



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

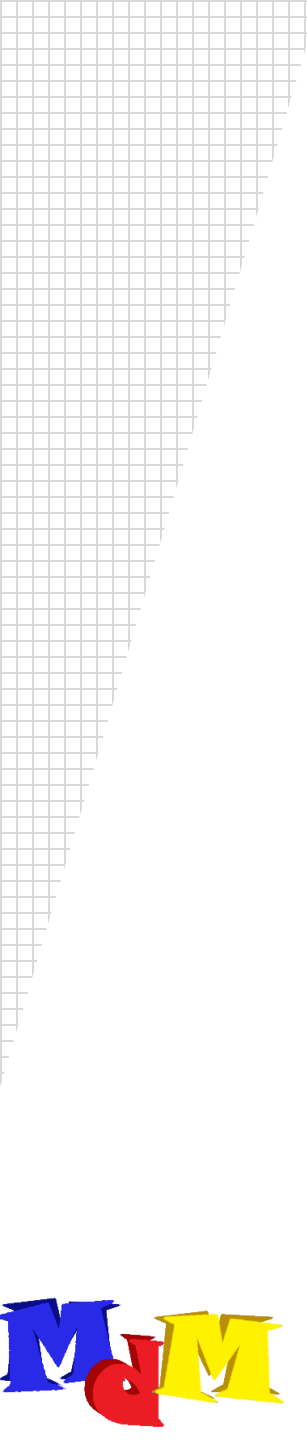


Università
di Catania



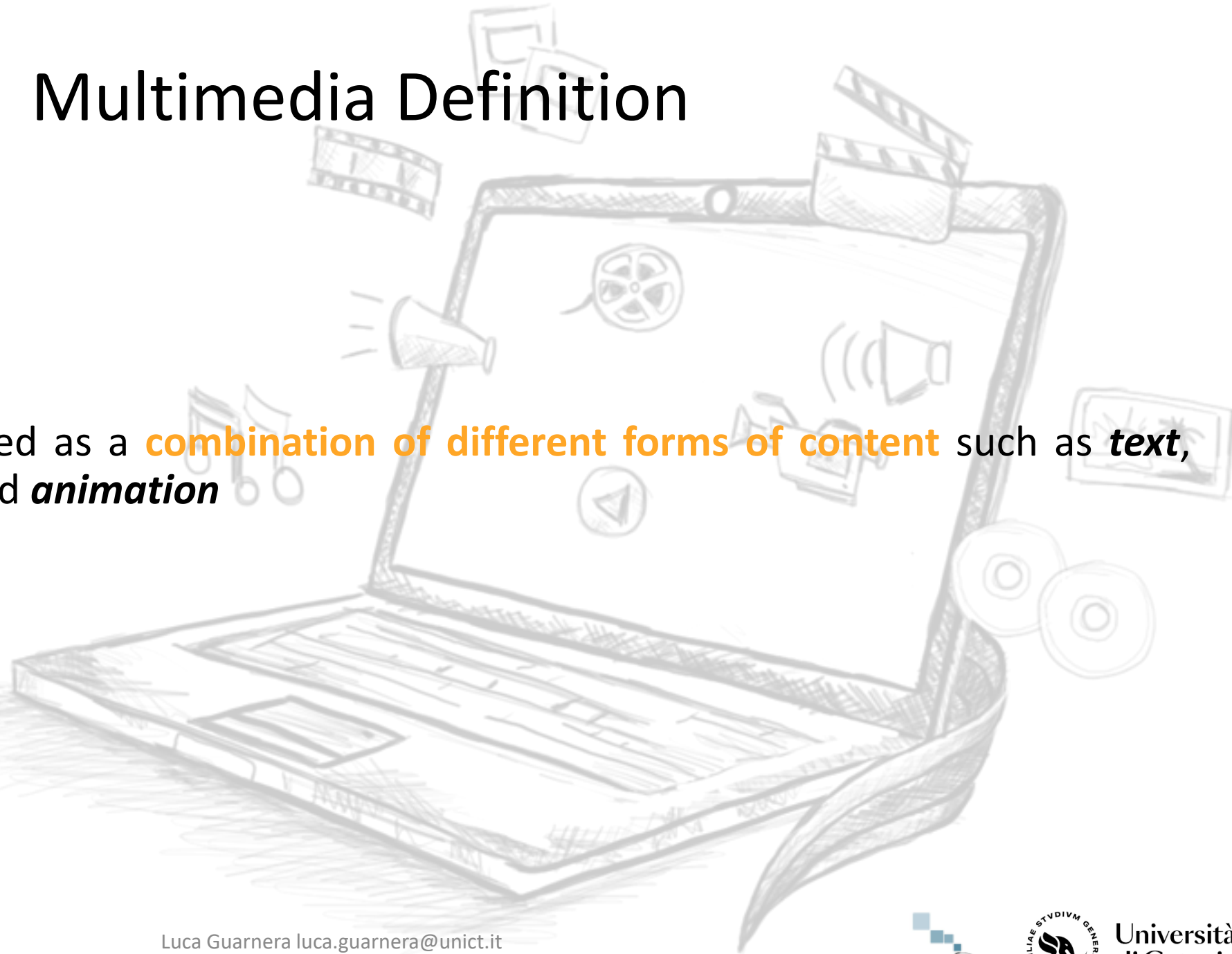
Multimedia Data Modelling

Luca Guarnera luca.guarnera@unict.it



Multimedia Definition

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

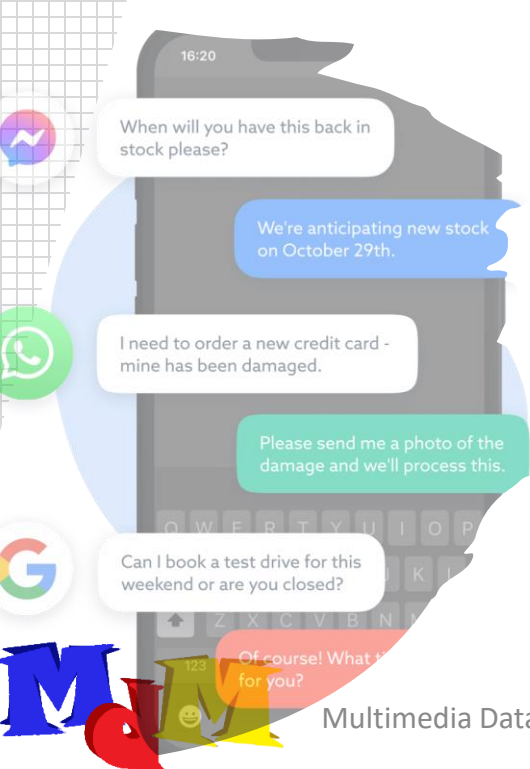


Exploring the different forms of multimedia



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**.

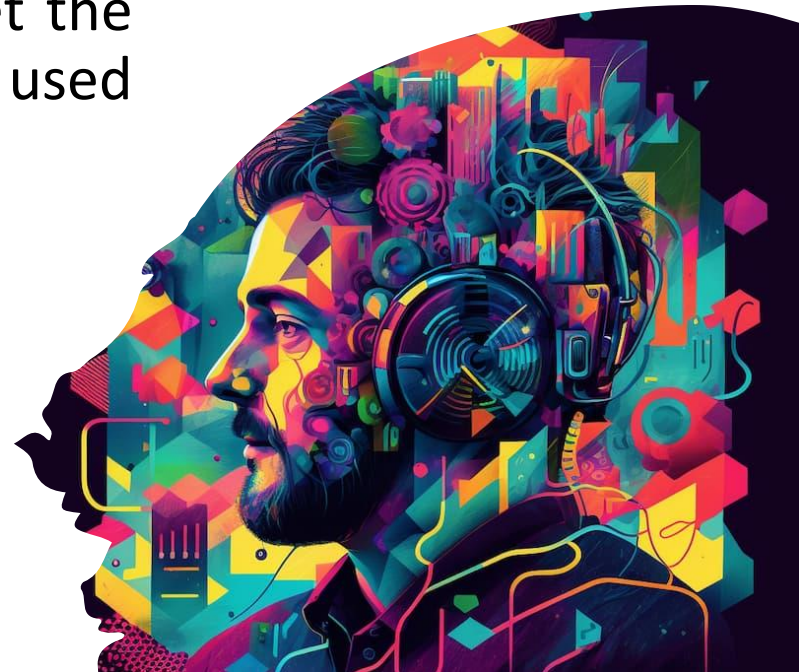
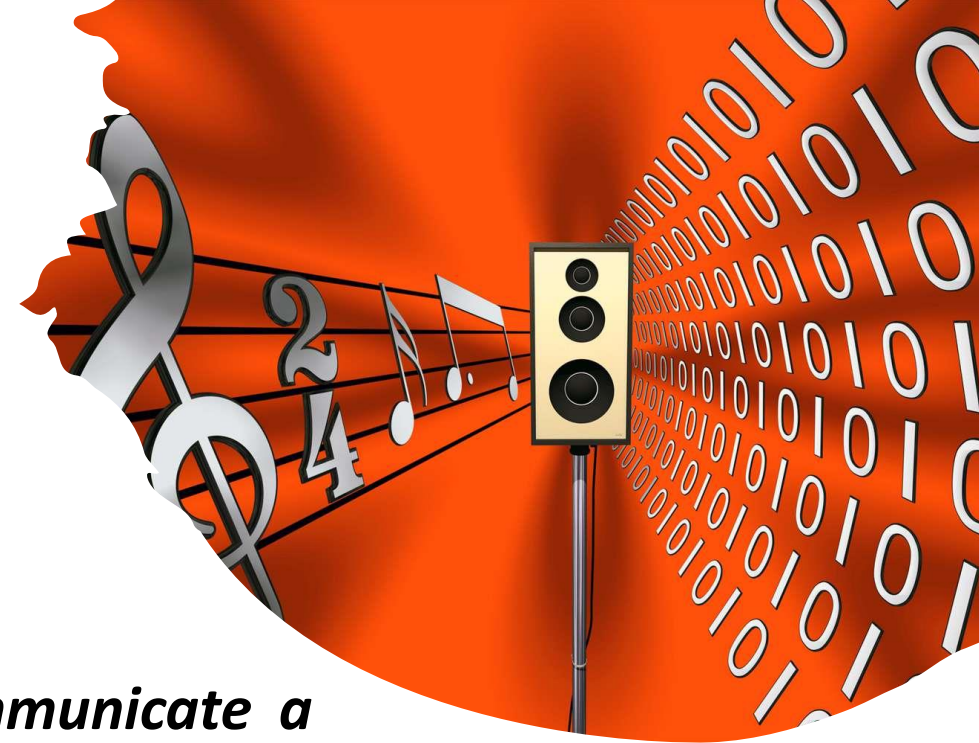


Animated Graphics



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

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.

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 - 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 natural environment and the real-world issues. The various educational contexts are explained in this article.

1. Introduction

In a very real sense, virtually every individual has had experience with multimedia systems of one type or another. Perhaps the most common multimedia experiences are reading the daily newspaper or watching television. For most of us, when we think about multimedia and the promise for future systems that combine video, graphics, animation with special effects (as seen in movies like 'Who Framed Roger Rabbit') and CD quality audio. On a more business oriented scale, we think about creating virtual meeting rooms with 3-dimensional whiteboards, computer applications, and perhaps even computer-generated Business Meeting Notes documenting the meeting in an efficient communications format. Other glamorous applications in which we learn and interact with instructors remotely over a broadband communication network. Virtual Library Access in which we instantly have access to all of the published material in

the world, in its original form and format, and can be displayed, printed, even modify the material with electronic and Living Books which supplement the written material with associated graphics with animations, and hyperlink access to associated resources.

Modern voice communications networks evolved around the turn of the century with a focus on creating a 'Universal Service' that would connect any telephone user with any other user. The first telephone network was defined before the turn of the century and consisted of a series of technological problems that had to be solved. The first vision became reality, including the telephone vacuum tube for amplification of the signal, mechanical switching to replace the plans and were used in most localities, numbering plans and signaling systems to route calls, etc. The first transcontinental call in the United States was completed in 1915, thereby ushering in the 'modern age of voice communications', an age that has been developed and improved upon for the past 80 or so years.

We are now in the midst of another revolution in communications, one which holds the promise of providing ubiquitous service in multimedia communications. The vision for this revolution is to provide seamless, easy-to-use, high quality, affordable multimedia communications between people and machines, anywhere, and anytime. There are three key aspects of the vision which characterize the changes that will occur in communications once this vision is achieved, namely:

- the basic currency of communications evolves from narrowband voice telephony to seamlessly integrated, high quality, broadband, transmission of multimedia signals;
- the basic access method changes from wireline connections to combinations of wired and wireless, including cable, fiber, cell sites, satellite, and even electrical power lines;
- the basic mode of communications expands from primarily involving people-to-people communications, to include people-to-machine communications.

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.

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 natural environment and the real-world issues. The various educational contexts are explained in this article.

Key Words - Multimedia; Interactive Media; Educational Multimedia; Multimedia Sign

1. Introduction

In the modern world, we use many tools, slides, overhead transparencies, etc., to assist in the teaching process inside the classroom. The word 'multimedia' describes a combination of different media. When computers became prevalent in society, people started using them to perform various tasks to make their life easier. Computers are a powerful resource that we can use for many things. Interactive multimedia is a reliable technological innovation, and it has the potential to modernize the way we read and explore educational things. Multimedia is an interactive as well as educational tool. These types of interactional tools are likely to make educational subjects more interesting. The objective of using multimedia as an educational tool is not to eliminate the teacher from the classroom; rather, it is just a tool, which makes it easier for both students and teachers to learn as well as teach particular subjects. Multimedia tools include well-designed programs that simulate the teacher's role by adding various educational elements to the cognitive process. The learning process involved in multimedia programs facilitates active learning and ensures that users are practicing a subject rather than merely reading about it. Actual communication between multimedia programmes and students constitutes a proper learning process. Interactive multimedia is a new technology that is

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.

introducing new methods in the learning process to the new generation. Teachers should be able to easily access these multimedia applications to monitor their students' progress and modify the application's contents according to the students' characteristics.

Park and Hannafin stated that most of the aspects of interactive multimedia systems, which are used for teaching, are questionable. Here, we can see the difference between computer-based education methods and traditional education. In multimedia education, there are many multimedia choices available to choose from, it is important that developers of these instructional multimedia applications understand the needs of their applications. Interactive multimedia applications are crucial, as they are the only reason for further investment and research in this field. The main implications of building an educational multimedia system. We will also discuss theories, which can be used to build a proper guidance system for developers to build and select the proper elements to build multimedia systems. Finally, we will review several different educational elements that can be used for educational purposes for multimedia systems.

