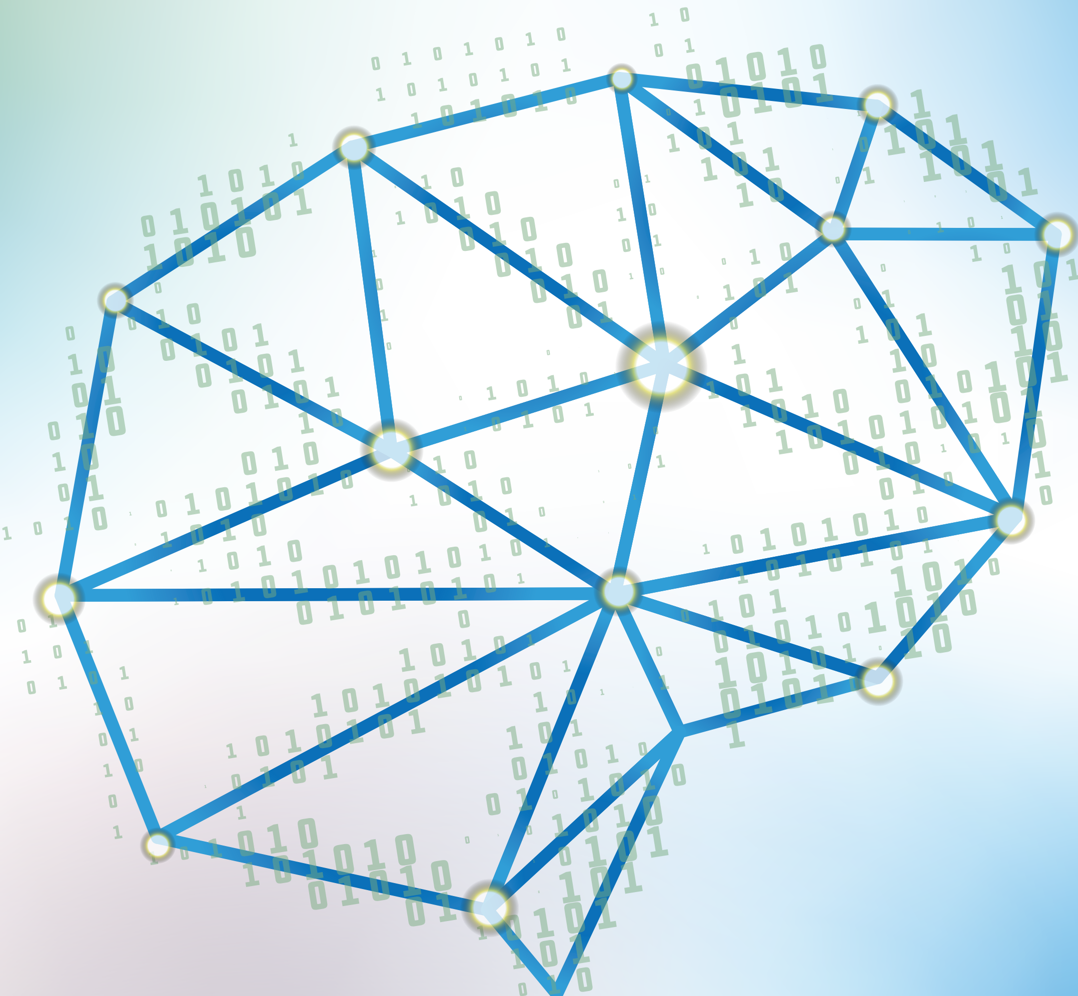
LTIMindtree



Deep Point of View

Generative AI

Table of Contents

01

Executive Summary

The term Artificial Intelligence (AI) was first coined by John McCarthy almost 60 years ago in 1956. AI is not just tech jargon in this digital age. The field of AI is evolving and breakthroughs are happening daily. We are developing complex algorithms and computing systems leveraging AI that can quickly process and analyze massive volumes of data, which would be impossible for an average human to complete in a single lifetime.

We have now focused on creating AI-powered machines that can generate images, texts, and similar multimedia content independently with minimal human intervention. To make this a reality, researchers and programmers have produced an innovative concept called **"Generative AI."**

Generative AI is an emerging technology that uses unsupervised learning algorithms to generate novel images, audio, video, text, or code. This next-gen AI discovers the underlying pattern associated with the input to build new, realistic artefacts representing the training data's properties. According to the MIT Technology Review, Generative AI is one of the most promising advancements in the field of artificial intelligence in the last decade.

By self-learning from each batch of data, Generative AI can create authentic artifacts that did not exist before using a wide array of inputs. Advancements in neural networks and machine learning algorithms developed specifically for data crunching and pattern analysis are key growth drivers in this domain. It is expected that deep research in Generative AI will open new avenues for bulk data evaluation and analysis. Generative AI is in research infancy, but early application trials have exhibited promising results that rival human competency.

At present, Generative AI applicability is confined to network models only. However, with consistent research, new models are being created daily. Generative Adversarial Network (GAN) is the most well-understood and heavily researched model of all the network models available today. It offers a plethora of use cases in the image and video processing domain.

This document aims to provide a high-level view of Generative AI, its building blocks, market analysis, and industry use cases. This document also presents some ongoing research on the generative models , thus providing a good starting point for anyone who wishes to dive deeper into this domain.

02 Introduction

In this age of artificial intelligence and deep neural networking, we often use computers to perform tasks beyond human capabilities. We rely on them for menial chores and to perform activities that require zero mistakes or errors.

Scientists wanted to develop technology that would liberate humans from mundane duties so that we could dedicate our lives to more creative and imagination-based work. Poetry, painting, craft, etc., are some creative options for human beings to excel at.

Our creativity went digital with the advent of powerful graphic cards. We transitioned from creating 3D models and digital photography to converting them into digital art and selling it as NFTs. Let us look at some of the best examples of digital art created by incredible 3D artists and photographers!

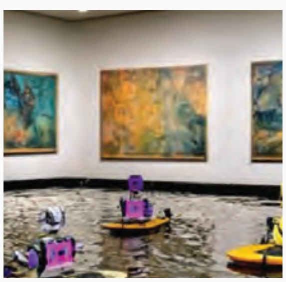
Fig.1: Digital Art

What if we told you that every single one of them was done by a single 3D artist and photographer?

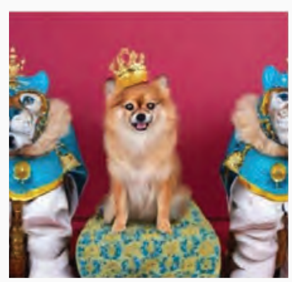
You would think she would be an artistic personality who grew up in a home that encouraged her creative side from an early age, right? Or that she has many designer friends and enjoys a nice cup of coffee while watching the sunset?

This artist, on the other hand, did not grow up in a creative home. She was raised in a laboratory. She has no designer friends. She has no friends. She only has researchers.

This artist is not a person at all. It is an AI and one of the greatest AIs ever produced in a Google lab called Google Imagen, and it takes a random bit of text and turns it into art.

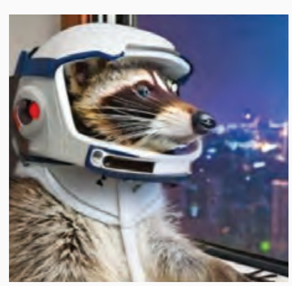


An art gallery displaying Monet paintings. The art gallery is flooded. Robots are going around the art gallery using paddle boards.



A Pomeranian is sitting on the Kings throne wearing a crown. Two tiger soldiers are standing next to the throne.

A giant cobra snake on a farm. The snake is made out of corn.



A photo of a raccoon wearing an astronaut helmet, looking out of the window at night.

Fig.2: Digital Art by Google Imagen

Google Imagen displays the power of what Generative AI can do with limited input. This tech is a game changer for image and language modeling. Such a Generative AI can read and analyze data inputs (in the form of text, pictures, audio, or video) and produce new and unique media while retaining the essence of the original data.

As discussed earlier, this tech employs unsupervised learning algorithms to create unique content from existing data. Simply put, it allows machines to recognize patterns in incoming text and utilize them to produce similar content.

As these models are provided with limited parameters during the training stage, the model creates its interpretation and judgment about the unique and vital properties of the data. Due to this, the results from these Generative AI models are free from human experience-based biases and mental processes. A significant drawback in wide-scale adoption of generative AI currently is cost. As this technology requires high processing power, the cost of deployment and operation is
more increased.

03 Common Techniques of Generative AI

Artificial intelligence techniques have traditionally been used to clean data, enhance predictive analysis, compress data, and decrease the dimensionality of datasets for other algorithms. Novel generative AI techniques like Variational autoencoders (VAEs), for example, push this a step further by reducing errors between the raw signal and the reconstruction.

Generative models are exceptionally good at producing near-original material with a little vector. It also enables us to create previously non existing material that may be used without licensing. Some Generative AI techniques are used when working with pictures or visual data. There are certain Generative AI models that perform better in signal processing applications such as anomaly detection for predictive maintenance or security analytics. Let's discuss some of these Generative AI techniques in this section.

• Generative Adversarial Networks

Ian Goodfellow and colleagues at the University of Montreal pioneered the use of Generative Adversarial Networks (GANs) in 2014. They have showed enormous potential in producing many forms of realistic data. Yann LeCun, Meta's chief AI scientist, called GANs and their variants "the most exciting topic in machine learning in the last ten years."

