**Attribute manipulation / Lip-Syncing / Face Reenactment / Indentity Swap / Entire Face Synthesis**

AttGAN – adversarial learning

Attribute manipulation

A diagram of a group of people

Description automatically generated

STGAN

Attribute manipulation

A person with blonde hair and a graph

Description automatically generated

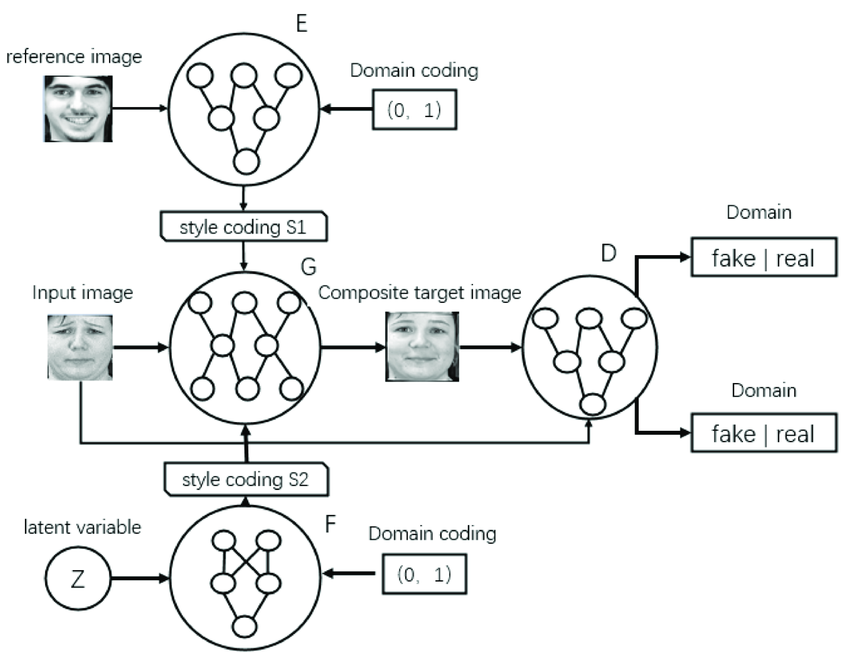
StyleGAN

Entire Face Synthesis – style transfer

A diagram of a computer

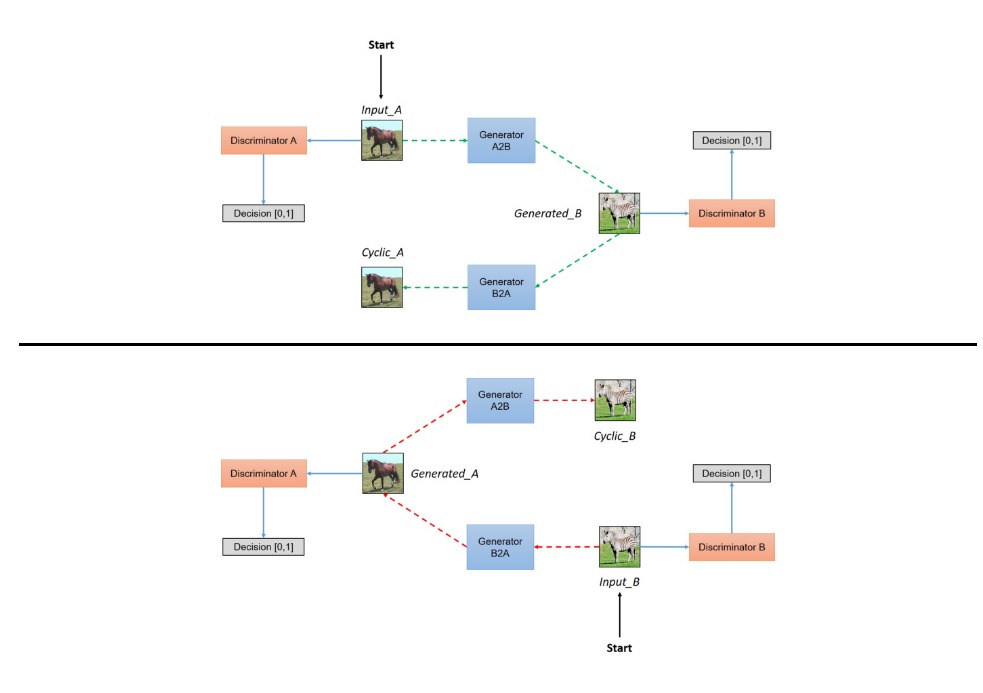
Description automatically generated

STARGANv2