II. Multimedia

Before starting discussion on our topic, we must clarify the definition of 'Multimedia'. The word 'Multimedia' is a reasonably new one in its field. It is used to describe several different mediums when they are merged together. We can define multimedia according to its common characteristics: texts, graphics, animations, video, and sound. These are all combine to create multimedia, but they can also be organized and presented differently. In other words, multimedia can be define as numerous media elements combined into one whole subject, which produces fruitful outcomes for its end user. All these media elements are making communication more organized and clear than ever before.

Several researchers have provided definitions of multimedia. Moore et al. defined multimedia as follows: the use of numerous media devices in a

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 boulder technology.



Metaverse

Brief introduction to the Metaverse



Metaverse



Metaverse: A first definition

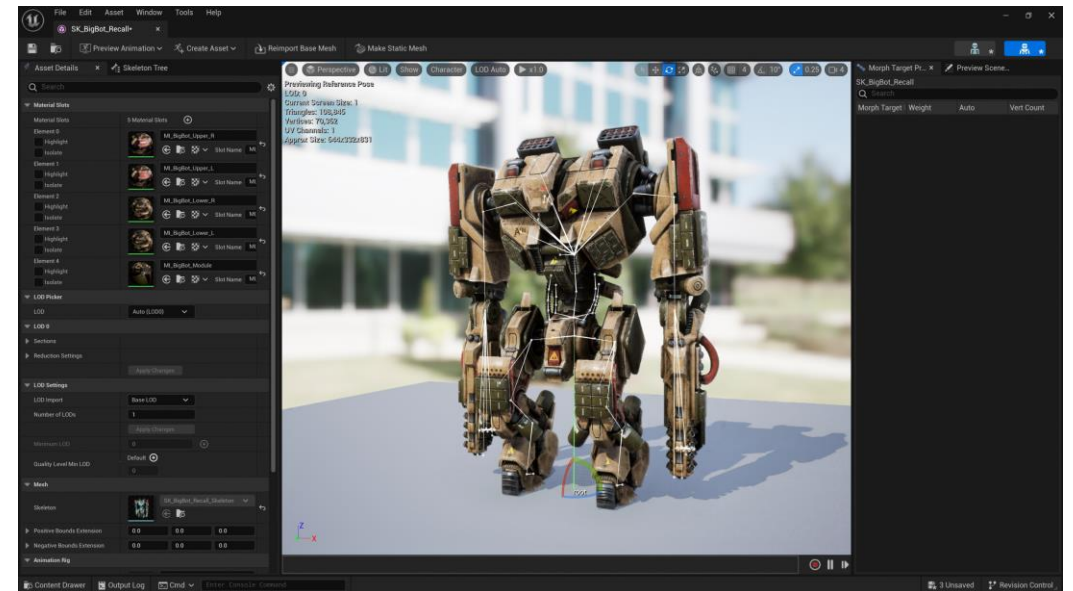
Metaverse, a term born in the cyberpunk world in 1992 and now in the limelight because it is used by Facebook, is the evolution of the Internet, but does not replace it. **This is a difficult concept to define exactly**, which prefigures a set of **interconnected virtual** and real worlds, populated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



**UNREAL
ENGINE**



Unity

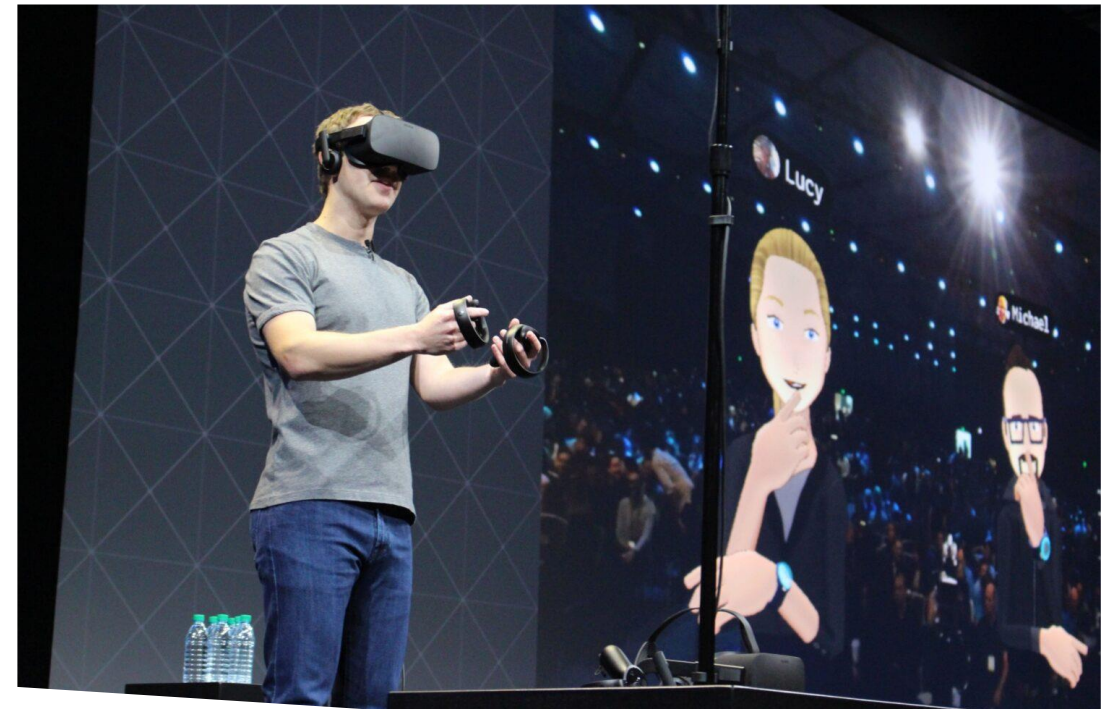


Multimedia Data Modelling

Luca Guarnera luca.guarnera@unict.it

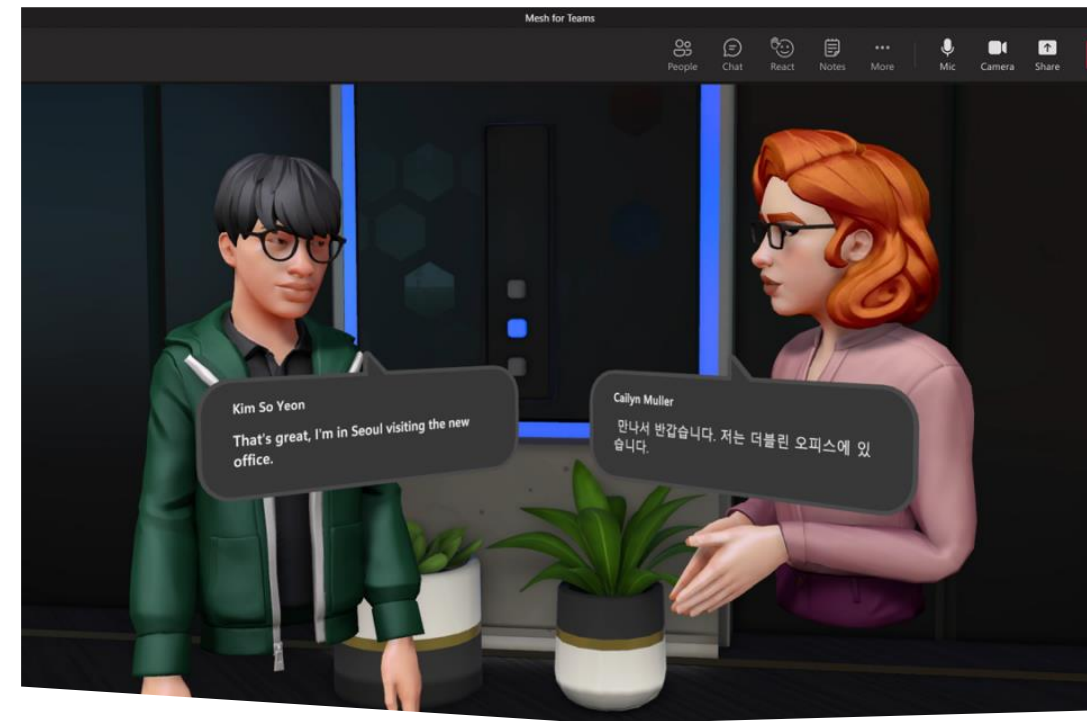


Università
di Catania



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.

A person is shown from the chest up, wearing a VR headset. The image is heavily stylized with digital glitch effects, featuring horizontal lines of various colors (blue, green, red, yellow) and pixelated distortions across the entire scene. The person's right hand is raised, with the index finger pointing upwards. The background is black.

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

- <https://about.meta.com/it/metaverse/impact/>



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."

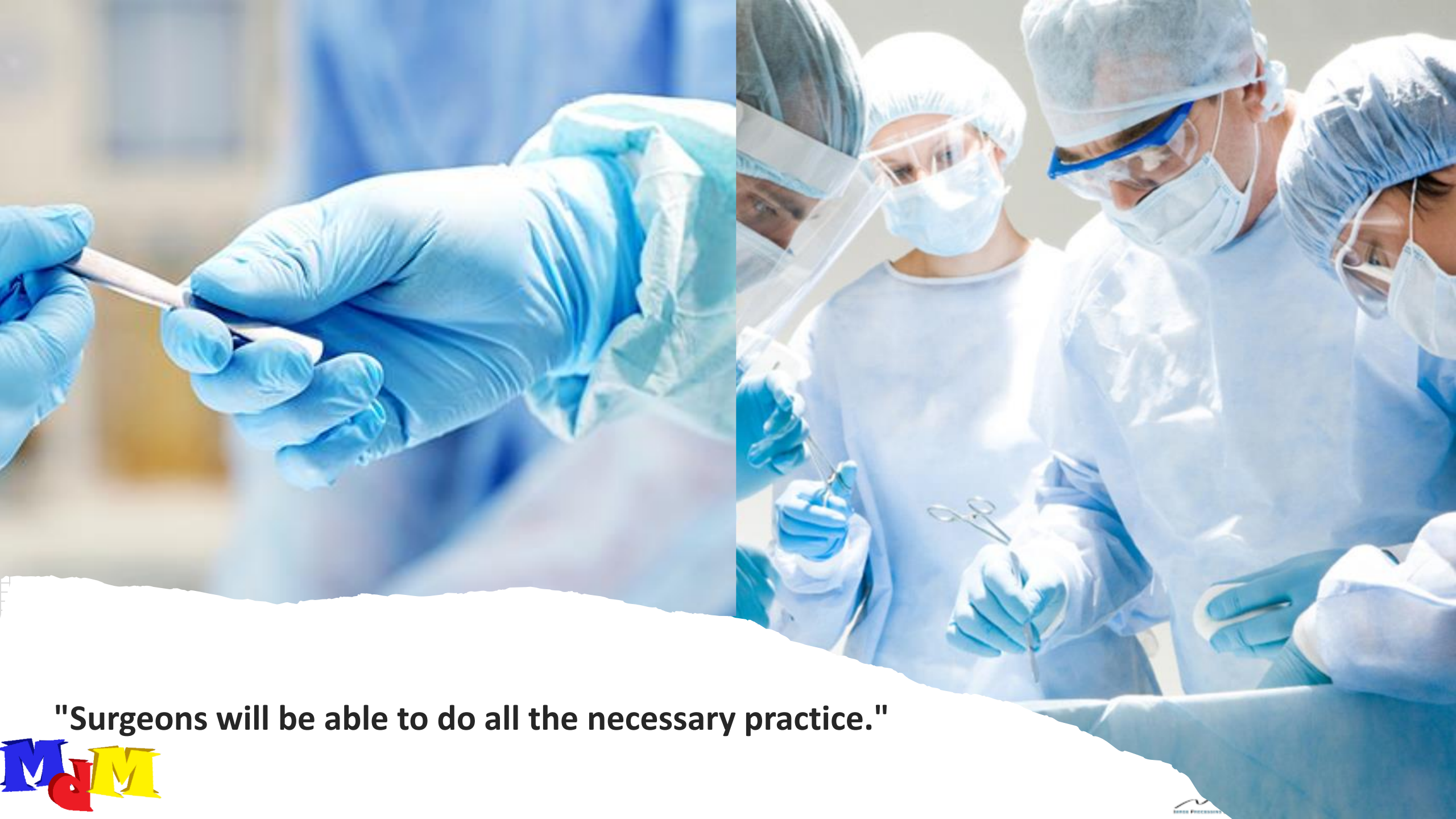






"One day doctors will be able to view diagnostic images like never before and make decisions more quickly."





"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***

Metaverse, how to access?

A woman with long dark hair is wearing a purple VR headset and looking upwards. She is in a music studio, with her hands on a keyboard and a mixing console. The background is a wood-paneled wall. Overlaid on the scene are digital elements: a large curved screen on the left showing a digital audio workstation (DAW) interface with multiple tracks and waveforms, and a large vertical screen on the right showing a mixing console with numerous faders and knobs. The overall lighting is dim, with the digital screens providing the primary light source.

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.

A man with a beard is looking out of a window at a city at night. The city lights are visible in the background, and the man is wearing a dark hoodie. The scene is dimly lit, with the city lights providing the main source of illumination.

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](#)



A person is holding a smartphone in their hand. The phone's screen shows a digital filter applied to a photo of a woman with dark, curly hair and large, expressive eyes. The background of the image is a blurred crowd of people, with some wearing blue and red clothing, suggesting a festive or public event. The word "Augmented reality" is written in a large, white, sans-serif font across the middle of the image.

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.



