Deep learning

Zhuoyan Xu

University of Wisconsin-Madison

zxu444@wisc.edu

Feb 6, 2020

Outline

Auto encoder

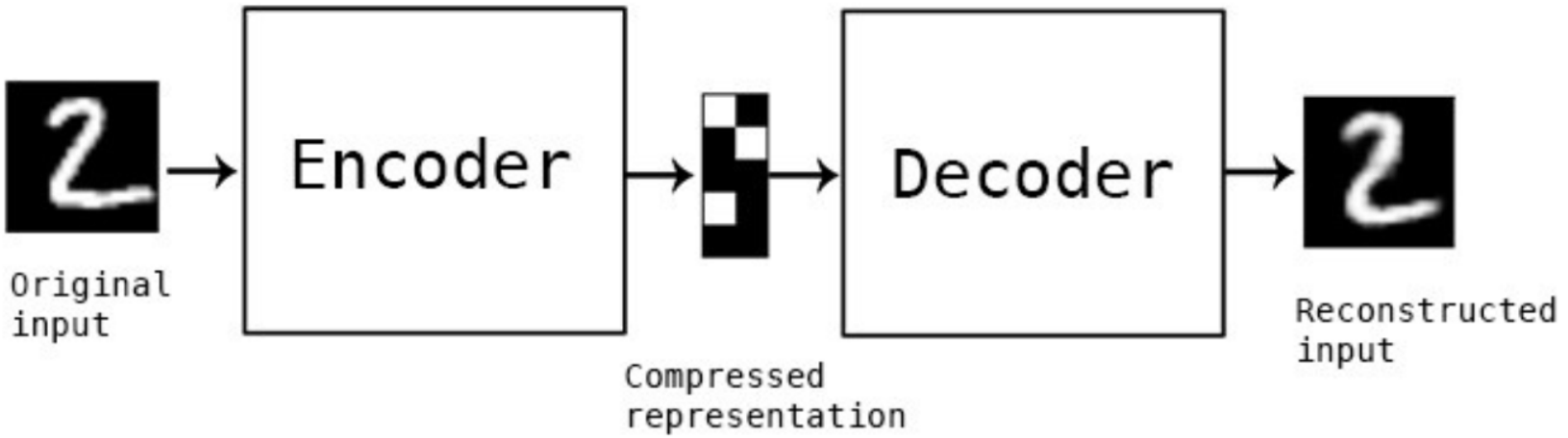


Figure: https://towardsdatascience.com/auto-encoder-what-is-it-and-what-is-it-used-for-part-1-3e5c6f017726

The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore signal "noise".

