Introduction to Multimedia Data Modelling

Luca Guarnera, Ph.D.

Research Fellow

luca.guarnera@unict.it

University of Catania
Dipartimento di Matematica e Informatica

Multimedia Data …

… Data Modelling

What are Multimedia?

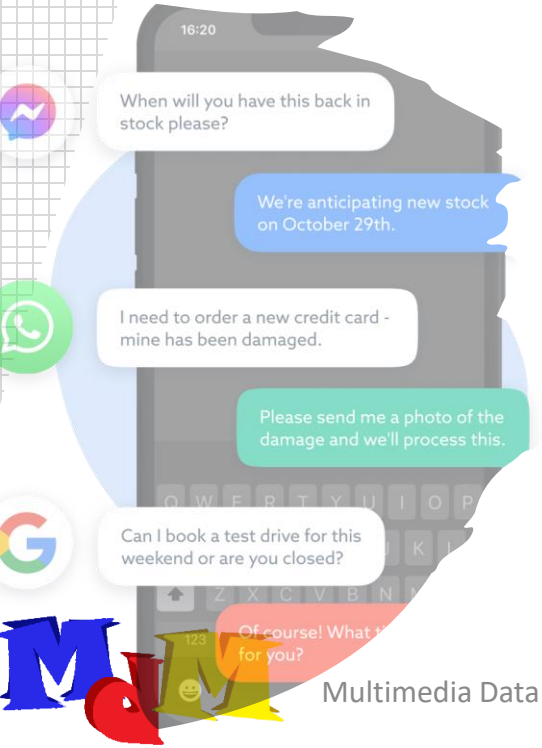
Multimedia Definition

Multimedia can be defined as a **combination of different forms of content** such as ***text***, ***graphics***, ***audio***, ***video***, and ***animation***

Exploring the different forms of multimedia

Text

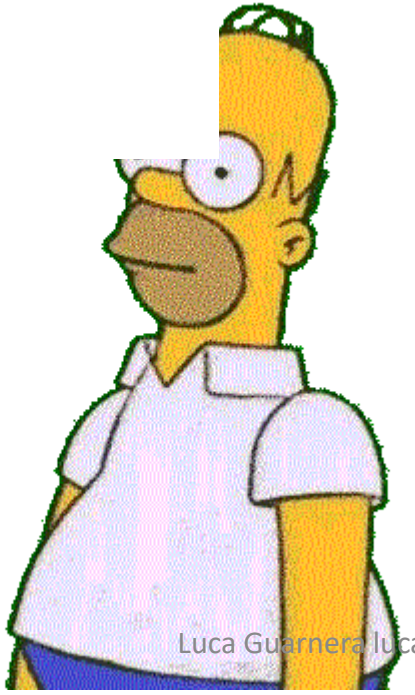
Text is the most **basic form of multimedia**. It is used to communicate information and ideas in a concise and structured way. Text can be used to explain complex concepts, tell stories, and provide instructions.



Graphics

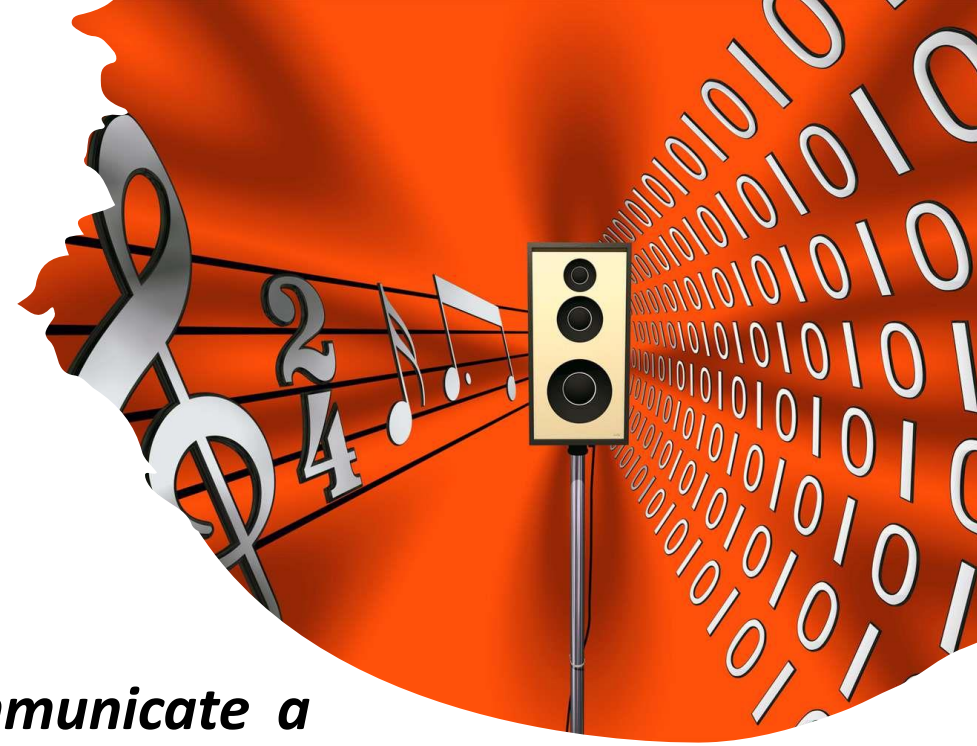
Graphics are ***visual elements that can be used to convey a message or an idea***. They can be used to illustrate concepts, highlight important information, and create an engaging experience. Graphics can be **static** or **animated**.

Animated Graphics

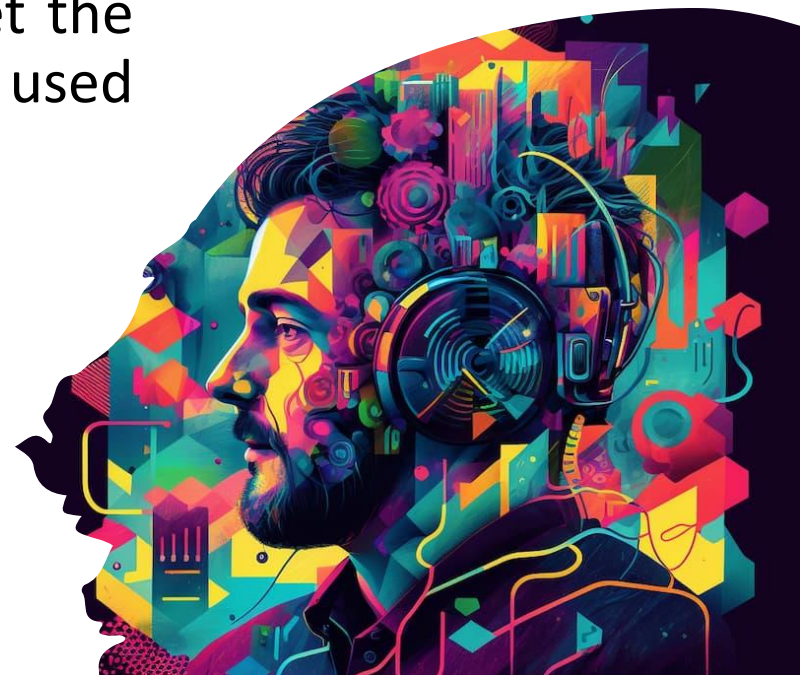


Luca Guarnera luca.guarnera@unict.it

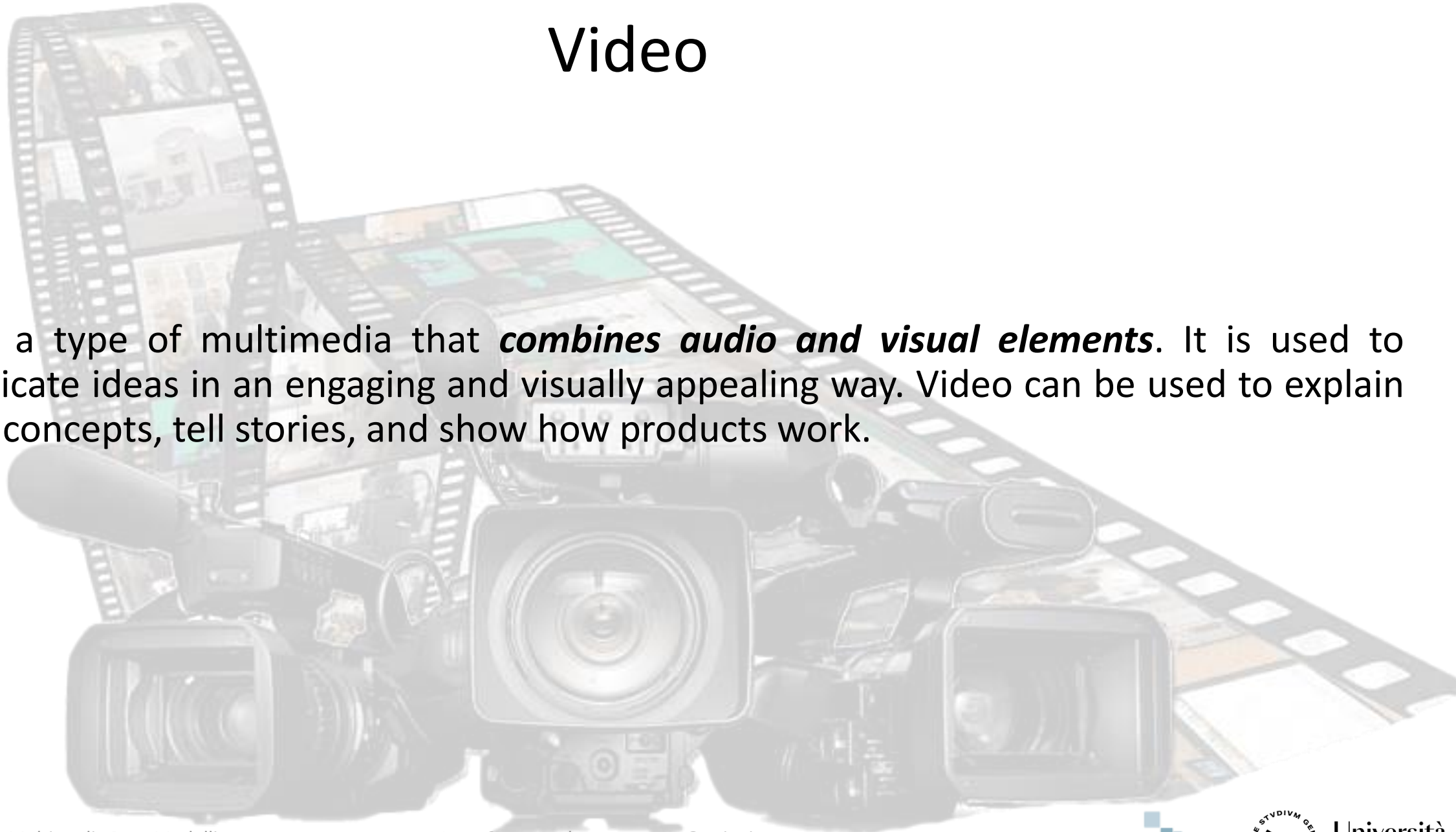
Audio



Audio is a type of multimedia that uses ***sound to communicate a message***. It can be used to create an immersive experience, set the tone for an experience, and convey emotions. Audio can also be used to provide instructions and feedback.



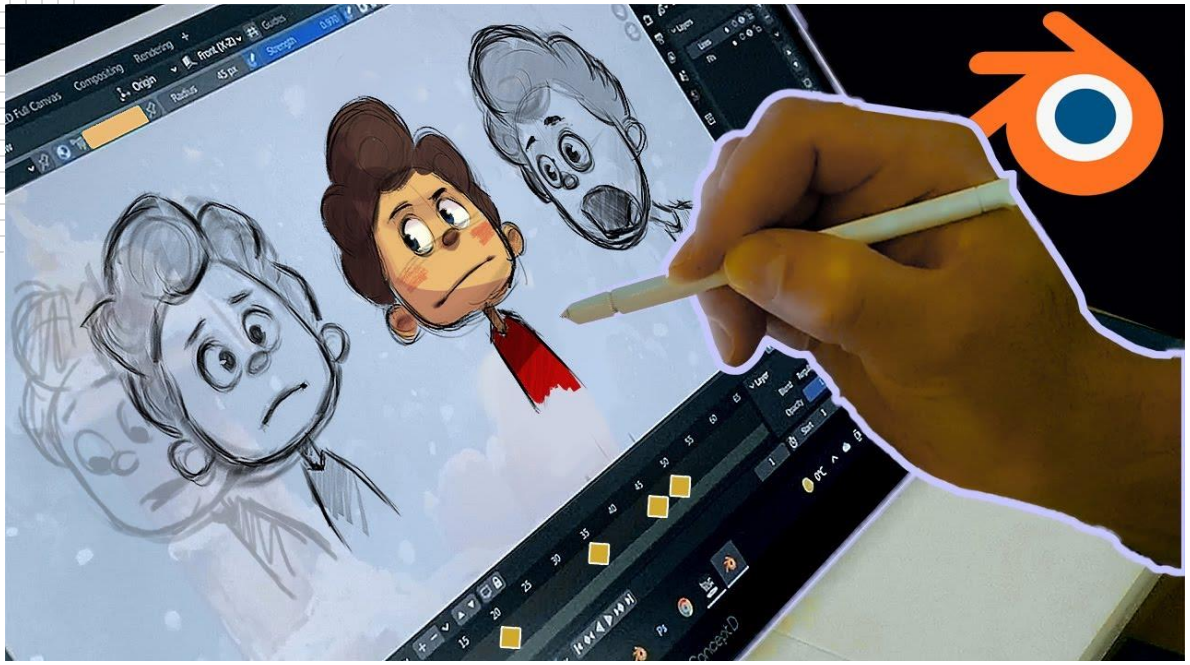
Video



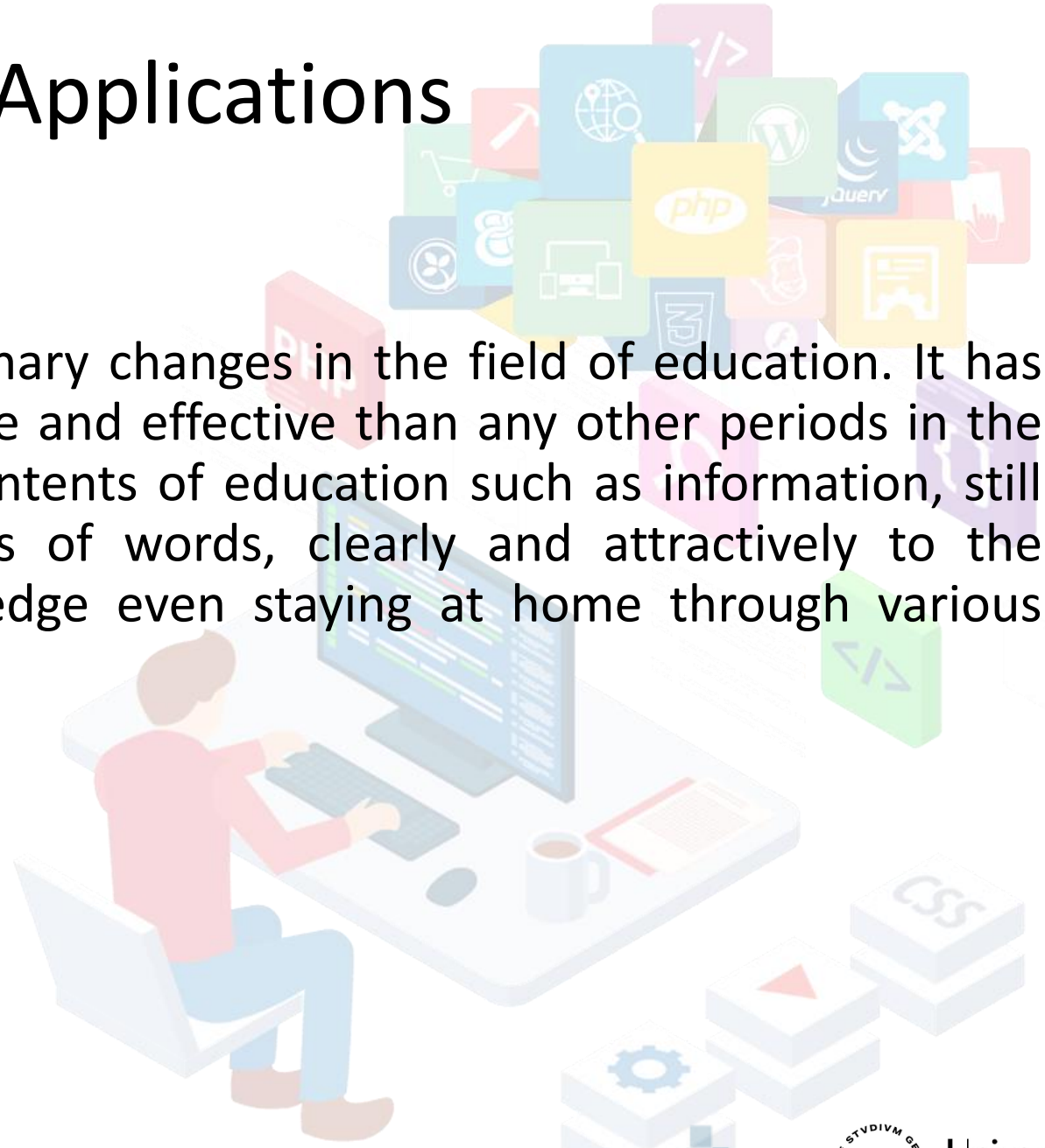
Video is a type of multimedia that ***combines audio and visual elements***. It is used to communicate ideas in an engaging and visually appealing way. Video can be used to explain complex concepts, tell stories, and show how products work.

Animation

Animation is a type of multimedia that ***combines graphics and audio***. It is used to create interactive experiences that are engaging and entertaining. Animation can be used to explain concepts, tell stories, and provide feedback.

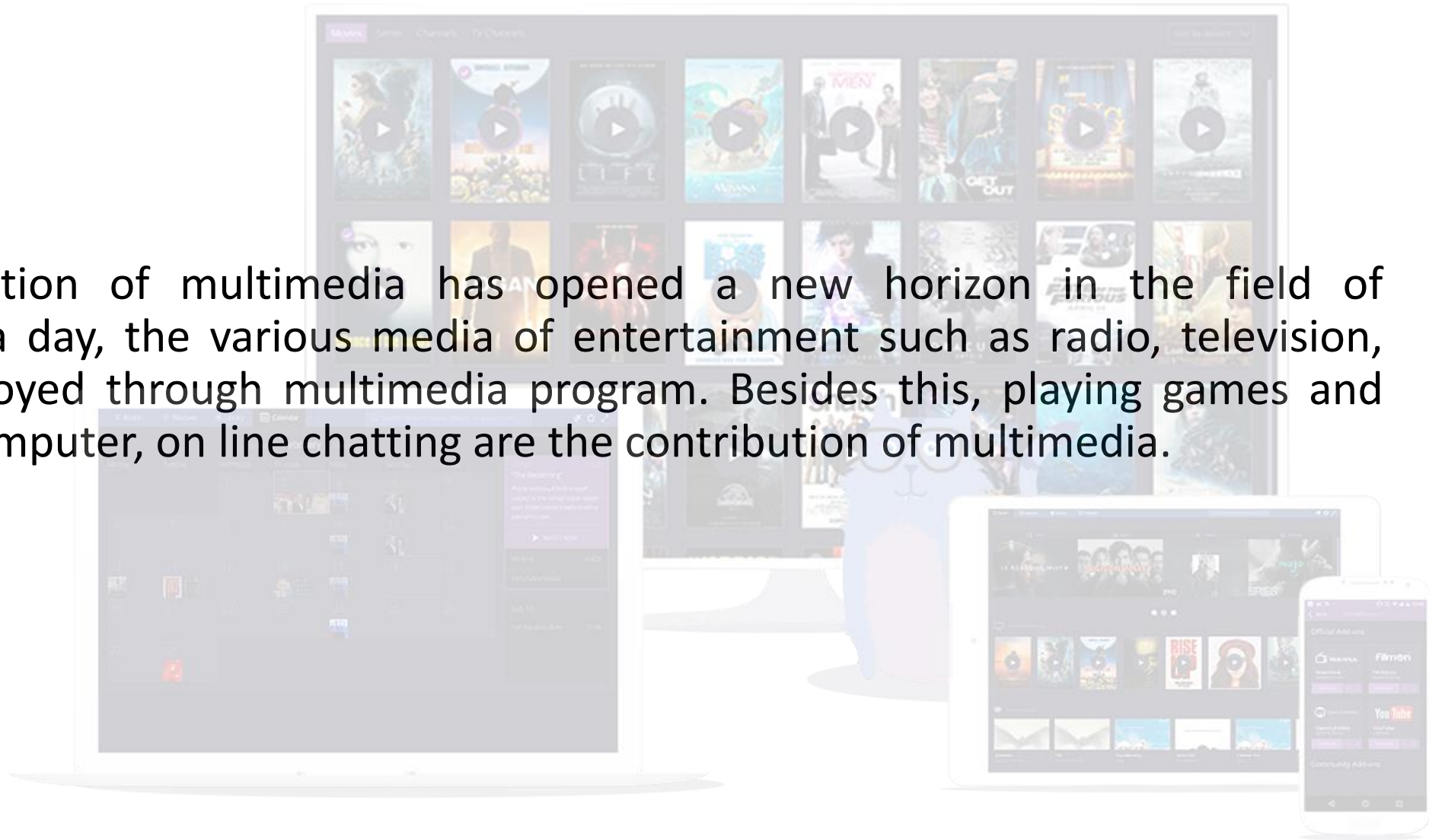


Multimedia Applications



Education: Multimedia has brought revolutionary changes in the field of education. It has made the system of education more attractive and effective than any other periods in the past. Now a day, multimedia presents the contents of education such as information, still and moving pictures, sounds pronunciations of words, clearly and attractively to the students. It is now possible to earn knowledge even staying at home through various multimedia software.

