Intelligent Inclusive
Interaction Design (I³D) Lab

Mixed Reality & Computer Vision

By Gyanig Kumar
Research Assistant, I3D Lab, IISc

Overview

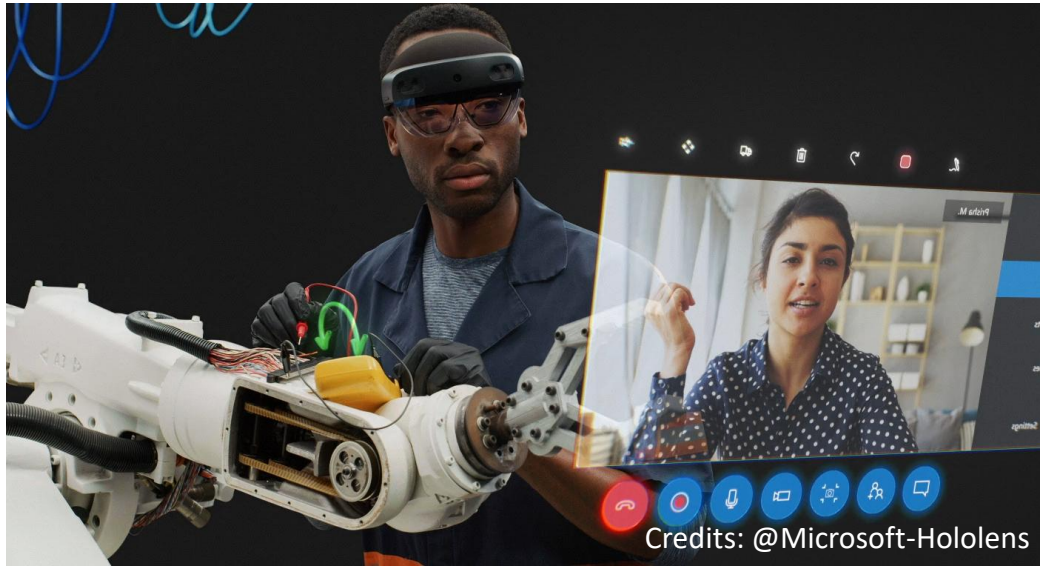
- Why Computer Vision is needed for Mixed Reality ?
- How do we integrate Computer Vision with Mixed Reality ?
- Practical Problems
- Role of Generative AI – synthetic data

Mixed-Reality Environment

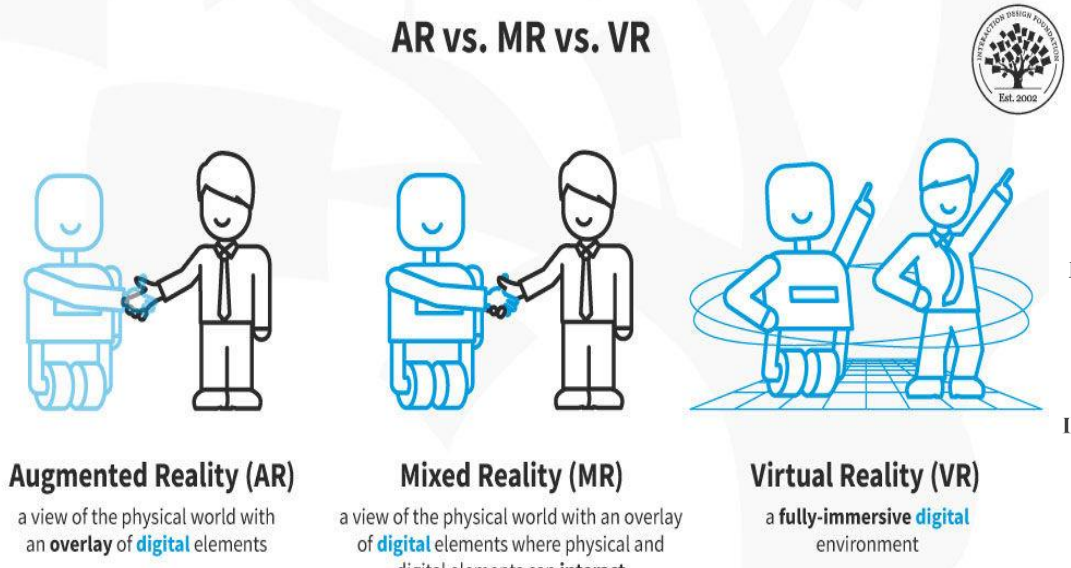
MR emerged with the creation of the MR environment. A real-world object interacts with a virtual object to execute practical scenarios for the user.