For starters, they have been utilized to make realistic speech by mimicking humans and matching voices and lip movements for better translations. They have also interpreted visuals, distinguished between night and day, and defined dancing motions between bodies. They are also used in conjunction with other AI approaches to increase security and create stronger AI classifiers.

GANs use two competing neural networks, a generator and a discriminator. The generator, also known as the generative network, is a neural network responsible for producing new data or content comparable to the original data. A discriminator, also known as a discriminative network, is a neural network that differentiates between source and produced data. The competition between these two networks is to develop their algorithms until they can create data indistinguishable from the original material.

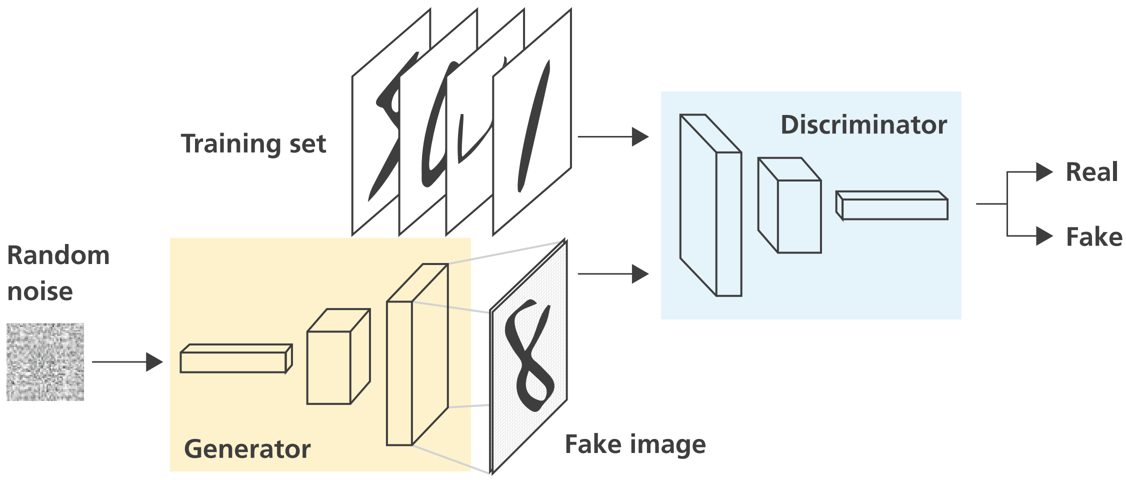


Fig 3: Generative Adversarial Networks (Thalles Silva)

• Models

DCGAN	ProGAN	BigGAN
(Deep Convolutional GAN)	(Progressively Growing GAN)	

• Transformer-based Models

Transformer-based models are mainly used to analyze data with a sequential structure (such as the sequence of words in a sentence). In modern times, transformer-based techniques have become a standard tool for modeling natural language.

The ability of the transformer models to attend to various positions of the input sequence to compute a representation of that sequence is core to their architecture.

Fig 5: Transformer-based Models (Source: GitHub)

• Models

BERT

(Bidirectional Encoder Representation from Transformers)

RoBERTA

(Robustly Optimized BERT)

• Autoregressive Convolutional Neural Networks

Autoregressive refers to self-regression. The word autoregression refers to forecasting future outcomes of a series based on previously observed effects of that sequence. AR- CNNs investigate systems that change over time and believe that the likelihood of specific data is only based on what has happened before. To create reliable new data, they rely on past time-series data. RNNs and casual convolutional networks are the most common autoregressive designs.

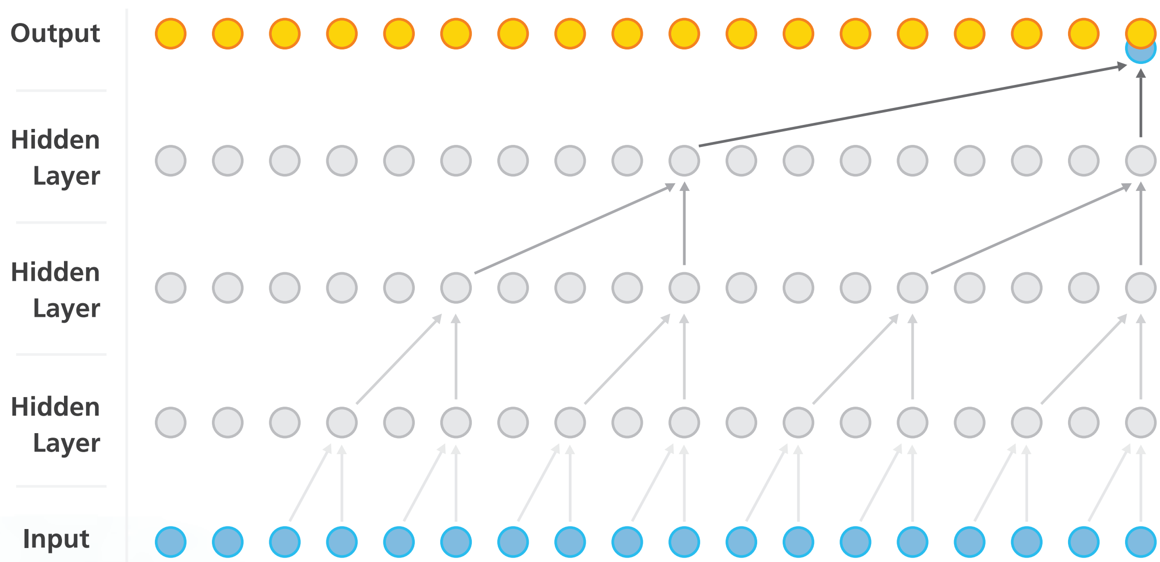


Fig 4: Autoregressive Convolutional neural networks (Source: deepmind.com)

• Models

PixelRNN

PixelCNN

WaveNet

- ## Other Nascent Techniques

Bayesian Network

Bayesian Network or Bayes Network is a generative probabilistic graphical model that allows efficient and effective representation of the joint probability distribution over a set of random variables. Bayes Network consists of two main parts, which are structure and parameters. The structure is a directed acyclic graph (DAG), and the parameters consist of conditional probability distributions associated with each node. This network can be used for various applications, such as time series prediction, anomaly detection, reasoning, etc.

Gaussian Mixture Model

Gaussian Mixture Model is a generative probabilistic model which assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. GMMs are commonly used as a parametric model of the probability distribution of features in a biometric system, which includes vocal tract-related spectral components in a speaker recognition system. Thus, a well-trained prior model estimates GMM parameters from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum a Posteriori (MAP) estimation.

Hidden Markov Model

A Hidden Markov Model (HMM) is a statistical model that can describe the evolution of observable events that depend on internal factors, which are not directly observable. The model is popularly known for its effectiveness in modeling the correlations between adjacent symbols, domains, or events They have been extensively used in various fields, especially in speech recognition and digital communication. A Hidden Markov Model consists of two stochastic processes: an invisible circle of hidden states and a visible process of observable symbols.

Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) is a generative probabilistic model with collections of discrete data such as text corpora. LDA is a three-level hierarchical Bayesian model in which each collection item is modeled as a finite mixture over an underlying set of topics. The model has applications for various problems, including collaborative filtering and content-based image retrieval.

Variational Autoencoders (VAEs)

Variational Autoencoders (VAEs) have been one of the most popular approaches to unsupervised learning of complicated distributions. They are built on top of standard function approximators, which are neural networks and can be trained with stochastic gradient descent. The application of VAEs includes generating various kinds of complex data, including handwritten digits, faces, CIFAR images, predicting the future from static images, and more.

04 Generative AI Market Potential

It is difficult to estimate the true market sense of Generative AI as the technology is in the experimental stage. As no practical uses and adoption models have yet been developed, it has become difficult to gauge the exact range of applications. Thus, it is too early to put a number on the market potential.

The most reliable information that we can use to derive some insights and assess the market scenario is by studying the Generative Design Market. This market is an offshoot of Generative AI, which deals with CAD-based automated design.

Currently, the Generative Design Market is expected to surpass USD 529 Mn by 2027 at a CAGR of 19.4%.

Drawing inspiration from the above statistics, we believe that the market potential for non-niche Generative AI will be 10x of the Generative Design market with a tentative timeline of 3-5 years in the future.

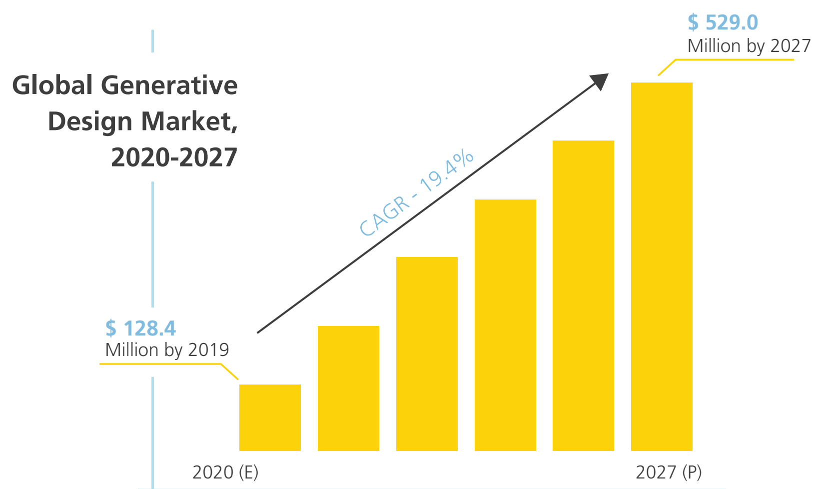


Fig 6: Global Generative Design Market Potential (Source: Verified Market Research)

05 Industry Use Cases of Generative AI

According to Gartner, the amount of digital data created by generative AI will be more than 10X by 2025, a dramatic increase from the 1% it currently accounts for. Let us find out what industries will benefit the most from implementing this technology.