Auto encoder

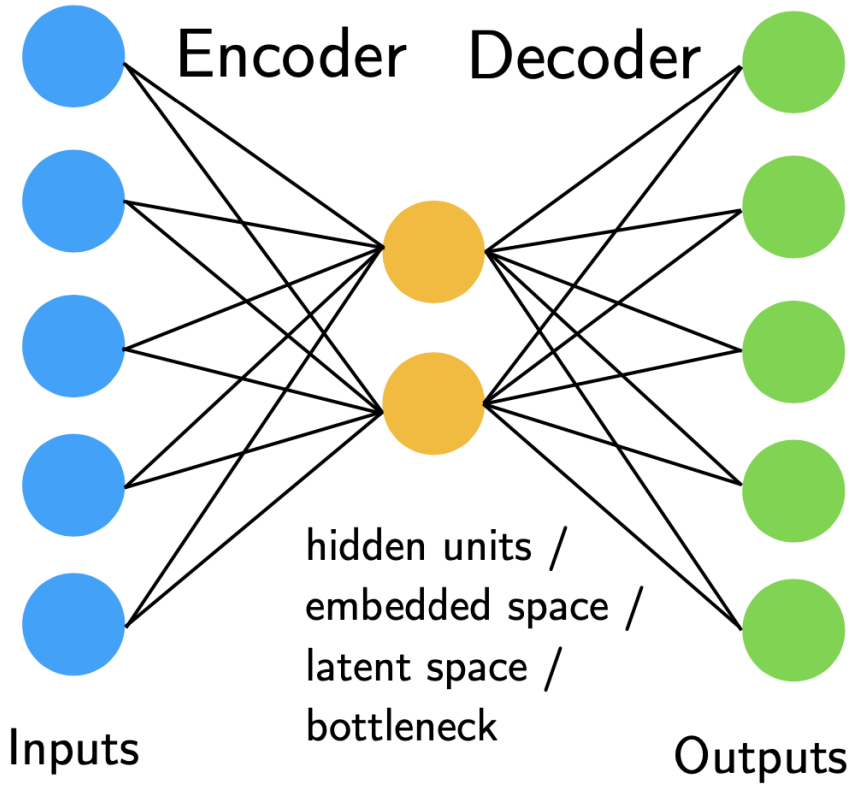


Figure: latent space

$$\mathcal{L}\left(\mathbf{x}, \mathbf{x}'\right) = \left\|\mathbf{x} - \mathbf{x}'\right\|_2^2 = \sum_i \left(x_i - x_i'\right)^2$$

Application

- Use embedding as input to classic machine learning methods.
- Latent space can also be used for visualization or clustering(聚类)).
- Image denoising.

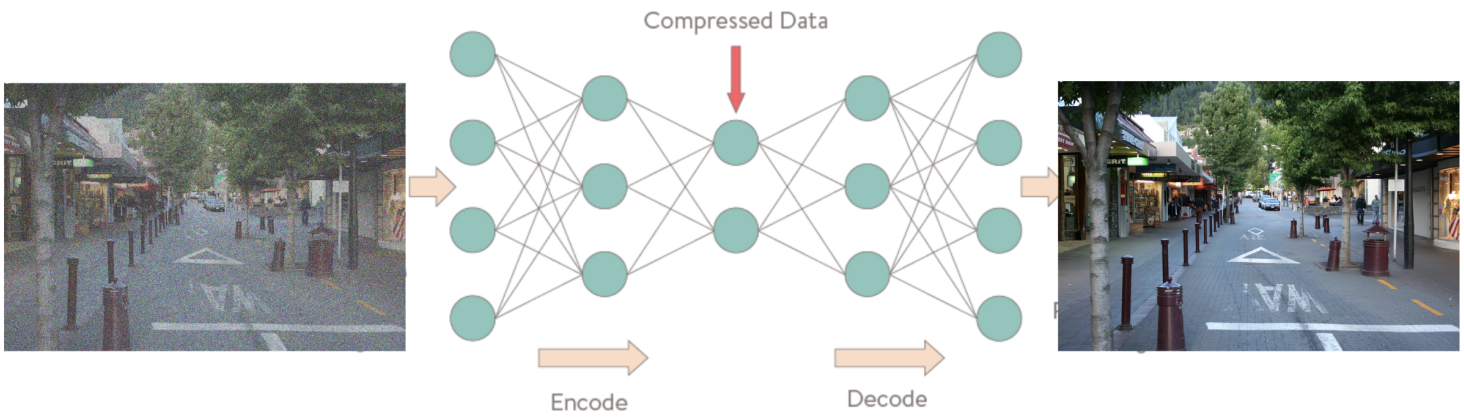


Figure: https://towardsdatascience.com/auto-encoder-what-is-it-and-what-is-it-used-for-part-1-3e5c6f017726

