



chist-era



CHIST-ERA Projects Seminar 2021

Topic XAI

starting 1st of May, 2021.

XAI face

Measuring and Improving Explainability for AI-based Face Recognition

Jean-Luc DUGELAY (EURECOM, Sophia Antipolis)

April 13, 2021



Programme co-funded by the
EUROPEAN UNION

❖ FRANCE: EURECOM

Jean-Luc Dugelay & Chiara GALDI

Expertise includes: Face recognition



❖ AUSTRIA: JOANNEUM RESEARCH

Martin Winter & Wemer Bailer

Expertise includes: Machine Learning



❖ PORTUGAL: Instituto de Telecomunicações

Fernando Pereira & João Ascenso

Expertise includes: AI-based coding



❖ AUSTRIA: U. Vienna

Erich Schweighofer & Jakob Zanol & Felix Schmautzer

Expertise includes: Privacy/ethics



❖ SWITZERLAND: EPFL

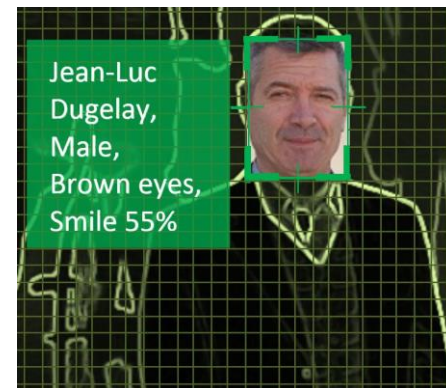
Touradj Ebrahimi

Expertise includes: Performance/evaluation metrics



Companies supporting XAlface

Face Recognition



- **Key technology** in our society, frequently used in multiple applications: access control, video surveillance, social media, ...
 - raises several security and **ethical issues**
- **Trust** is a central issue in (AI-based) Face Recognition (FR)
 - various **weaknesses and biases** depending on datasets, training protocol, network architecture,



- ❖ CNT1: develop clear **legal guidelines** on the use and design of AI-based FR following the privacy-by-design approach;
- ❖ CNT2: disentangling **demographic information** (age, gender, ethnicity) from the overall face representation in order to understand the impact of such traits on FR but also to develop demographic free face recognition;
- ❖ CNT3: address fairness and **non-discrimination** issues by following the idea of de-biasing during the training;
- ❖ CNT4: optimize the **tradeoff** between **interpretability** and **performance**;
- ❖ CNT5: create **tools** that will allow assessment and measurement of performance and explanation of decisions of AI-based FR systems;
- ❖ CNT6: understand how future **AI-based coding** solutions may be different from a recognition explainability point of view.