• Healthcare

In this sector, Generative AI serves a dual purpose. Primarily, it has the potential to improve patient care. Second, it can enhance patient data privacy. Here fabricated and under-represented data is used to train and improve the Generative AI model. GANs, for example, can provide numerous viewpoints of an X-ray picture to show potential tumor development outcomes. It can also identify cancerous developments by comparing images of healthy organs from a databank to damaged ones. The second use of the technology focuses on data de-identification, which aids in the security of the reversal process, which is far from impenetrable.

• Life Science

Here Generative AI can help with drug discovery. The technology can produce molecular structures of medications used to treat various diseases. When this technology performs quick database search of substances for this purpose, the treatment of novel ailments can also be facilitated. This automated technology is faster than the manual procedure. Gartner predicts that generative AI will be used in 50% of drug development activities by 2025.

LTIMindtree

Media and Entertainment

Movie Restoration

Many vintage movies and classic Disney cartoons are treasures of world culture, but their quality sometimes falls short of the demands of our day. Generative AI can upscale them to 4k and beyond, create 60 frames per second of the standard 23, reduce noise, and convert black-and-white to color.

Generation of Animated Models

Along with film, the video game business relies on moving pictures, and generative AI can aid. When AI algorithms build 3D models in computer games, software developers' efforts are lightened, and development time is significantly decreased. Such models might be completely new or derived from previously inputted 2D photos. Additionally, the system can produce 2D images for usage in a certain game and animation genres, such as anime.

Audio synthesis

Generative AI is about more than just images. Its application can also improve the field of sound. This knowledge may be used in cinematography and video gaming to create foley components, ambient noises, voiceovers, and other audio effects that are an essential aspect of a movie or video game that spectators love. With Generative AI, it is possible to create voices that resemble humans. The computer-generated voice helps develop video voiceovers, audible clips, and narrations for companies and individuals.

• Retail and e-commerce

People express their feelings and evaluate the things they purchase and the organization supplies services when engaging with items. AI algorithms may be trained to assess consumer-generated texts, audio samples, and facial expressions that provide insight into clients' attitudes on the item in issue.
Other generative AI techniques can monitor online consumers' web activity and evaluate user data to determine how enjoyable the UX is or how effective an advertising or the overall marketing campaign was. Such information may then be used in client segmentation to identify different consumer groups and map out focused promotional programs, enhancing upselling and cross-selling potential.

• Finance

Fraud detection

Several businesses already use automated fraud-detection practices that leverage the power of AI. These practices have helped them locate malicious and suspicious actions quickly and with accurately. AI is now detecting illegal transactions through preset algorithms and rules and is making detecting theft identification easier.

Trend evaluation

ML and artificial learning technology help predict the future. These technologies provide valuable insights into the trends beyond conventional calculative analysis.

• IT Industry

Software development

Generative AI has also influenced the software development sector by automating manual coding. Rather than coding the software completely, IT professionals now have the flexibility to quickly develop a solution by explaining the AI model about what they are looking for. For instance, a model-based tool GENIO can enhance a developer's productivity multifold compared to a manual coder. The tool helps citizen developers, or non-coders, develop applications specific to their requirements and business processes and reduces their dependency on the IT department.

Data Synthesis and augmentation

Data unavailable in the real world can be generated using generative AI. This may be used for research, such as testing new machine learning algorithms or deep learning architectures. The artificial data set generated by generative AI can be supplemented with original data to improve neural network performance, especially deep learning. Data augmentation using generative AI can be used to enhance the quality of data. Generative AI can help with the task of tuning the neurons in neural networks by automatically finding the best set of connections.

Algorithm Invention

One application of generative AI is to help researchers invent new machine learning algorithms. This process has far been done mostly by hand, but with the help of generative AI, it can be automated.

• Other use cases

Artificial General Intelligence (AGI)

Artificial general intelligence (AGI) consists of algorithms that can successfully perform any intellectual task a human being can do. Humans have used tools for thousands of years to solve problems and create new things. Now we need to automate this process as well! Generative AI is a crucial step towards building AI that can design better machine learning algorithms and other forms of AI.

NFT Development

Non-fungible tokens are all the rage in today's digitally driven society, with sales exceeding $25 billion last year. NFT art is prevalent in the niche, with cartoons, memes, and paintings dominating. Generative AI technologies are unrivaled in their ability to create such art creations that may bring in large sums of money for their creators.

Text, Image, and Music Generation

AI text generators can create summaries of articles, generate product descriptions, write blog posts, and paraphrase text to prevent plagiarism. Generative AI-created images can be used for research work and in computer graphic applications. Generative AI can also be used to create music. It is even possible to use generative AI algorithms to listen to the generated music and find the parts they like most, like Pandora or Spotify. This can also be used to improve the existing music experience.

Artificial Creativity

Artificial creativity is a subfield of generative AI where the main goal is not generating new data but creating something that did not exist before, for example, generating an abstract painting or a novel story without human input.

06 Generative Adversarial Networks (GAN) Based Research Work

• Image Generation Using Datasets

Ian Goodfellow, et al., in their 2104 paper "Generative Adversarial Networks", used GANs to generate new plausible sample images for the MNIST handwritten digit dataset, the CIFAR-10 small object photograph dataset, and the Toronto Face Database.

Along the same line of thought, Alec Radford, et al. in their 2015 paper titled "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (DCGAN)" demonstrated how to train stable GANs at scale. They showed models for generating new examples of bedrooms.

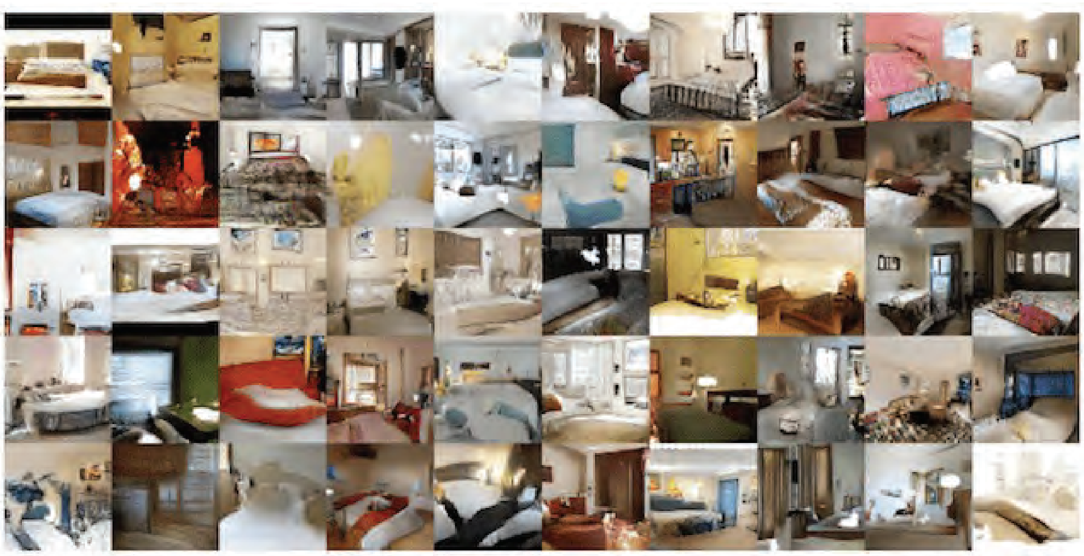


Fig 7: Example of GAN-Generated Photographs of Bedrooms (Source: Arxiv)

Importantly, this paper demonstrates the ability to perform vector arithmetic with the input to the GANs (in the latent space) both with generated bedrooms and with generated faces.