GAN

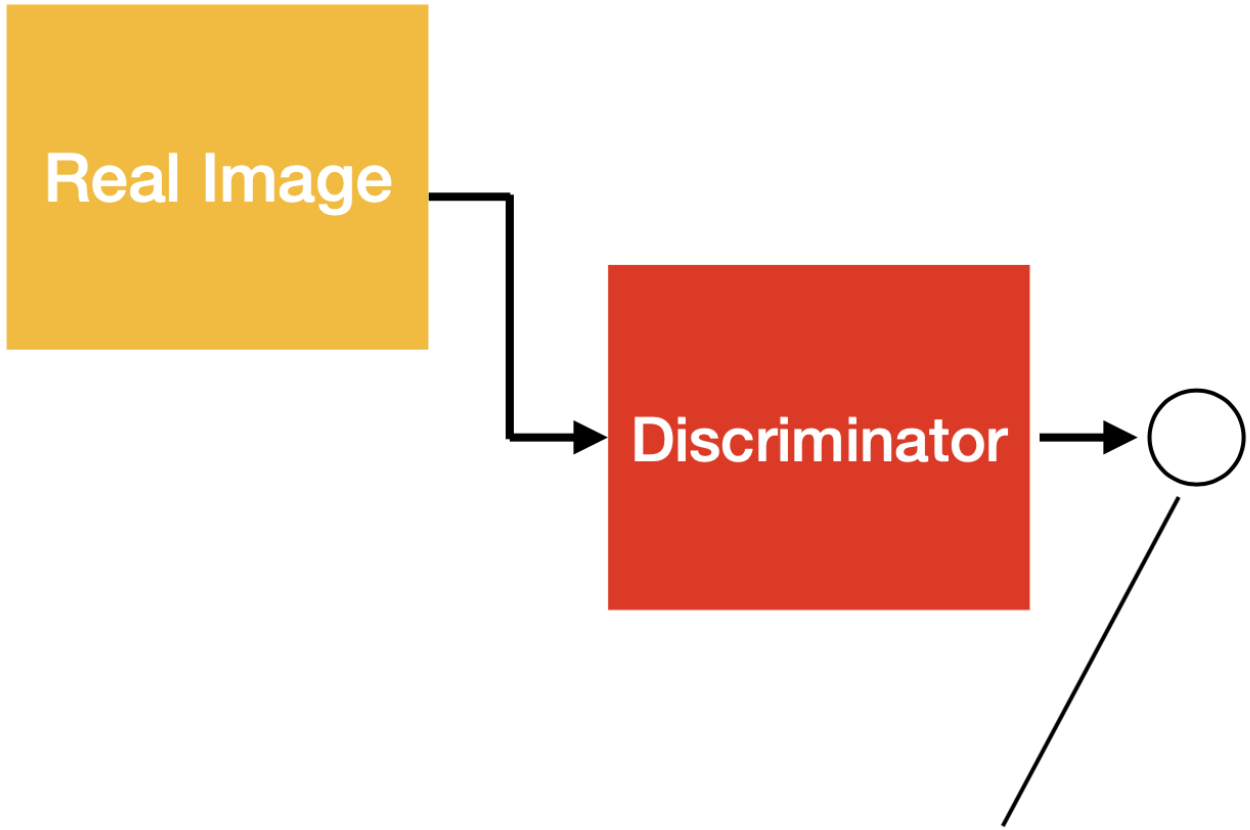
In a 2016 seminar, Yann LeCun described GANs as "the coolest idea in machine learning in the last twenty years".