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

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

1

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

2

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

3

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

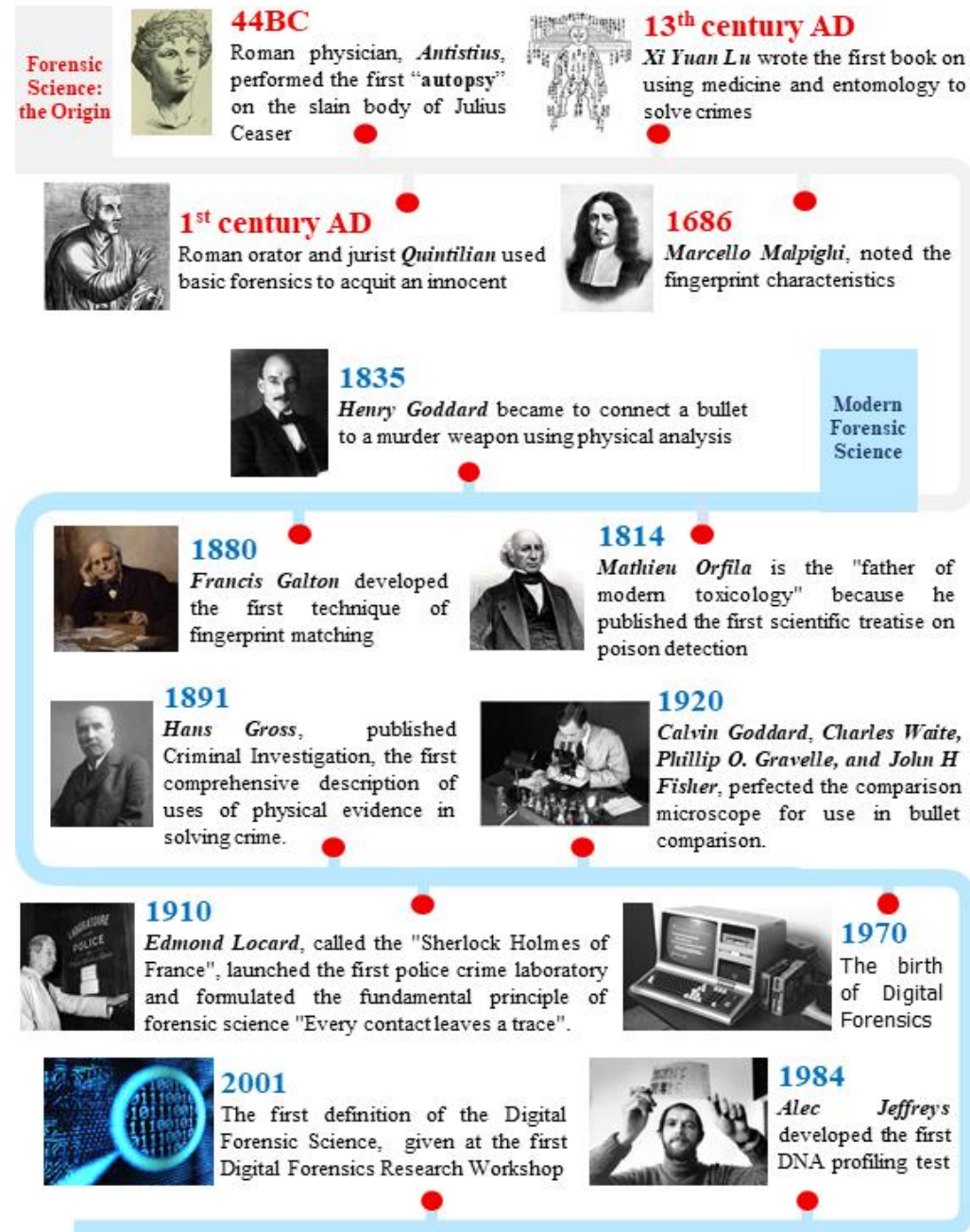


Image Forensics



Multimedia Forensics

Computer Forensics

Multimedia Forensics

Forensics

Network Forensics

Mobile Forensics

Database Forensics



What can we say about this image?



What can we say about this image?



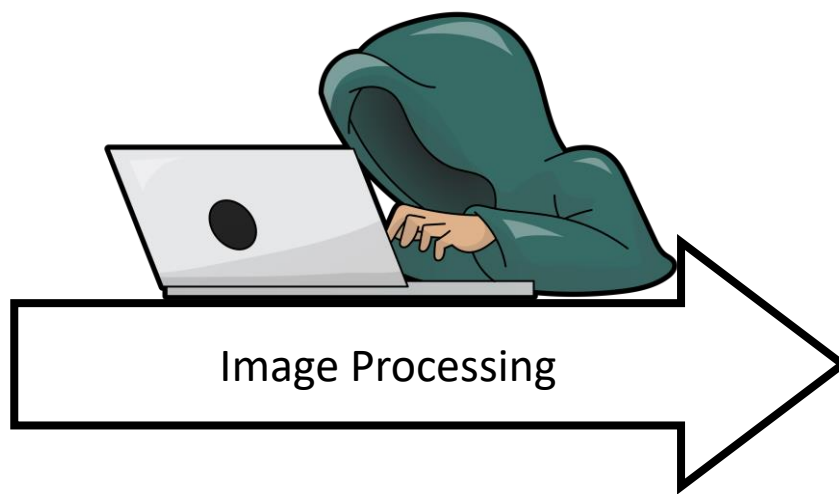
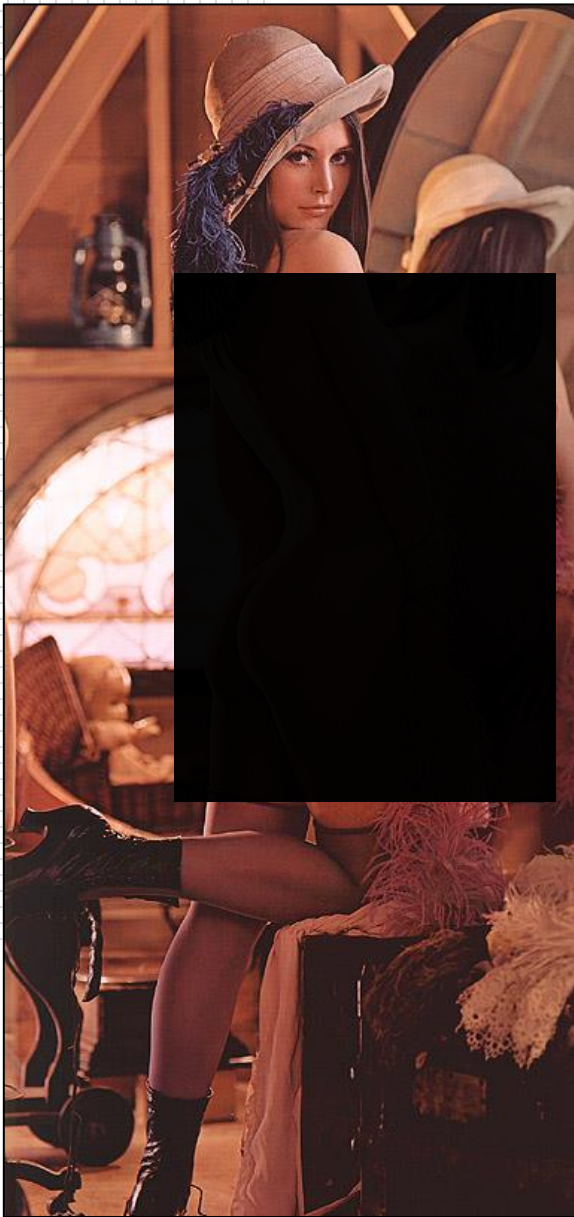
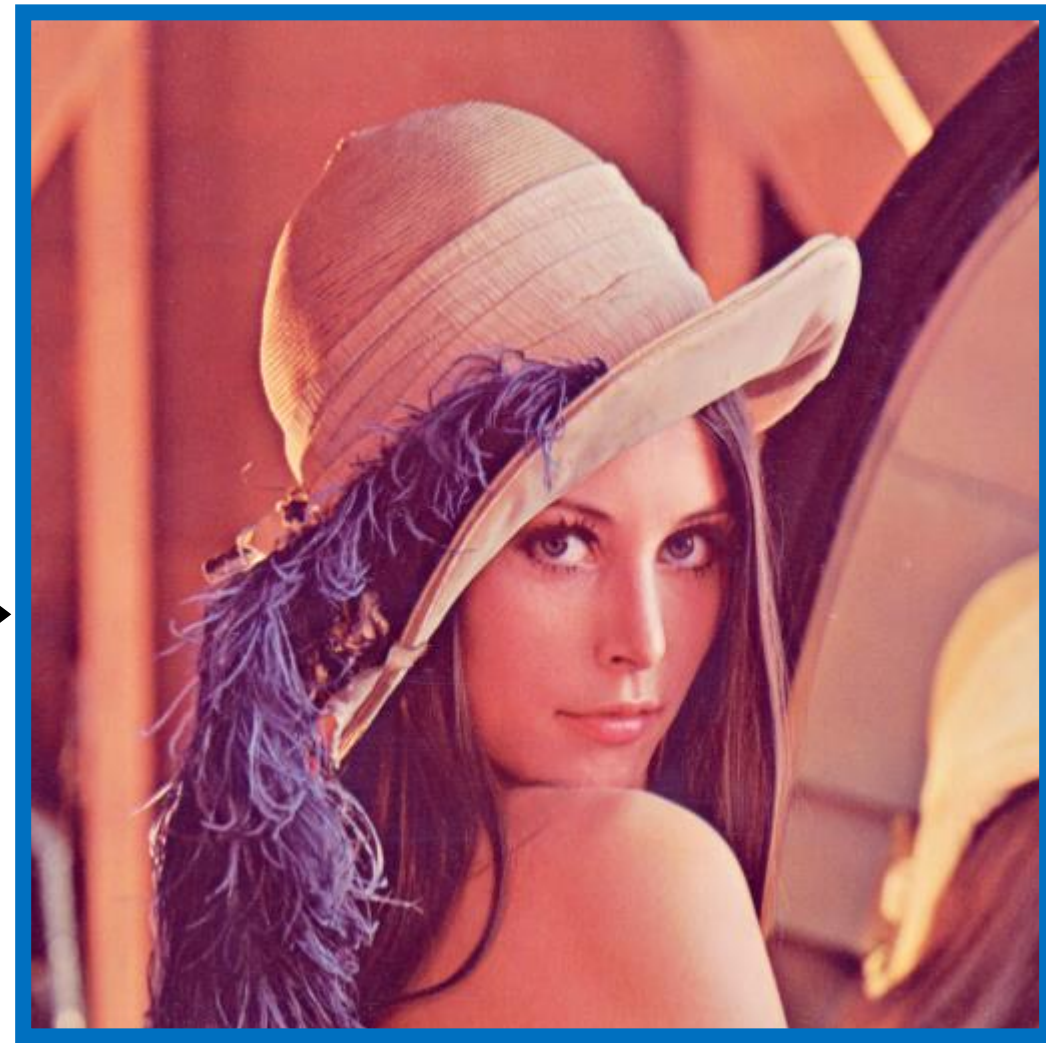


Image Processing



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

≠

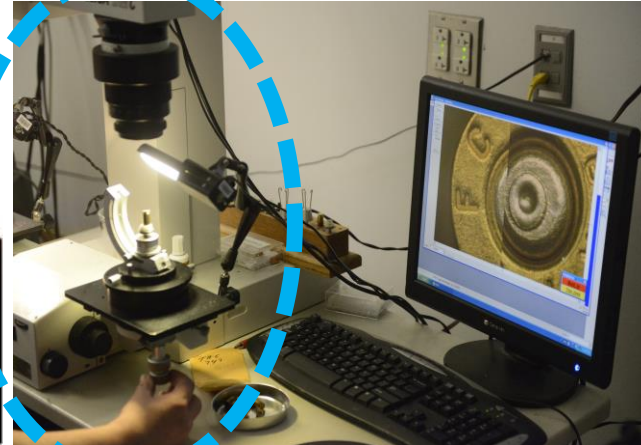
INTEGRITY

**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](#)

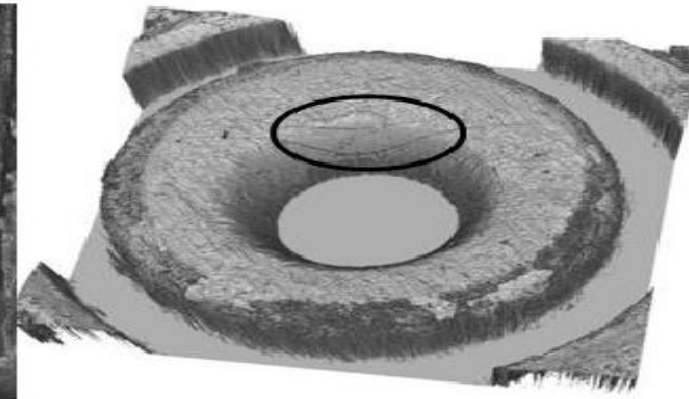
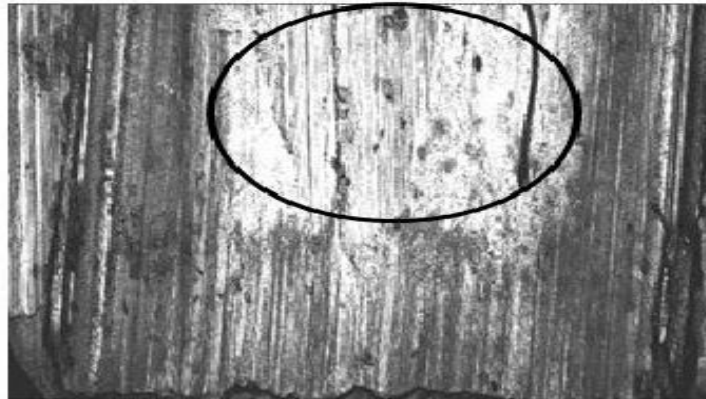
Forensics Examination of Firearms