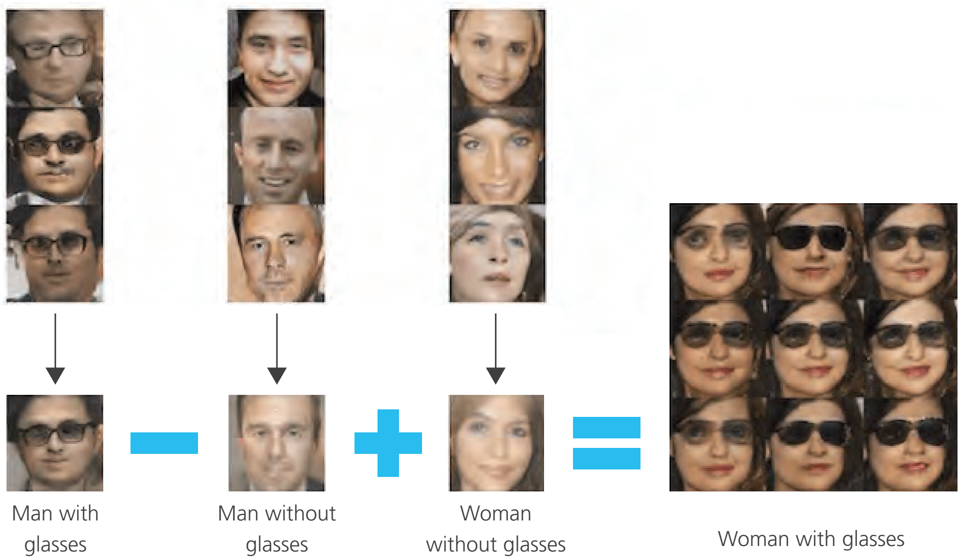


Fig 8: Example of Vector Arithmetic for GAN-Generated Faces (Source: Arxiv)

- ## Generate Photographs of Human Faces

Tero Karras, et al. in their 2017 paper titled "Progressive Growing of GANs for Improved Quality, Stability, and Variation" demonstrated the generation of plausible realistic photographs of human faces. The images were real, and the results were promising. Application based on this research received a lot of media attention. The face generations were trained on celebrity examples, meaning that there are elements of existing celebrities in the generated faces, making them seem familiar, but not entirely.

Fig 9: Examples of Photorealistic GAN-Generated Faces. (Source: Arxiv)

Examples from this paper were also used in a 2018 report titled "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation" to demonstrate the rapid progress of GANs from 2014 to 2017.

Fig 10: Example of the Progression in the Capabilities of GANs from 2014 to 2017 (Source: Arxiv)

• Generate Realistic Photographs

Andrew Brock, et al. in their 2018 paper titled "Large Scale GAN Training for High Fidelity Natural Image Synthesis" demonstrate the generation of synthetic photographs with their technique BigGAN that are indistinguishable from real photos.

Fig 11: Example of Realistic Synthetic Photographs Generated with BigGAN (Source: Arxiv)

• Generate Cartoon Characters

Yanghua Jin, et al., in their 2017 paper titled "Towards the Automatic Anime Characters Creation with Generative Adversarial Networks" demonstrate the training and use of a GAN for generating faces of anime characters (i.e., Japanese comic book characters). Inspired by the anime examples, several people have tried to create Pokemon characters, such as the pokeGAN project and the Generate Pokemon with DCGAN project, with limited success.

Fig 12: Example of GAN-Generated Pokemon Characters (Source: PokeGAN project)

- ## Image-to-Image Translation

This is a bit of a catch-all task, for those papers that present GANs that can do many image translation tasks. Phillip Isola, et al., in their 2016 paper titled "Image-to-Image Translation with Conditional Adversarial Networks" demonstrate GANs, precisely their pix2pix approach for many image-to-image translation tasks.

Examples include translation tasks such as:

- Translation of semantic images to photographs of cityscapes and buildings.
- Translation of satellite photographs to Google Maps.
- Translation of photos from day to night.
- Translation of black and white photographs to color.
- Translation of sketches to color photographs.

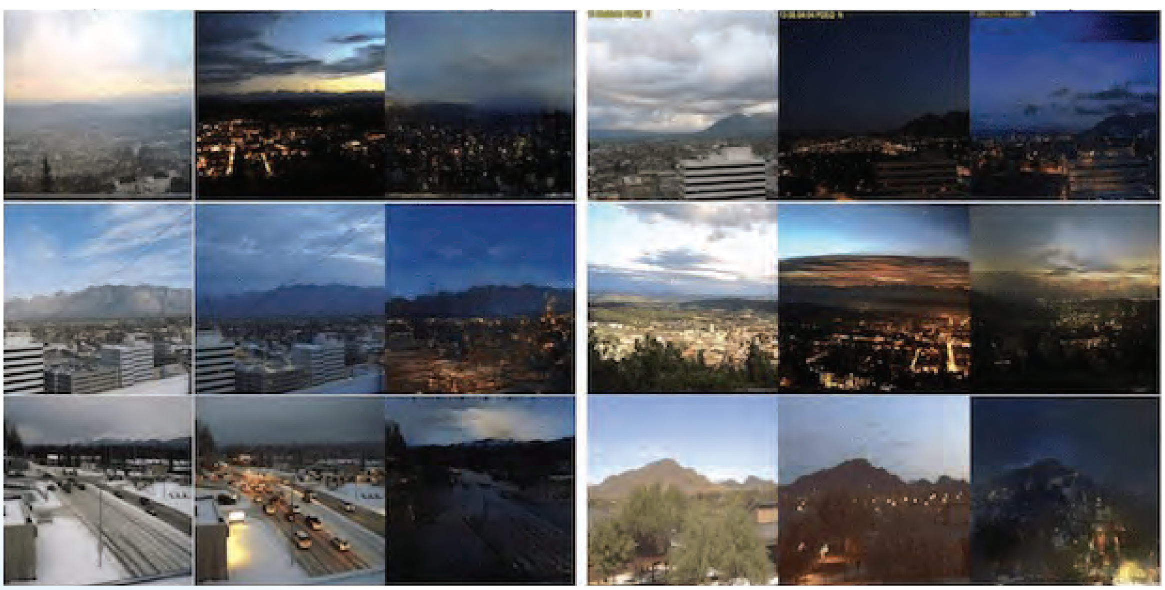


Fig 13: Example of Photographs of Daytime Cityscapes to Night-time With pix2pix (Source: Arxiv)

Similarly, Jun-Yan Zhu in their 2017 paper titled "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks" introduces their famous Cycle GAN and a suite of impressive image-to-image translation examples.

The example below demonstrates four image translation cases:

Translation from photograph to artistic painting style.	Translation of horse to zebra.	Translation of photographs from summer to winter.	Translation of satellite photographs to Google Maps view.

Fig 14: Example of Four Image-to-Image Translations Performed with CycleGAN (Source: Arxiv)

The paper also provides many other examples, such as:

- Translation of painting to photograph.
- Translation of sketch to photograph.
- Translation of apples to oranges.
- Translation of photographs to artistic painting.

• Text-to-Image Translation (text2image)

Han Zhang, et al., in their 2016 paper titled "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks" demonstrate the use of GANs, specifically their StackGAN to generate realistic-looking photographs from textual descriptions of simple objects like birds and flowers.

The small bird has a red head with feathers that fade from red to gray from head to tail

Stage-1 images

Stage-2 images

Fig 15: Example of Textual Descriptions and GAN-Generated Photographs of Birds (Source: Arxiv)

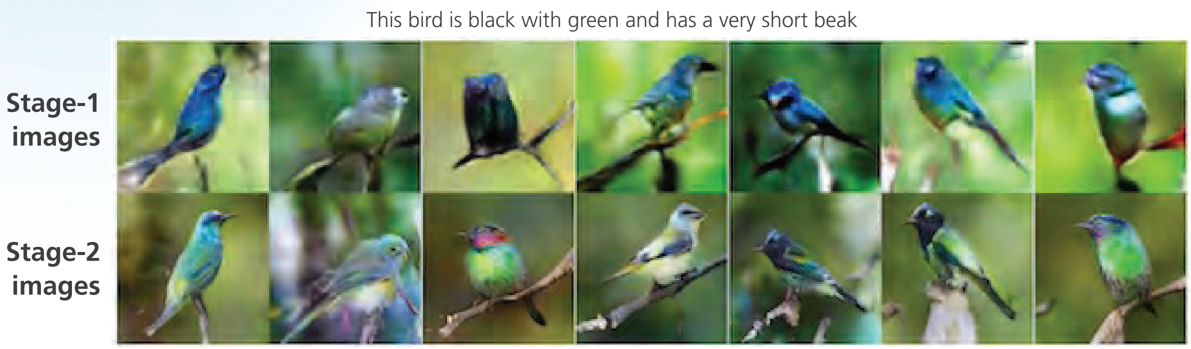


Fig 15: Example of Textual Descriptions and GAN-Generated Photographs of Birds (Source: Arxiv)

Scott Reed, et al., in their 2016 paper titled "Generative Adversarial Text to Image Synthesis" also provide an early example of text-to-image generation of small objects and scenes including birds, flowers, and more. Ayushman Dash, et al. gave even more examples on the same dataset in their 2017 paper titled "TAC-GAN – Text Conditioned Auxiliary Classifier Generative Adversarial Network". Scott Reed, et al. in their 2016 paper titled "Learning What and Where to Draw" expand upon this capability and use GANs to both generate images from text and use bounding boxes and key points as hints as to where to draw a described object, like a bird.