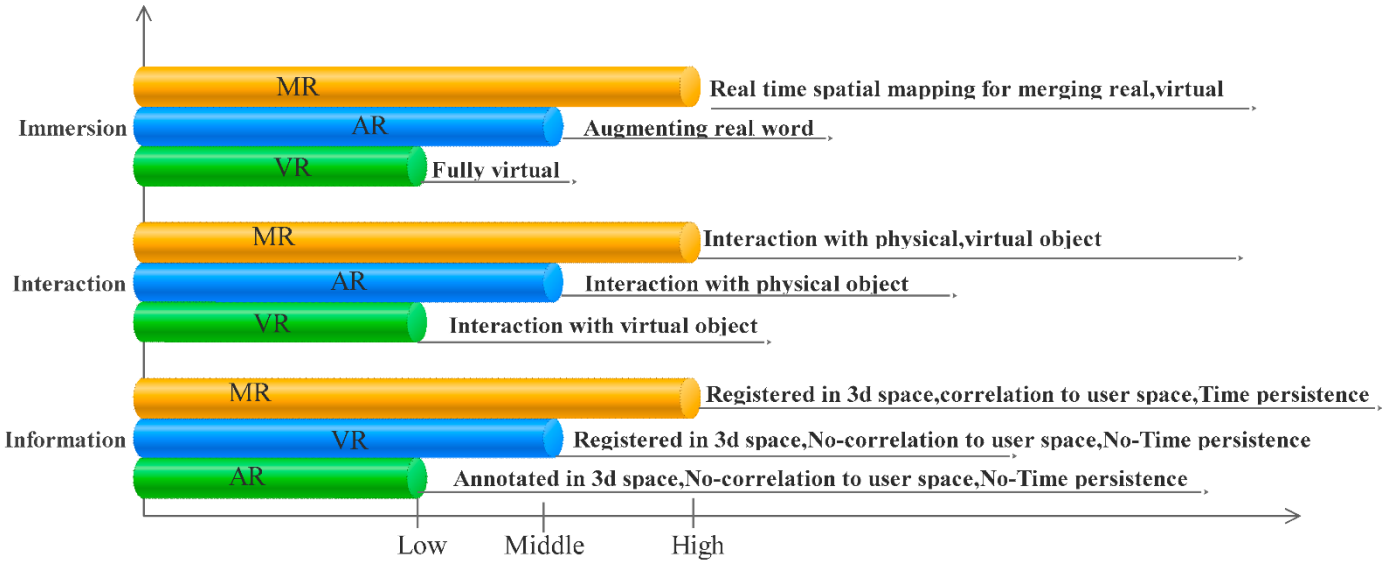
There are three important features of any MR system:

(1) combining the real-world object and the virtual object.
(2) interacting in real-time.
(3) mapping between the virtual object
and the real object to create interactions between them.



Why MR and not VR or AR?