Weapon serial numbers

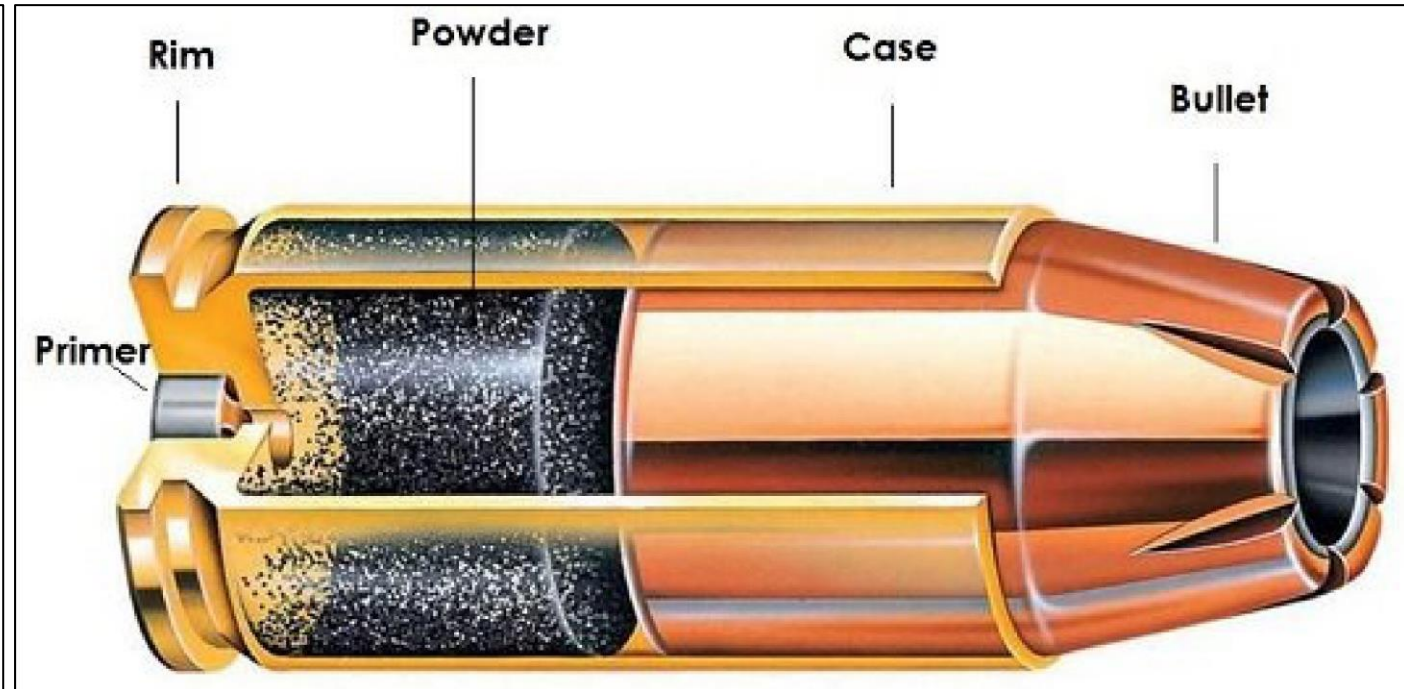
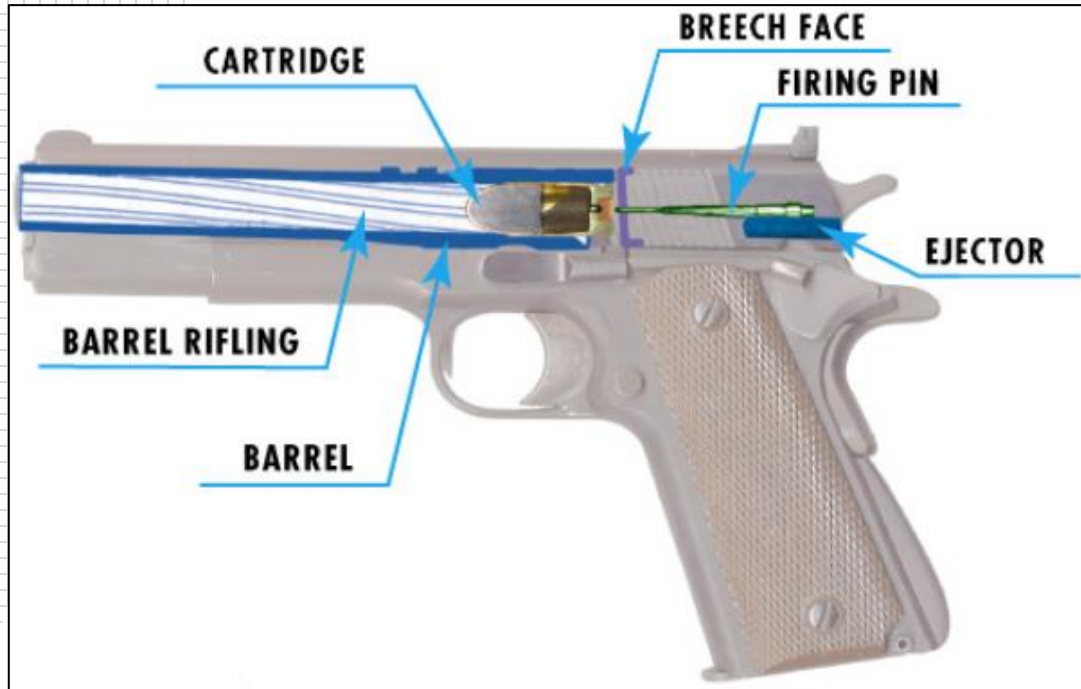


Fingerprints on the weapon's surface



Ballistic striations

Firearm and Cartridge components



Weapon firing process



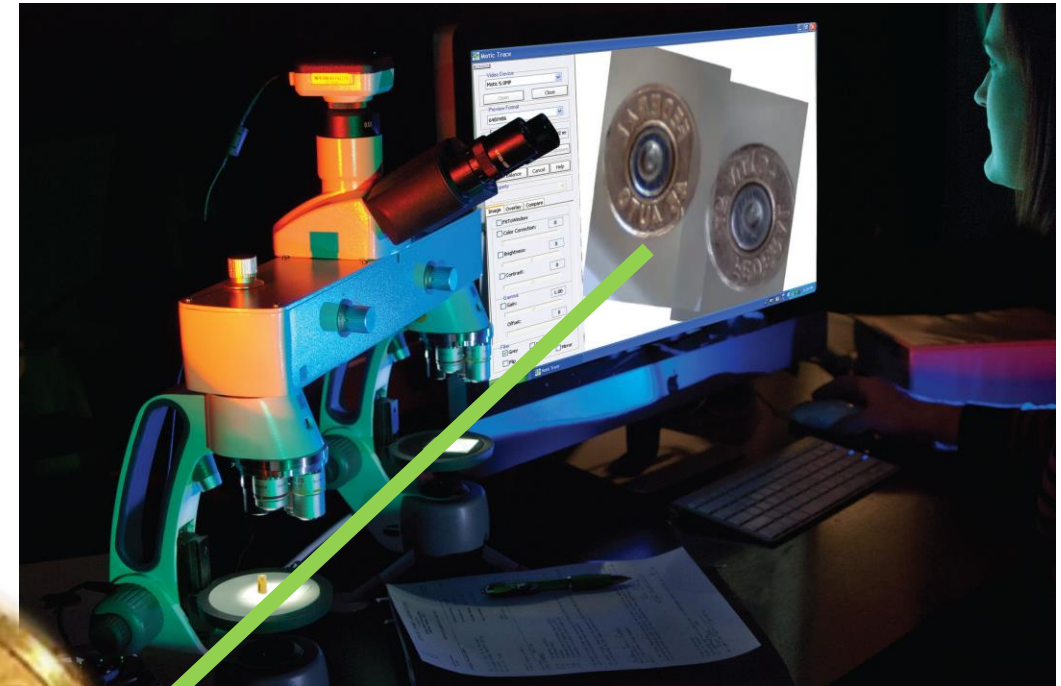
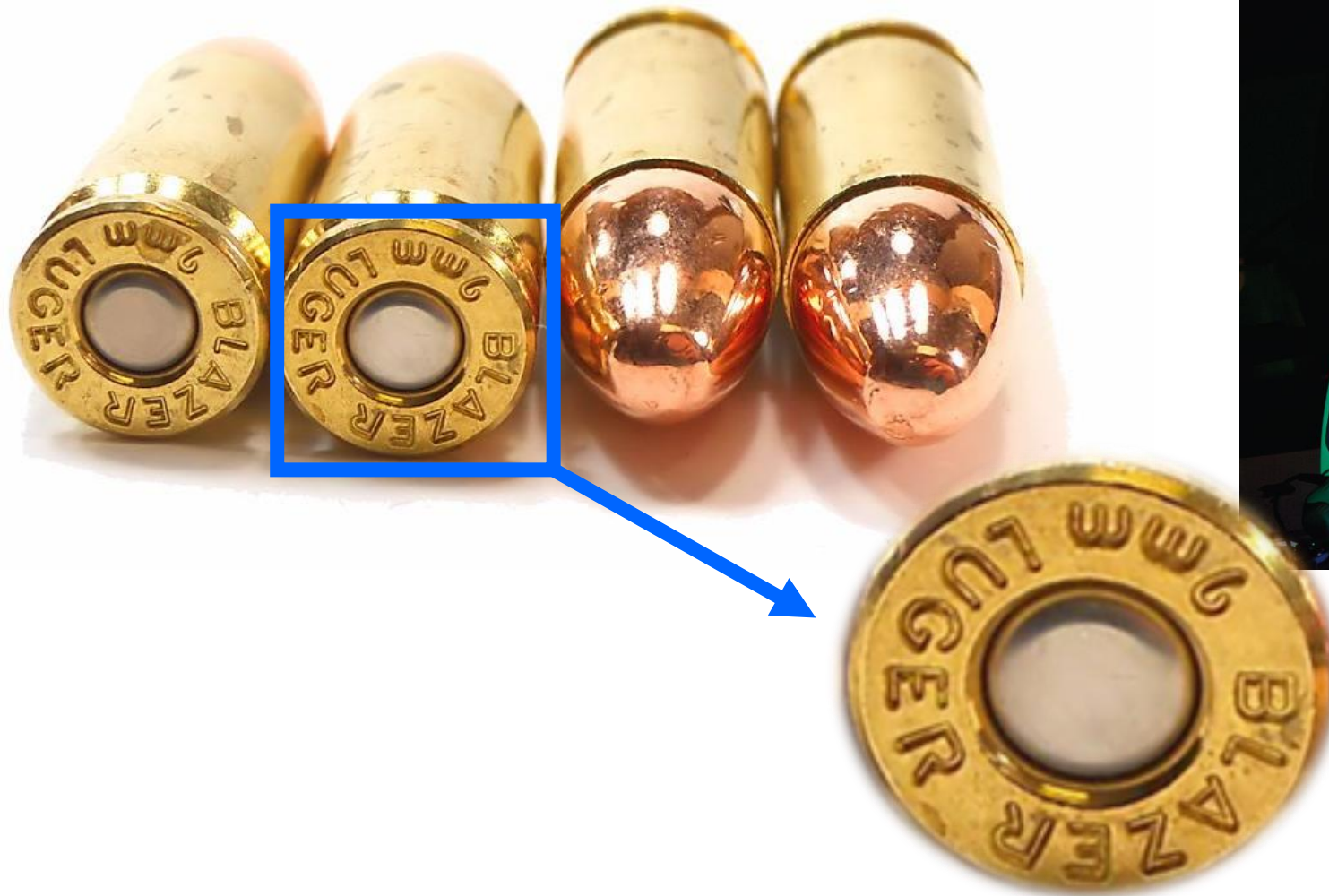
The weapon **loading**

The **shot**

The cartridge **ejection**



Cartridge Analysis



DEMO - 3D Forensic Ballistics Comparison



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



Università
di Catania

University of
Hertfordshire **UH**



iCtLab
Digital Forensics



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

Luca Guarnera¹, Oliver Giudice², Salvatore Livatino³, Antonino Barbaro Paratore⁴, Angelo Salici⁵, Sebastiano Battiato¹

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

² Applied Research Team, IT Department, Banca d'Alitalia, Rome, Italy - giudice@dmf.unict.it

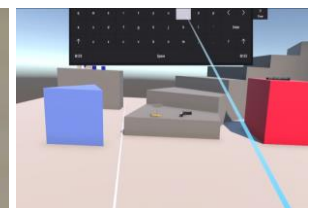
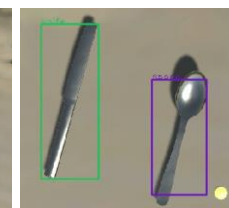
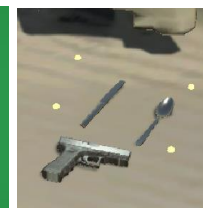
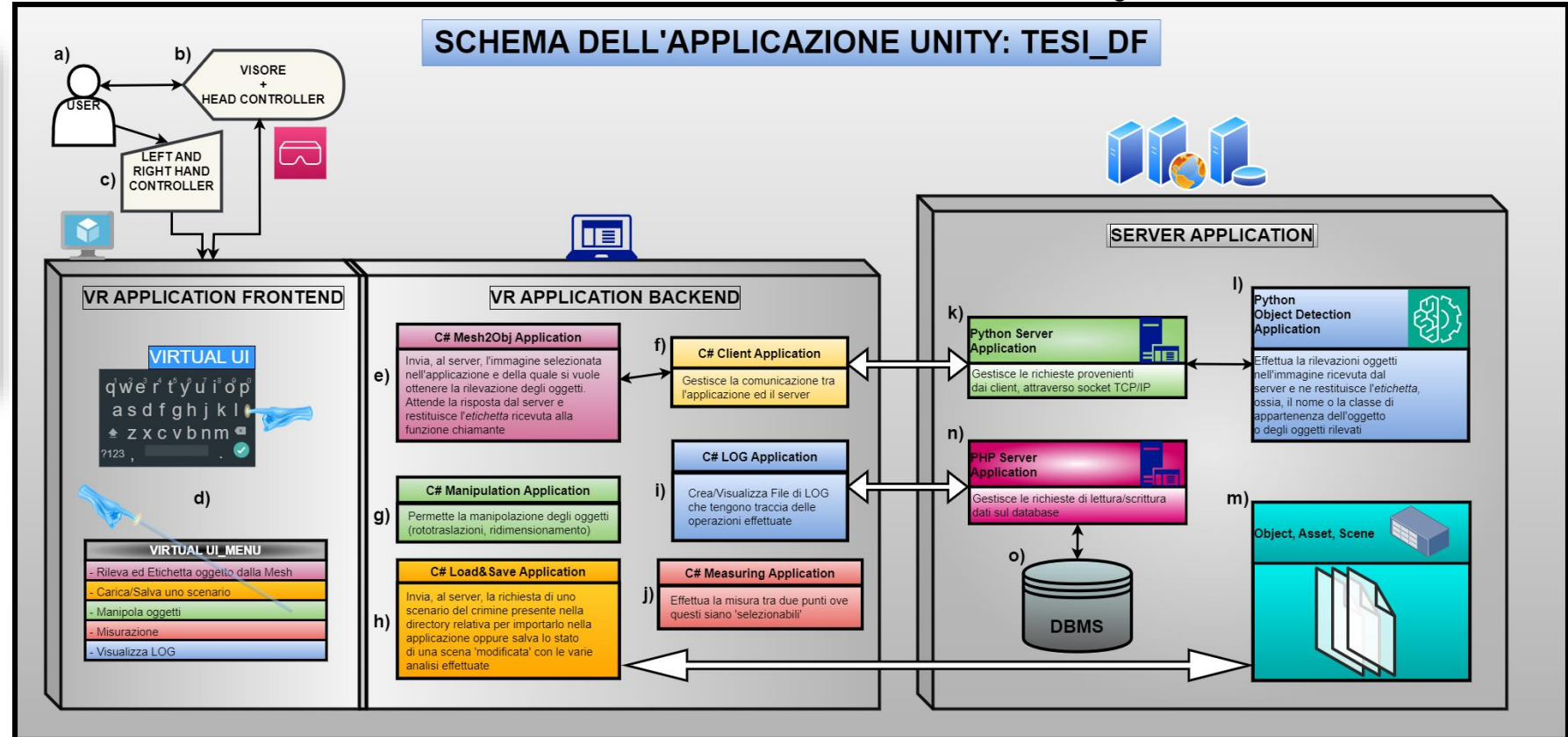
³ School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK - s.livatino@herts.ac.uk