Multimedia Applications



Entertainment: Invention of multimedia has opened a new horizon in the field of entertainment. Now a day, the various media of entertainment such as radio, television, VCR, VCD can be enjoyed through multimedia program. Besides this, playing games and drawing pictures in computer, on line chatting are the contribution of multimedia.

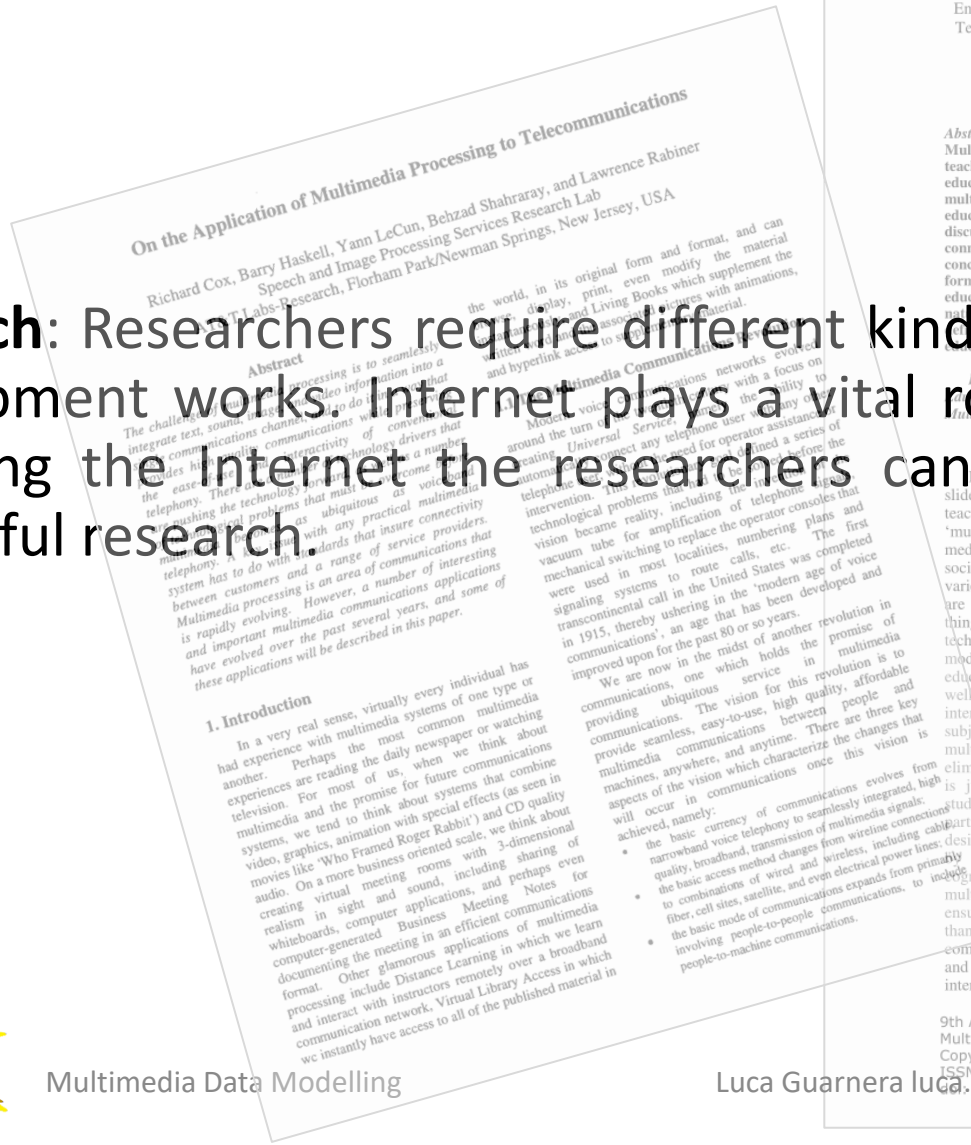
Multimedia Applications

Marketing: In order to increase sales, large companies develop CD Rom software with different information like prices, uses, terms of sales of their products and services and distribute them to the probable customers. Therefore, the buyers can buy their required products by knowing the prices, terms of sale etc. from the CD even staying at home or office.

Multimedia Applications

Communication: Today, multimedia has brought audio and video conference. These computer and electronic based communication creates dynamism in business, social, political, economic and international activities.

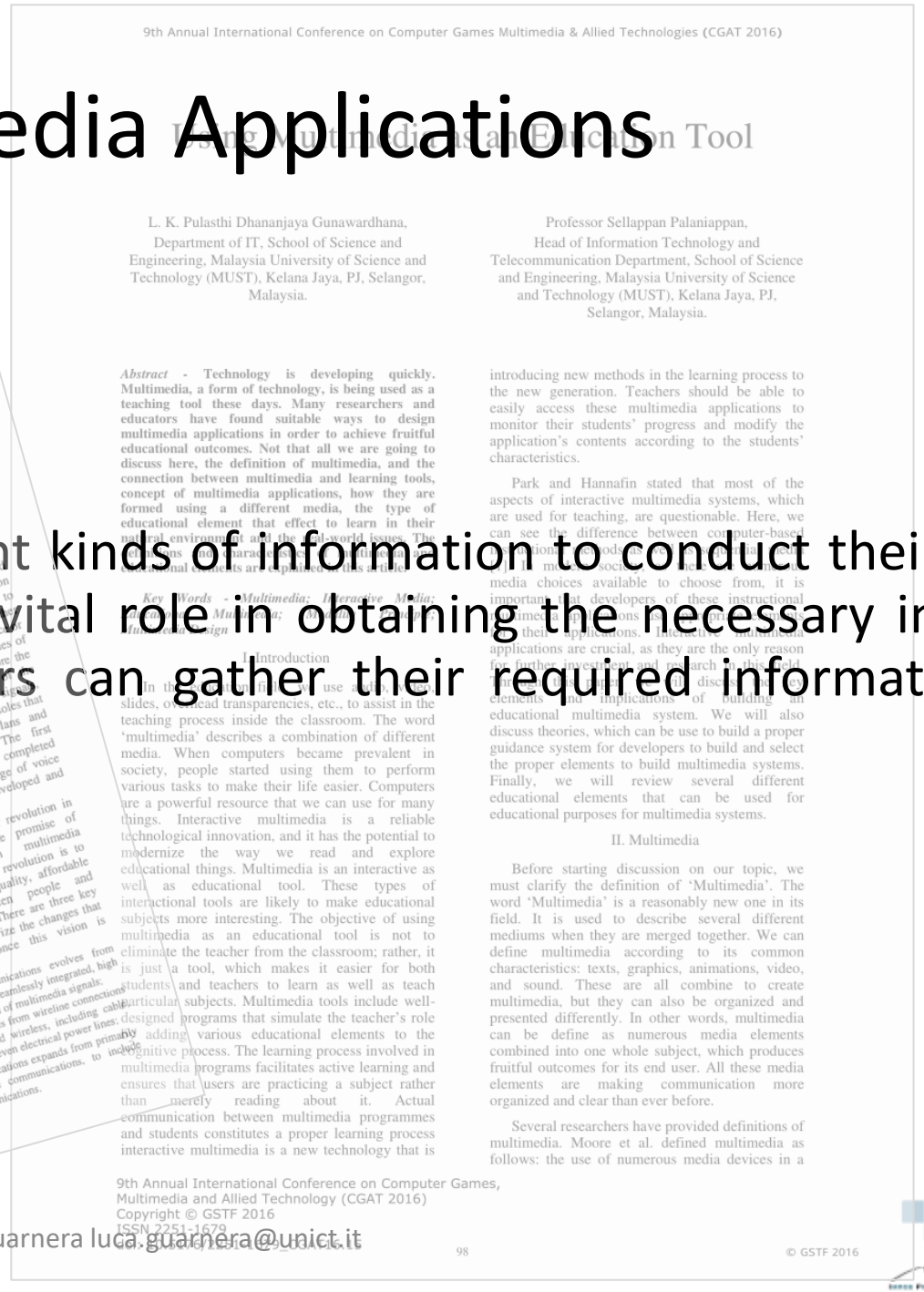
Multimedia Applications


On the Application of Multimedia Processing to Telecommunications

Richard Cox, Barry Haskell, Yann LeCun, Behzad Shahraray, and Lawrence Rabiner
Speech and Image Processing Services Research Lab
AT&T Labs-Research, Florham Park/Newman Springs, New Jersey, USA

Abstract

1. Introduction


9th Annual International Conference on Computer Games Multimedia & Allied Technologies (CGAT 2016)

Interactive Multimedia as an Education Tool

L. K. Pulasthi Dhananjaya Gunawardhana,
Department of IT, School of Science and Engineering, Malaysia University of Science and Technology (MUST), Kelana Jaya, PJ, Selangor, Malaysia.

Professor Sellappan Palaniappan,
Head of Information Technology and Telecommunication Department, School of Science and Engineering, Malaysia University of Science and Technology (MUST), Kelana Jaya, PJ, Selangor, Malaysia.

Abstract - **Technology is developing quickly. Multimedia, a form of technology, is being used as a teaching tool these days. Many researchers and educators have found suitable ways to design multimedia applications in order to achieve fruitful educational outcomes. Not that all we are going to discuss here, the definition of multimedia, and the connection between multimedia and learning tools, concept of multimedia applications, how they are formed using a different media, the type of educational element that effect to learn in their**

Key Words - Multimedia; Interactive Media;

I. Introduction

II. Multimedia

9th Annual International Conference on Computer Games, Multimedia and Allied Technology (CGAT 2016)
Copyright © GSTF 2016
ISSN 2251-1679

98

© GSTF 2016

Research: Researchers require different kinds of information to conduct their research and development works. Internet plays a vital role in obtaining the necessary information. By searching the Internet the researchers can gather their required information and make successful research.

Multimedia Applications

Virtual Reality: Virtual reality is a technology which allows a user to interact with computer-simulated environment. The simulated environment can be similar to the real world. For example: simulation for pilot or combat or surgery training. This simulated environment can also differ significantly from reality, as in virtual reality games. This virtual reality is actually based on multimedia bounder technology.

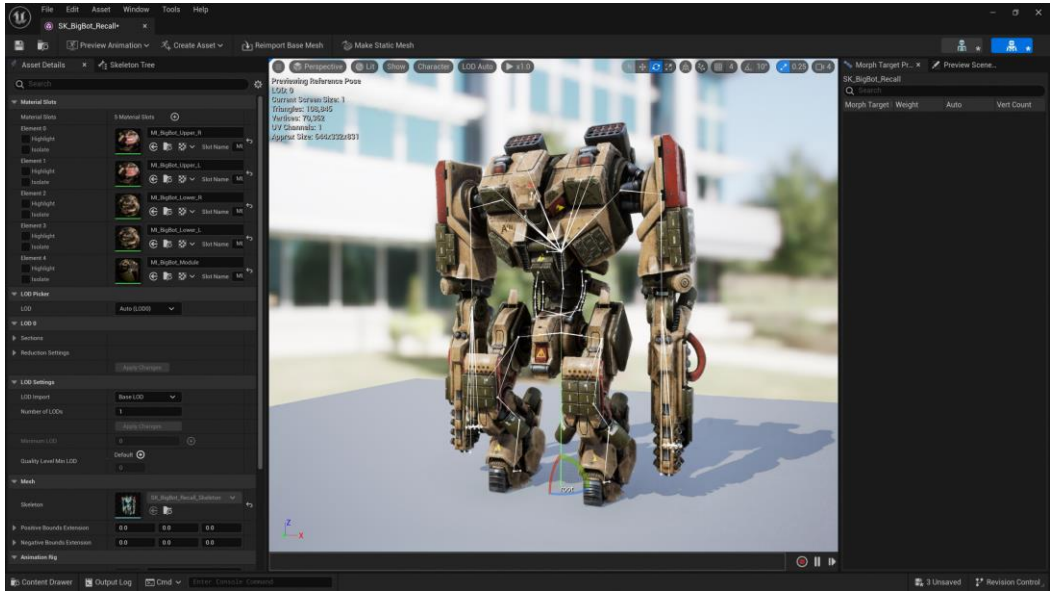
Metaverse

Brief introduction to the Metaverse

Metaverse

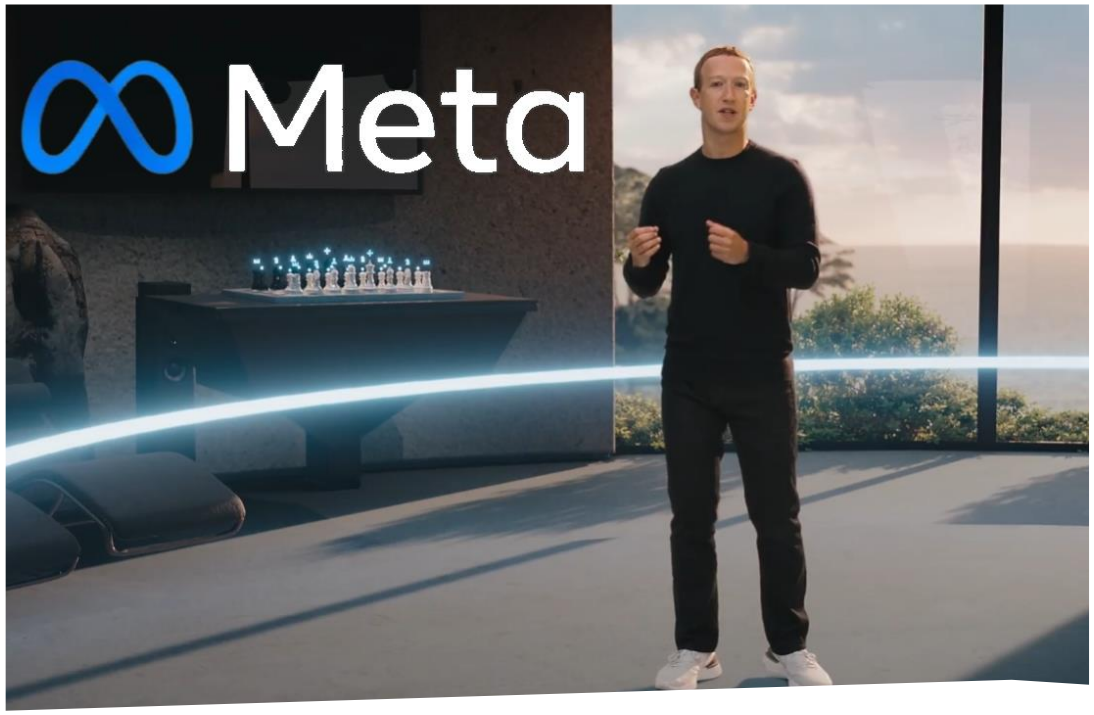
Metaverse: A first definition

Metaverse, a term born in the cyberpunk world in 1992 and now in the limelight because it is used by Facebook, is the evolution of the Internet, but does not replace it. **This is a difficult concept to define exactly**, which prefigures a set of **interconnected virtual** and real worlds, populated by **avatars**. There are many questions still open



UNREAL ENGINE

Unity



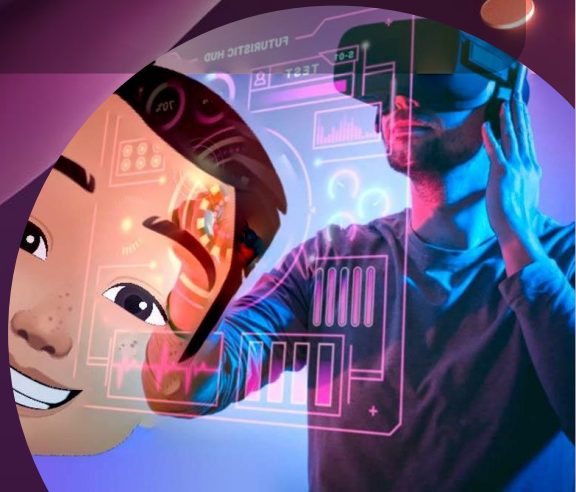
Meta = Metaverso & Facebook

The Metaverse is not new, but it rose to prominence with the announcement of Facebook, which decided to name the Group's holding company (which controls the Facebook, Whatsapp, Instagram and Oculus platforms) "Meta" and to start a project with this name, a project about which little is known yet.

Mash = Metaverso & Microsoft

A few days later, Microsoft announced that from 2022 it will integrate the Metaverse into the Teams platform with a feature called Mash: users will be able to create an avatar with which to participate in work meetings.

Metaverse is the future of the Internet, a set of virtual spaces crossed by avatars, a step forward compared to virtual reality.



Luca Guarnera luca.guarnera@unict.it

… The metaverse is a virtual space, but its impact will be real …

- **https://about.meta.com/it/metaverse/impact/**

Meta

The benefits of the metaverse to society are immense

The metaverse will connect people to a new range of experiences: from immersive sessions in education and training, to incredible possibilities for healthcare and workplaces.

"With the metaverse, one day firefighters will have the tools to rescue more people in less time."

"One day doctors will be able to view diagnostic images like never before and make decisions more quickly."

"Surgeons will be able to do all the necessary practice."

"Students will take a field trip to the Ice Age to rediscover the past."

"Urban planners will be able to run simulations to help reduce traffic."

Ways to enter the metaverse

Eyeglasses Smart

Augmented reality

Virtual reality

2

3

Metaverse, how to access?

Università di Catania

Virtual reality

Virtual reality allows you to discover new worlds and new shared experiences. Live new experiences with your friends and family, even if you are far apart.

Immersed app requires compatible PC. Images for illustrative purposes and are not formatted to fit your screen.

Luca Ardito, luca.ardito@polito.it, luca.ardito@unict.it

Example of Virtual Reality

https://www.youtube.com/watch?v=YaRastZmucQ

L'avventura ti attende con Meta Quest 2

Ora con due fantastici giochi inclusi (valore fino a € 43).
Divertimento infinito. L'esperienza completa. Termina il 3/6.*

Acquista Meta Quest 2 › Scopri di più

Augmented reality

Through photos and videos, AR allows you to improve shared experiences with fun virtual effects that are activated simply by swiping your finger on the screen, leaving you the freedom to express yourself with the most important people in your life

Augmented reality for plant maintenance

https://www.youtube.com/watch?v=tIWu5g2_6XA

Multimedia: Film & Cinema

Film & Cinema

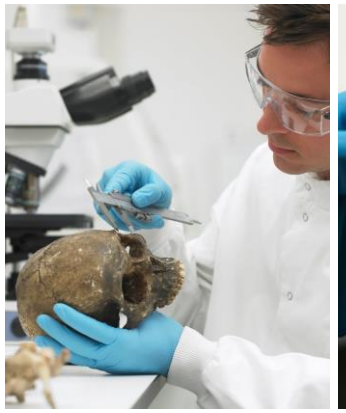
Film means single media or multimedia entertainment content for distribution or exhibition to the general public by any means and media in any digital media format, film, or videotape, including, but not limited to, a motion picture, a documentary, a television series, a television miniseries, a television special, interstitial television programming, long-form television, interactive television, music videos, interactive games, video games, commercials, internet programming, an internet video, a sound recording, a video, digital animation, or an interactive website.