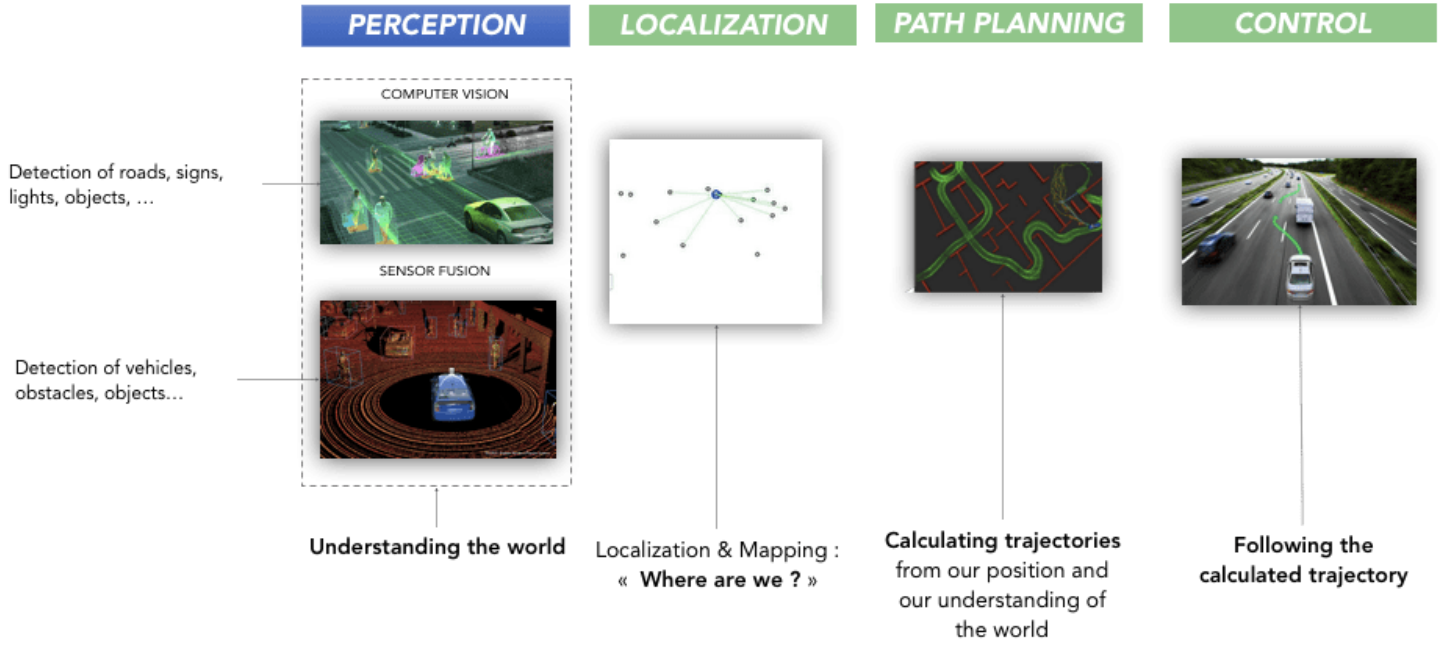


Rokhsaritalemi, Somaiieh, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. 2020. "A Review on Mixed Reality: Current Trends, Challenges and Prospects" *Applied Sciences* 10, no. 2: 636. https://doi.org/10.3390/app10020636

Interaction Between Real world Objects and Mixed Reality Holograms

Intelligent Inclusive Interaction Design (I³D) Lab

Computer Vision



Computer Vision is a tool for Mixed reality Research.

Application Areas of Computer Vision are :

- **Computer Vision applications in Autonomous Vehicles**
  - Lane Tracking
  - Vehicle Detection
  - Pedestrian Detection
- **Computer Vision in Agriculture**
  - Automating Animal Counting and Tracking
  - Crop Disease Detection
- **Computer Vision in Security**
  - Facial Authentication
  - Public Safety Monitoring
  - CCTV Cameras for Unusual Activity Tracking
- **Computer Vision in HealthCare**
  - Breast Cancer Detection
  - More Precise Diagnosis
  - Interactive Medical Imaging
- **Computer Vision in Education Industry**
  - Engagement Detection in Online Classes
  - Security in Educational Institutes
  - Cheating Detection in Exams
- **Computer Vision in Manufacturing**
  - Predictive Maintenance
  - 3D Vision Inspection
  - Packaging Inspection