Figure: `https://www.youtube.com/watch?v=9reHvktowLY`

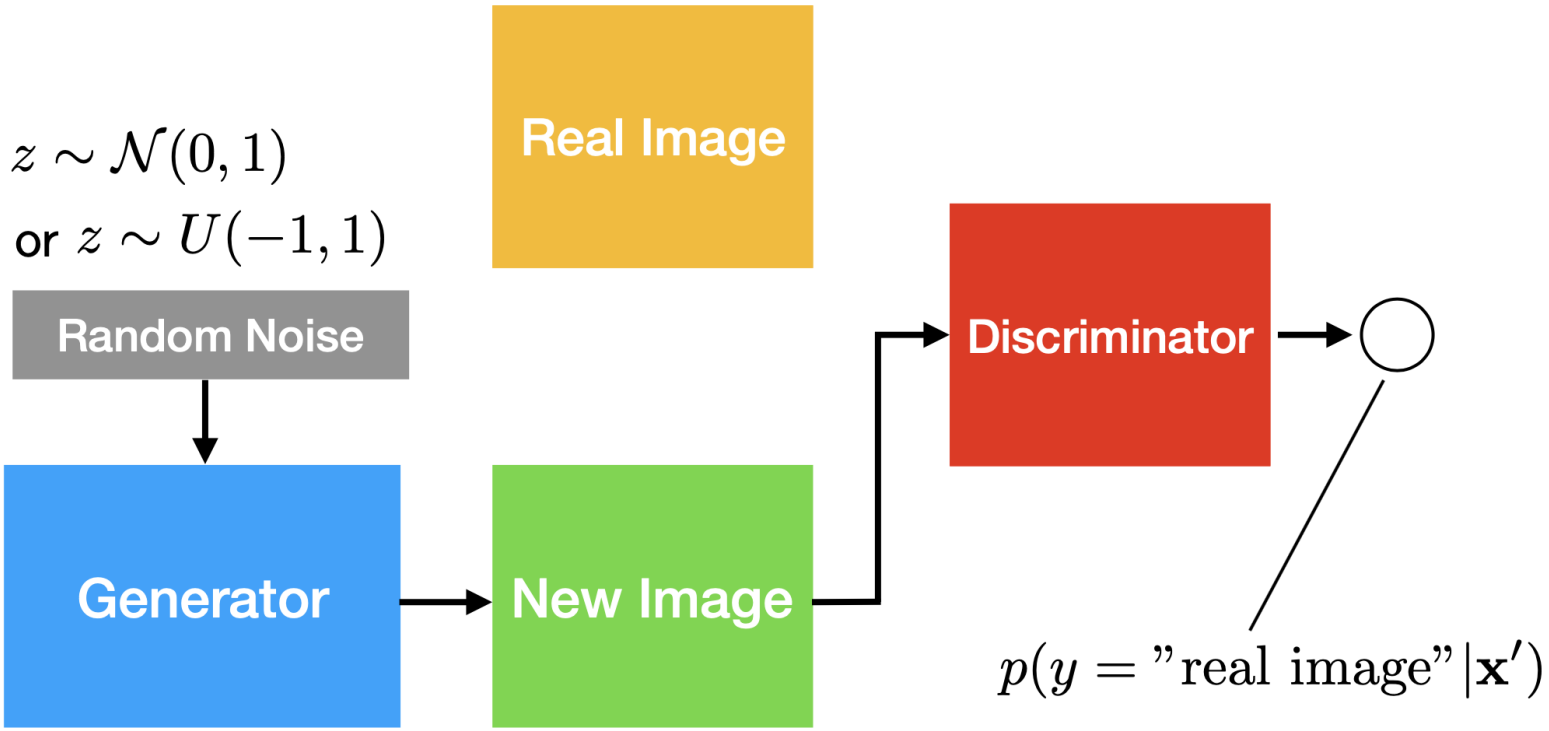
GAN

- The original purpose is to generate new data.
- Classically for generating new images, but applicable to wide range of domains.
- Learns the training set distribution and can generate new images that have never been seen before.
- GANs generate the whole output all at once.