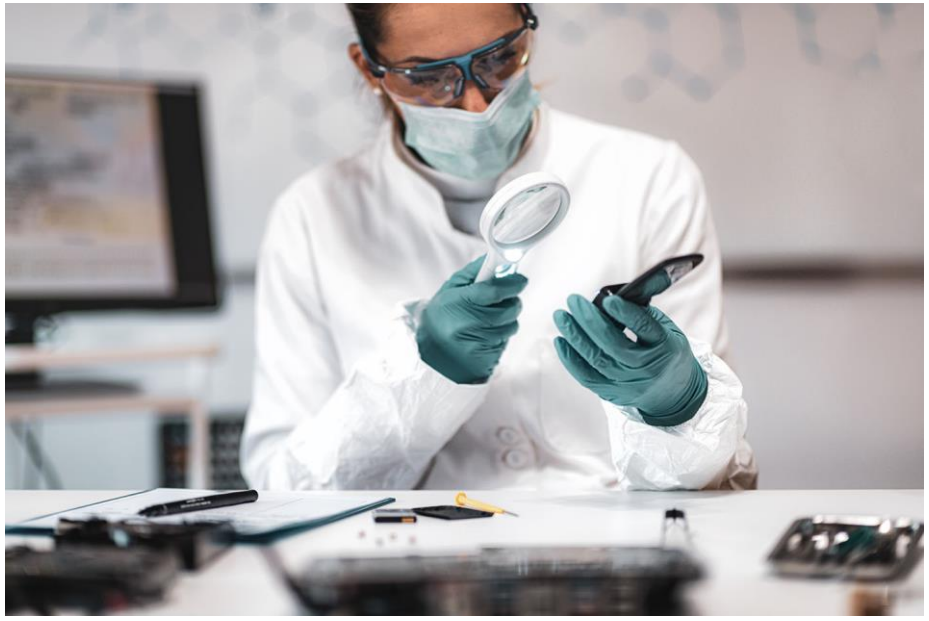
Multimedia:
Application in Digital Forensics

Forensic Science

(sometimes shortened to Forensics) is the application of technical and scientific methods to the justice, investigation and evidence discovery domain.




Digital Forensics History

1. The first notion of digital forensics came in the **1970s**, when the ***Federal Rules (US)*** regulations began talking about **DIGITAL EVIDENCE**

2. The **DIGITAL FORENSICS INVESTIGATIONS** started in the mid-to-late 1980s, when federal agents began to study how to find digital evidence in computers

3. The **first research groups** in the area of Forensic Science started around **2000 and 2001**

4. The first definition of the Digital Forensic Science was given at the first Digital Forensics Research Workshop (DFRWS), in 2001

The use of scientifically derived and proven methods toward the **preservation**, **collection**, **validation**, **identification**, **analysis**, **interpretation**, **documentation** and **presentation** of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations

But there is much more to the origins of forensic science

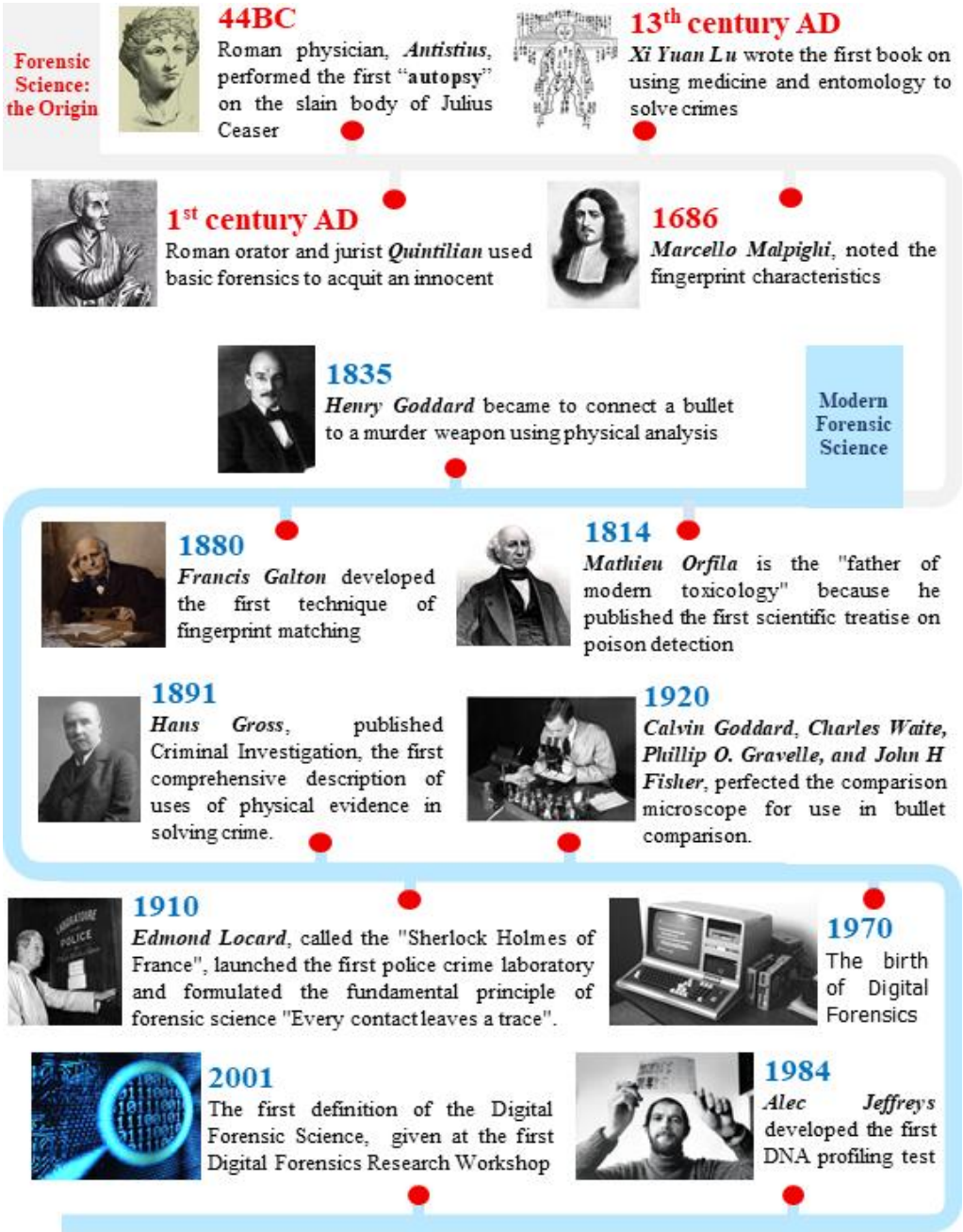
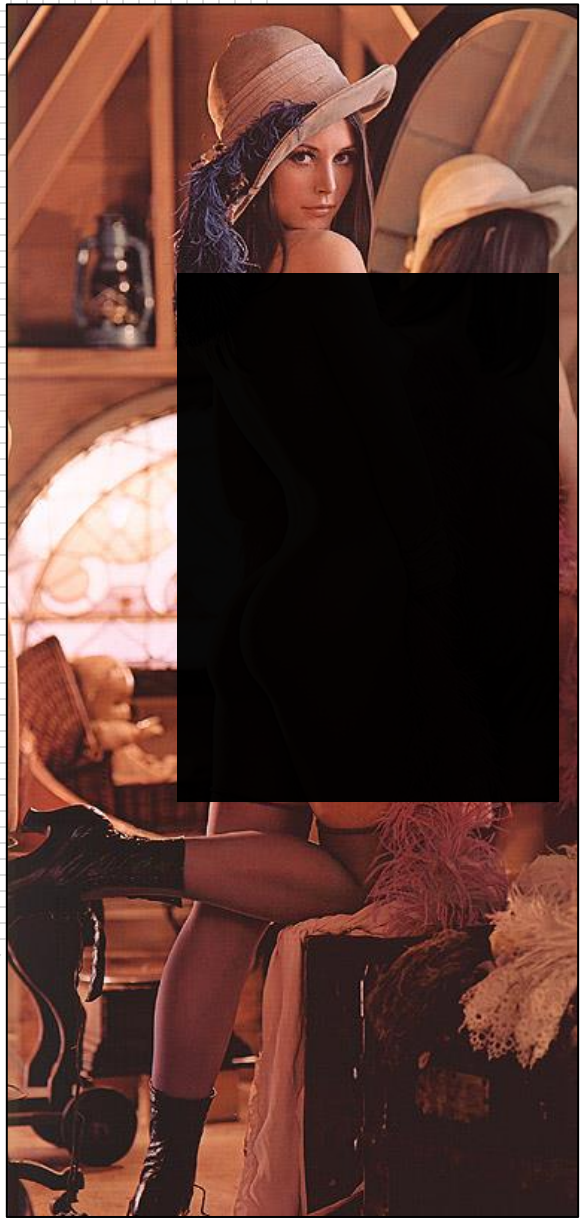


Image Forensics

What can we say about this image?

What can we say about this image?



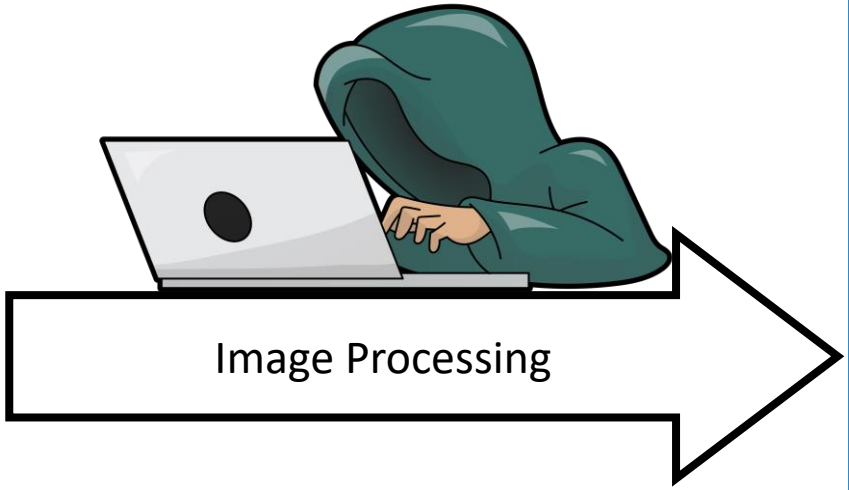


Image Processing

What can we say about this image?

What can we say about this image?

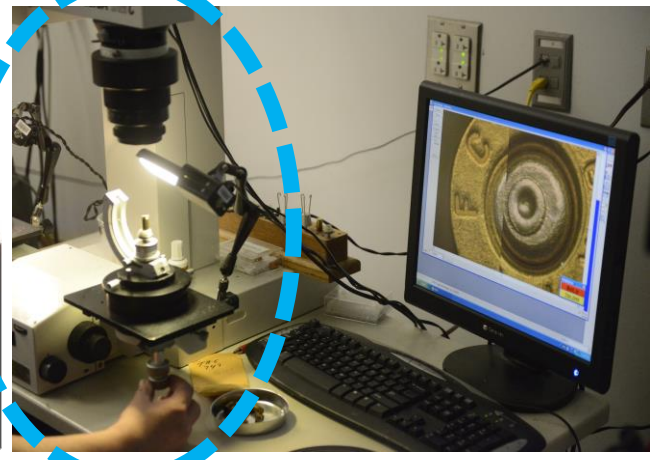
Authenticity VS Integrity

First and foremost: **Authenticity is not Integrity**

AUTHENTICITY	≠	INTEGRITY
Image is an accurate representation of the original event		**Information is unaltered from the time of acquisition until its final disposition**

Forensic firearm ballistics

Forensic firearm ballistics is the science of examining the characteristics of **firearms**, **cartridges**, and/or **bullets** found at a crime scene.

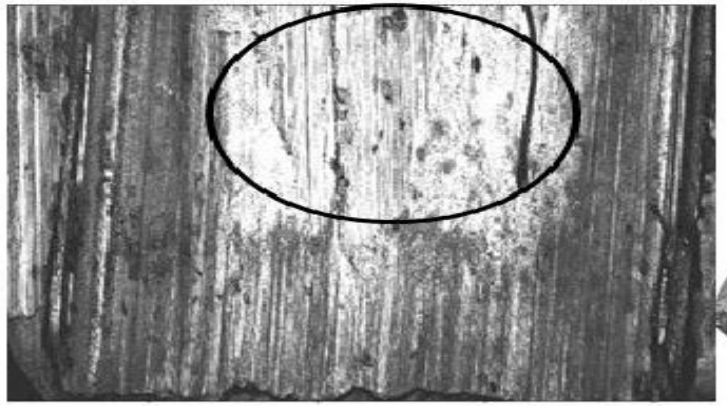



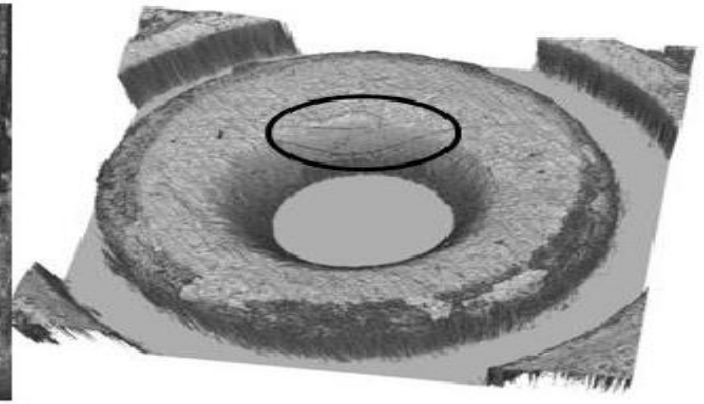
- Ballistic Forensics Specialists are tasked with tying this evidence to crime suspects.
- Comparison Microscope