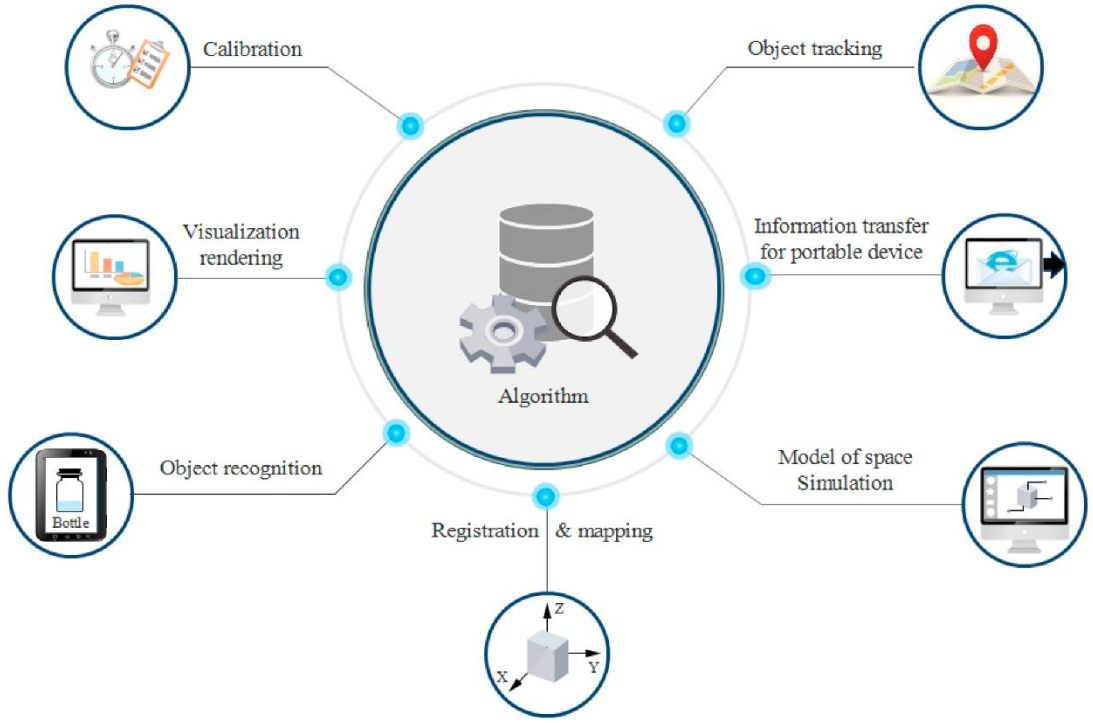
Many more…..

Why Computer Vision is needed for Mixed Reality ?

Computer vision and Mixed Reality technology encompass a wide range of capabilities, including

- Ego-centric object recognition,
- Hand and user tracking,
- Activity recognition,
- SLAM,
- 3D reconstruction,
- Scene understanding,
- Sensor-based localization,
- Navigation,

and numerous other advanced functionalities.