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 D'Arrigo, 5, Messina, 98122, Italy - angelo.salici@carabinieri.it



VR Application with Deep Learning Techniques for Forensic Crime Scene Analysis



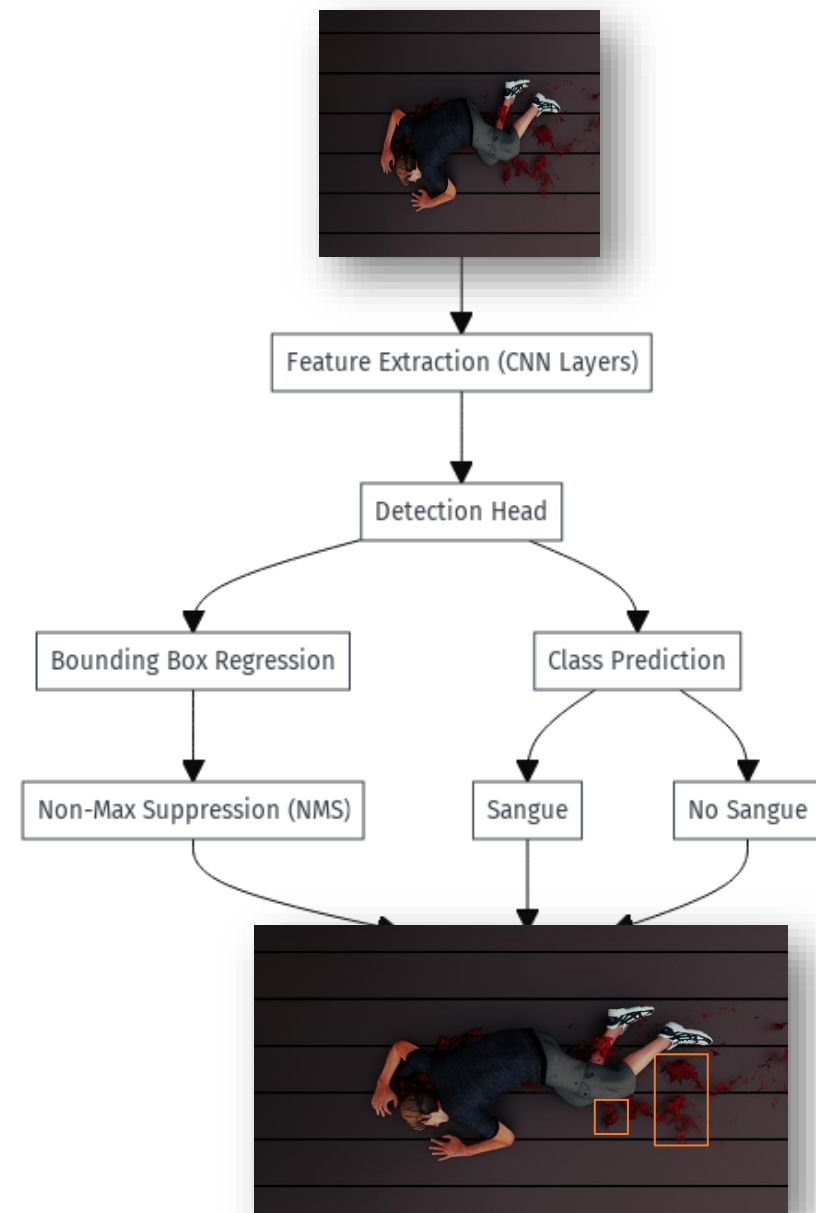
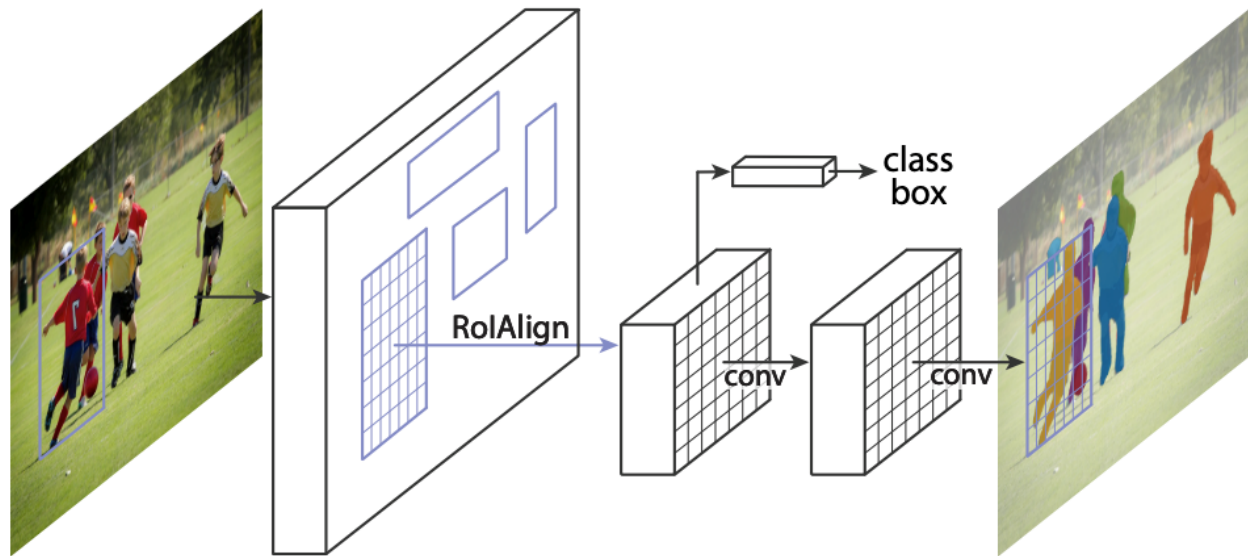
Manipolazione degli
oggetti

Rilevamento ed etichettatura automatica
e/o manuale degli oggetti

Blood Pattern Analysis



Yolov9 Architecture





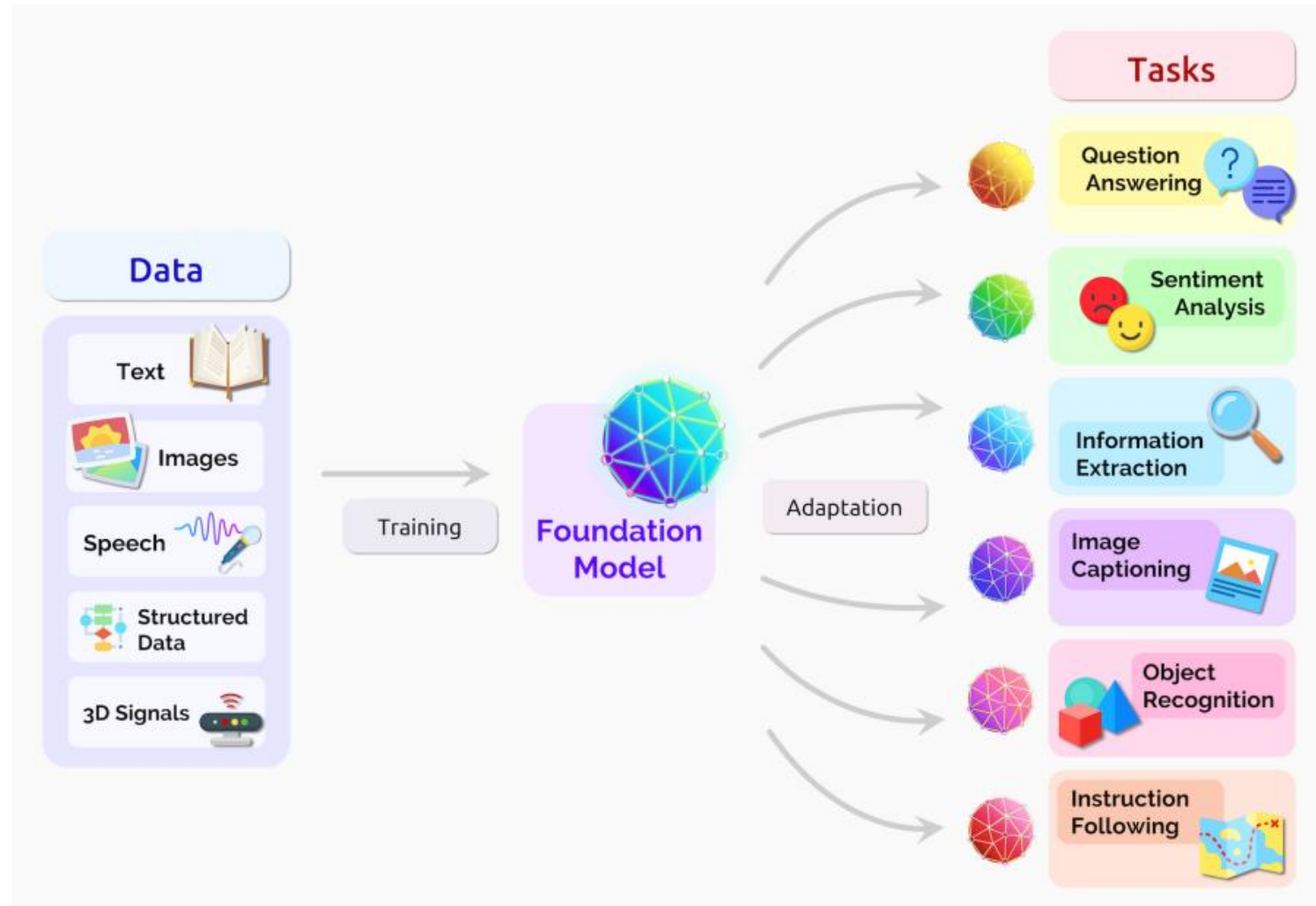
DETECT BLOOD

Detect

Multimedia based on AI



Foundation Models





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 ■



GenAI: Image

“ A cat and a female character in a spaceship exploring a hidden galaxy. With detailed backgrounds, expressive characters, including magical elements, illustration made by hand. ”