Forensics Examination of Firearms

Weapon serial numbers

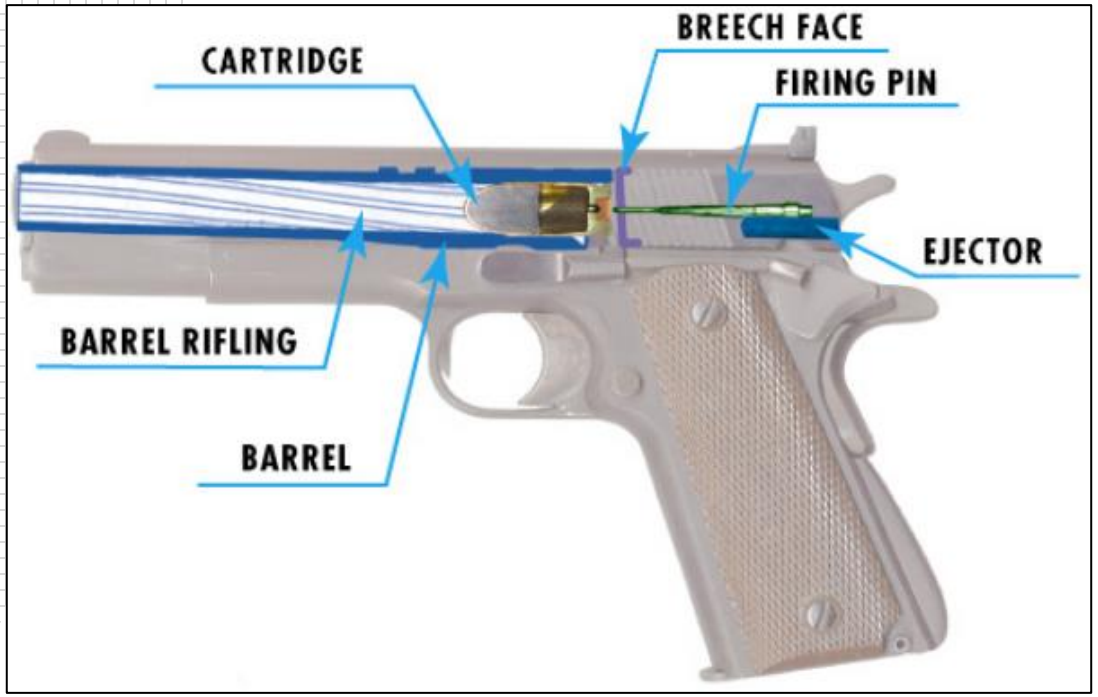
Fingerprints on the weapon's surface

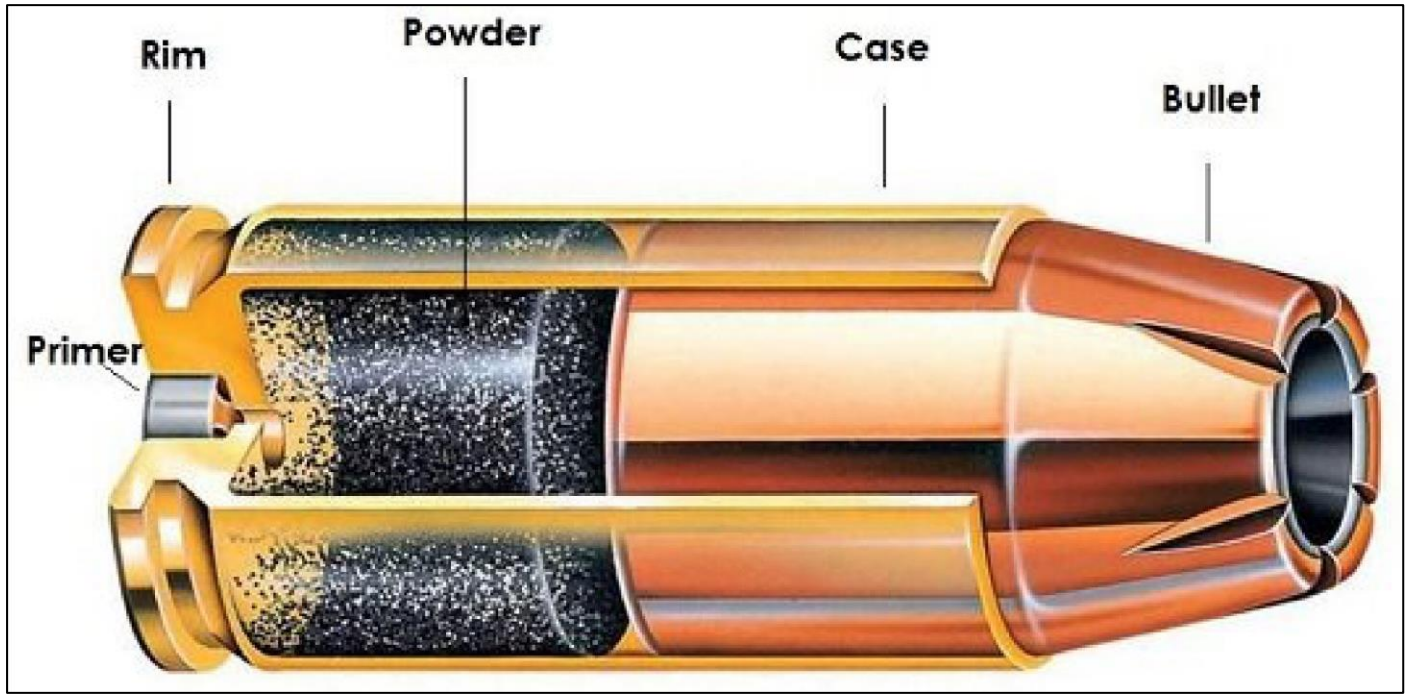




Ballistic striations

Firearm and Cartridge components





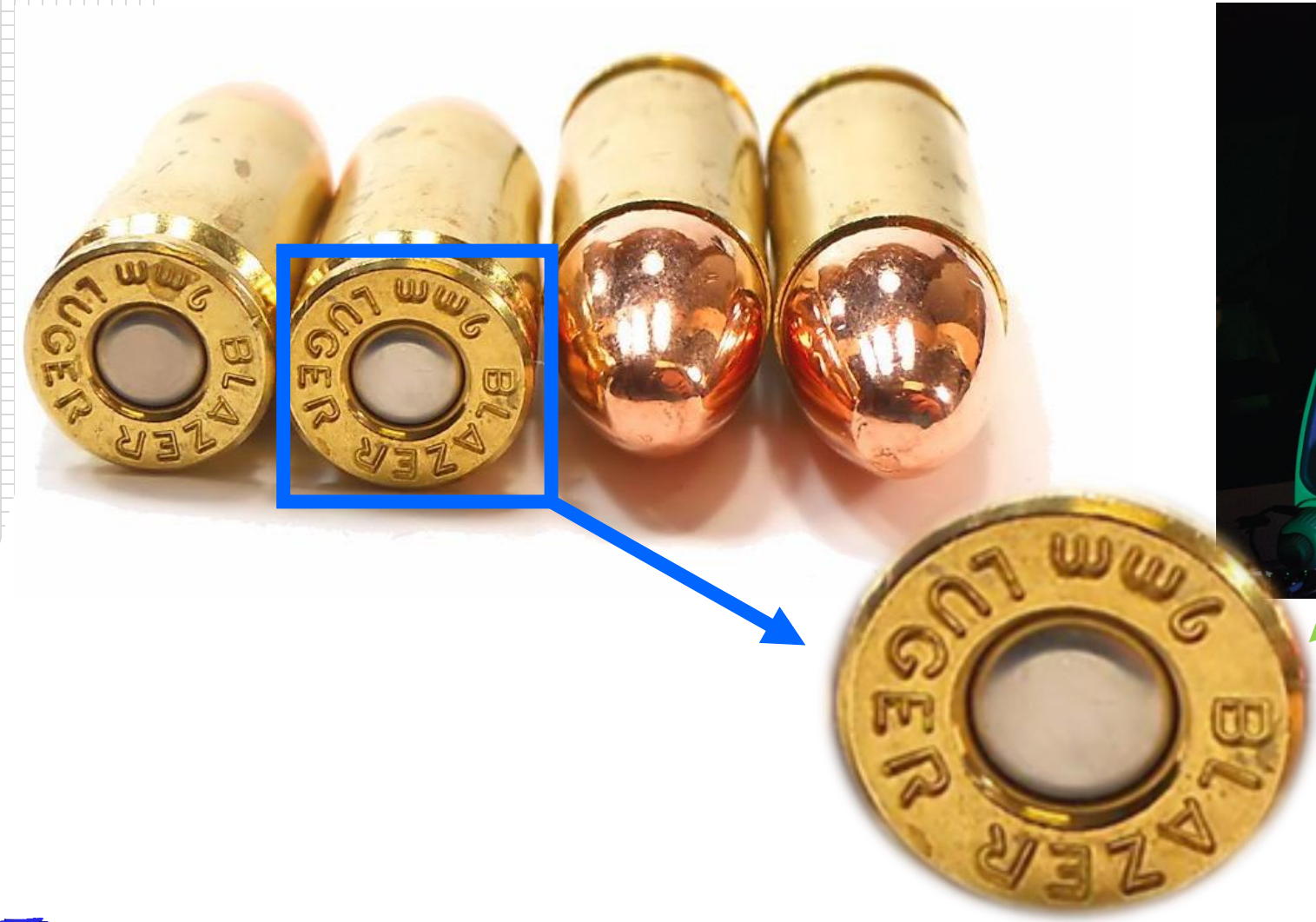
Weapon firing process

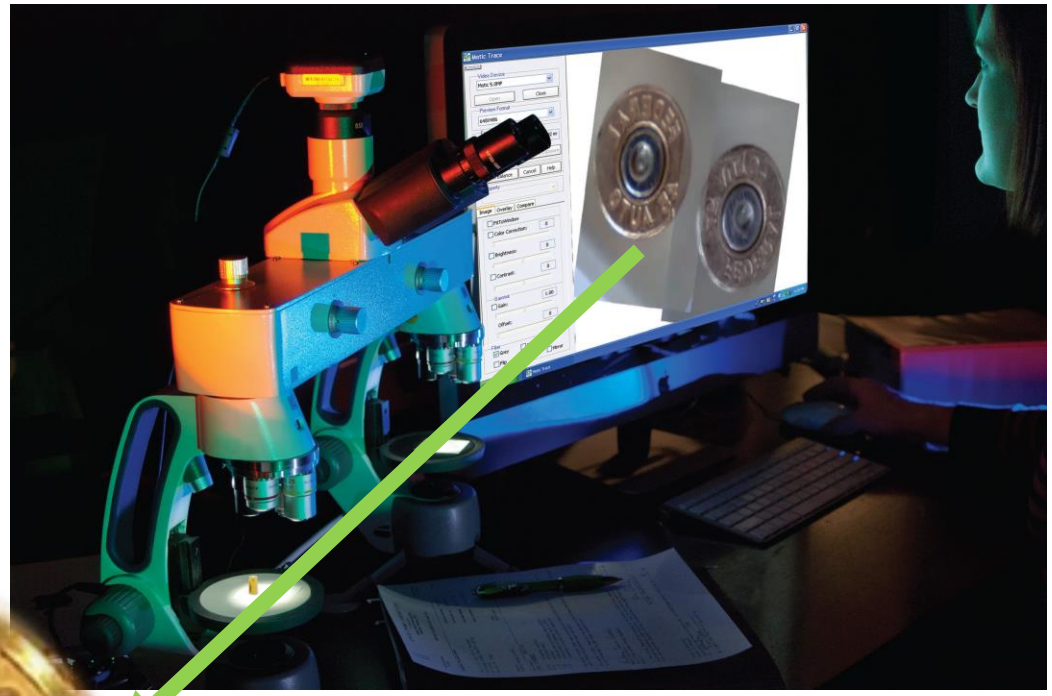
The weapon **loading**

The **shot**

The cartridge **ejection**

Cartridge Analysis





DEMO - 3D Forensic Ballistics Comparison

https://iplab.dmi.unict.it/mfs/Forensic-Firearms-Ballistics-VR/

Assessing Forensic Ballistics Three-Dimensionally through Graphical Reconstruction and Immersive Observation

Luca Guarnera[1], Oliver Giudice[2], Salvatore Livatino[3], Antonino Barbaro Paratore[4], Angelo Salici[5], Sebastiano Battiato[1]

[1] Dipartimento di Matematica e Informatica, University of Catania, Viale Andrea Doria, 6, Catania, 95126, Italy - *luca.guarnera@unict.it, battiato@dmi.unict.it*

[2] Applied Research Team, IT Department, Banca dâ€™Italia, Rome, Italy - *giudice@dmi.unict.it*

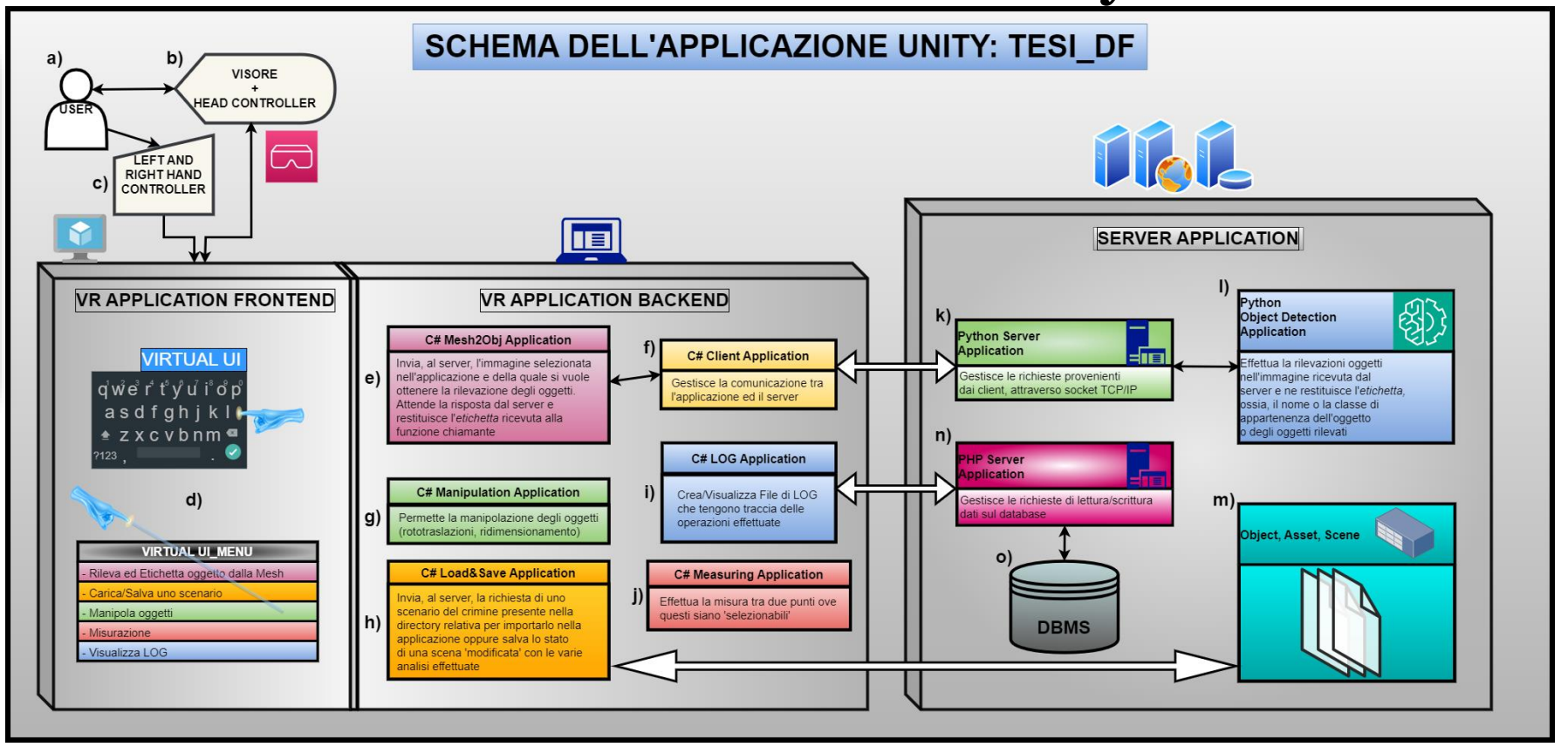
[3] School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK - *s.livatino@herts.ac.uk*

[4] iCTLab Spinoff of University of Catania, Viale Andrea Doria, 6, Catania, 95126, Italy - *antonino.paratore@ictlab.srl*

[5] Carabinieri Scientific Investigations Department of Messina, Via Monsignor Dâ€™Arrigo, 5, Messina, 98122, Italy - *angelo.salici@carabinieri.it*

TESI TRIENNALE
INGEGNERIA INFORMATICA

VR Application with Deep Learning Techniques for Forensic Crime Scene Analysis

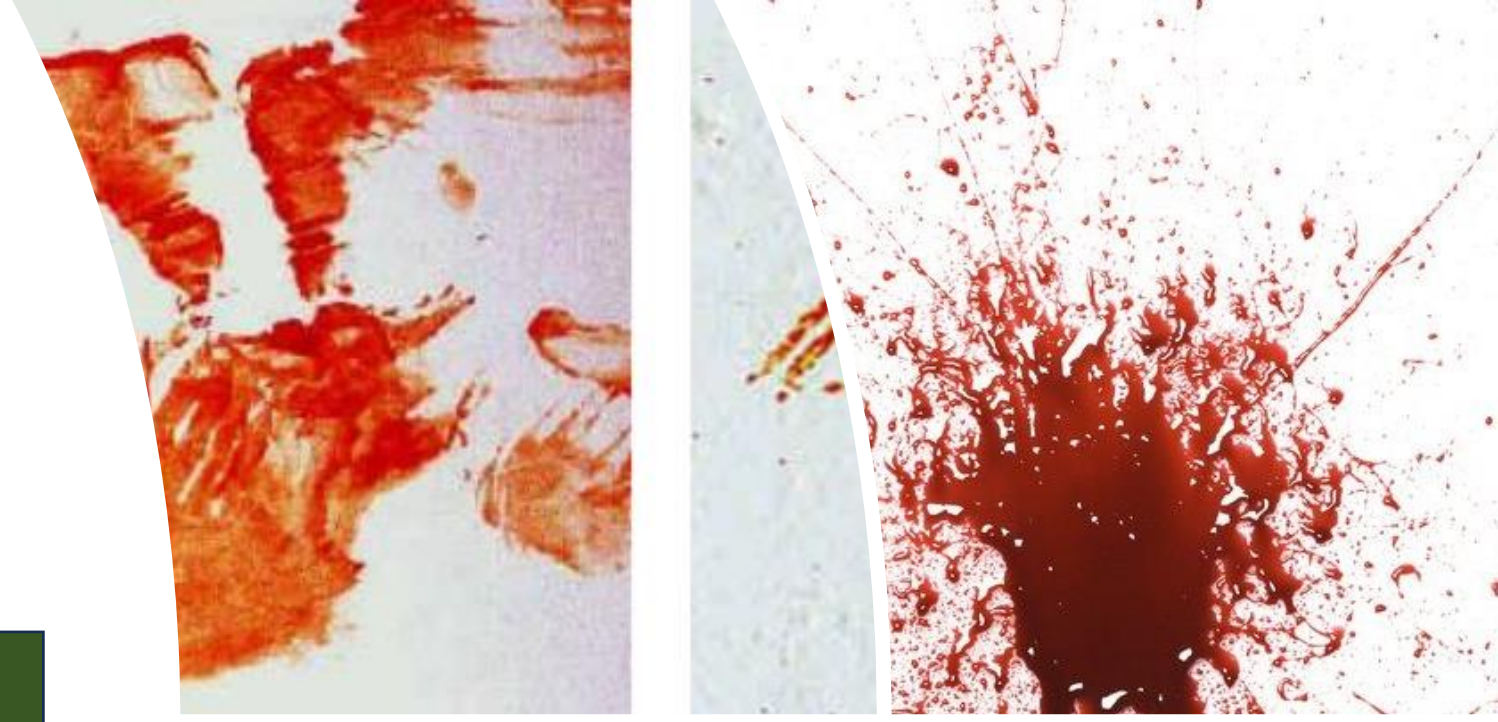


Manipolazione degli oggetti

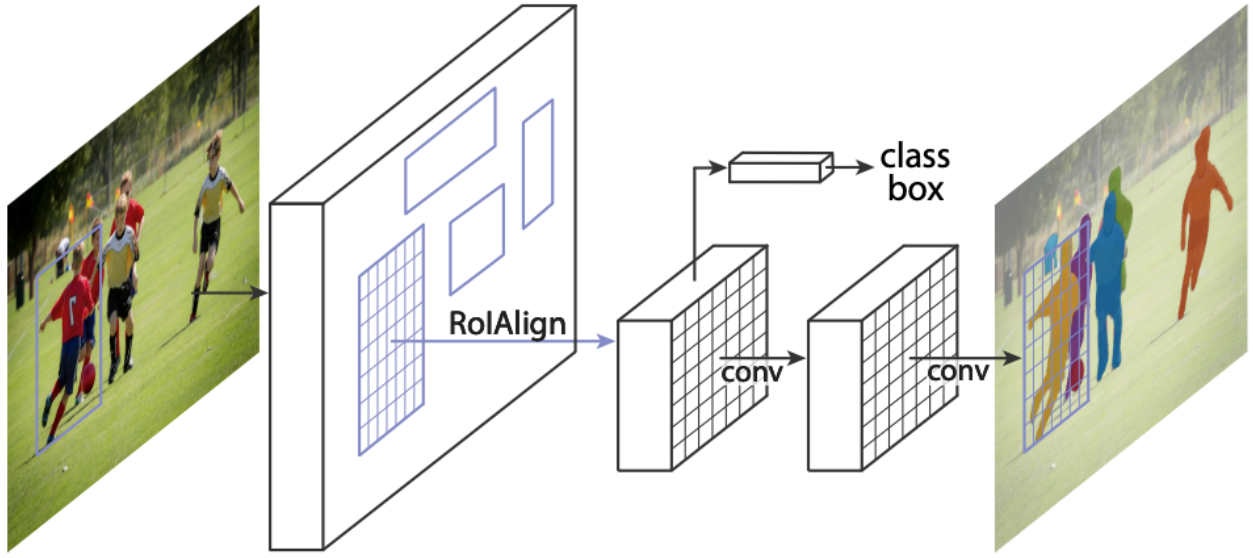
Rilevamento ed etichettatura automatica e/o manuale degli oggetti