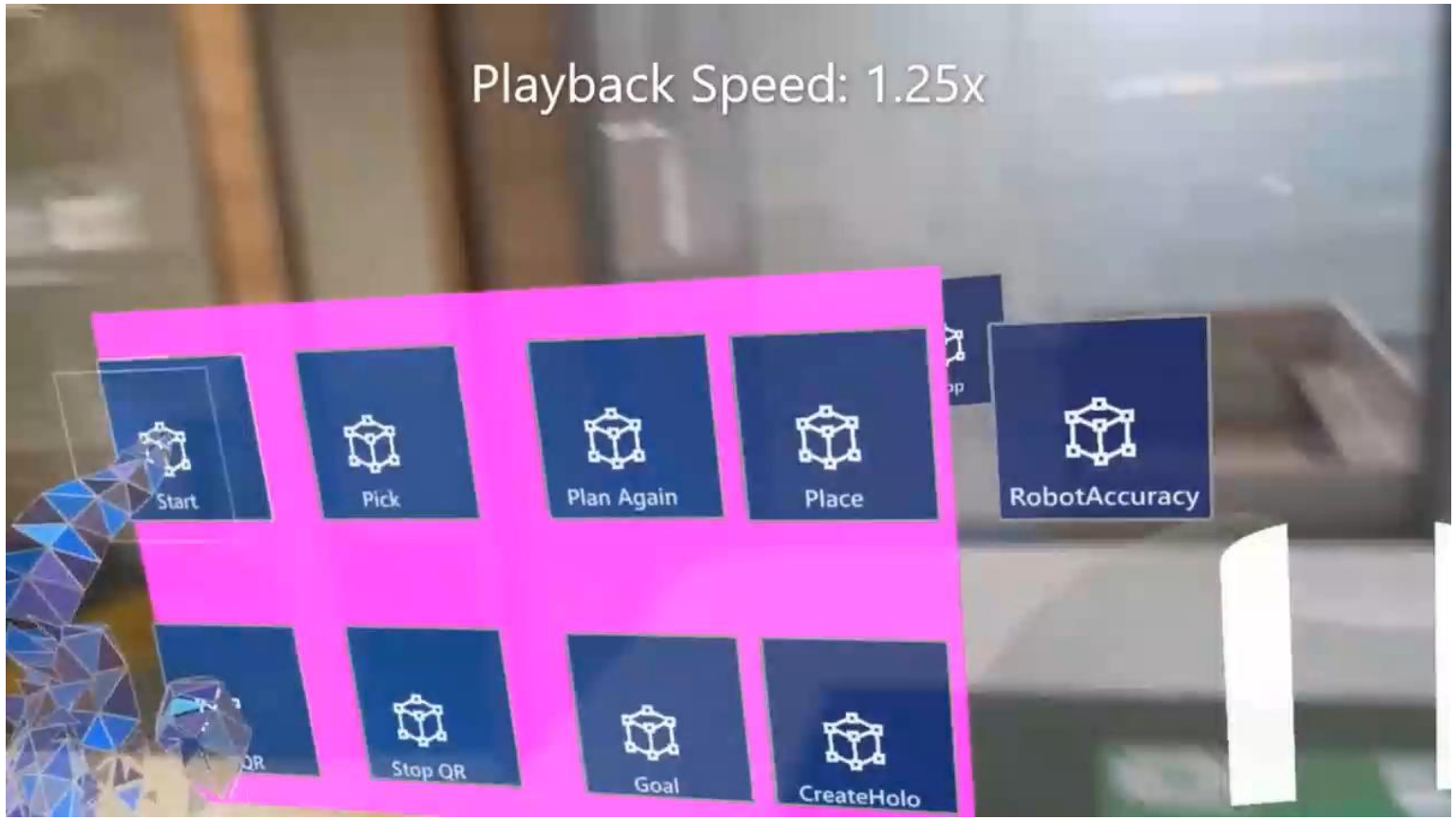


Rokhsaritalemi, Somaiieh, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. 2020. "A Review on Mixed Reality: Current Trends, Challenges and Prospects" *Applied Sciences* 10, no. 2: 636. https://doi.org/10.3390/app10020636

Areas of Mixed Reality and Computer Vision

- Immersive Experiences
- Training and Education
- Architectural Visualization
- Remote Collaboration
- Industrial Maintenance
- Healthcare
- Retail and Marketing
- Gaming
- Research and Innovation

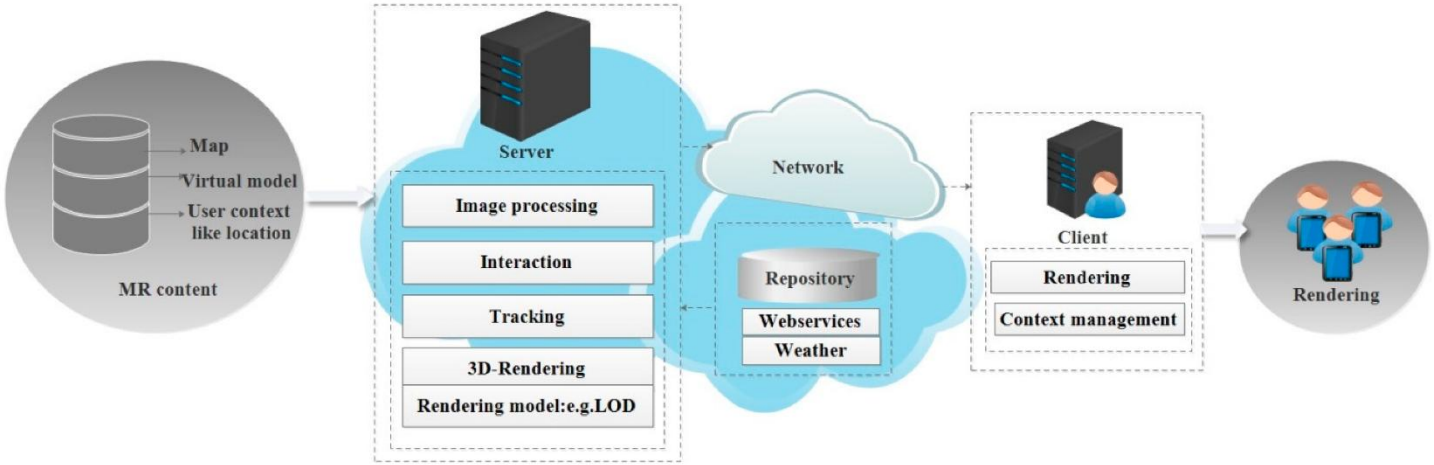
Robotic Task Completion in Mixed Reality