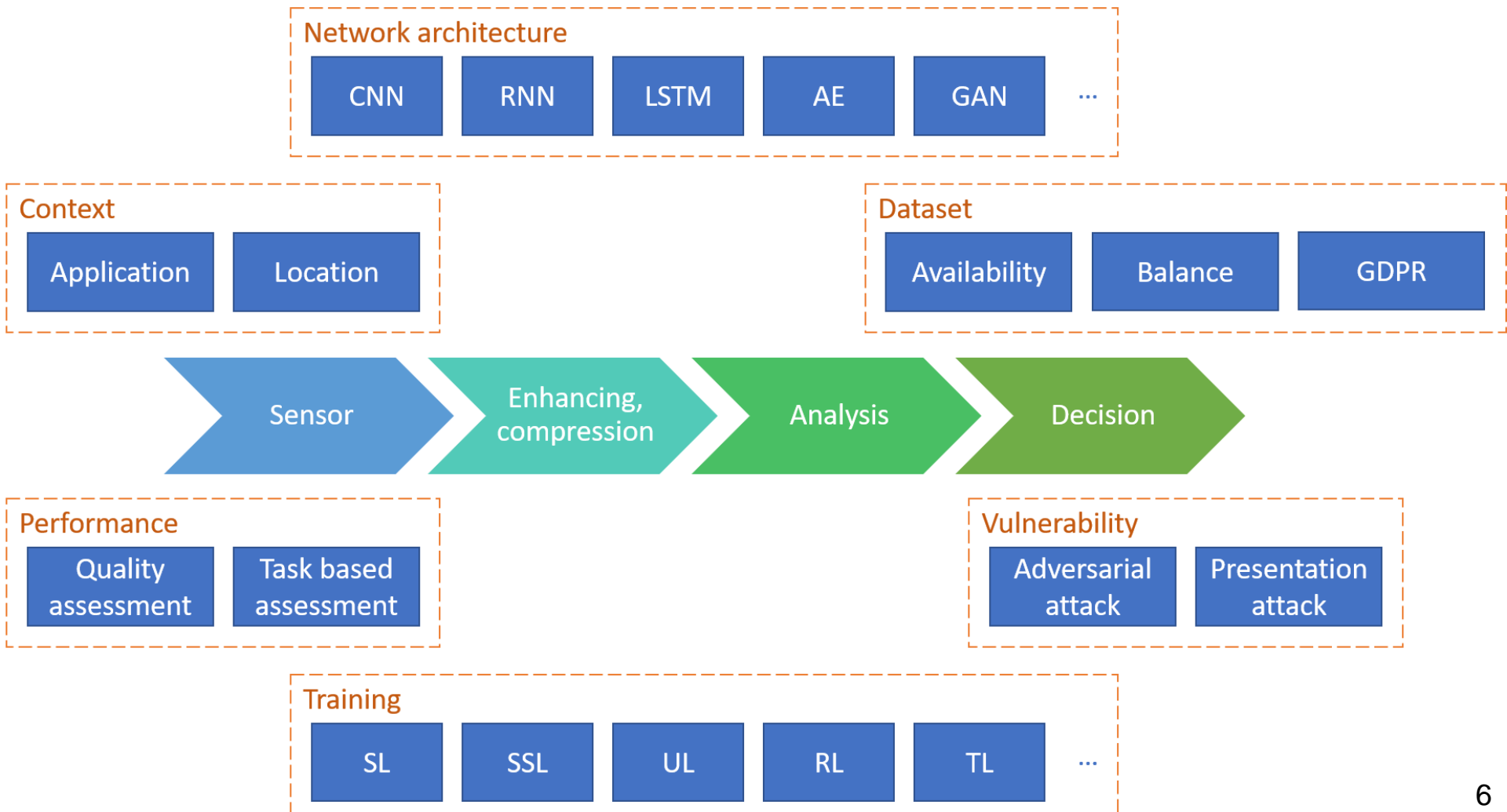
- ❖ OBJ1: To Identify and describe **key influencing factors** that impact performance of FR based on ML.
- ❖ OBJ2: To design **metrics and protocols** that measure the impact of the influencing factors on FR systems based on ML and AI.
- ❖ OBJ3: To quantify and understand the **interaction between influencing factors** and create an explainable FR system based on ML.
- ❖ OBJ4: To address **fairness as well as legal and ethical** obligations for FR systems based on ML and AI and develop a systematic approach to ensure fairness and trustworthiness, informational self determination and acceptance of the persons concerned/data subjects.
- ❖ OBJ5: **To disseminate** the results of the project to the scientific community, the industry, as well as to a larger audience.

XAI face

Measuring and Improving Explainability for AI-based Face Recognition

Approach and research method

to break down the problem into the analysis of single or small groups of factors at a time



WP1: Management

WP2: Influencing factors, performance metrics, and protocols

WP3: Acquisition, databases, and data protection

WP4: Explainable approach design and implementation

WP5: End-to-end face recognition system

WP6:
Dissemination
and
exploitation



- **Deeper understanding** of the AI based FR mechanisms and capability to express them using well identified key influencing factors; this would represent the end of the ‘black box’ negative stigma;
- Availability of **new FR solutions beyond the SoA** where AI explainability is a key feature, notably allowing to optimize the trade-off between interpretability/explainability and recognition performance;
- Availability of **non gender, age, or ethnicity biased FR solutions**, able to disentangle this information from the overall face representation, thus offering truly inclusive technical approaches;
- Availability of **novel, efficient metrics and protocols** for explainability assessment of AI-based FR systems to be made publicly available for researchers and end users; this will also contribute to the definition of standards for AI explainability assessment;



- **Increased general trust and social acceptance** of AI-based FR by providing effective instruments to explain the decisions made by such systems, notably explanatory videos and papers appropriate for the general user;
- Increased fulfilment of the objectives outlined in the GDPR notably regarding the “**right to explanation**” in order the legal obligations for providing explanations to AI decisions may be fulfilled;
- Availability of **solid technical literature** on influencing factors, assessment metrics and protocols, and FR systems driven by AI explainability;
- Strengthened **training of graduate students** for their future careers by using the results of the project in several teaching programs in European universities.



chist-era

XAI face

Measuring and Improving Explainability for AI-based Face Recognition

Thank you for your attention

jean-luc.dugelay@eurecom.fr / @jld2eurecom 