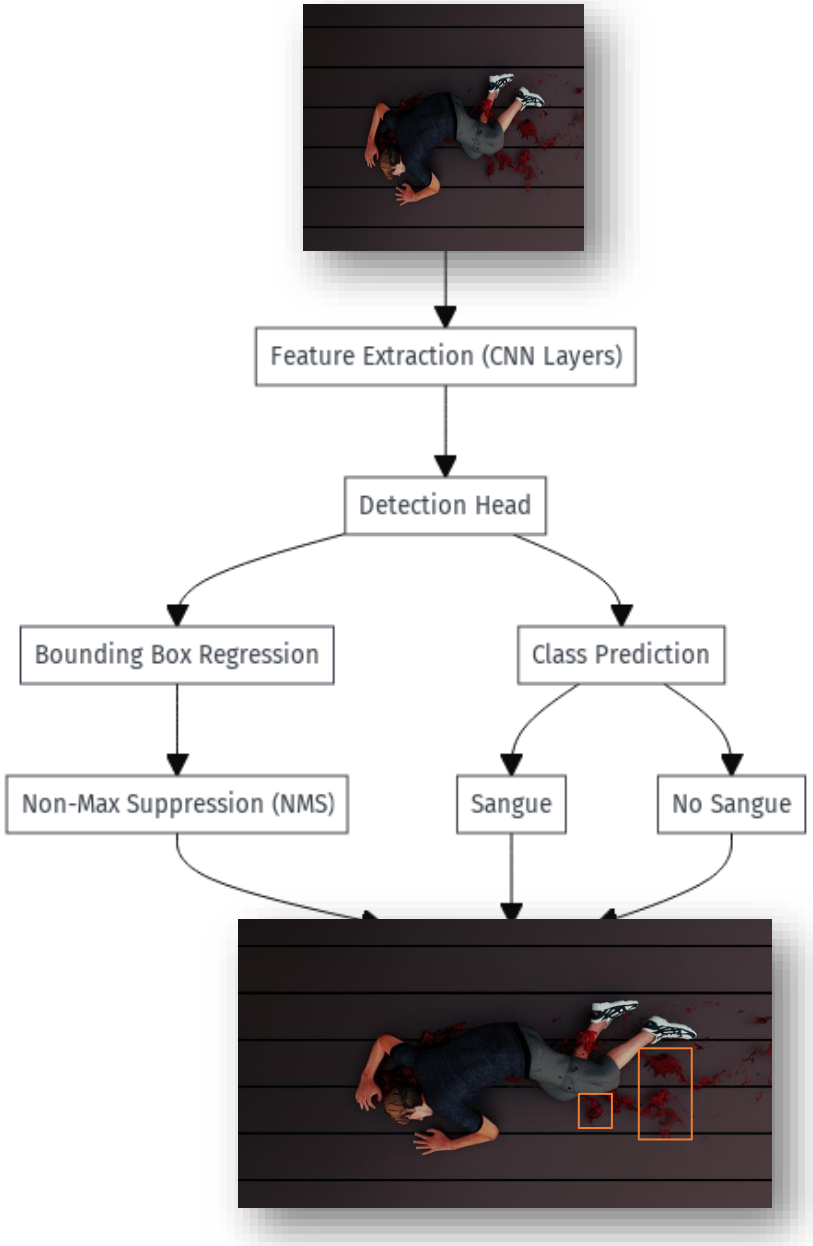
Università di Catania

Blood Pattern Analysis



Yolov9 Architecture

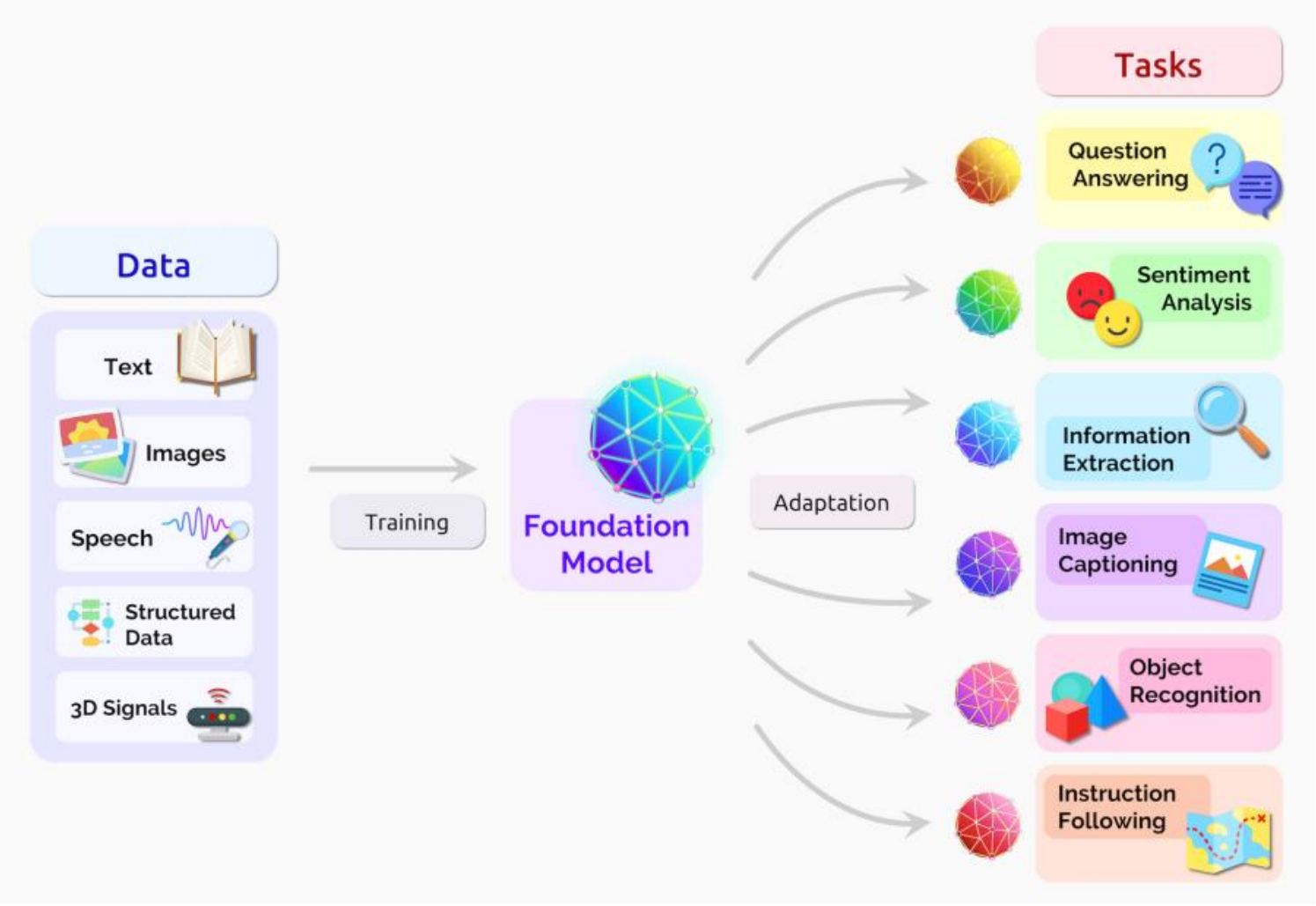




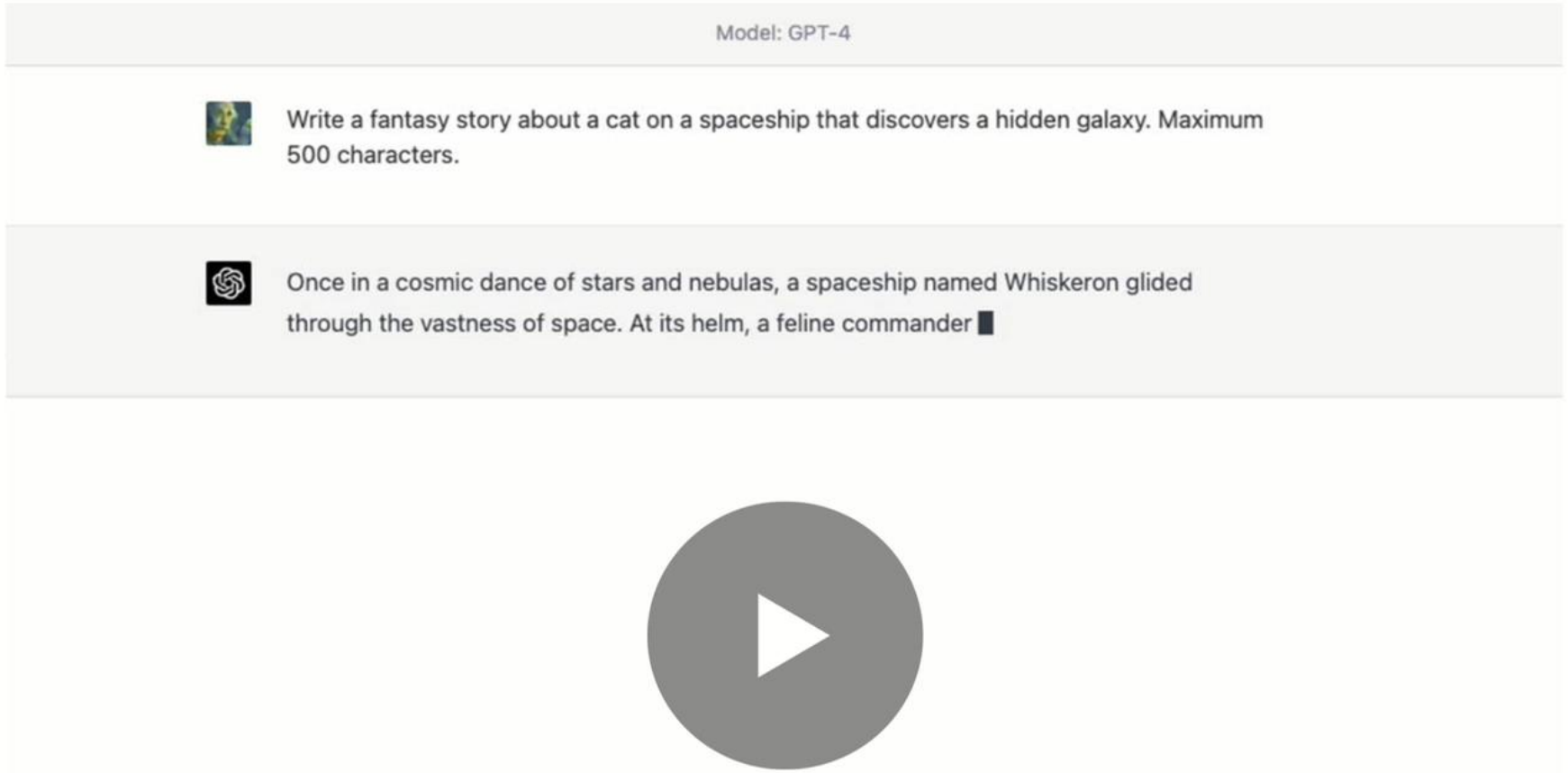

DETECT BLOOD
Detect

Multimedia based on AI

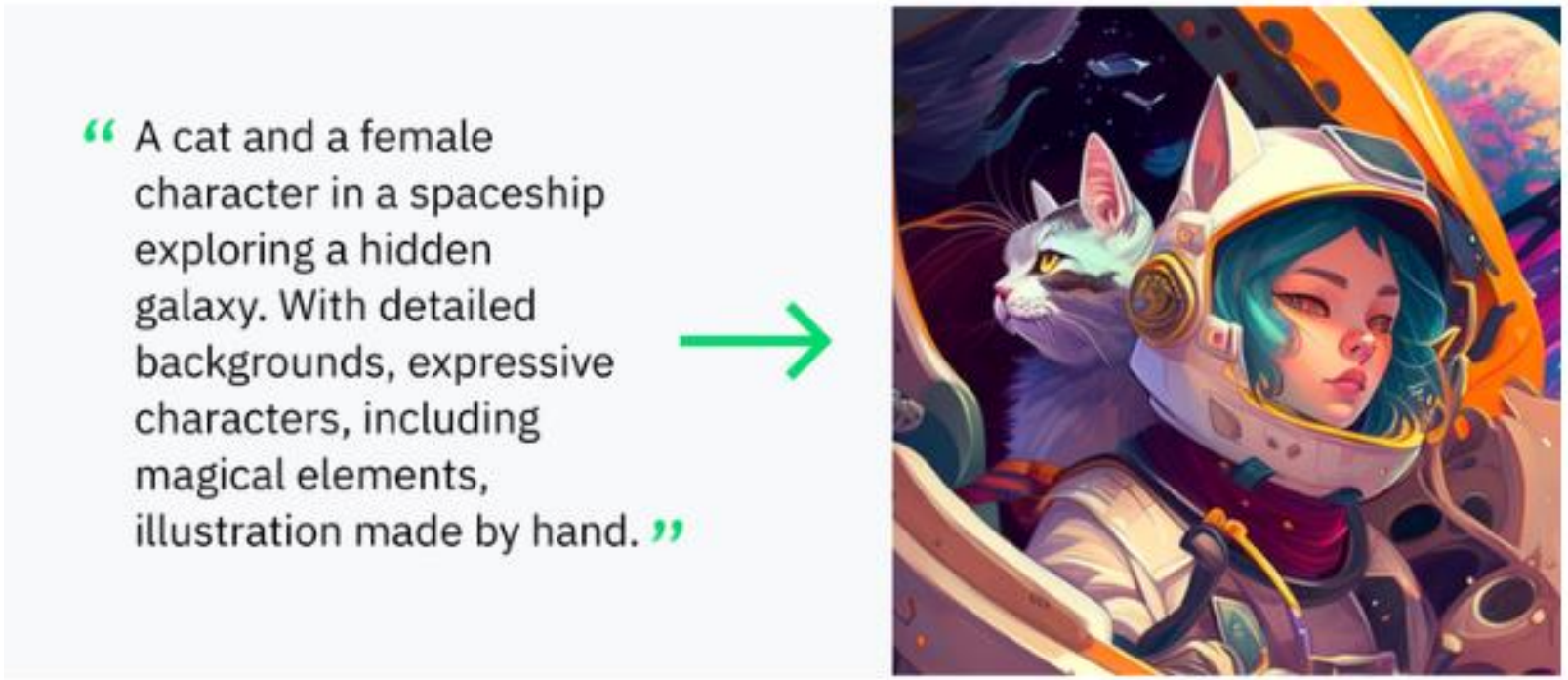
Foundation Models



On the Opportunities and Risks of Foundation Models – Bommasani et al. 2022


Model: GPT-4
Write a fantasy story about a cat on a spaceship that discovers a hidden galaxy. Maximum 500 characters.
Once in a cosmic dance of stars and nebulas, a spaceship named Whiskeron glided through the vastness of space. At its helm, a feline commander
00:00 / 00:46
Stop generating

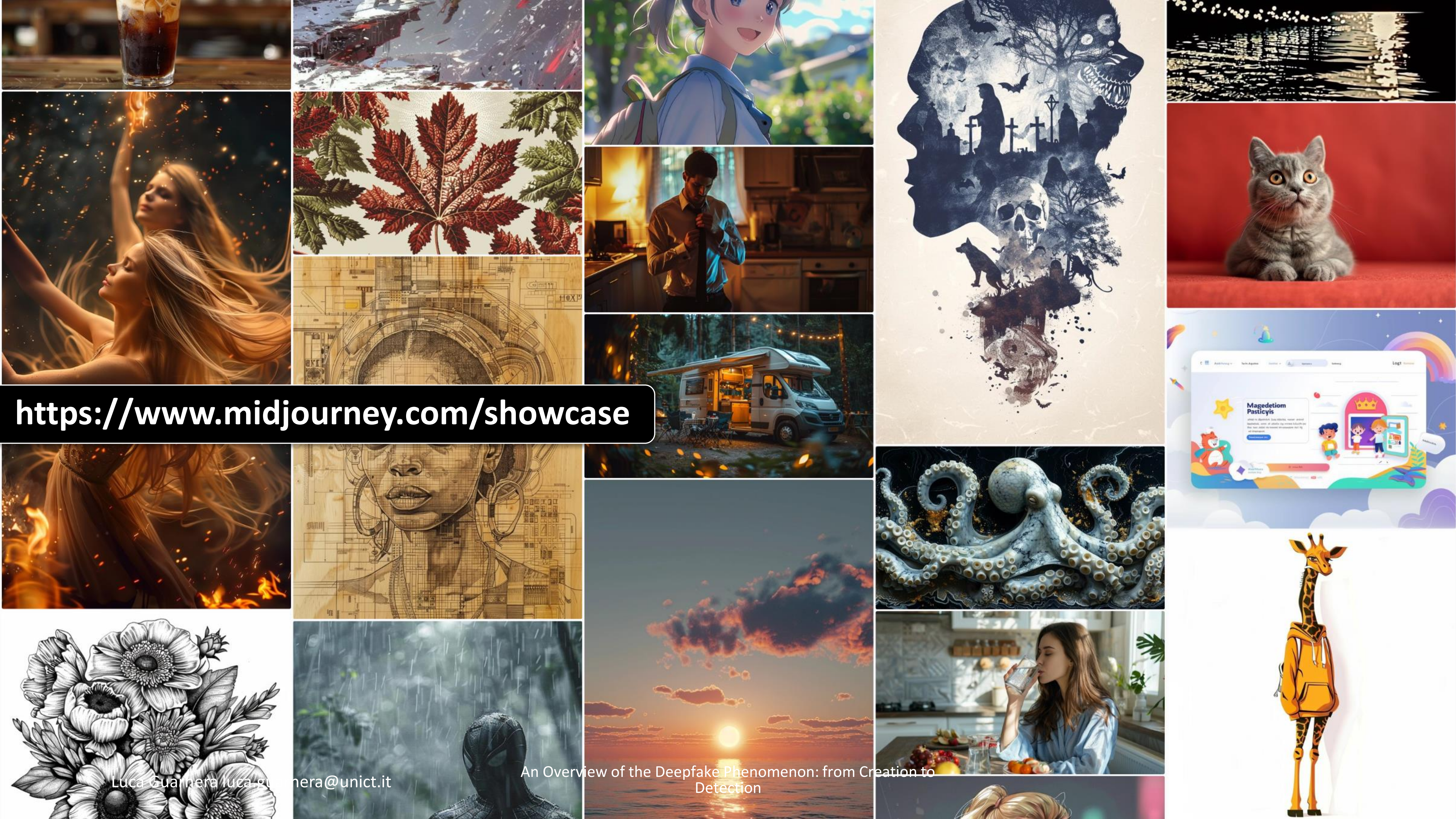
GenAI: Image



Application: Very successful platforms such as MidJourney and DALL-E have become a popular choice for anyone seeking to generate realistic images through Image Generation techniques.

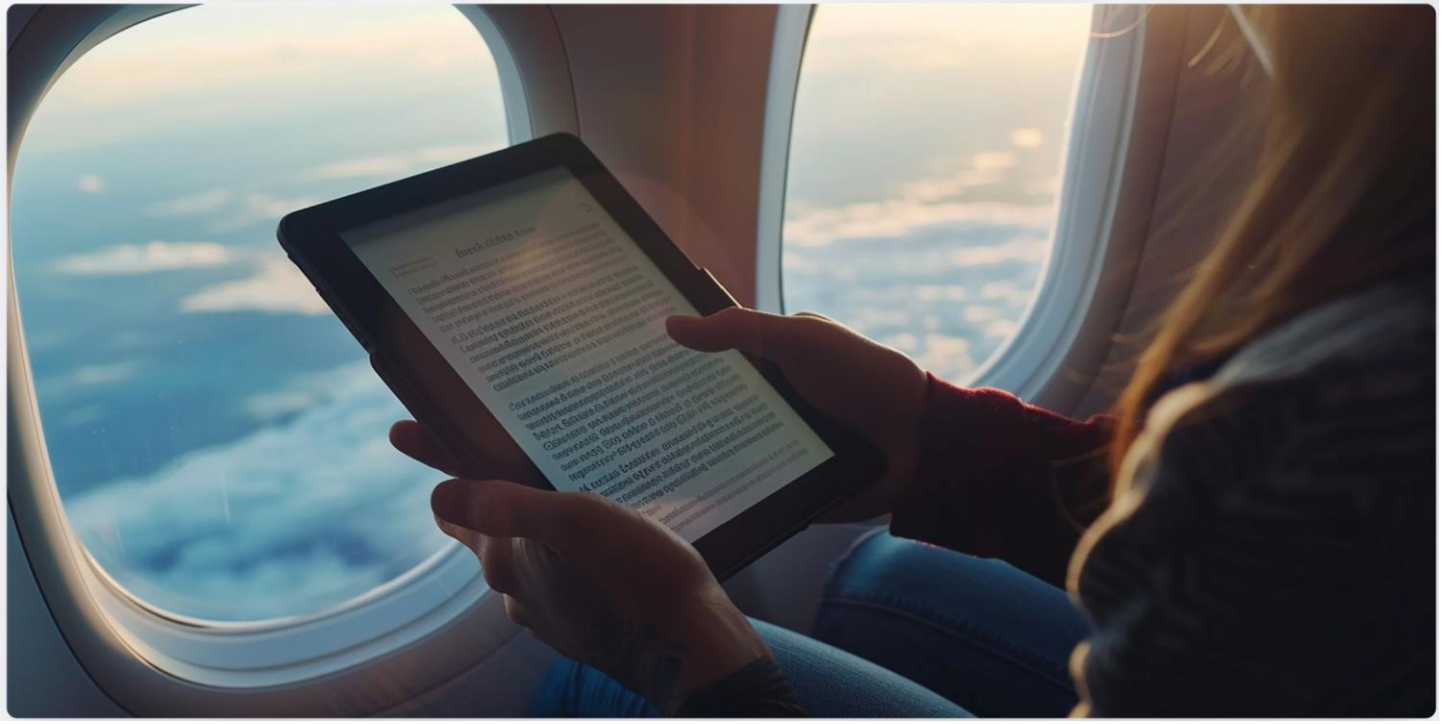

Midjourney v1
March, 2022
Midjourney v2
April, 2022
Midjourney v3
July, 2022
Midjourney v4
November, 2022
Midjourney v5
March, 2023
rubenhassid.

GenAI:Image



https://www.midjourney.com/showcase

Cow, colorful graffiti with inscription on his body, hyper-realistic, movie poster, full body, 4k



A person reading an e-book on a tablet during a flight

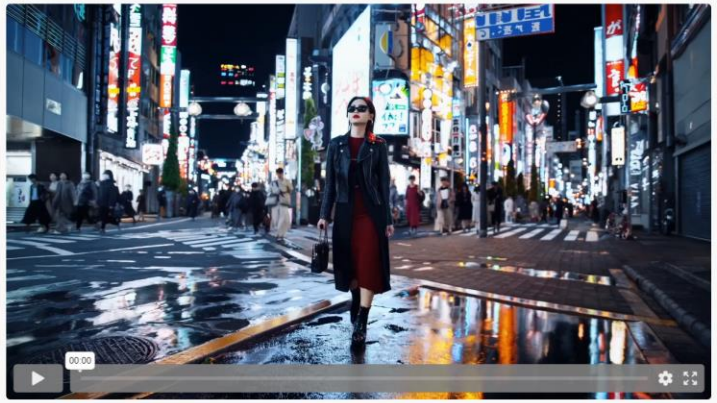
A young, energetic woman in her late 20s to early 30s, wearing a bright blue short-sleeved T-shirt and black leggings, with a fit and muscular build, riding a bicycle generator in a park. Her hair is neatly tied in a ponytail, emphasizing an active and healthy image. She has a bright and lively expression, pedaling vigorously. The rear wheel of the bicycle is connected to a generator, with wires leading to a blender filled with fruits and vegetables, showing it in operation. The background is a green lawn or park, creating a nature-friendly atmosphere. The setup clearly shows the connection between the generator and the blender, conveying an eco-friendly message, detailed and realistic depiction, realistic photography style, Canon EOS 5D Mark IV, 50mm lens, f/1.8, natural daylight

a creative digital collage to represent Africa climate

Sora: an AI model capable of creating realistic and fantasy scenes from textual instructions.

https://openai.com/index/sora/

Prompt: A young man at his 20s is sitting on a piece of cloud in the sky, reading a book.



Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

Prompt: Several giant wooly mammoths approach treading through a snowy meadow, their long wooly fur lightly blows in the wind as they walk, snow covered trees and dramatic snow capped...

Prompt: A Chinese Lunar New Year celebration video with Chinese Dragon.

Prompt: The camera follows behind a white vintage SUV with a black roof rack as it speeds up a steep dirt road surrounded by pine trees on a steep mountain slope, dust kicks up from it's tires...

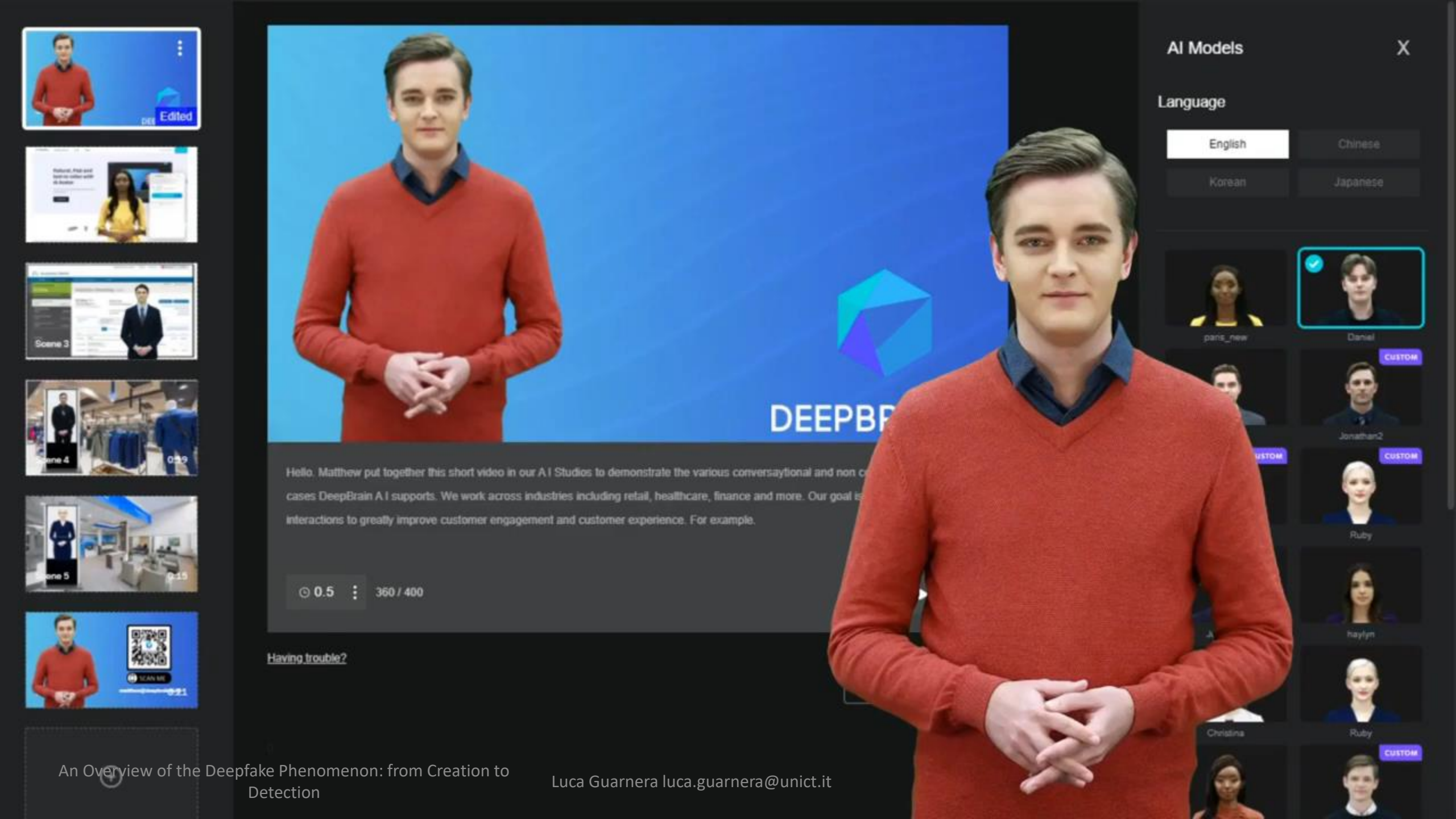
Prompt: Reflections in the window of a train traveling through the Tokyo suburbs.

GenAI: Speech/Video

Video Generation involves deep learning methods such as **GANs** and **Video Diffusion** to generate new videos by predicting frames based on previous frames. Video Generation can be used in various fields, such as entertainment, sports analysis, and autonomous driving. Video Generation can be often seen in use with Speech Generation. The models used for speech generation can be powered by **Tranformers**. Speech Generation can be used in text-to-speech conversion, virtual assistants, and voice cloning.

🪄 Application: Platforms such as DeepBrain and Synthesia utilize Video and Speech Generation to create realistic video content, that appears as if a human was speaking on camera.

Source: https://generativeai.net/



Luca Guarnera luca.guarnera@unict.it

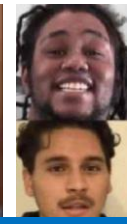
Luca Guarnera luca.guarnera@unict.it

What are Deepfakes?





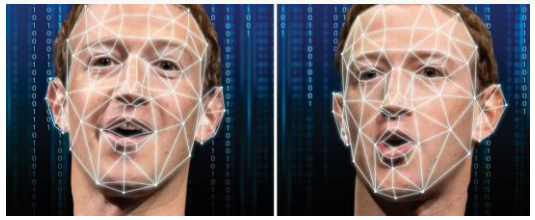




Deepfakes refers to all those multimedia contents synthetically altered or created by exploiting machine learning generative models

"[…] In autumn 2017, an **anonymous Reddit user** under the pseudonym "**deepfakes**" posted several **porn videos on the Internet** […]".



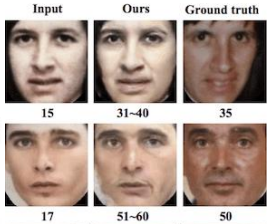



Luca Guarnera luca.guarnera@unict.it

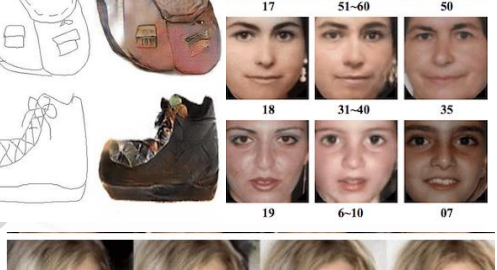
Università di Catania









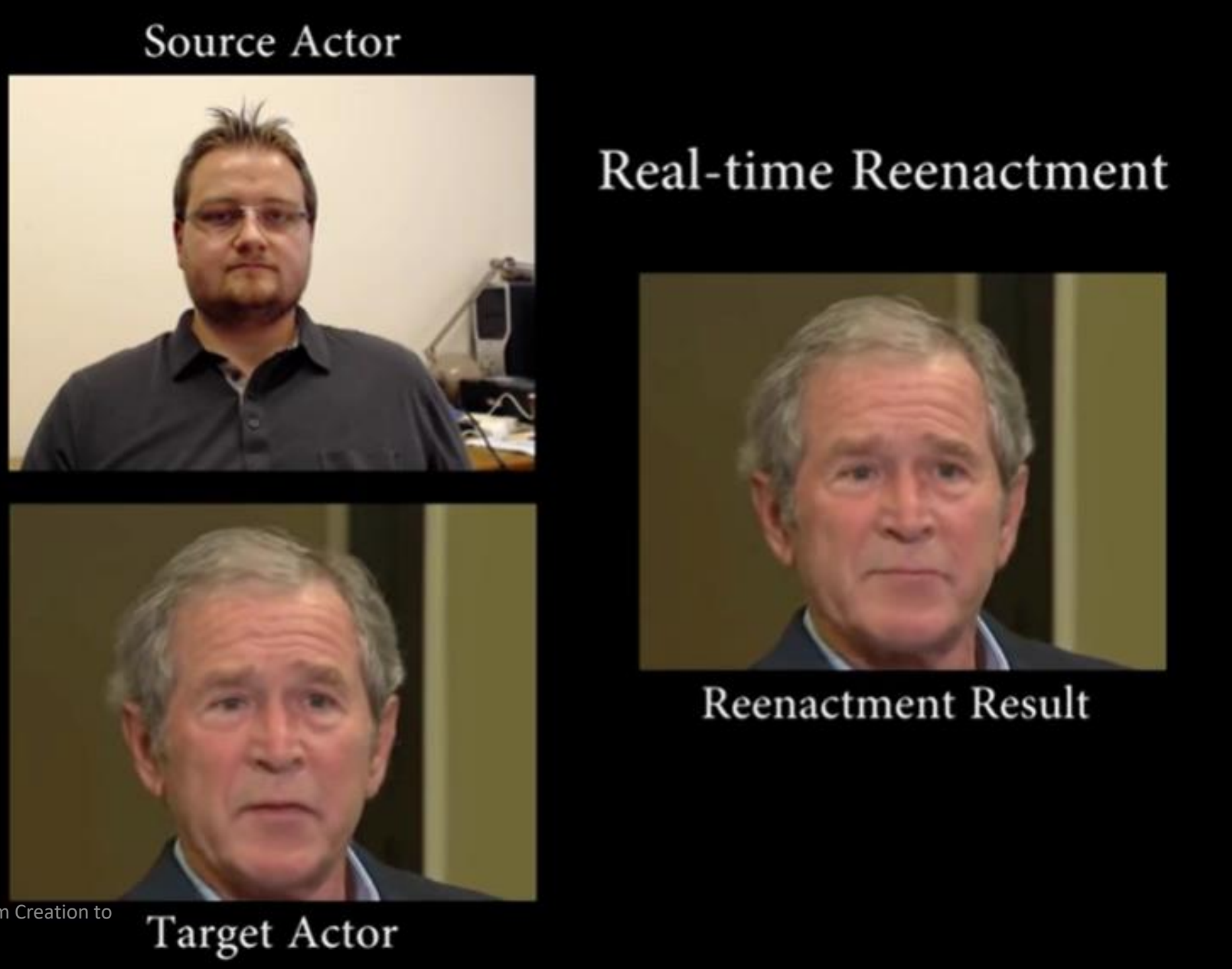




Luca Guarnera luca.guarnera@unict.it

Luca Guarnera luca.guarnera@unict.it

Face Reenactment

Thies, J., Zollhofer, M., Stamminger, M., Theobalt, C., & Nießner, M. (2016). **Face2face: Real-time face capture and reenactment of rgb videos.** In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2387-2395).

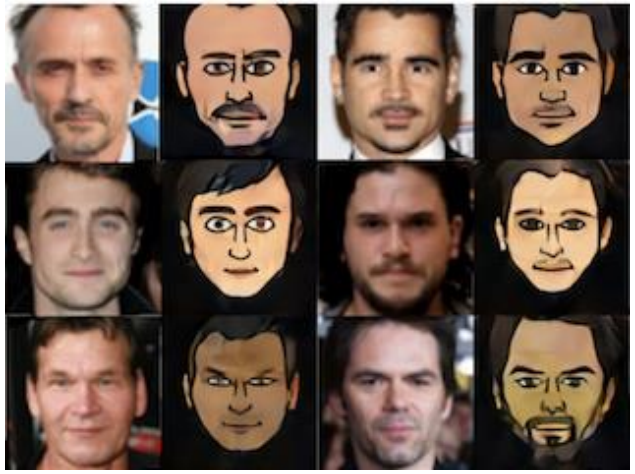


Luca Guarnera
luca.guarnera@unict.it

Impressive Applications of Generative Models

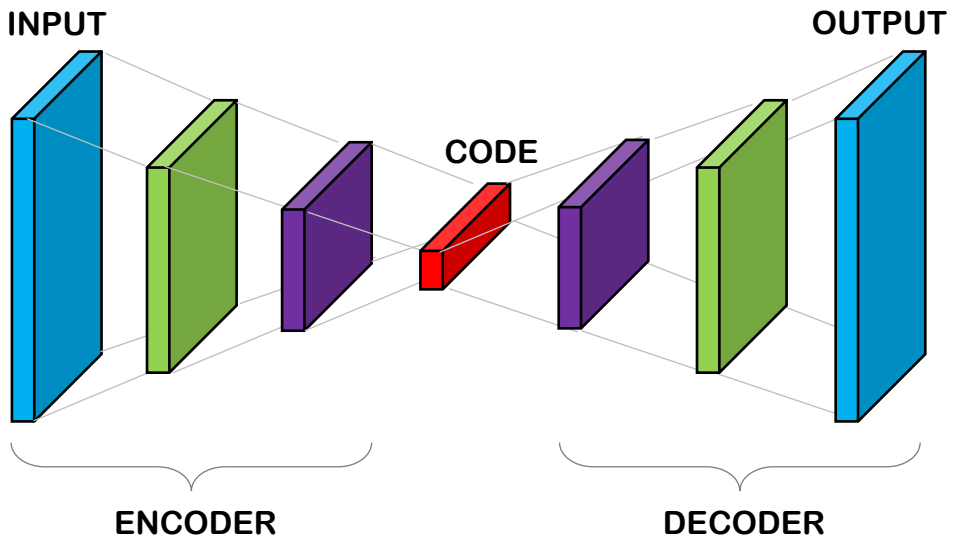
- Generate Human Faces
- Generate Cartoon Characters
- Image-to-Image Translation
- Text-to-Image Translation
- Semantic-Image-to-Photo Translation
- Generate New Human Poses
- Face Aging
- Super Resolution
- Photo Inpainting
- …