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

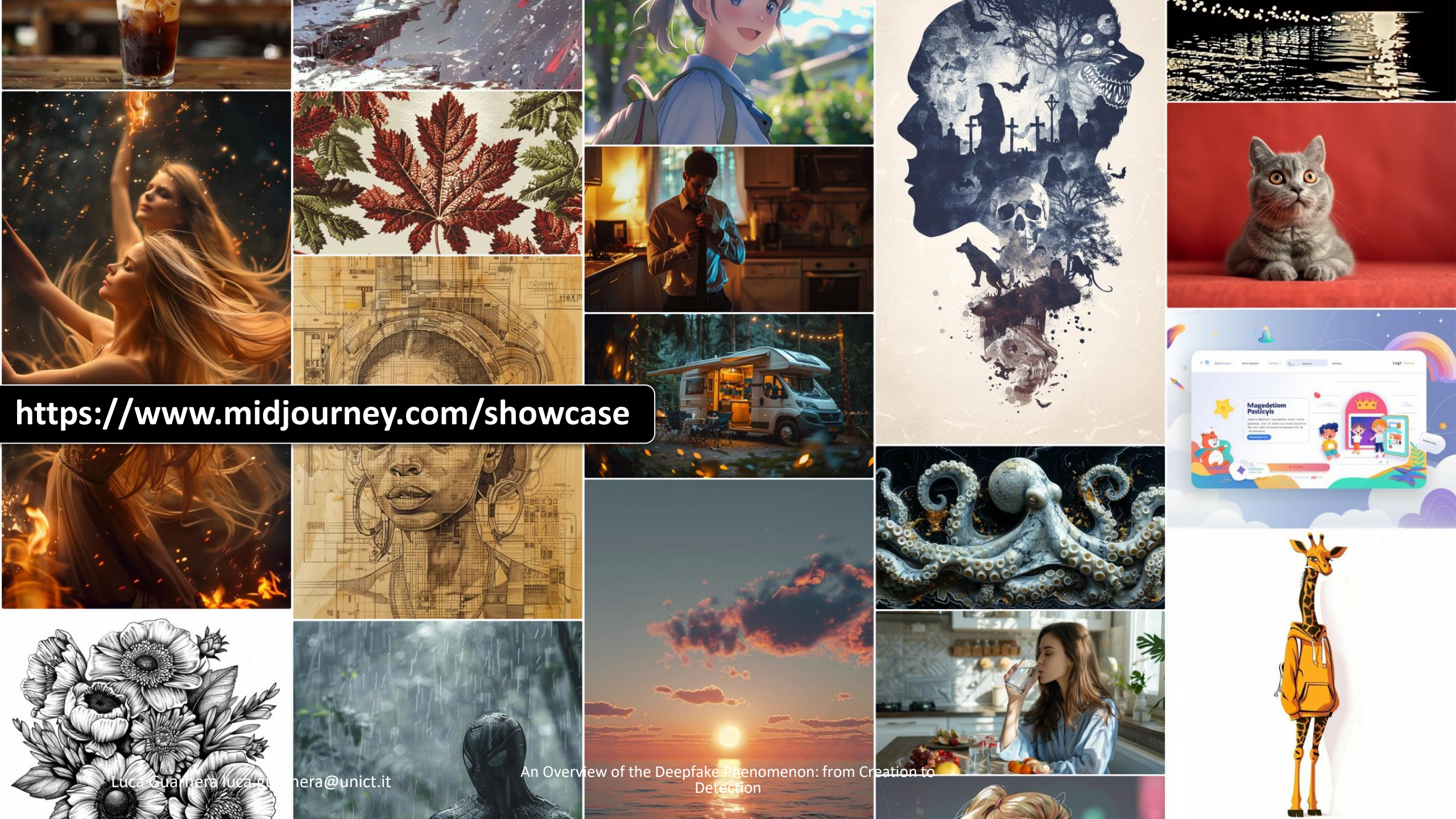


Midjourney v5

March, 2023

GenAI:Image





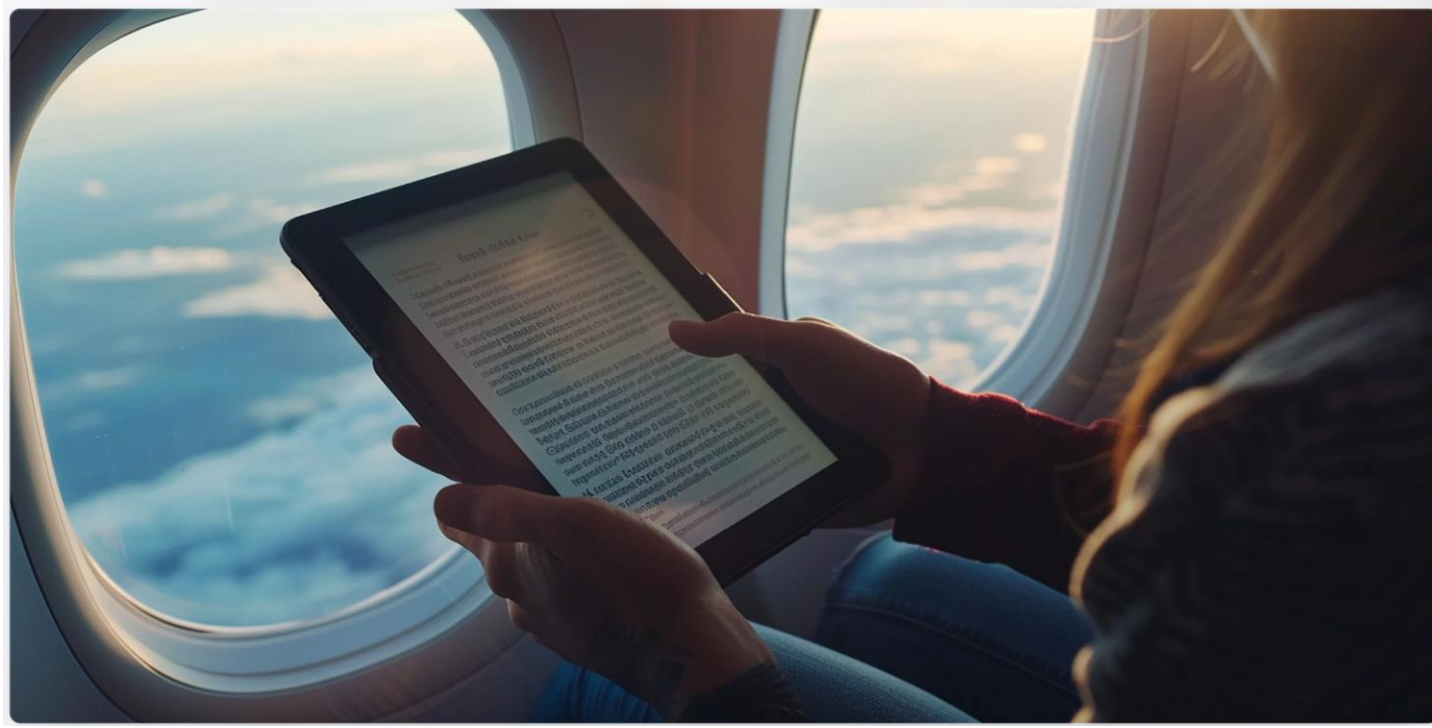
<https://www.midjourney.com/showcase>

Luca Guarniera luca.guarniera@unict.it

An Overview of the Deepfake Phenomenon: from Creation to Detection



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 woolly mammoths approach treading through a snowy meadow, their long woolly 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 its 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 **Transformers**. 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/>





Hello. Matthew put together this short video in our AI Studios to demonstrate the various conversational and non conversational cases DeepBrain AI supports. We work across industries including retail, healthcare, finance and more. Our goal is to use AI powered interactions to greatly improve customer engagement and customer experience. For example.

© 0.5 360 / 400

[Having trouble?](#)

AI Models



Language

English

Chinese

Korean

Japanese



paris_new



Daniel



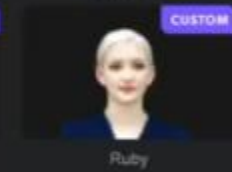
CUSTOM



Jonathan2



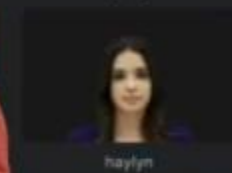
CUSTOM



Ruby



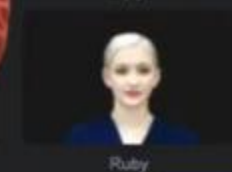
CUSTOM



haylyn



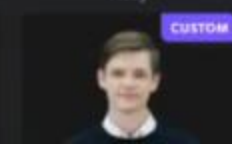
Christina



Ruby



CUSTOM



CUSTOM

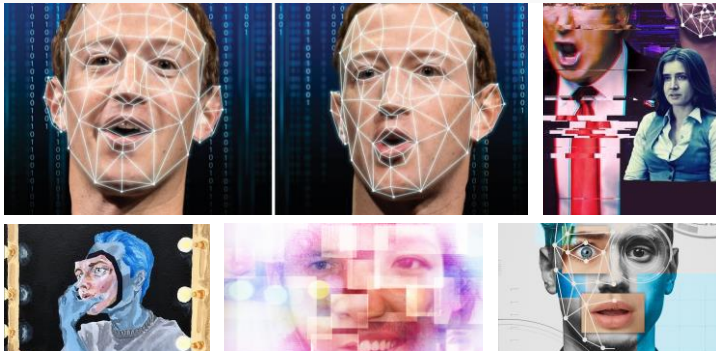


What are Deepfakes?



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 [...]”.





An Overview of the Deepfake Phenomenon: from Creation to Detection

Luca Guarnera luca.guarnera@unict.it



thispersondoesnotexist.com

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).

Source Actor



Real-time Reenactment



Reenactment Result



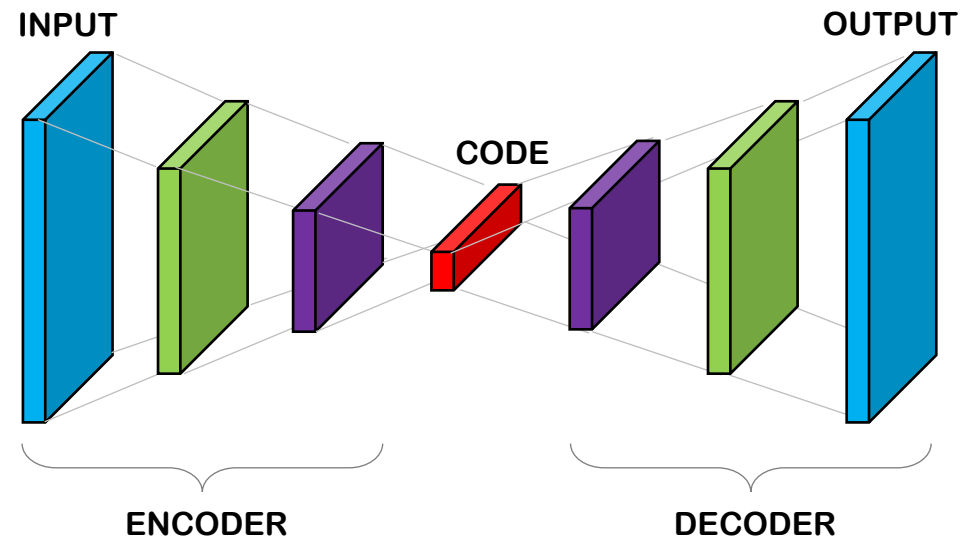
Target Actor

Impressive Applications of Generative Models

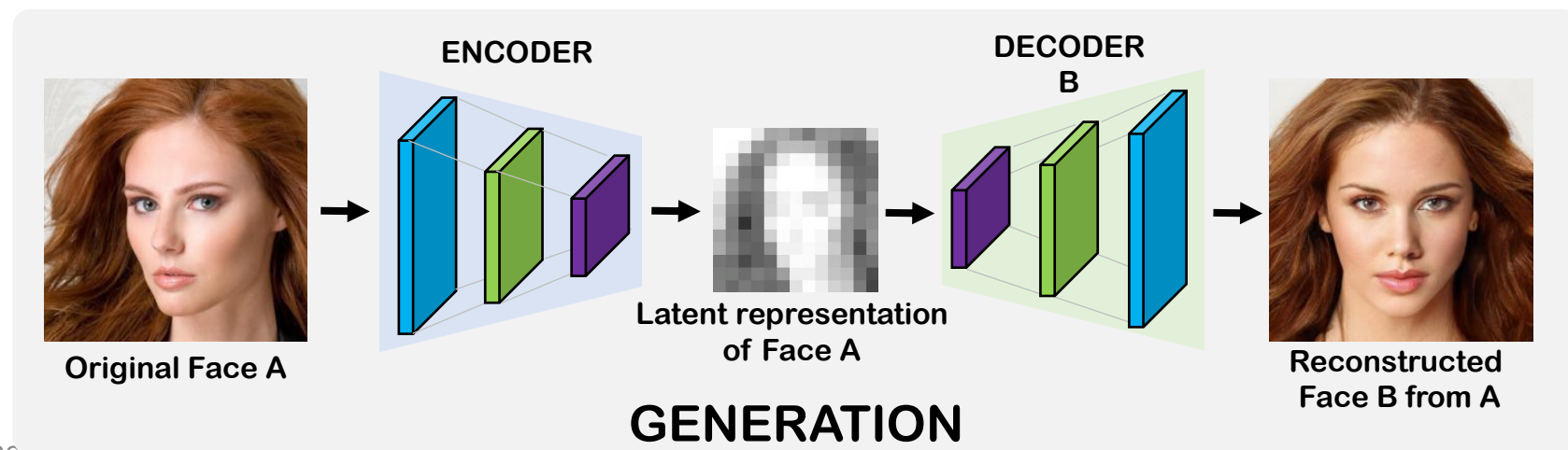
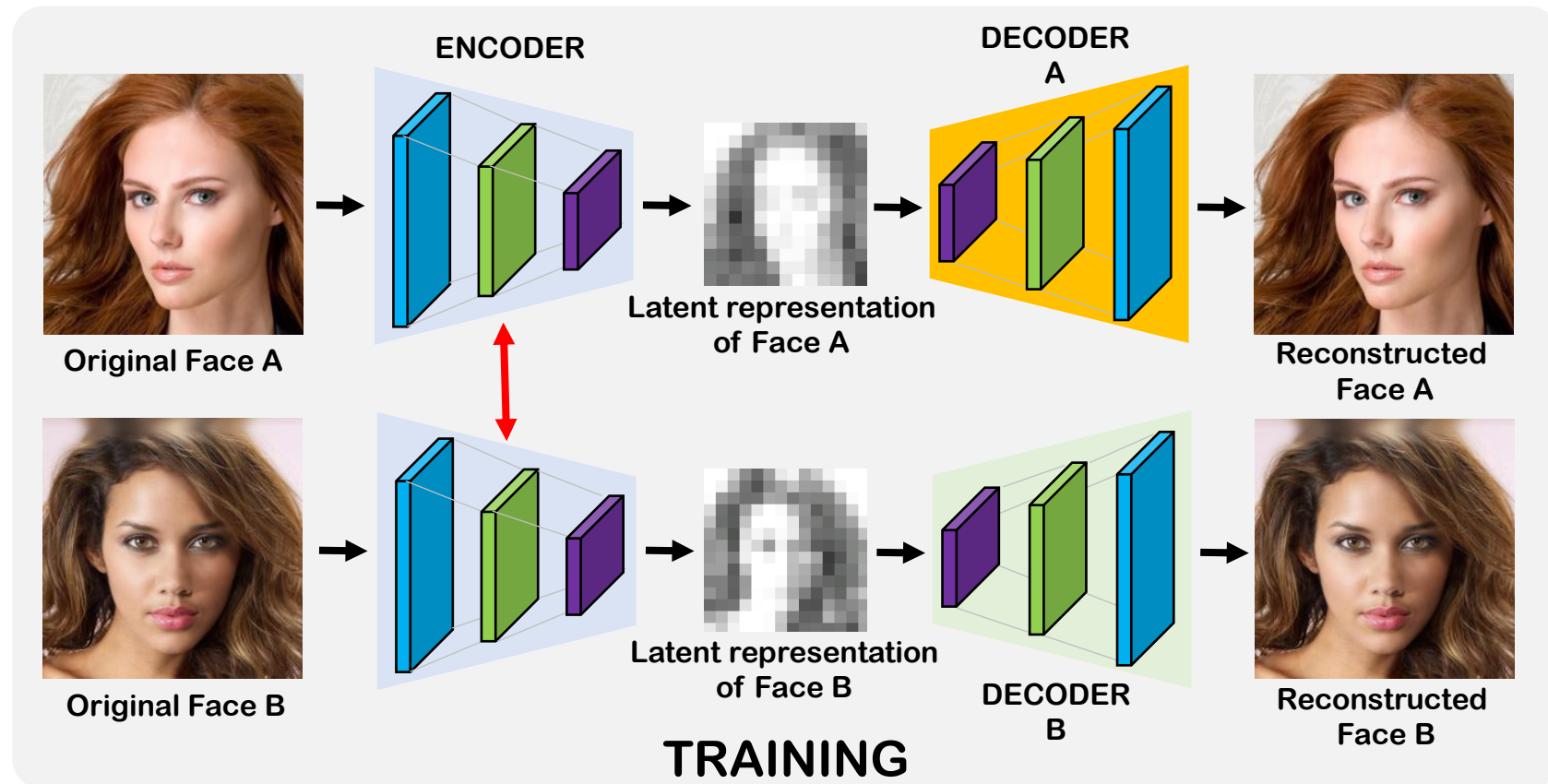
- Generate Human Faces
- Generate Cartoon Characters
- Image-to-Image Translation
- Text-to-Image Translation
- Semantic-Image-to-Photo Translation
- Generate New Human Poses
- Face Aging
- Super Resolution
- Photo Inpainting
- ...