multi domain style transfer

CycleGAN – style transfer – multi domain

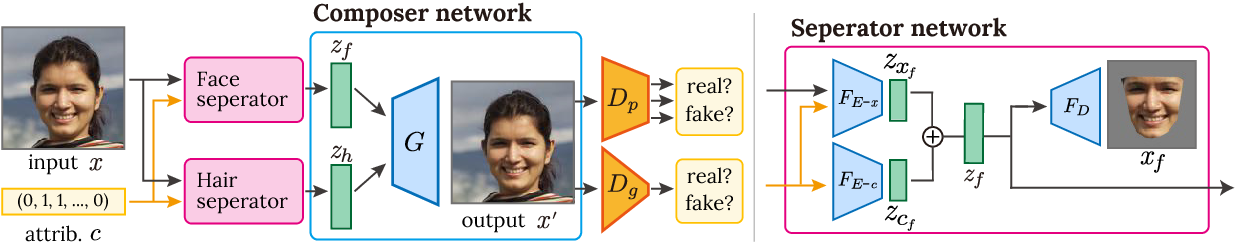
Entire face synthesis – face swapping



RSGAN

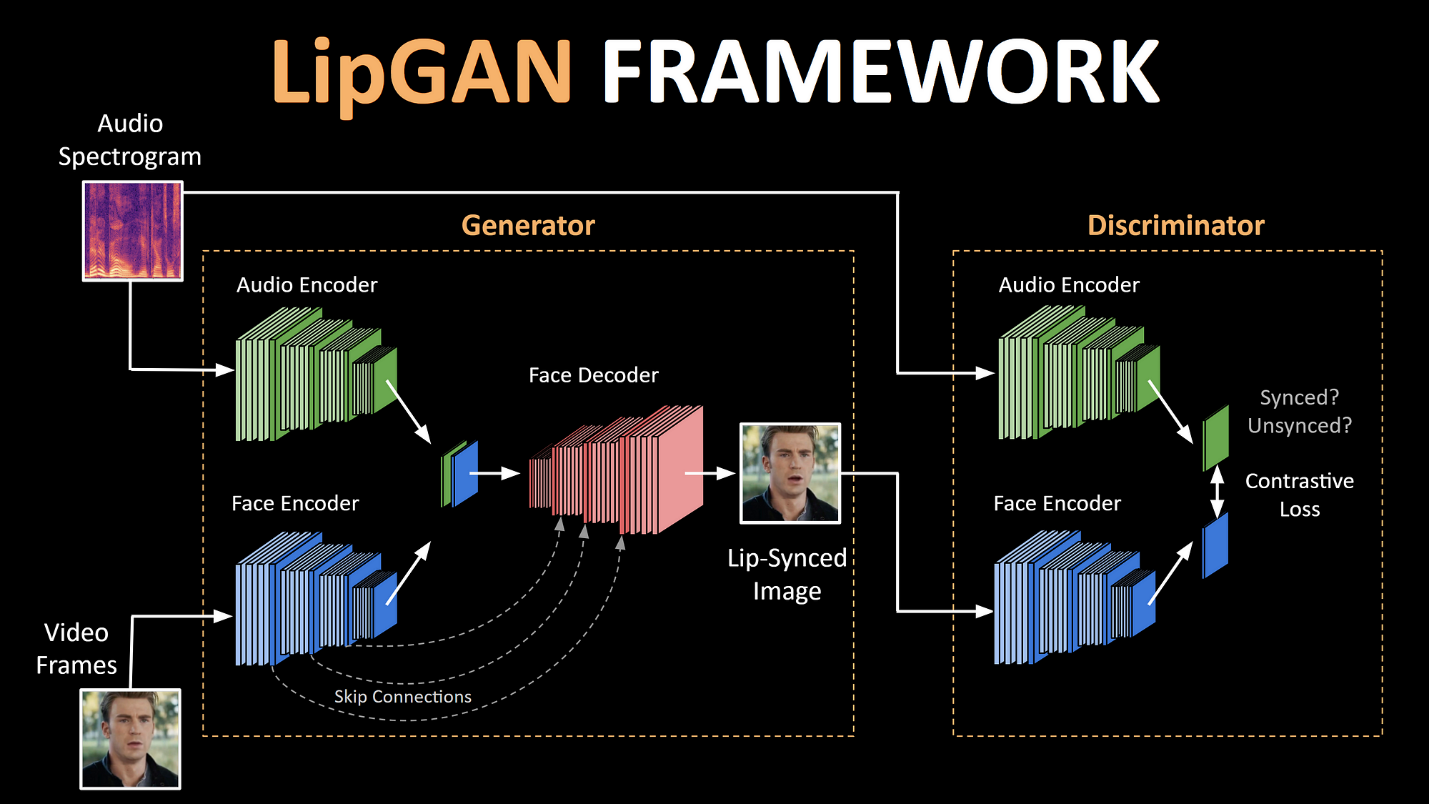
Attribute manipulation / face swapping

Hair and face only



LIPGAN

Lip synching



A diagram of different colors

Description automatically generated with medium confidence

DETECTION

3 main types

**Spatial**

- Image Forensics

- chrominance domain

- Photo response non uniformity – noise pattern caused by light sensor in the camera