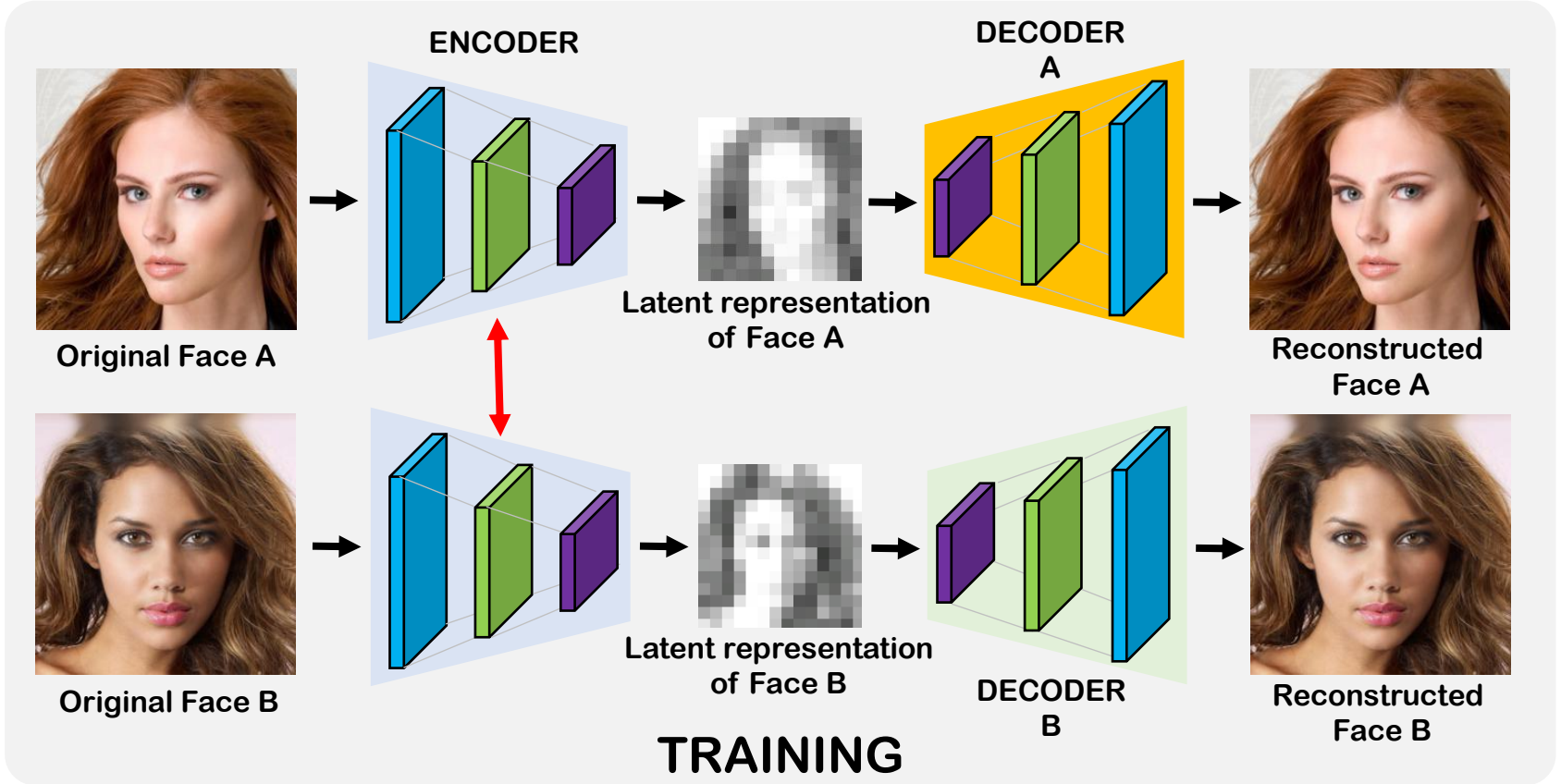


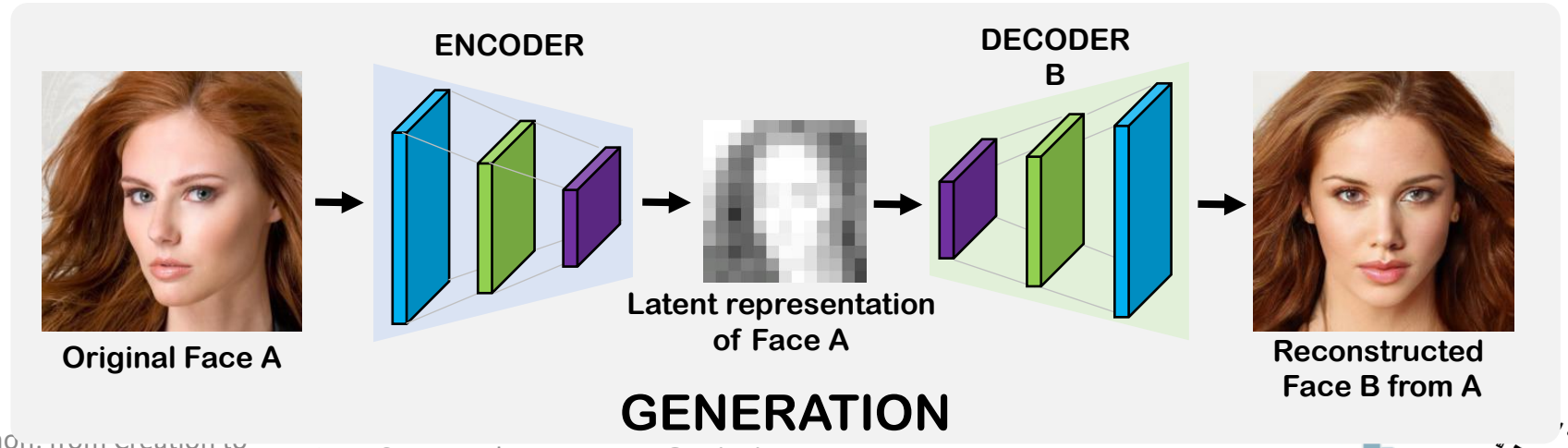

Luca Guarnera luca.guarnera@unict.it

Autoencoder



Deep autoencoders





2005: SCIgen
An Automatic CS Paper Generator

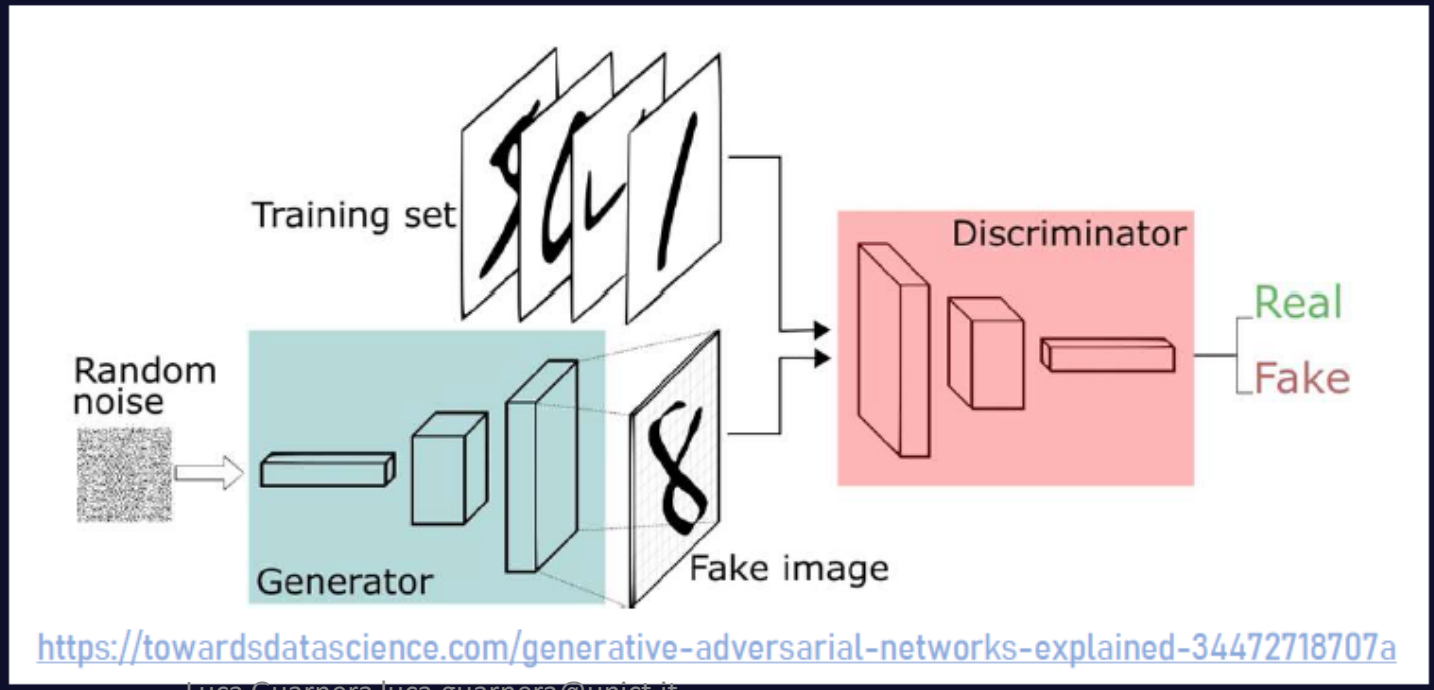
2012: Deep learning revolution
"ImageNet Classification with Deep Convolutional Neural Networks" paper

2014: GAN revolution
"Generative Adversarial Networks" paper

Generative Adversarial Nets

Abstract

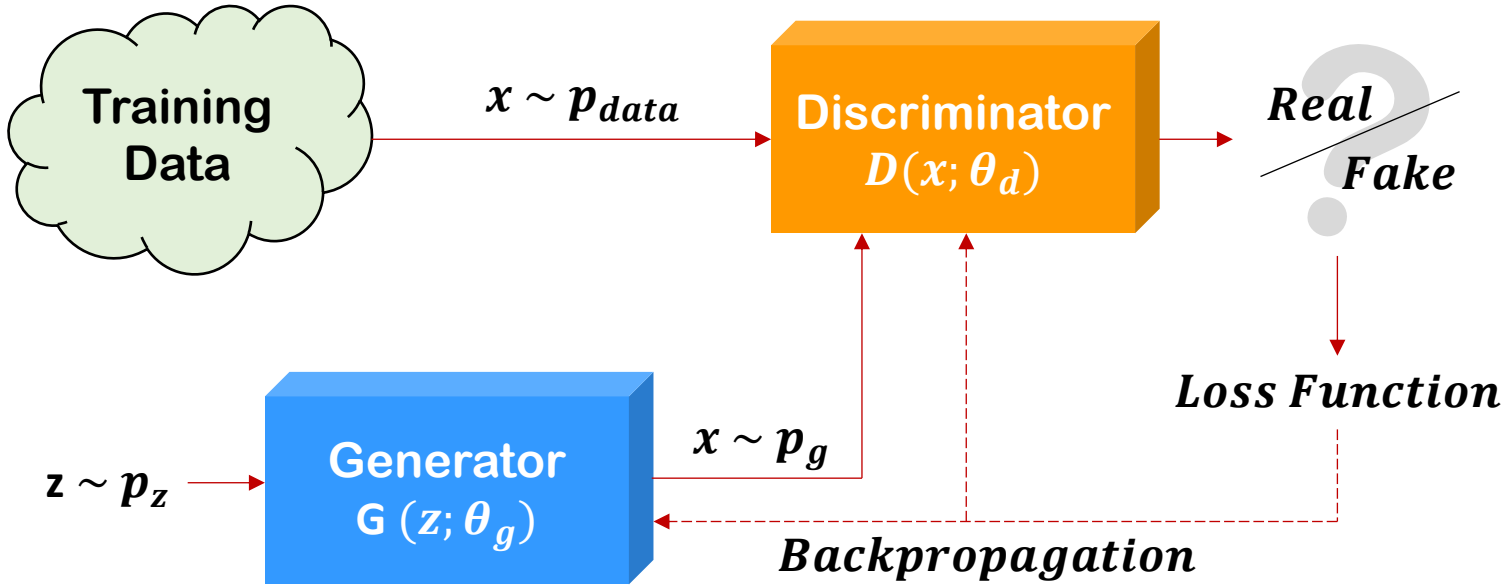
1 Introduction



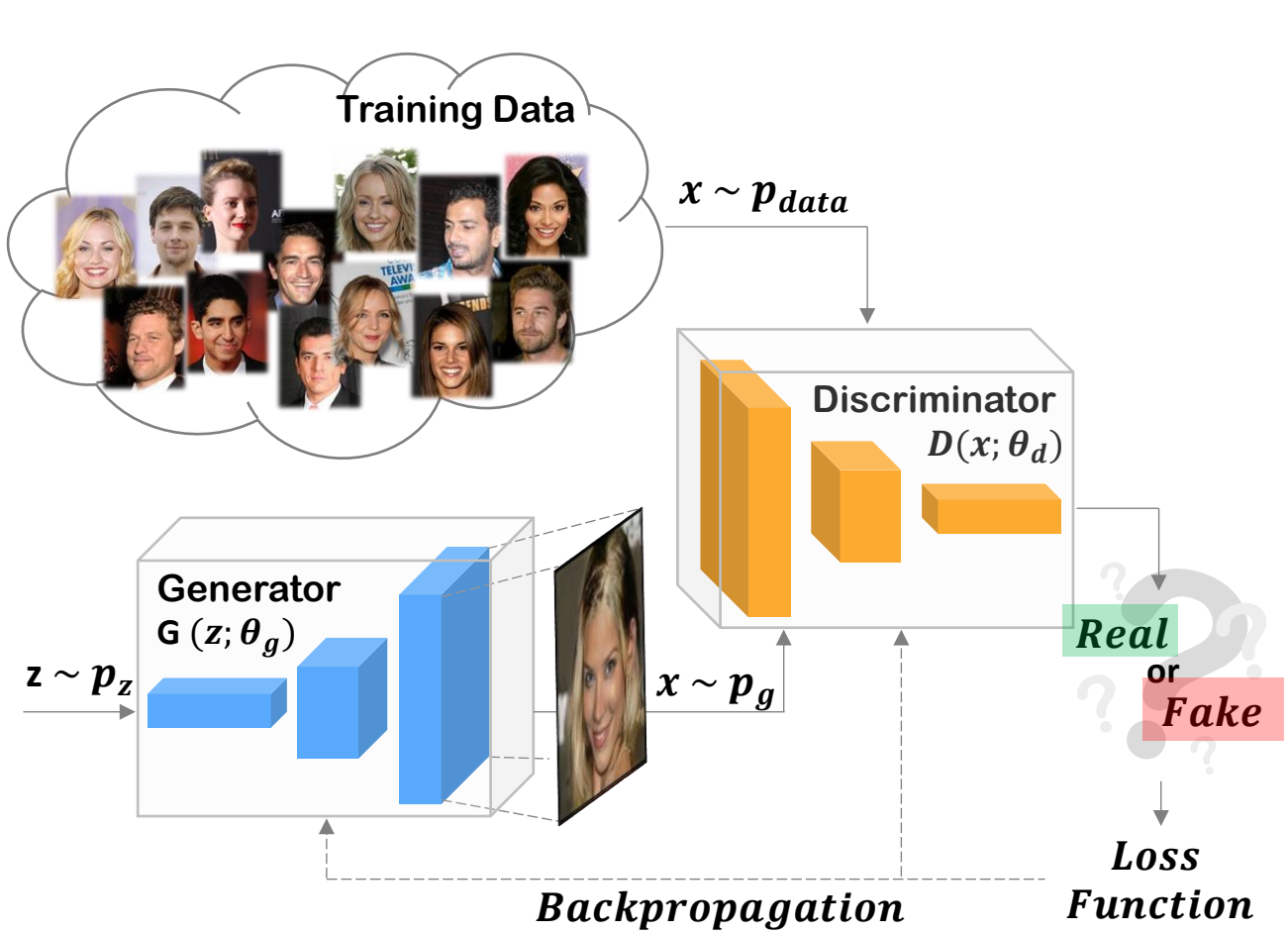
https://towardsdatascience.com/generative-adversarial-networks-explained-34472718707a

https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

Generative Adversarial Networks



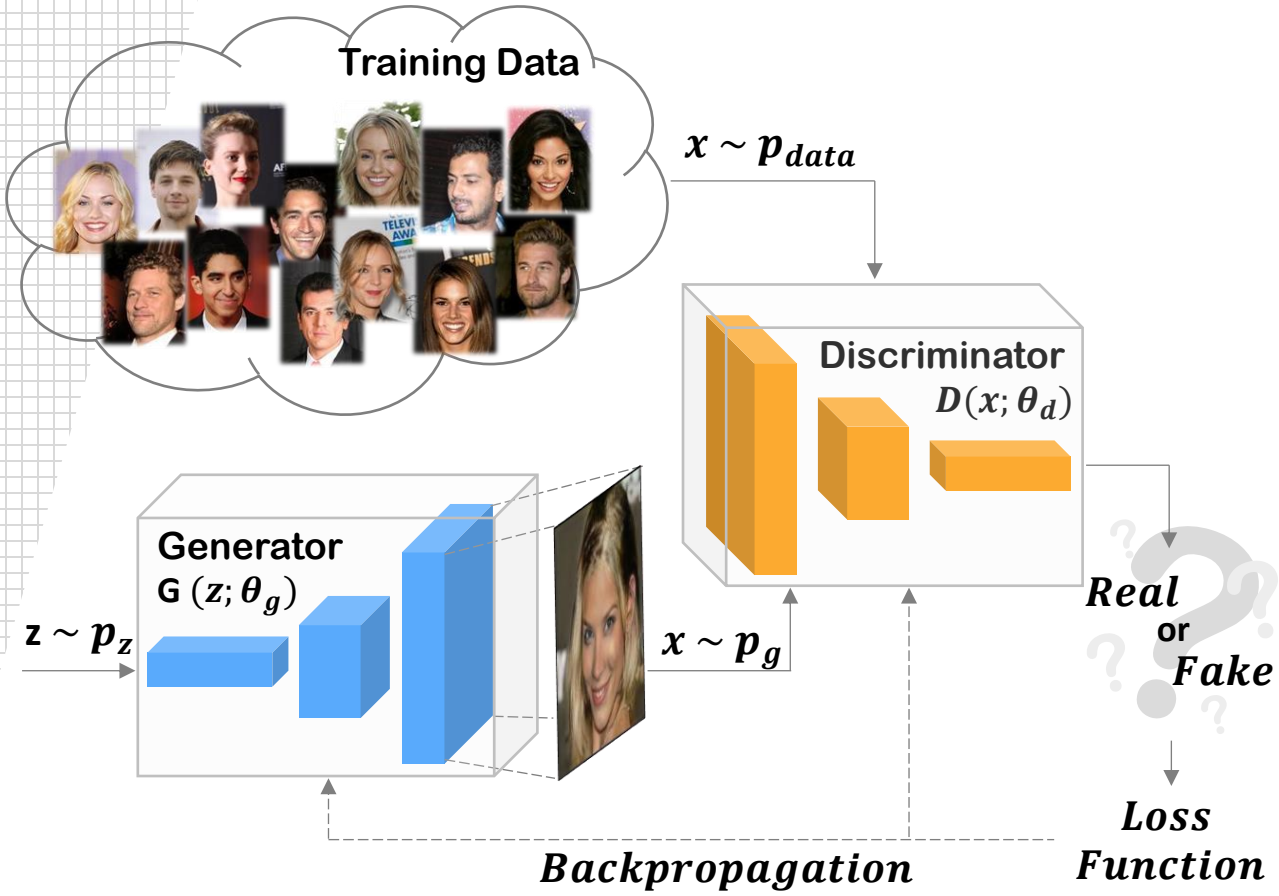
GAN example for human faces generation



But the image was created with GAN architecture. The image is a DEEPFAKE!

Luca Guarnera luca.guarnera@unict.it

GAN example for human faces generation

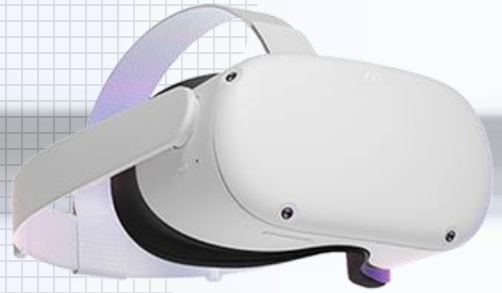


The evolution of GAN models for the creation of synthetic multimedia contents

Image source: Salehi, Pegah, Abdolah Chalechale, and Maryam Taghizadeh. **"Generative Adversarial Networks (GANs): An Overview of Theoretical Model, Evaluation Metrics, and Recent Developments."** *arXiv preprint arXiv:2005.13178* (2020).

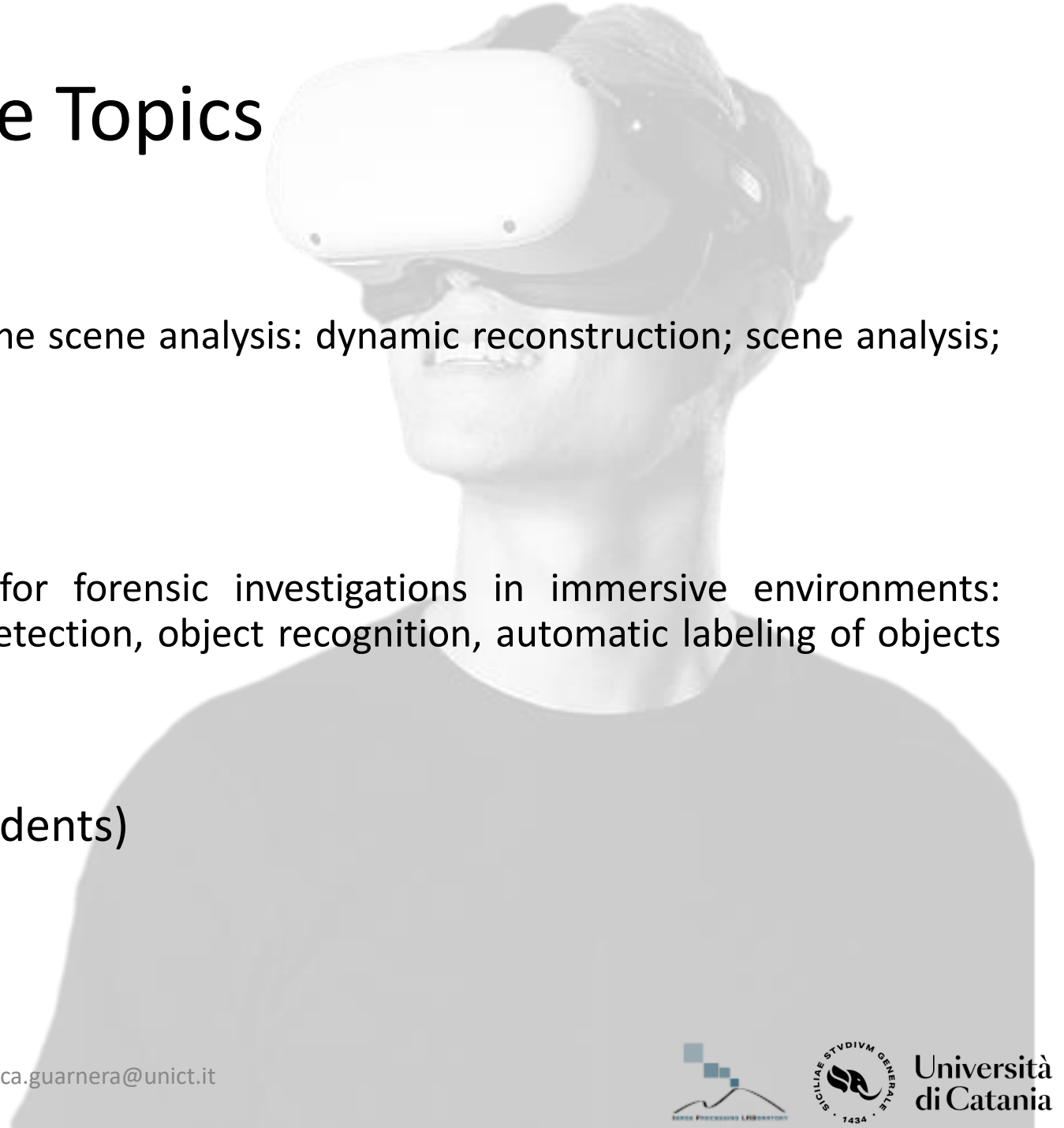
Opportunities for Students

Thesis



Students will be able to use advanced technologies in order to implement applications in the immersive environment through tools such as Meta Quest.