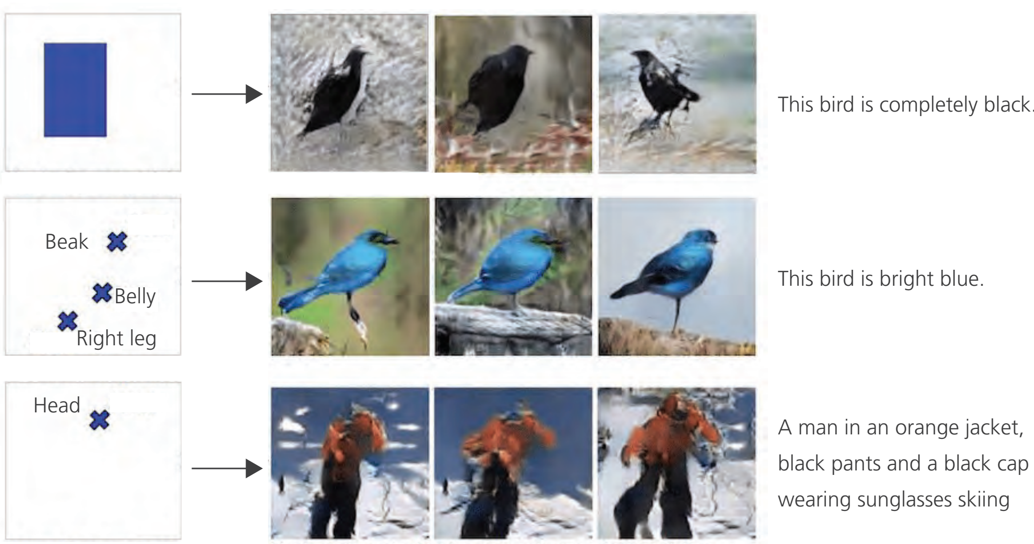


Fig 16: Example of Photos of Object Generated from Text and Position Hints with a GAN (Source: Arxiv)

• Semantic-Image-to-Photo Translation

Ting-Chun Wang, et al. in their 2017 paper titled "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs" demonstrate the use of conditional GANs to generate photorealistic images given a semantic image or sketch as input.

Fig 17: Example of Semantic Image and GAN-Generated Cityscape Photograph (Source: Arxiv)

Specific examples included:

- Cityscape photograph, given semantic image.
- Bedroom photograph, given semantic image.
- Human face photograph, given semantic image.
- Human face photograph, given sketch.

• Face Frontal View Generation

Rui Huang, et al. in their 2017 paper titled "Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis" demonstrate the use of GANs for generating frontal-view (i.e., face-on) photographs of human faces given photographs taken at an angle. The idea is that the generated front-on photos can then be used as input to a face verification or face identification system.

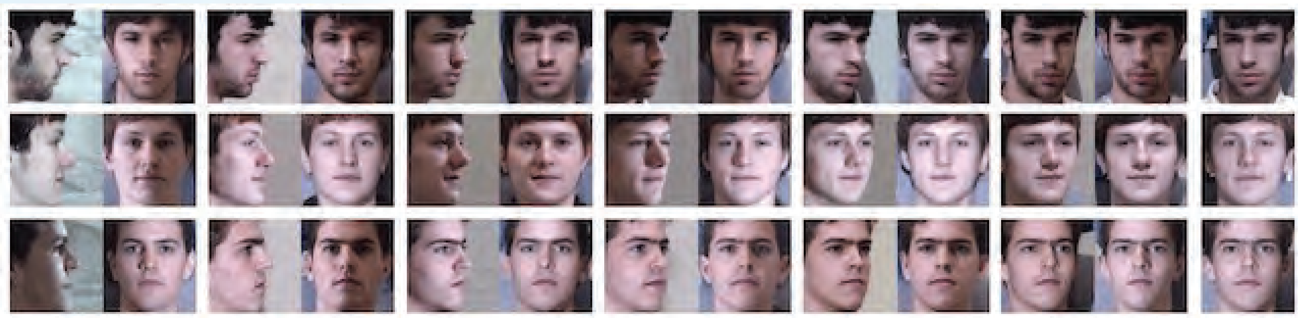


Fig 18: Example of GAN-based Face Frontal View Photo Generation (Source: Ieeexplore)

• Generate New Human Poses

Liqian Ma, et al. in their 2017 paper titled “Pose Guided Person Image Generation” provide an example of generating new photographs of human models with new poses.

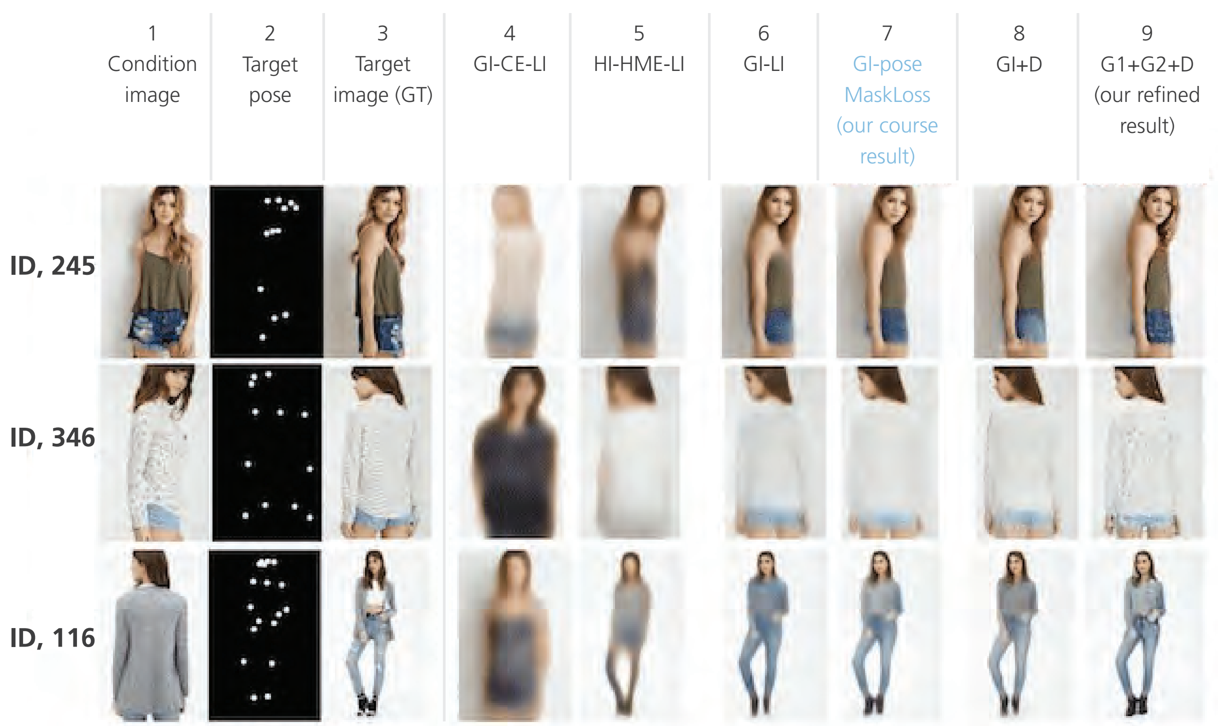


Fig 19: Example of GAN-Generated Photographs of Human Poses (Source: Arxiv)

• Photos to Emojis

Yaniv Taigman, et al. in their 2016 paper titled "Unsupervised Cross-Domain Image Generation" used a GAN to translate images from one domain to another, including from street numbers to MNIST handwritten digits, and from photographs of celebrities to what they call emojis or small cartoon faces.

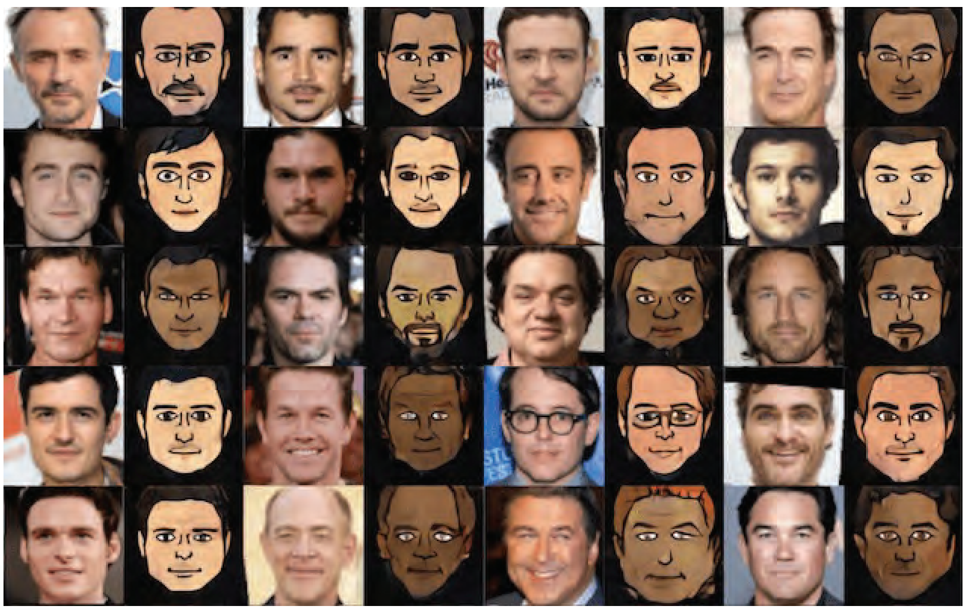


Fig 20: Example of Celebrity Photographs and GAN-Generated Emojis (Source: Arxiv)

• Photograph Editing

Guim Perarnau, et al. in their 2016 paper titled "Invertible Conditional GANs For Image Editing" use a GAN, specifically their IcGAN, to reconstruct photographs of faces with specific specified features, such as changes in hair color, style, facial expression, and even gender.