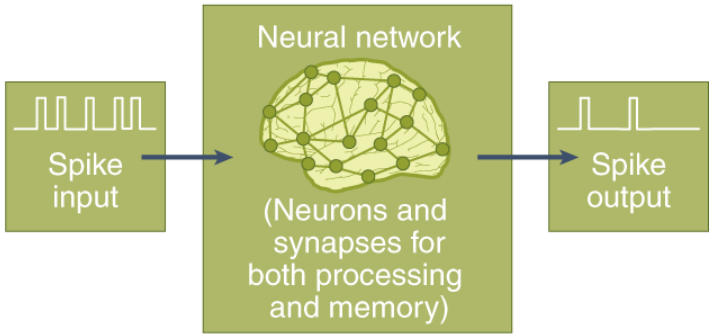
How do we Integrate Computer Vision in Mixed Reality?

What kind of Data we focus ?

- Realtime Camera
- Tracking Environment Cameras
- Eye Tracking



Rokhsaritalemi, Somaiieh, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. 2020. "A Review on Mixed Reality: Current Trends, Challenges and Prospects" *Applied Sciences* 10, no. 2: 636. https://doi.org/10.3390/app10020636



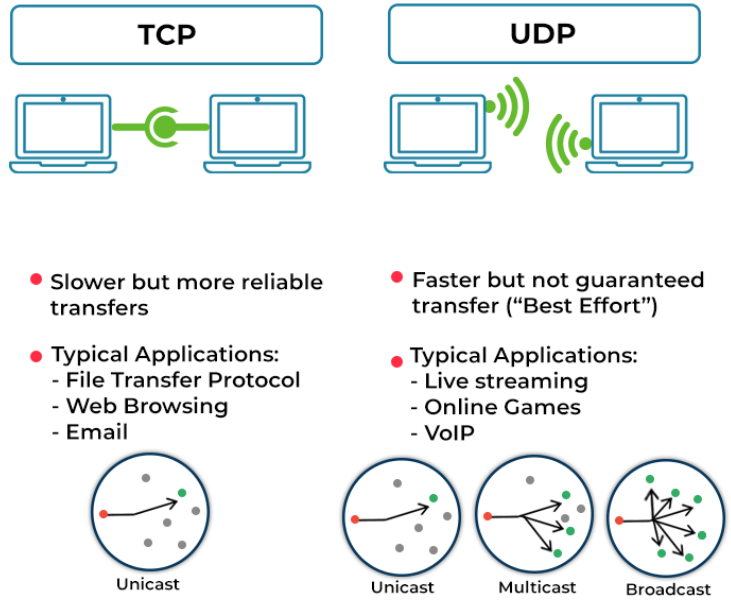
Operation	→	Massively parallel processing
Organization	→	Collocated processing and memory
Programming	→	Spiking neural network
Communication	→	Spikes
Timing	→	Asynchronous (event-driven)

We need High Computation for most Computer Vision Algorithms. By convention, we process data on a dedicated system and deploy on Mixed Reality Devices.
There has been little development on running such algorithms on MR devices.

Schuman, C.D., Kulkarni, S.R., Parsa, M. *et al.* Opportunities for neuromorphic computing algorithms and applications. *Nat Comput Sci* **2**, 10–19 (2022). https://doi.org/10.1038/s43588-021-00184-y

So, Communicate with dedicated Systems….