Autoencoder



Deep autoencoders



2005: SCLGen
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

John J. Goodfellow,¹ Jean-Pierre Arnaud, Mehdi Mirza, Bing Xu, David Warde-Farley,
Sherjil Ozair, Aaron Courville, Yoshua Bengio
¹Département d'Informatique et de Sciences de l'opinion
Université de Montréal
Montréal, QC, H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G . The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions G and D , a simple variational autoencoder can be trained with backpropagation. There is no need for any Markov chains or model-specific approximate inference networks during training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

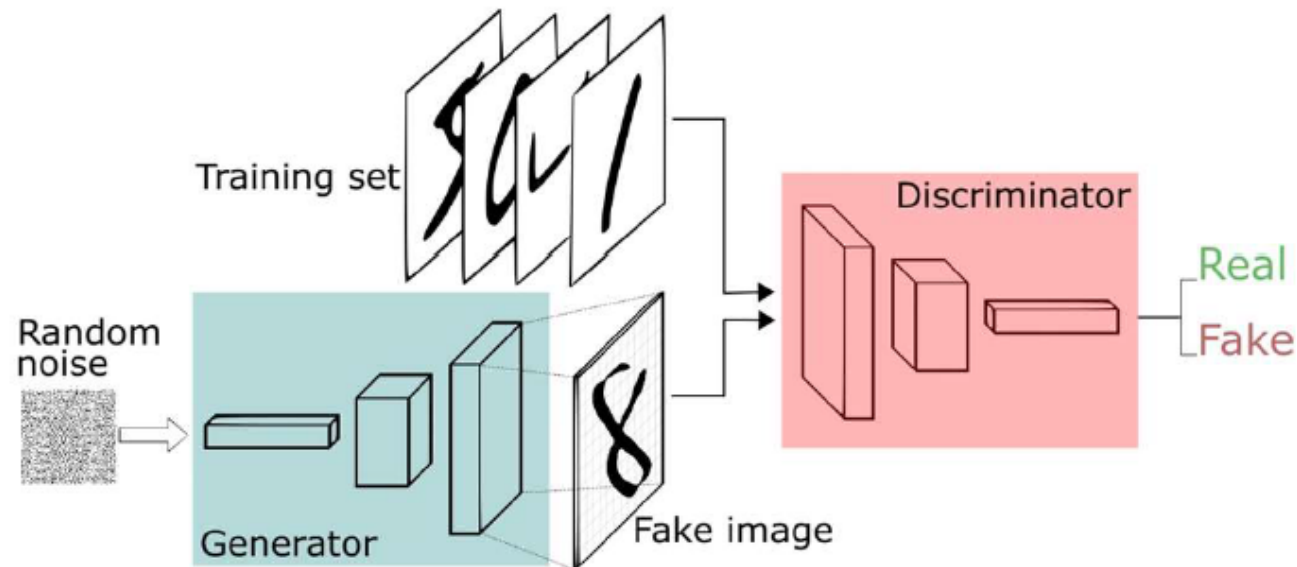
1 Introduction

The premise of deep learning is to discover rich, hierarchical models [2] that represent probability distributions over the kinds of data encountered in artificial intelligence applications, such as natural images, audio waveforms containing speech, and speech in natural language corpora. So far, the most striking success in deep learning have involved discriminative models, usually those that map a high-dimensional, rich sensory input to a class label [14, 20]. These striking successes have primarily been based on the backpropagation and dropout algorithms, using powerful linear units [17, 8, 9] which have a particularly well-behaved gradient. Deep generative models have had less of an impact, due to the difficulty of approximating many forms of probabilistic computation that arise in maximum likelihood estimation and related strategies, and due to difficulty of leveraging the benefits of generative models for the generative context. We propose a new generative model estimation procedure that addresses these difficulties.¹

In the proposed adversarial network, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as an imitator or a counterfeiter, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles.

¹For conditions to learn a generative model of images, but this does not require a GAN model.
²John J. Goodfellow is now a research scientist at Google, but this work was done as a PhD student.
³David Warde-Farley is a research scientist at Microsoft, but this work was done as a PhD student.
⁴Yoshua Bengio is a CIFAR Senior Fellow.

⁵All code and supplementary materials available at <https://github.com/jgoodfellow/gan>.

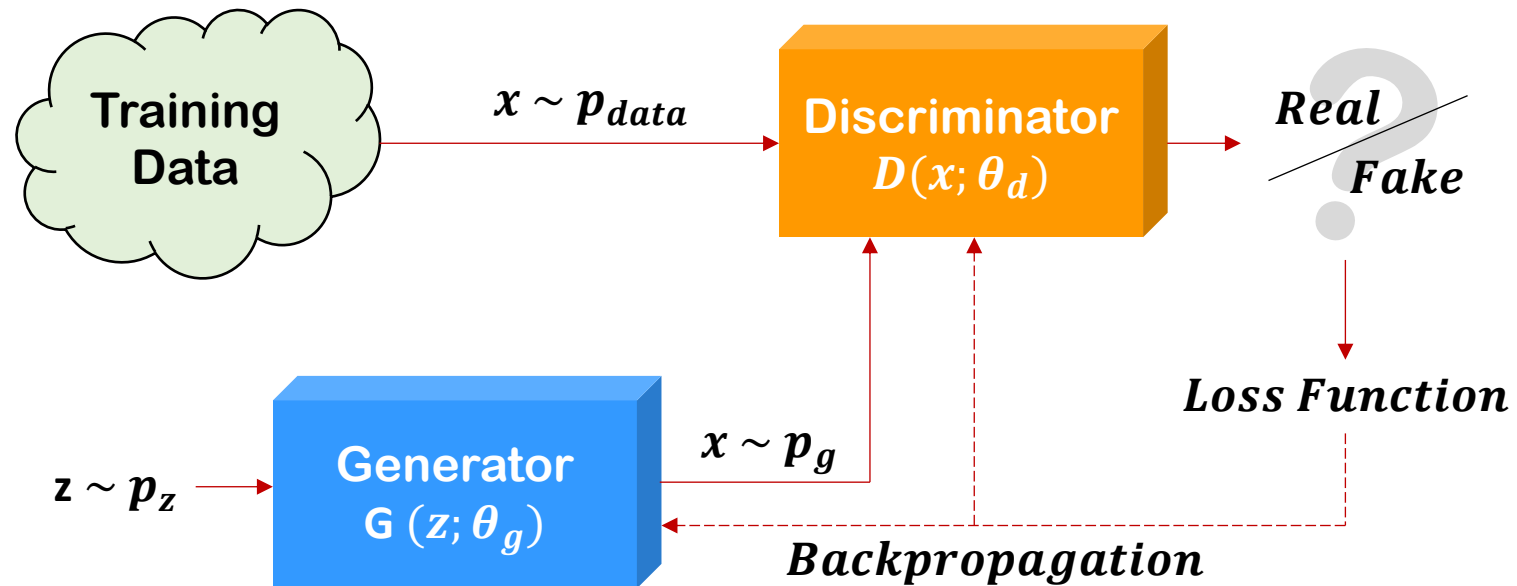


<https://towardsdatascience.com/generative-adversarial-networks-explained-34472718707a>

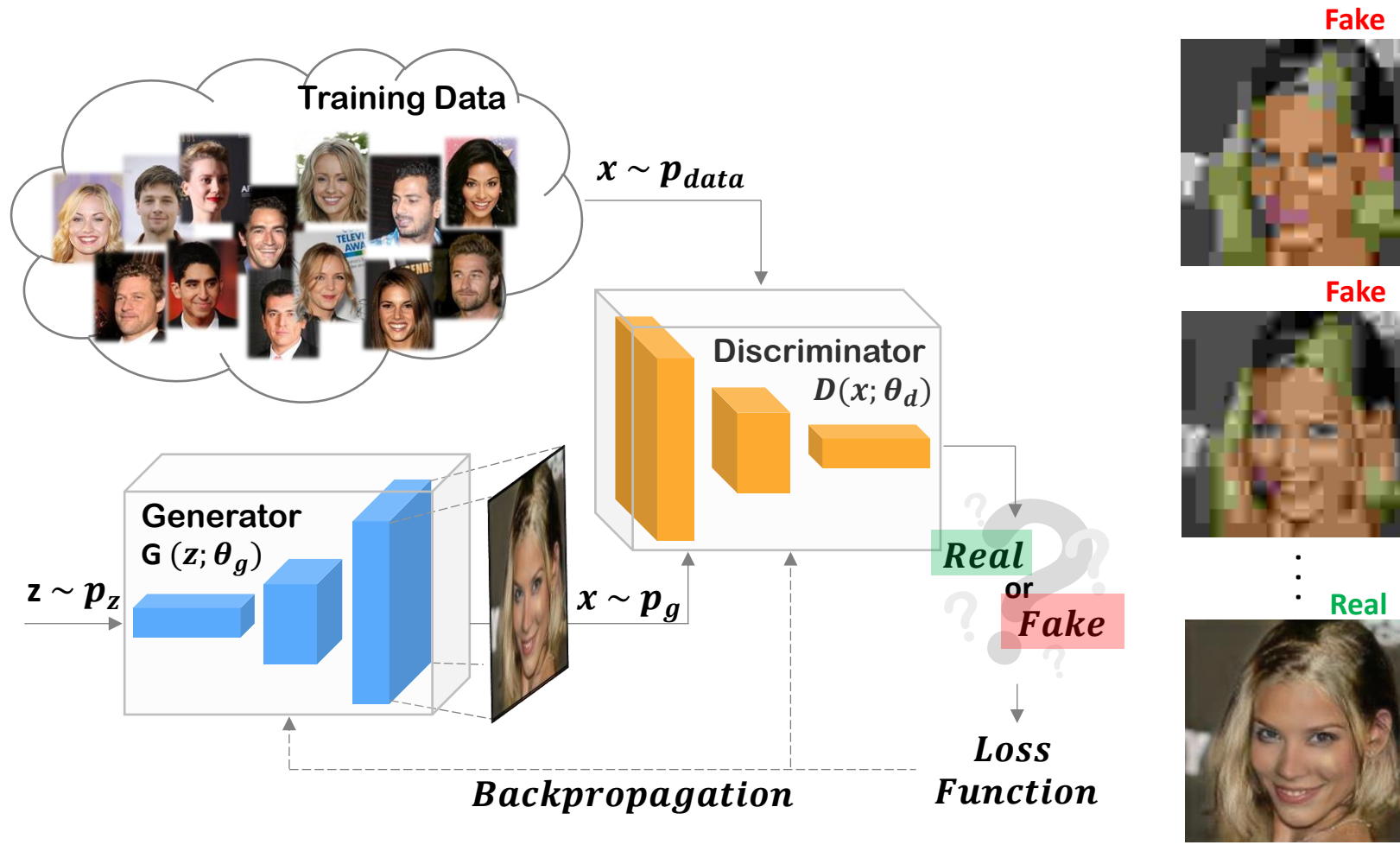
Luca Guarnera luca.guarnera@unict.it

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

Generative Adversarial Networks

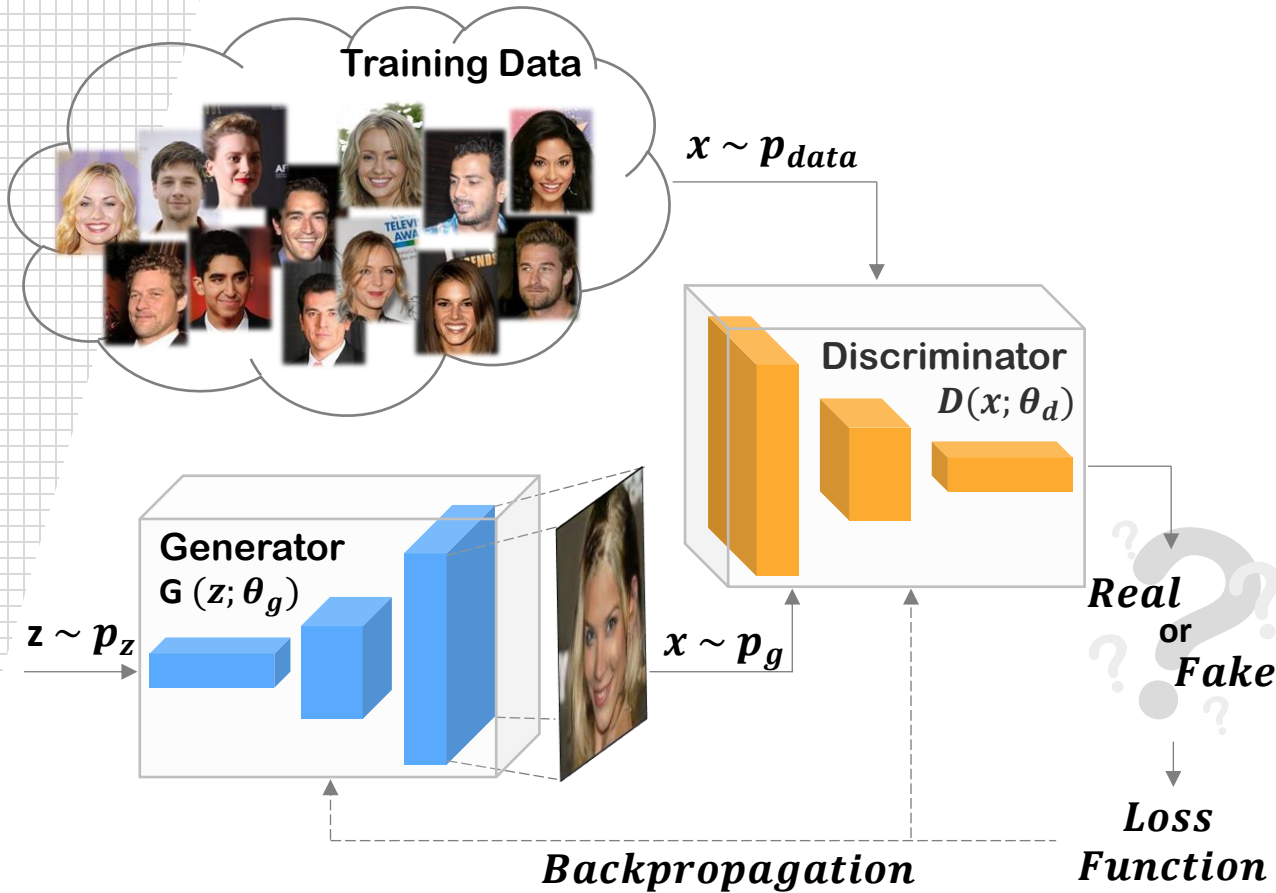


GAN example for human faces generation



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

GAN example for human faces generation



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



Image source: Salehi, Pegah, Abdollah 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 Learning 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