GAN



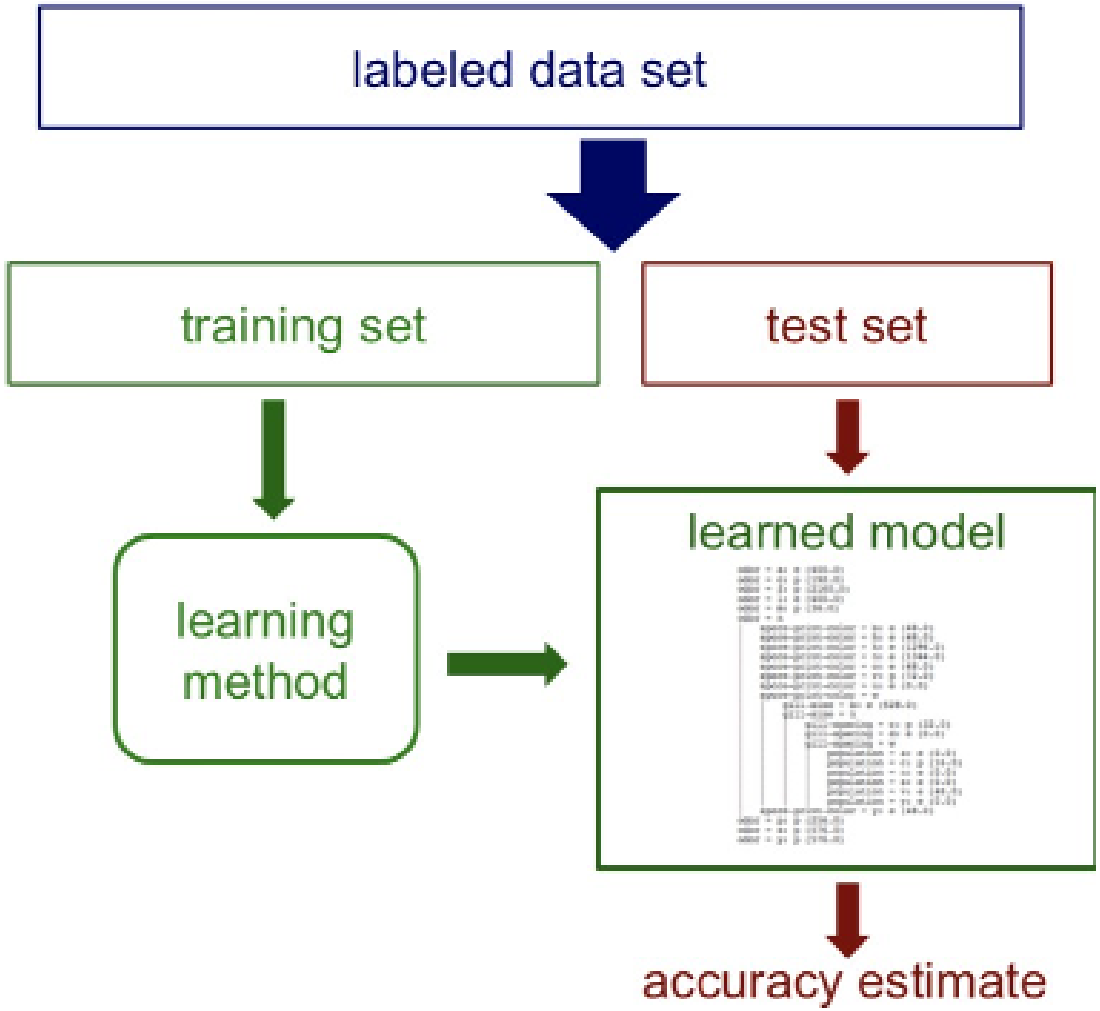
GAN


$z \sim \mathcal{N}(0, 1)$
or $z \sim U(-1, 1)$
Random Noise
Real Image
Generator
New Image
Discriminator
$p(y = \text{"real image"}|\mathbf{x}')$

GAN convergence

- Converges when Nash-equilibrium (Game Theory concept) is reached in the minmax (zero-sum) game.
- Nash-Equilibrium in Game Theory is reached when the actions of one player won't change depending on the opponent's actions.
- Here, this means that the GAN produces realistic images and the discriminator outputs random predictions (probabilities close to 0.5).