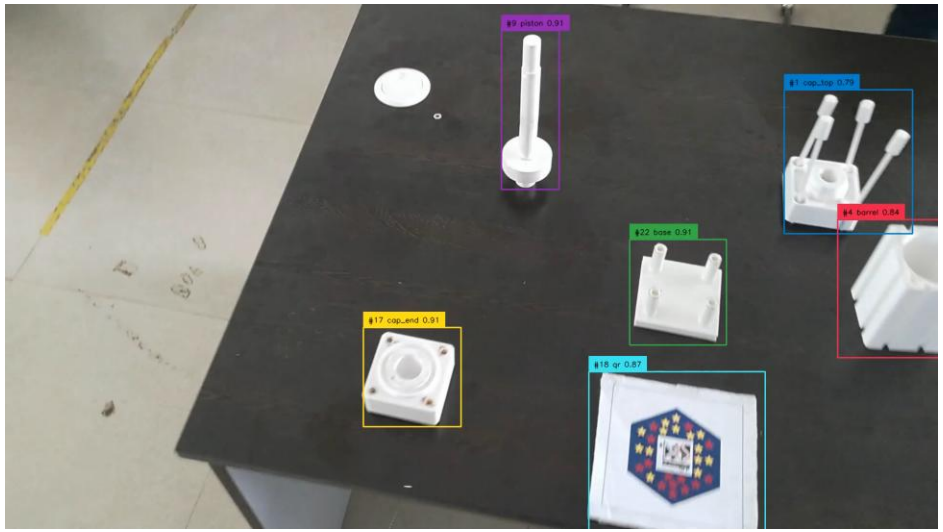
- WebSocket
- MQTT (Message Queuing Telemetry Transport)
- UDP (User Datagram Protocol)
- TCP (Transmission Control Protocol)
- REST API (Representational State Transfer)
- WebRTC (Web Real-Time Communication)
- CoAP (Constrained Application Protocol)
- UDP Multicast
- Cloud APIs



https://www.spiceworks.com/tech/networking/articles/tcp-vs-udp/

Practical Problems

- Latency
  - Realtime scenarios
  - Synchronization
- Object Recognition
  - Environmental Variability
  - Lack of Dataset
- Calibration
  - Regression Based
- Interactivity
  - Multi-Object Tracking
  - 3D object Tracking
- Content Creation
- User Experience
- Integration with Real World



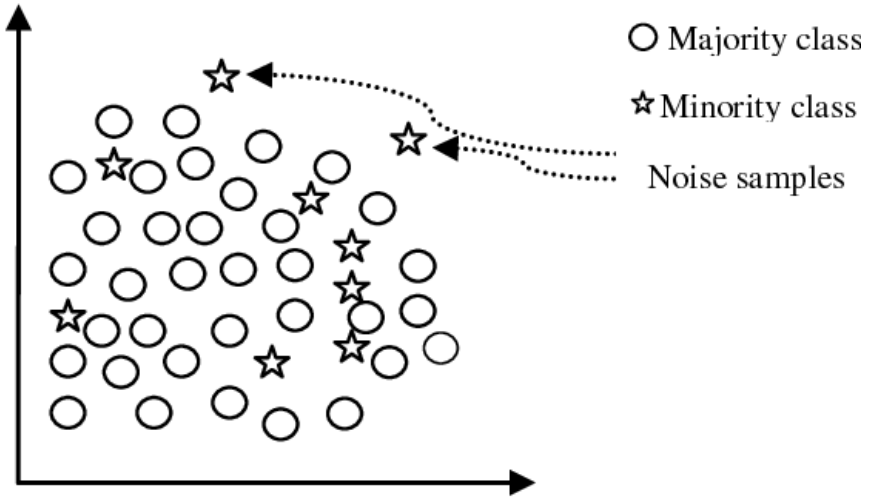


Fig 1. The data set having a between-class imbalance

Maheshwari, Satyam et al. “A Review on Class Imbalance Problem: Analysis and Potential Solutions.” (2017).

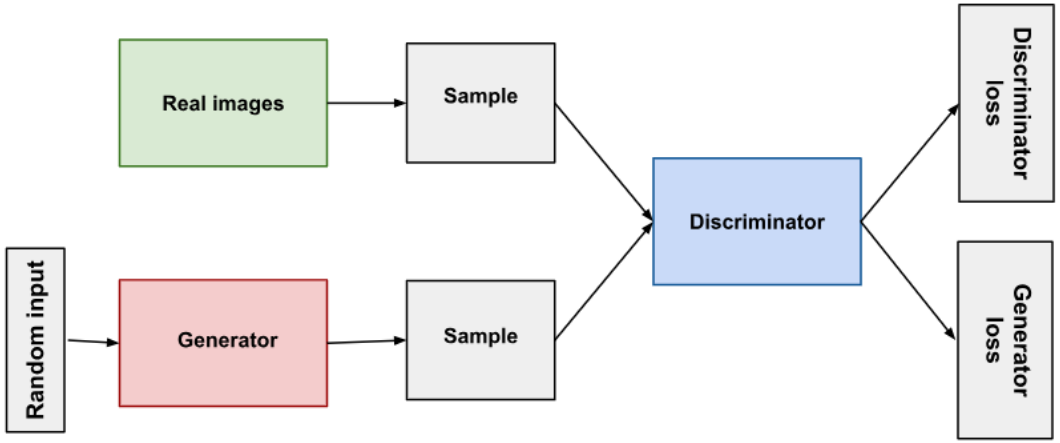
Why do we to generate better Datasets with Generative AI?