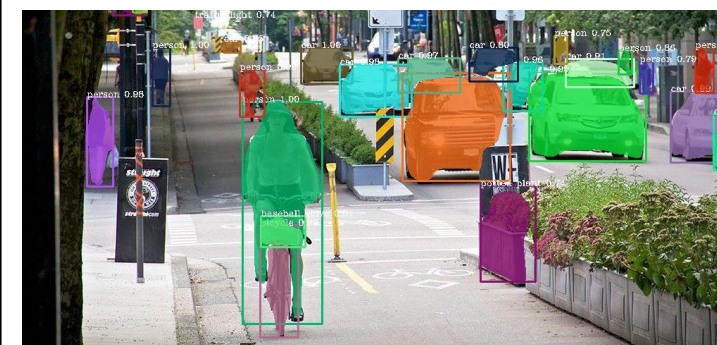
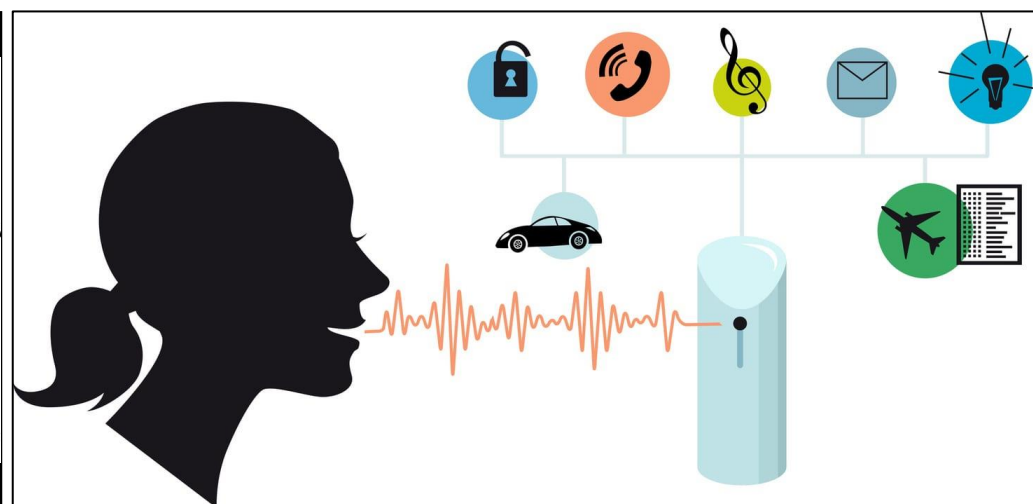
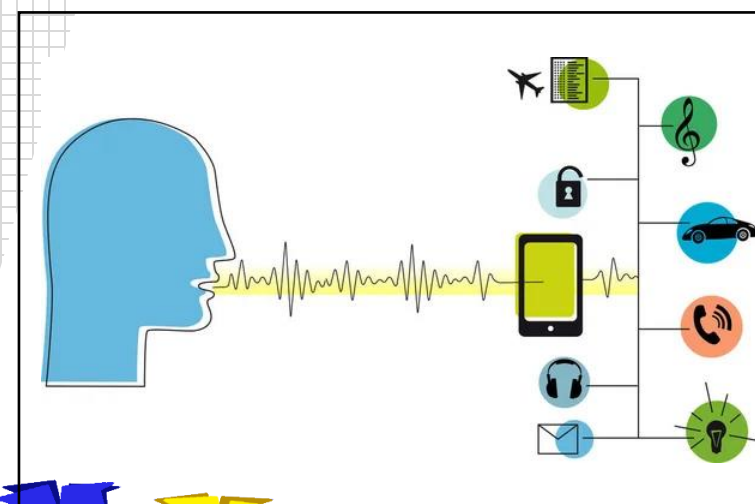
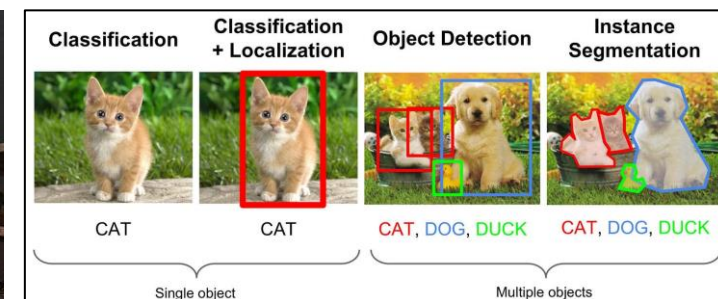
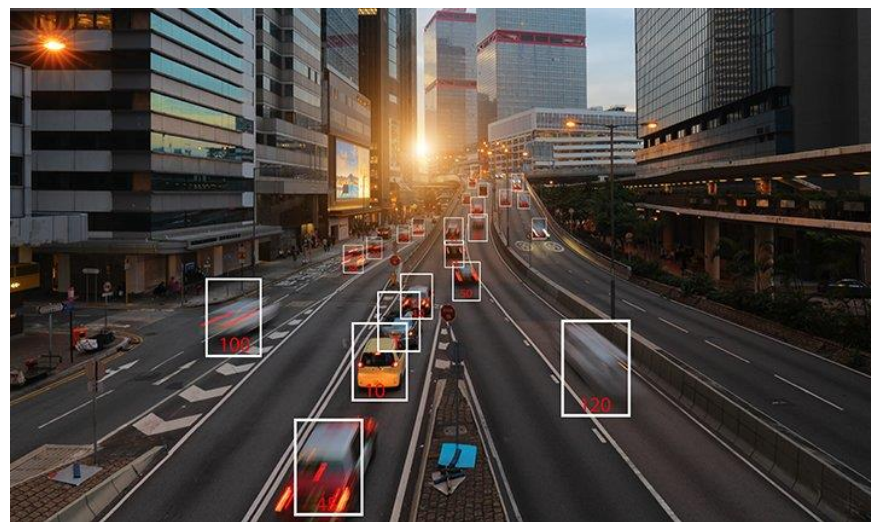
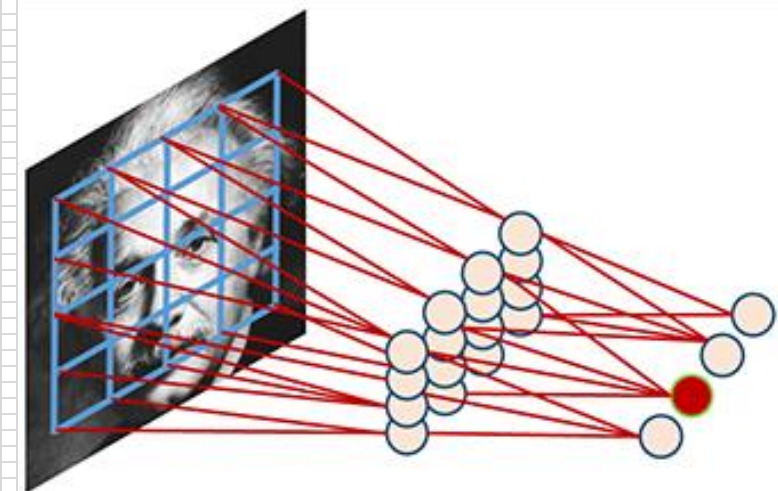
Depending on the application to be made, it is preferable that the student also knows

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

Multimedia Data Data Modelling



Applications



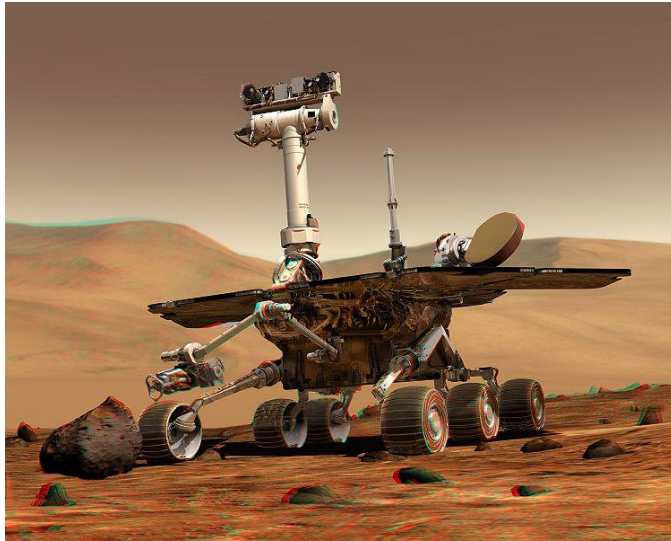
What is Computer Vision?

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.

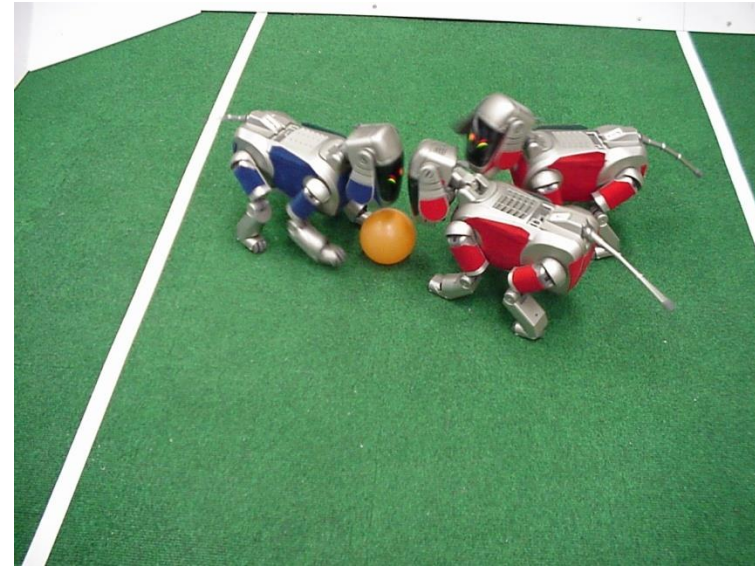


CV Applications: Mobile robots

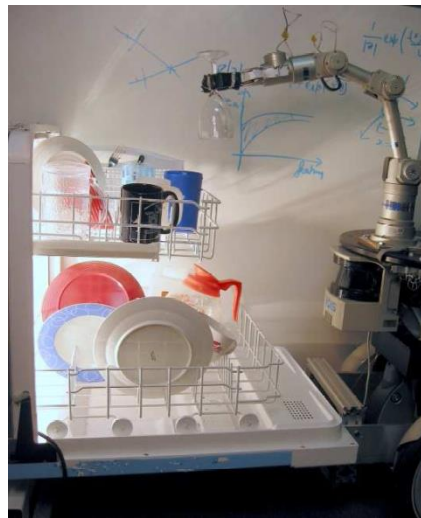


NASA's Mars Spirit Rover

http://en.wikipedia.org/wiki/Spirit_rover

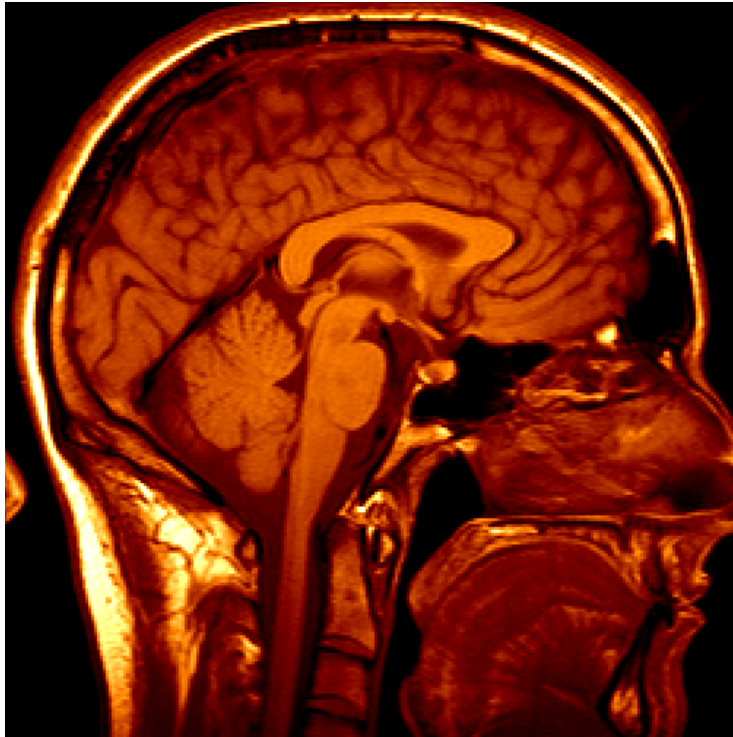


<http://www.robocup.org/>



Saxena et al. 2008
STAIR at Stanford

CV Applications: Medical imaging

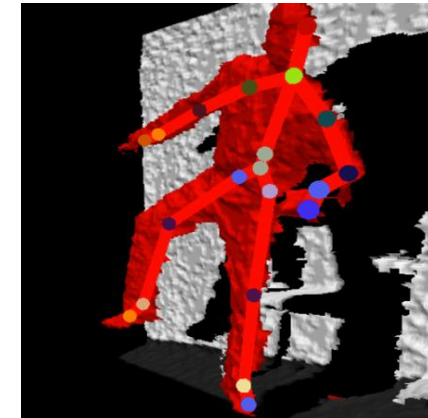
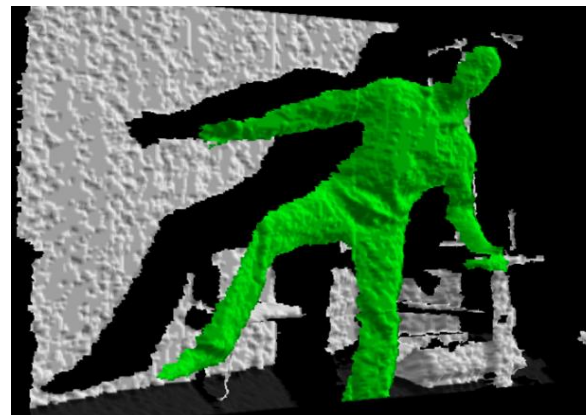
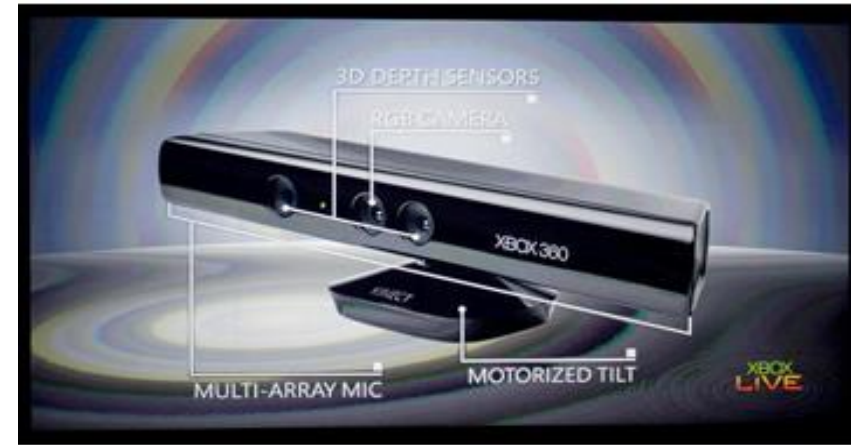


3D imaging
MRI, CT



Image guided surgery
Grimson et al., MIT

Vision-based interaction: Xbox Kinect



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

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



Università
di Catania

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