Accuracy estimation

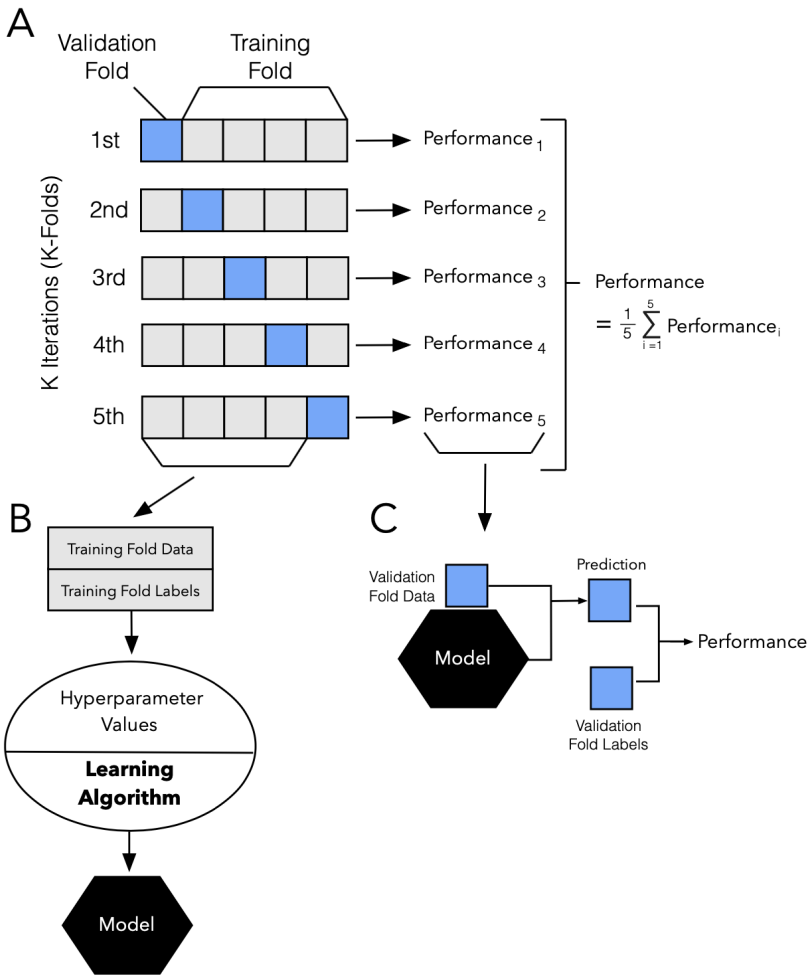


Bootstrap

Bootstrapping is the practice of estimating properties of an estimator (such as its variance) by measuring those properties when sampling from an approximating distribution. Bootstrap can also be used when sample is limited.

Cross validation

Techniques for assessing how the results of a statistical analysis will generalize to an independent data set.



confusion matrix

A specific table layout that allows visualization of the performance of an algorithm.

2x2 Confusion Matrix

		Predicted class P	**Predicted class** N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

confusion matrix

ex: Assuming a sample of 13 animals — 8 cats and 5 dogs — the resulting confusion matrix could look like the table below:

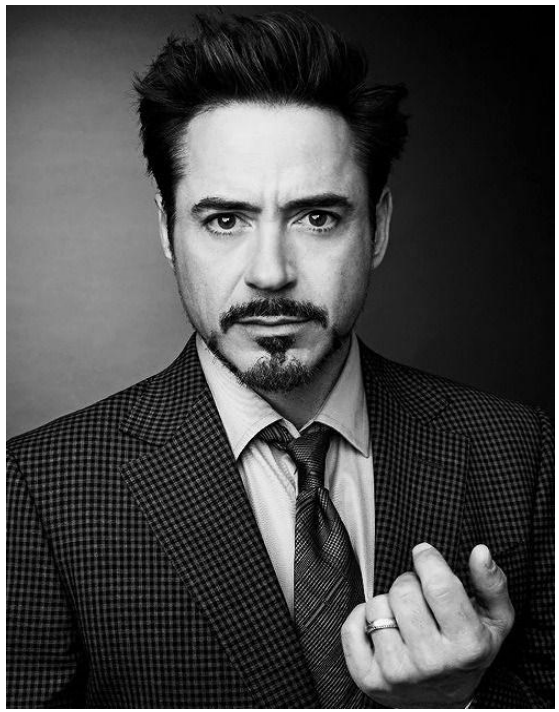
		Actual class	
		Cat	**Dog**
Predicted class	**Cat**	**5**	2
	Dog	3	**3**