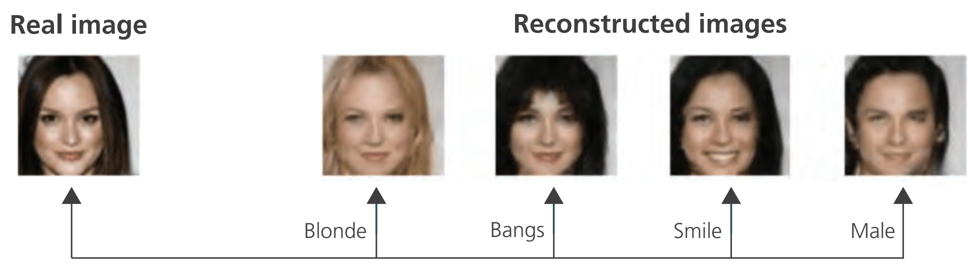


Fig 21: Example of Face Photo Editing with IcGAN (Source: Arxiv)

Ming-Yu Liu, et al. in their 2016 paper titled "Coupled Generative Adversarial Networks" also explores the generation of faces with specific properties such as hair color, facial expression, and glasses. They also explore the generation of other images, such as scenes with varied color and depth.

Fig 22: Example of GANs used to Generate Faces with and Without Blond Hair (Source: Arxiv)

Andrew Brock, et al. in their 2016 paper titled "Neural Photo Editing with Introspective Adversarial Networks" present a face photo editor using a hybrid of variational autoencoders and GANs. The editor allows rapid realistic modification of human faces including changing the hair color, hairstyles, facial expression, pose, and adding facial hair.

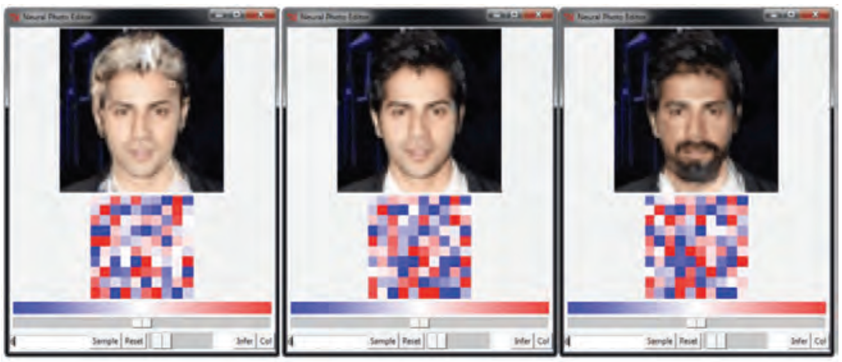


Fig 23: Example of Face Editing Using the Neural Photo Editor Based on VAEs and GANs (Source: Arxiv)

He Zhang, et al. in their 2017 paper titled "Image De-raining Using a Conditional Generative Adversarial Network" use GANs for image editing, including examples such as removing rain and snow from photographs.

(a) (b) (c) (d)

Fig 24: Example of Face Editing Using the Neural Photo Editor Based on VAEs and GANs (Source: Arxiv)

• Face Aging

Grigory Antipov, et al. in their 2017 paper titled "Face Aging with Conditional Generative Adversarial Networks" use GANs to generate photographs of faces with different apparent ages, from younger to older.

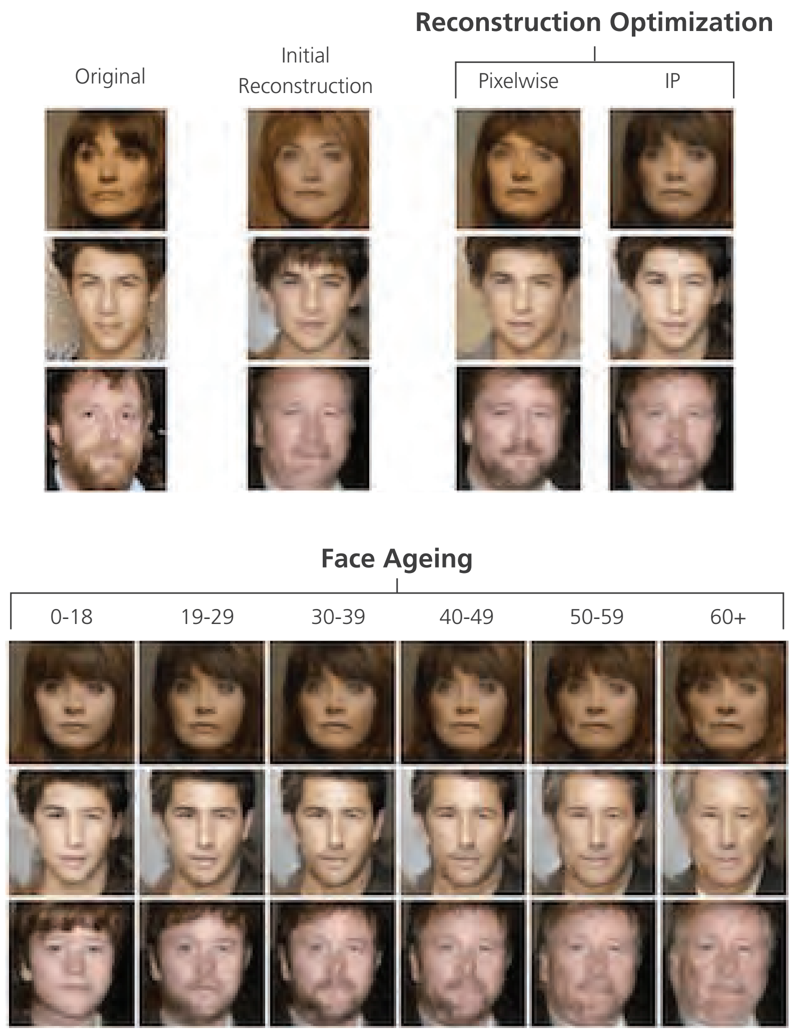


Fig 25: Example of Photographs of Faces Generated with a GAN With Different Apparent (Source: Arxiv)

• Photo Blending

Huikai Wu, et al. in their 2017 paper titled "GP-GAN: Towards Realistic High-Resolution Image Blending" demonstrates the use of GANs in blending photographs, specifical elements from different photographs such as fields, mountains, and other large structures.

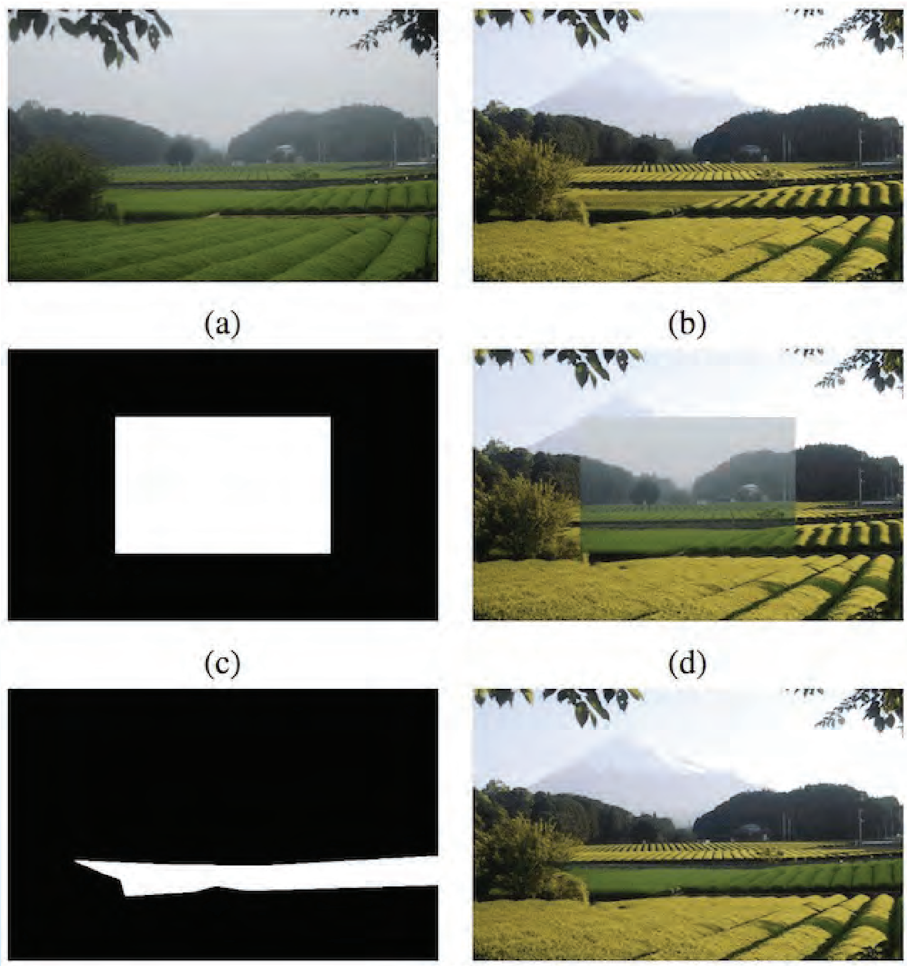


Fig 26: Example of GAN-based Photograph Blending (Source: Arxiv)

- ## Super Resolution

Christian Ledig, et al. in their 2016 paper titled "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network" demonstrate the use of GANs, specifically their SRGAN model, to generate output images with higher, sometimes much higher, pixel resolution.