- leveraging the local motion features

- DNN

- Improving generalization

- GRAM-net – global facial features

- ResNet-50 backbone trained on PGGAN

- OC – FakeDect trained on real faces with VAE

- Investigating artifacts clues

- CLRNet – LSTM – temporal information in consecutive frames

- SSTNet – spatial steganalysis and temporal features, XceptionNet

- DPNet – temporal dynamic features DNN with XAI

- CNN

- FDFtNet – SqueezeNet, ShallowNetV3,ResNetV2 and Xception improvement using a fine tune transformer

- deep transfer learning

- customized cnn

- cnn with local features

- 3d cnn

- cnn with optical flow

- etc

- Obvious artifact clues

- Detection and localization

- FakeLocator – Encoder Decoder architecture

- Face X-ray – discrepancies across the blending boundary of existing background

- ManTra-Net – end to end cnn

- Facial image preprocessing

-FakeSpotter – monitoring third-party dnn based neuron behaviors

- EN algorithm – extracting local features to represent convolutional traces

- ARENnet – Adaptive residual extraction network suppresses Image content to learn prediction residuals

**Biological**

* Visual-audio inconsistencies
* Visual inconsistencies
* Biological signs in video
  + DeepRhythm – heartbeat frequency in video

**Frequency**

* GAN based artifacts
  + Frequency of saturated and under-exposed pixels
  + AutoGAN - unique artifact in GANs due to under sampling design of common GANs
  + GAN fingerprints
* Frequency Domain
  + FGPD-FA – statistical, oriented gradient and blob
  + F^3 – Net – Local frequency statistics and frequency-aware pattern from frequency-aware image decomposition

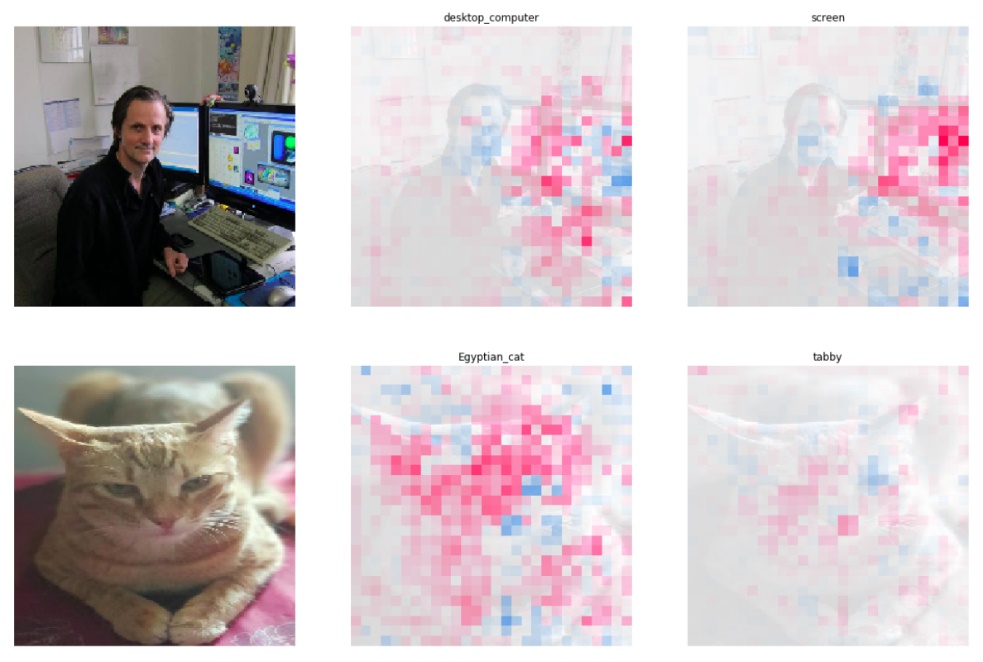
Explainable frameworks

Two types

Local

Global

SHAP (shapley additive explanations) – Uses shapley values, generally used for credit allocation, both local and global. Member of the additive feature attribution methods class



LIME (Local interpretable model-agnostic explanations) Preturbs data points of the black box model by generating syntetic data and using the black box model evaluation as the training set for the glass box model. Designed to be applied locally

A bird with blue arrows

Description automatically generated with medium confidence

Permutation Importance – Remove features and check scores like accuracy and F1, Applied globally

Partial dependence plot – Marginal effect of one or two features on the predicted outcome, Applied globally

Morris sensitivity Analysis – One input value adjusted per run, generally used for screening, Applied globally

Accumulated local Effects – Model agnostic global explanations for classification and regression models on tabular data, applied globally

Integrated gradients – Applied locally, easy to interpret saliency mask that attributes importance value to each feature

A purple and blue owl

Description automatically generated with medium confidence

XRAI – region based attribution method that builds on IG, local

A group of people dancing

Description automatically generated

GRAD-CAM

GUIDED BACK PROP

GUIDED GRAD CAM

BLUR IG

All local saliency maps