Outline

Dataset

We use this dataset to compare the results of some ML methods and CNNs:
Use icrawler to scrape the images by ourselves. Totally 6000 images of 15 celebrities, such as Gal Gadot, Robert Downey jr etc.

Gal Gadot



Robert Downey jr

Scarlet

Figure: Google image search

Data Augmentation

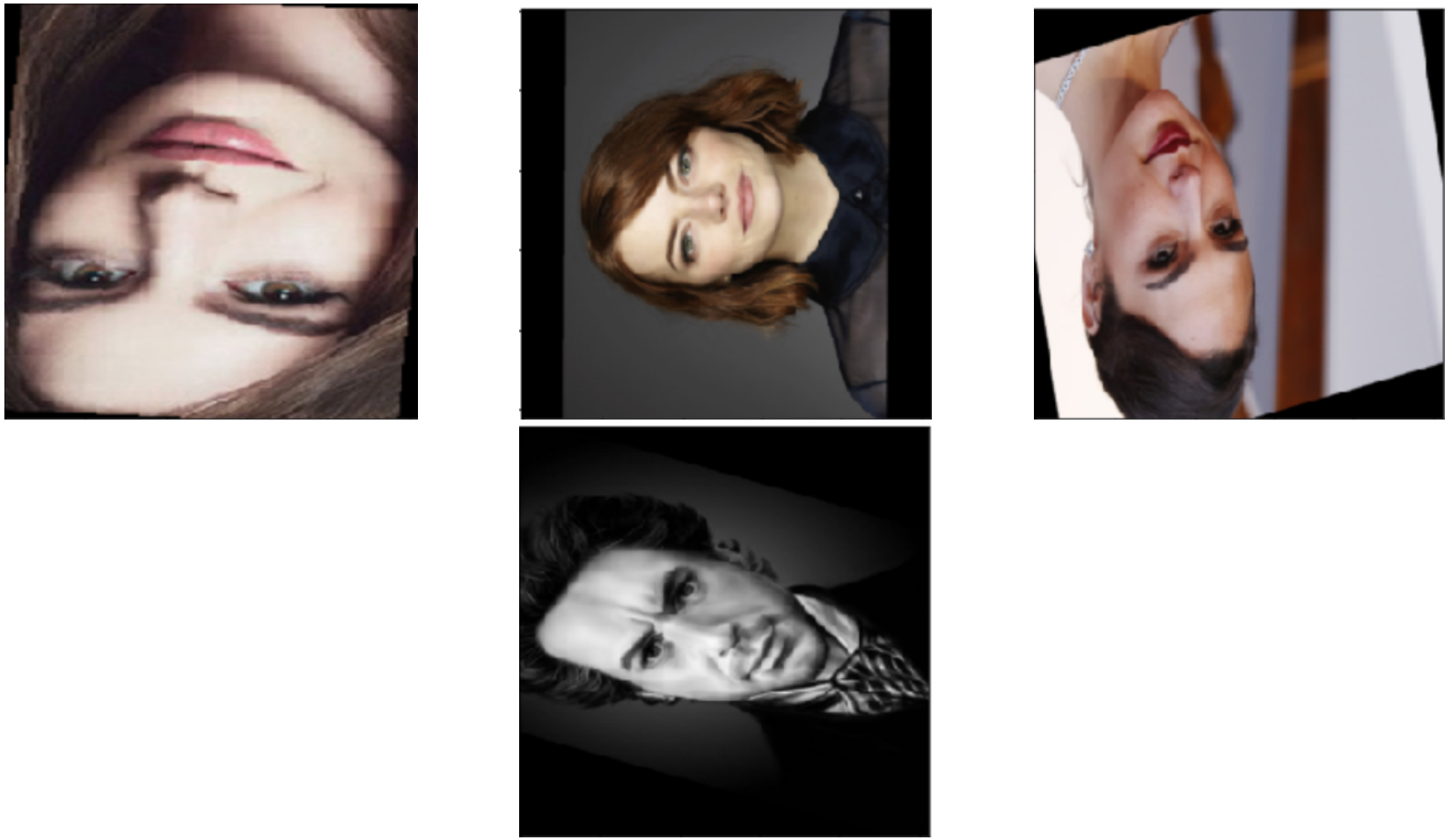


Figure: Data Augmentation

Transfer learning

- VGG16: Freeze all the Conv layers + 3 FC layers, train 2 FC layers + an additional output layer.
- VGG19: Freeze all the Conv layers + 5 FC layers, train 1 FC layers + an additional output layer.
- ResNet50: Freeze all Conv layers, train the output layer.
- ResNet101: Freeze all Conv layers, train the output layer.

Comparison

Model	Test accuracy
SVM	25.02%