Some Topics



- Applications in Forensics:
  - Virtual laboratory (basic and manual) for crime scene analysis: dynamic reconstruction; scene analysis; educational laboratory; etc....
- Applications in AI Forensics:
  - Machine and Deep Leargning algorithms for forensic investigations in immersive environments: Intelligent Forensic Laboratory with object detection, object recognition, automatic labeling of objects in the scene, and more...
- Applications in other fields (chosen by students)

Main skills/requirements

- Basic knowledge of object-oriented programming

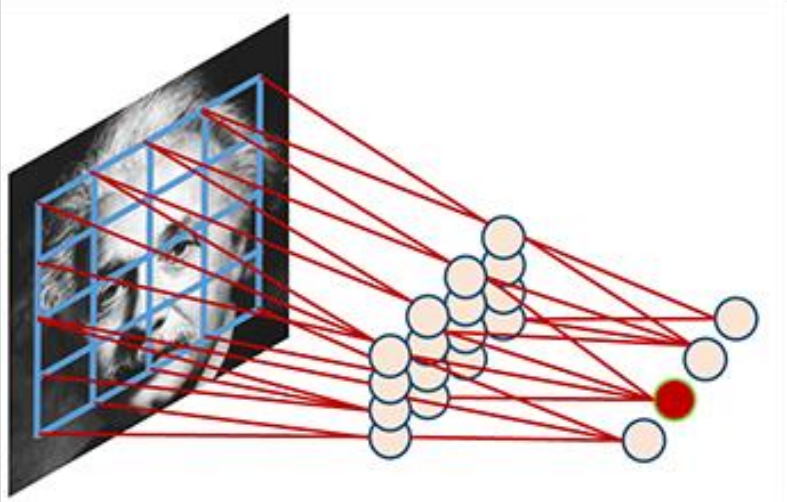
Depending on the application to be made, it is preferable that the student also knows

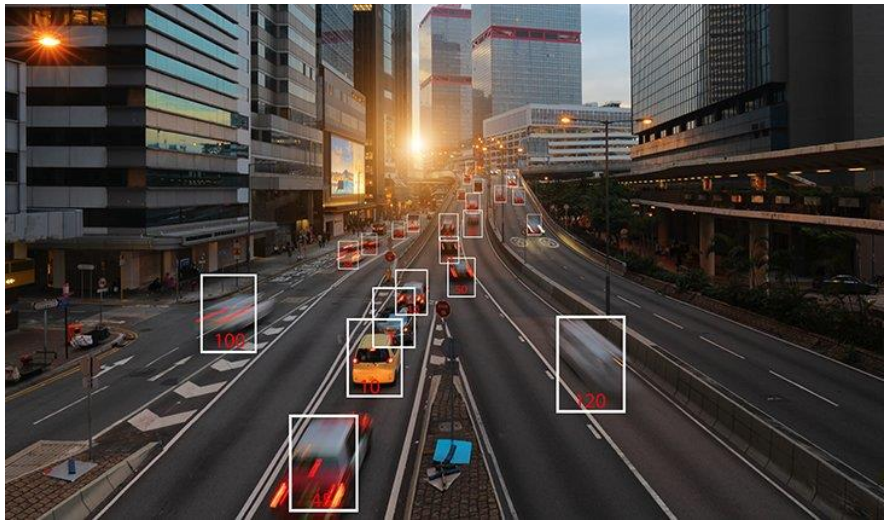
- Basic Knowledge of Computer Graphics, Unity, C# language (or Unreal and C++ language)

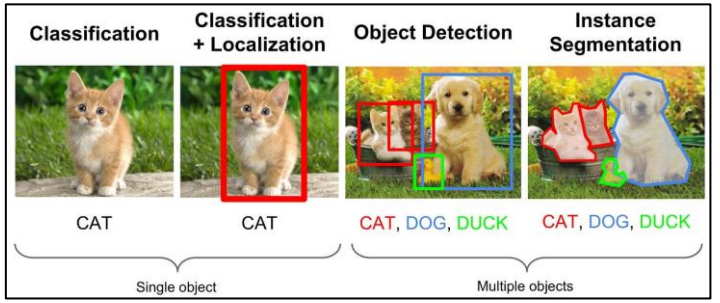
Multimedia Data …

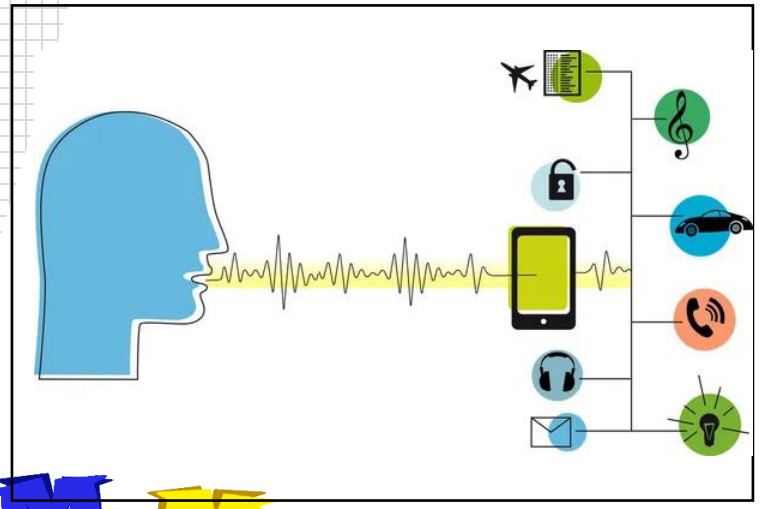
… Data Modelling

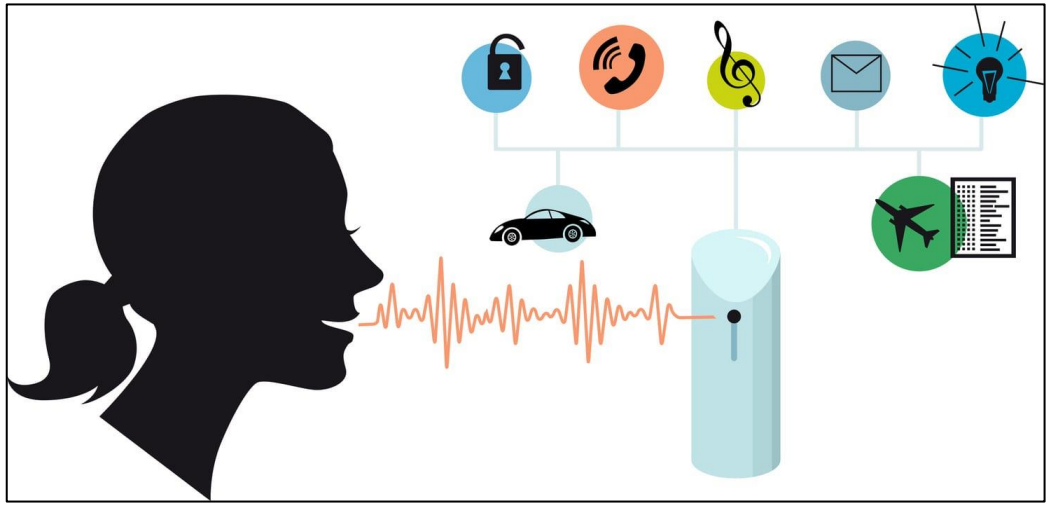
Applications

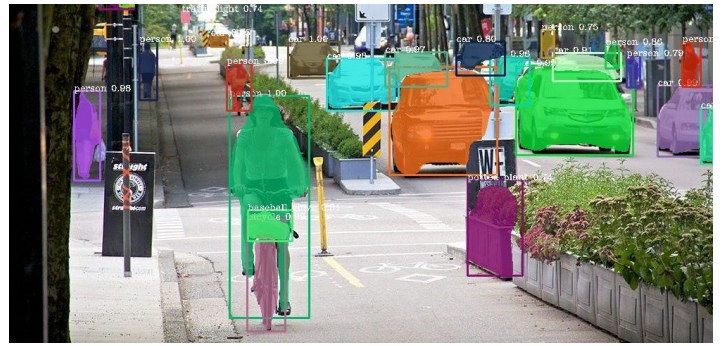










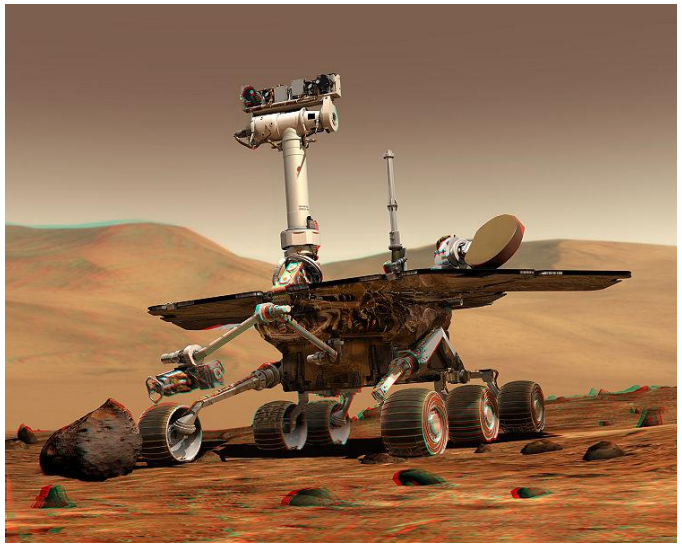


What is Computer Vision?

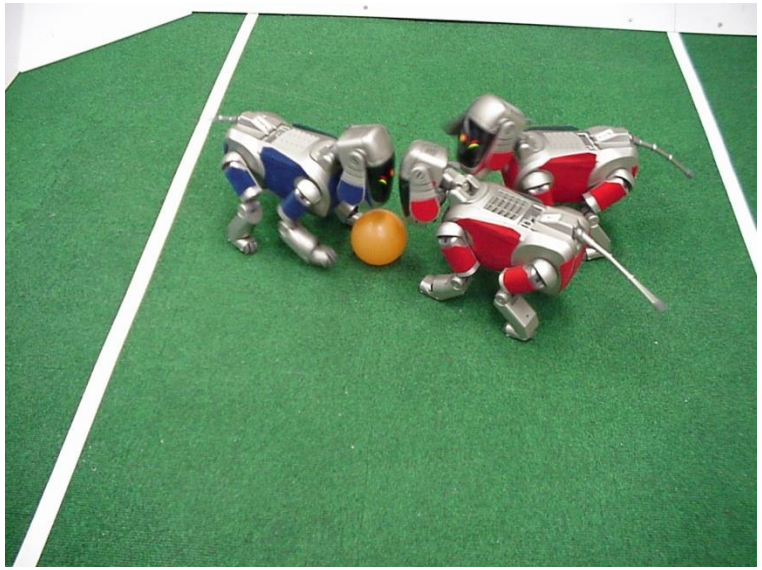
Vision is perhaps the **most important sense that humans possess**. It enables us to infer the three-dimensional world, recognize and locate objects in a scene, perceive rapid changes in the environment, etc.

Computer Vision is the discipline that studies how to enable computers to understand and interpret visual information present in images or videos.

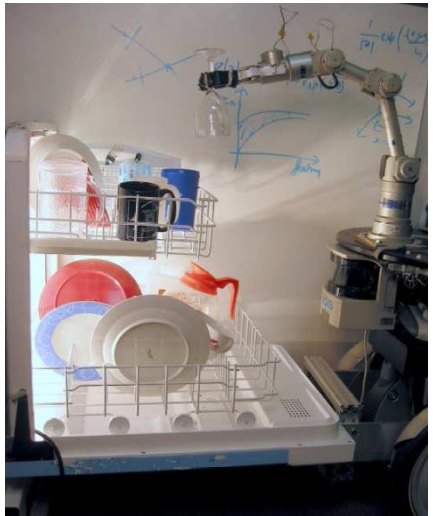
CV Applications: Mobile robots



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

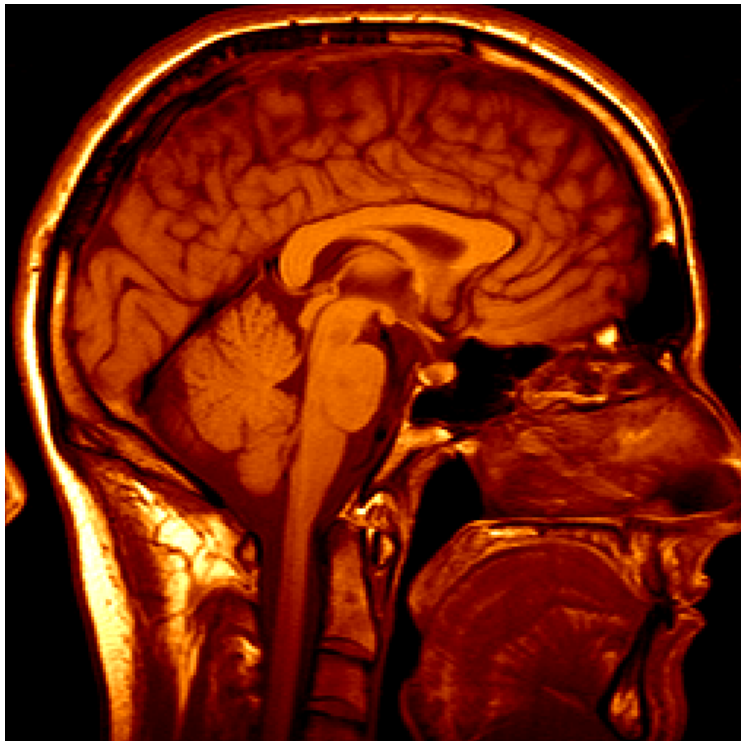


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

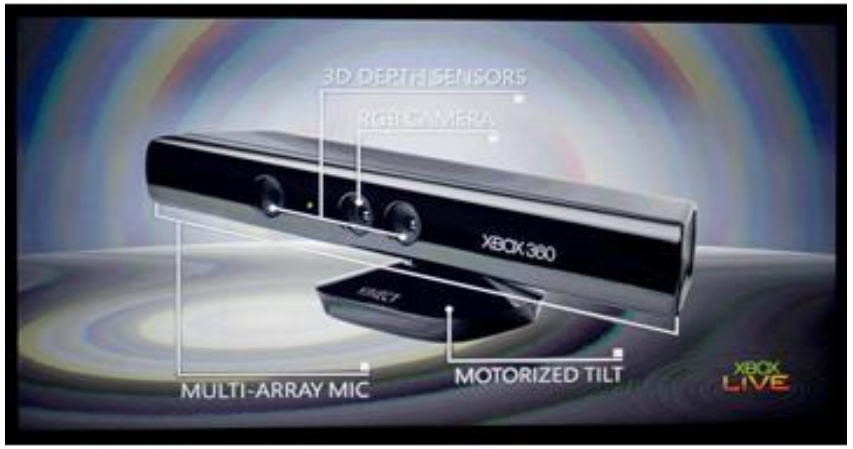
CV Applications: Medical imaging

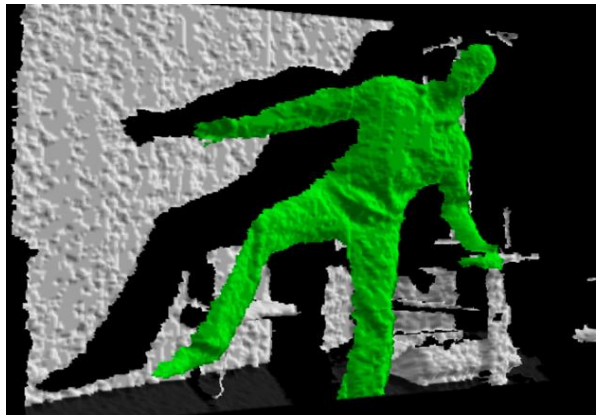


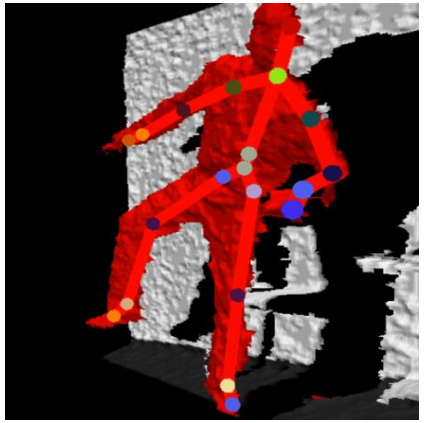
3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Vision-based interaction: Xbox Kinect







http://electronics.howstuffworks.com/microsoft-kinect.htm

slide credit: Lazebnik

CV Applications: Safety and Security

Goal:

Camera Stereo – Real time monitoring of dangerous conflicts (car, pedestrian, etc.)

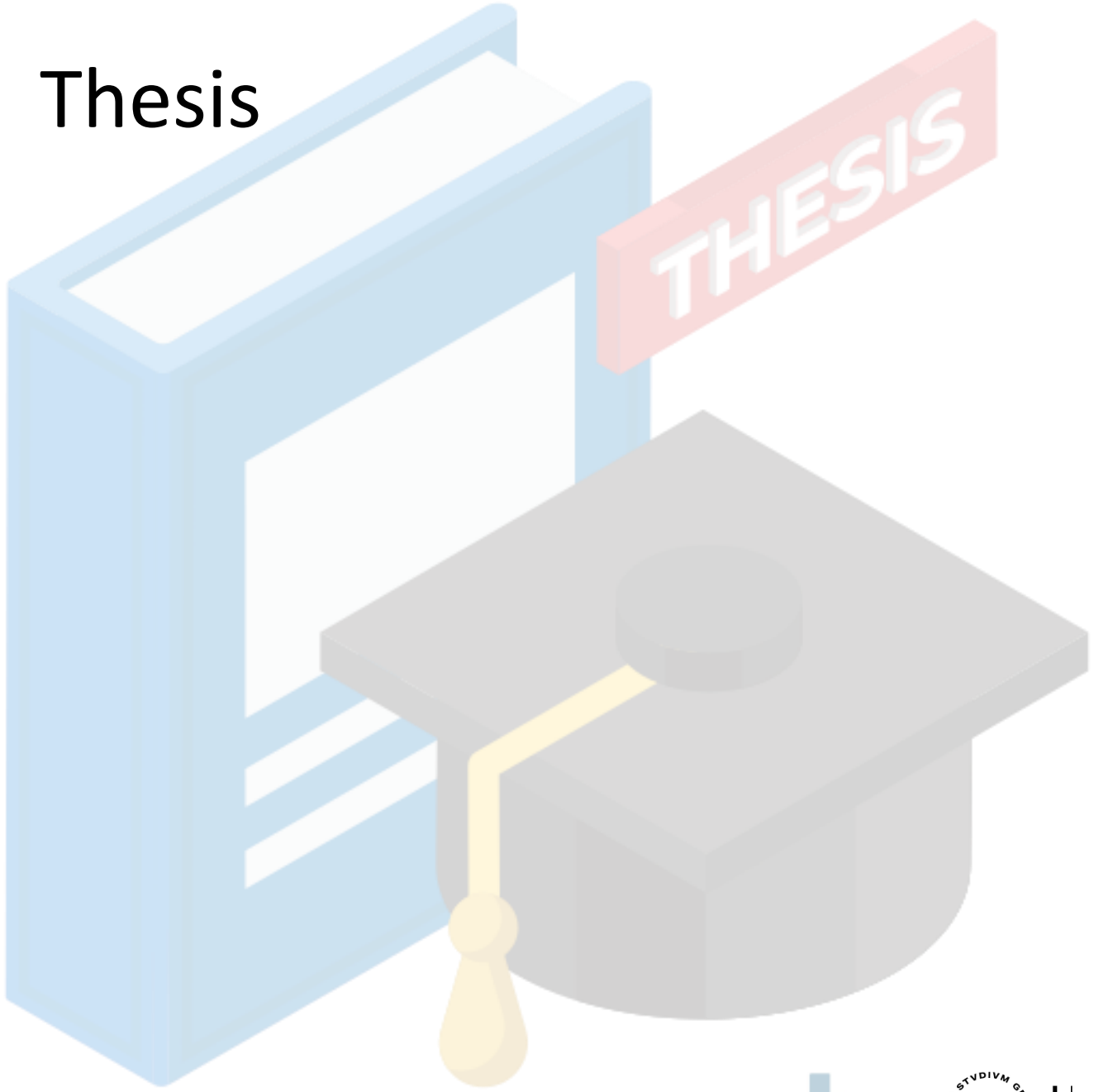
Real-time traffic monitoring

Computer Vision Goals

- Build systems capable of making decisions starting from a description of the scene extrapolated from images/videos;
- Infer the 3D world from digital images;
- Recognition of objects, scenes, context from digital images.
- ….

Opportunities for Students (2)

Thesis



- Deepfakes:
  - Creation
  - Detection
  - Adversarial
- Digital Forensics
- Computer Vision Task
- …

Questions?

Luca Guarnera, Ph.D.

Research Fellow

luca.guarnera@unict.it

University of Catania
Dipartimento di Economia e Impresa

Let's start with the course program …

Details

- **The first part of the course is about digital images**
- The second part of the course is about digital video
- The third part of the course covers Low-Level Vision
- The fourth part of the course concerns modeling and processing of digital data

Introduction to Multimedia Data Modelling

Luca Guarnera, Ph.D.

Research Fellow

luca.guarnera@unict.it

University of Catania
Dipartimento di Matematica e Informatica