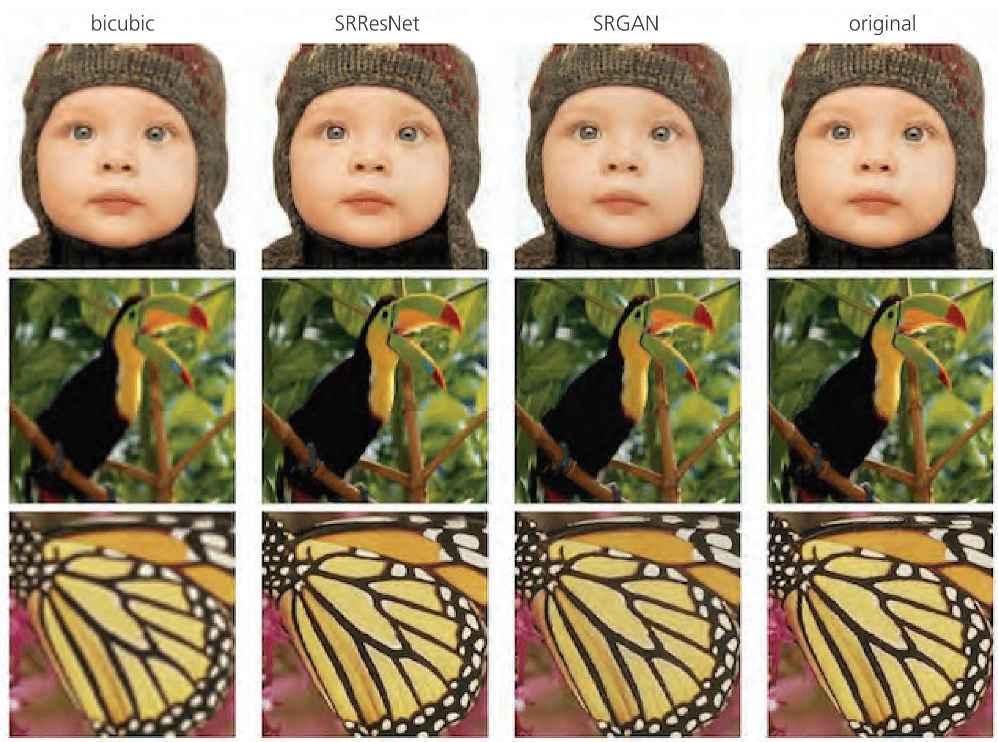


Fig 27: Example of GAN-Generated Images with Super Resolution (Source: Arxiv)

Subeesh Vasu, et al. in their 2018 paper titled "Analysing Perception-Distortion Trade-off using Enhanced Perceptual Super-Resolution Network" provide an example of GANs for creating high-resolution photographs, focusing on street scenes.

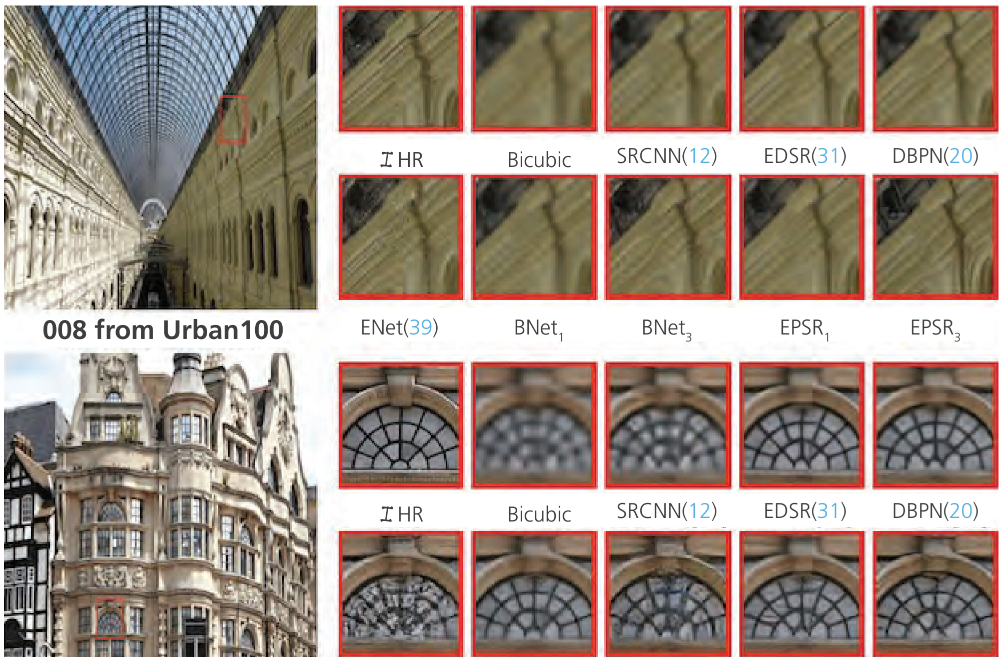


Fig 28: Example of High-Resolution GAN-Generated Photographs of Buildings (Source: Arxiv)

• Photo Inpainting

Deepak Pathak, et al. in their 2016 paper titled "Context Encoders: Feature Learning by Inpainting" describe the use of GANs, specifically Context Encoders, to perform photograph inpainting or hole filling, that is filling in an area of a photograph that was removed for some reason.

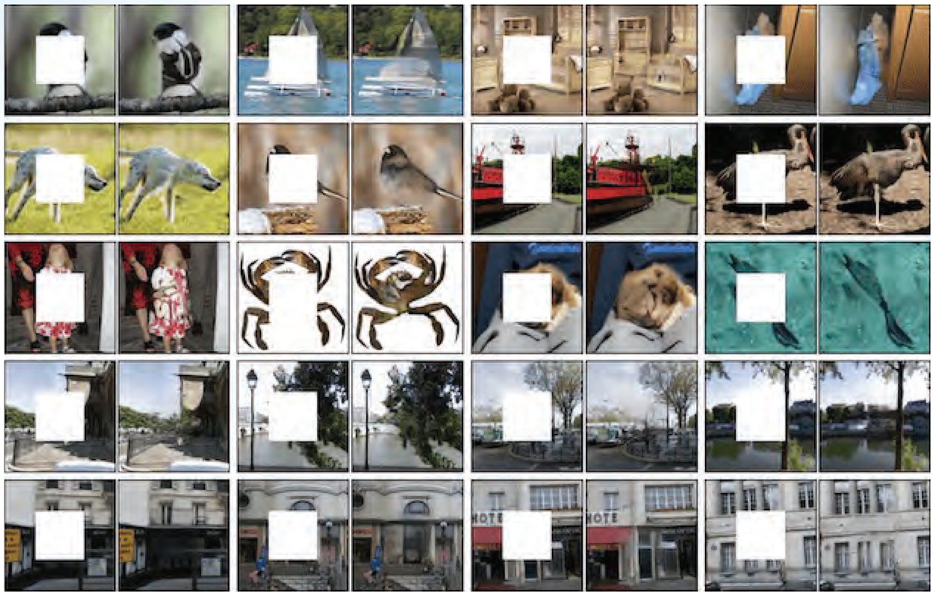


Fig 29: Example of GAN-Generated Photograph Inpainting Using Context Encoders (Source: Arxiv)

Raymond A. Yeh, et al. in their 2016 paper titled "Semantic Image Inpainting with Deep Generative Models" use GANs to fill in and repair intentionally damaged photographs of human faces.

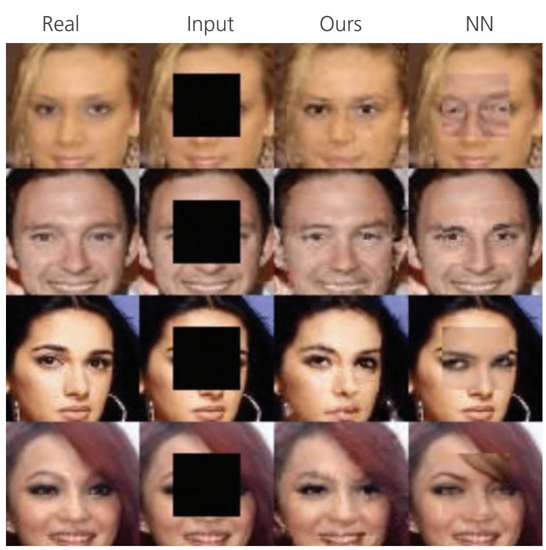


Fig 30: Example of GAN-based Inpainting of Photographs of Human Face (Source: Arxiv)

• Video Prediction

Carl Vondrick, et al. in their 2016 paper titled "Generating Videos with Scene Dynamics" describe the use of GANs for video prediction, specifically predicting up to a second of video frames with success, for static elements of the scene.