GBDT	20.13%
VGG16	38.12%
VGG19	34.26%
ResNet101	35.88%
ResNet50	40.00%

Neural style transfer

manipulate digital images, or videos, to adopt the appearance or visual style of another image:

(a) original

(b) add random noise

Figure: Gal Gadot

Neural style transfer

(a) Starry night Style

(b) Starry night result

Figure: Gal Gadot

Neural style transfer

(a) sandstone

(b) sandstone result

Figure: Gal Gadot

Neural style transfer

(a) Monet style

(b) monet result

Figure: Gal Gadot

Neural style transfer

(a) Skrik

(b) Skrik result

Figure: Gal Gadot

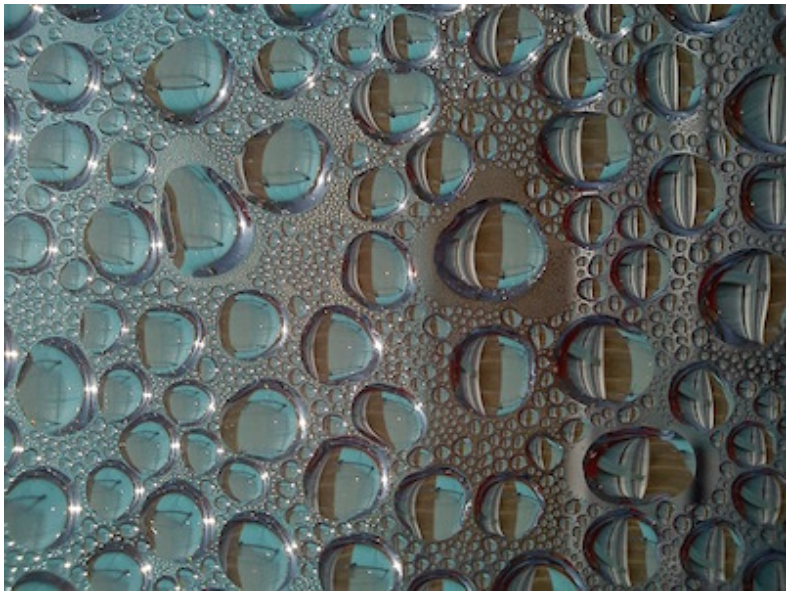
Neural style transfer

Guernica

Guernica Result

Figure: Guernica Style

Neural style transfer



waterdrop

watedrop Result

Figure: Waterdop style