Generative AI in Computer Vision and Mixed Reality

- It is an exciting recent innovation in machine learning.
- GANs are *generative* models:
  - Create new data instances that resemble your training data.
  - Achieve this level of realism by pairing a **generator** and with a **discriminator**

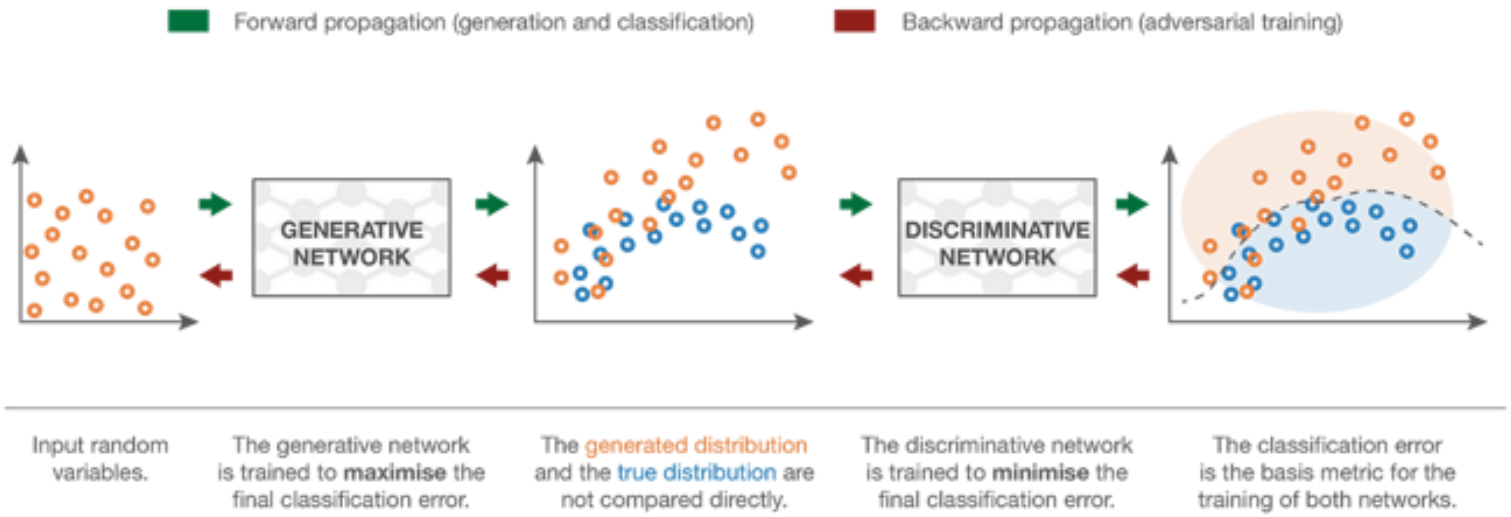
Fig. Images generated by a GAN (Source: NVIDIA)

Overview of GAN Structure



- **Generator** learns to generate plausible data. The generated instances become negative training examples for the discriminator.
- **Discriminator** learns to distinguish the generator's fake data from real data. The discriminator penalizes the generator for producing implausible results.
- Generator output is connected directly to the discriminator input.

More on GANs.....



Two networks can then be trained jointly (at the same time) with opposite goals:

- the goal of the generator is to fool the discriminator, so the generative neural network is trained to **maximise the final classification error** (between true and generated data)

- the goal of the discriminator is to detect fake generated data, so the discriminative neural network is trained to **minimise the final classification error**

Generation of Synthetic Data

The Use-Cases of GANs

- Virtual Environments
- Style Transfer
- Object Detection
- 3D Object Generation
- Data Imbalance Correction
- Image Augmentation

Intelligent Inclusive
Interaction Design (I³D) Lab

Thank You.

By Gyanig Kumar
Research Assistant, I3D Lab, IISc