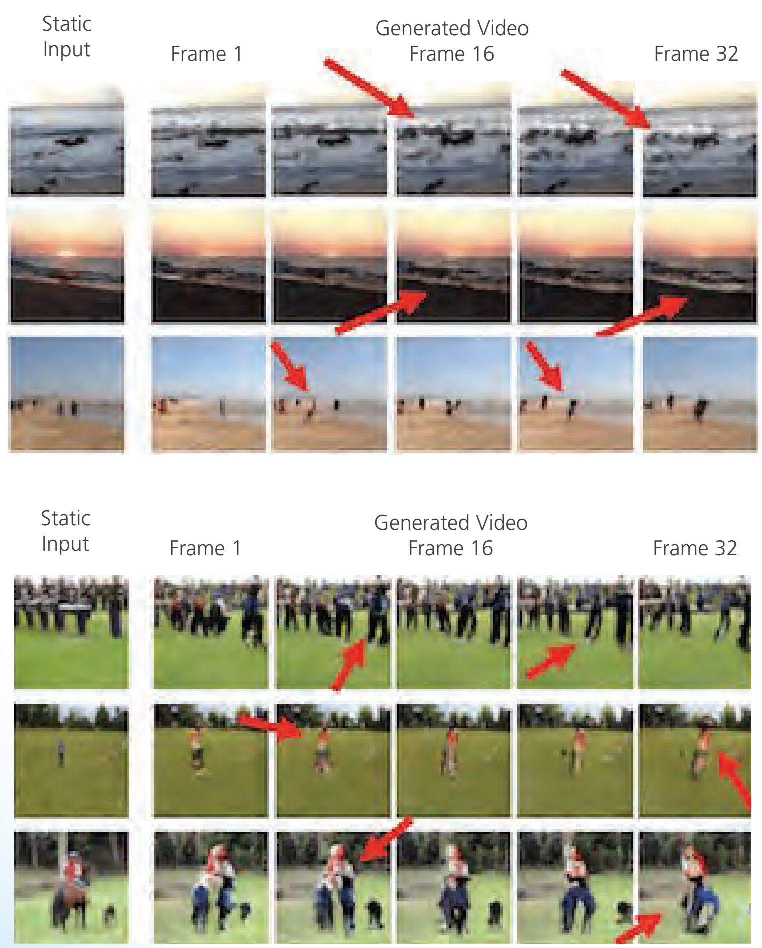


Fig 31: Example of Video Frames Generated with a GAN (Source: Arxiv)

- ## 3D Object Generation

Jiajun Wu, et al. in their 2016 paper titled "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modelling" demonstrates a GAN for generating new three-dimensional objects (e.g., 3D models) such as chairs, cars, sofas, and tables.

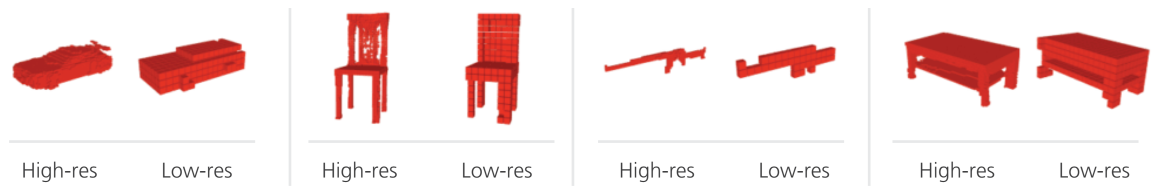


Fig 32: Example of GAN-Generated Three-Dimensional Objects (Source: Arxiv)

Matheus Gadelha, et al. in their 2016 paper titled "3D Shape Induction from 2D Views of Multiple Objects" use GANs to generate three-dimensional models given two-dimensional pictures of objects from multiple perspectives.

Fig 33: Example of Three-Dimensional Reconstructions of a Chair from Two-Dimensional Images (Source: Arxiv)

07 Challenges in Generative AI

While implementing the best practise of Generative AI, keep in mind potential bottlenecks and misconceptions. Some of these include:

- ## Safety

Utilization of Generative AI for cyber theft and criminal activities like scamming and identity theft are the biggest problems that emerge when it comes to Generative AI implementation. It has been reported that people are using this technology for scamming and theft.

- ## Highly limited abilities

Generative AI algorithms require extensive training and a high amount of data to perform tasks like creating digital art. Despite this, the content generated is not 100% new. Instead, these models can only mix and match and sequence the data in the best way possible.

- ## Unpredictable outcomes

Accuracy of the result is another challenge that crops up while implementing this technology. GAN's processes remain unstable and difficult to regulate, with the potential to produce completely unexpected results. While adopting some models, it is easier to manage the behavior of Generative AI, but in heavy applications, they yield erroneous and unexpected results.

- ## Data Security

Verticals like healthcare and defense are reluctant to adopt Generative AI, as there are no parameters available for data moderation, and Generative AI-based applications may create data security and privacy issues.

- ## Massive Data Sets Requirement

You cannot rely on generative AI algorithms to work well unless they have a substantial quantity of input content. This program can do miracles, but only within the constraints set by the training data. It cannot generate fresh text or images out of thin air.

08 Concluding Notes

It is evident that Generative AI is an extremely nascent technology with limited industrial use cases currently available. Due to this, clear market directives and revenue areas are yet to be identified.

Despite the buzz around technology and huge market capitalization potential, it is too early to predict the direction in which the market is heading. The timeline for Generative AI technology and its industry use cases seems to be farther away on the horizon.

If a company is looking to gain a competitive advantage in this domain, it should focus on understanding the fundamental AI models, delving deeper into the research as it stands today, and striving to develop a proof of concept.

By following the above suggestions, you would be ready to capture a significant market share when the market demands a full solution.

09

Authors

Sachin Jain
Head - Crystal & Deep POV

Bharat Trivedi
Principal – Enterprise Architecture

Manish Potdar
Head- Incubation & Industrialization

Hakimuddin Bawangaonwala
Research Analyst – Crystal & Deep POV

Soham Patankar
Lead Strategist- Incubation

10 References

[1] https://imagen.research.google/

[2] https://www.techopedia.com/definition/34633/generative-ai

[3] https://dresma.ai/what-is-generative-ai/

[4] https://arxiv.org/pdf/2205.11487.pdf

[5] https://www.researchgate.net/publication/230708329_Generative_Artificial_Intelligence

[6] https://medium.com/@bks46/introduction-to-generative-ai-models-6a5ebcebc168

[7] https://analyticsindiamag.com/7-types-of-generative-models-for-your-next-machine-learning-project/

[8] https://www.deepmind.com/blog/wavenet-a-generative-model-for-raw-audio

[9] https://jalammar.github.io/illustrated-transformer/

[10] https://www.marketdataforecast.com/market-reports/generative-design-market

[11] https://www.ventureradar.com/startup/Generative%20AI

[12] https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

[13] https://alexrachnog.medium.com/gans-beyond-generation-7-alternative-use-cases-725c60ba95e8

[14] https://sunverasoftware.com/10-use-cases-for-generative-ai/

[15] https://coe-dsai.nasscom.in/generative-ai-exploring-the-world-of-possibilities/

[16] https://www.analyticsinsight.net/what-is-generative-ai-its-impacts-and-limitations/

[17] https://www.e-spincorp.com/what-are-the-challenges-of-generative-ai-adoption/

[18] https://www.tcs.com/bridging-the-human-machine-divide

[19] https://www.infosys.com/iki/perspectives/advanced-trends-ai.html

[20] https://www.cognizant.com/us/en/ai/evolutionary-ai

[21] https://www.hcltech.com/blogs/applied-ai-time-lead-front

[22] https://www.hcltech.com/next-ai

[23] https://research.ibm.com/interactive/generative-models/

[24] https://www.microsoft.com/en-us/ai/ai-lab-gen-studio

[25] https://aws.amazon.com/blogs/startups/synthesis-ais-fules-computer-vision-innovation/

[26] https://ai.google/

[27] https://arxiv.org/abs/1406.2661

[28] https://arxiv.org/abs/1511.06434

[29] https://arxiv.org/abs/1710.10196

[30] https://arxiv.org/abs/1802.07228

[31] https://arxiv.org/abs/1809.11096

[32] https://arxiv.org/abs/1708.05509

[33] https://github.com/moxiegushi/pokeGAN

[34] https://github.com/kvpratama/gan/tree/master/pokemon

[35] https://arxiv.org/abs/1611.07004

[36] https://arxiv.org/abs/1703.10593

[37] https://junyanz.github.io/CycleGAN/

[38] https://arxiv.org/abs/1612.03242

[39] https://arxiv.org/abs/1605.05396

[40] https://arxiv.org/abs/1703.06412

[41] https://arxiv.org/abs/1610.02454

[42] https://arxiv.org/abs/1711.11585

[43] https://arxiv.org/abs/1705.09368

[44] https://arxiv.org/abs/1611.02200

[45] https://arxiv.org/abs/1611.06355

[46] https://arxiv.org/abs/1606.07536

[47] https://arxiv.org/abs/1609.07093

[48] https://arxiv.org/abs/1701.05957

[49] https://ieeexplore.ieee.org/document/8296650

[50] https://arxiv.org/abs/1702.08423
[51] https://arxiv.org/abs/1703.07195
[52] https://arxiv.org/abs/1609.04802
[53] https://arxiv.org/abs/1707.00737
[54] https://arxiv.org/abs/1811.00344
[55] https://arxiv.org/abs/1604.07379
[56] https://arxiv.org/abs/1607.07539
[57] https://arxiv.org/abs/1704.05838
[58] https://arxiv.org/abs/1603.07442
[59] https://arxiv.org/abs/1609.02612
[60] https://arxiv.org/abs/1610.07584
[61] https://arxiv.org/abs/1612.05872
[62] https://www.itbusinessedge.com/data-center/what-is-generative-ai/#Challenges_of_Generative_AI
[63] https://nix-united.com/blog/pushing-the-technological-envelope-with-generative-ai/