精細な生成が可能に