www.perseo.eu

TRAINING EVENTS

NEWSLETTER
#2 December 2022

@PerseoProject

ESR POST **BLOG**

NIYATI RAWAL - ESR7
ESTEVE VALLS MASCARO' -ESR6
FRANCESCO VIGNI - ESR12

ESR PUBLICATIONS

SUBSCRIBE

PERSEO project

PERSEO

European Training Network on PErsonalized Robotics as SErvice Oriented applications

The **European Training Network on PErsonalized Robotics as SErvice Oriented applications – PERSEO –** aims at training and creating a new generation of interdisciplinary researchers and professionals with an entrepreneurship ethos and transferrable skillset to face both societal and technological research challenges of this **forthcoming market of personal robots**. These consist respectively in the development of personalized and acceptable robot's perceptual and interaction capabilities to be distributed as **Cloud Services** personal robot applications as companions, in rehabilitation, and for edutainment. The PERSEO research and training program is organized into **three Research Themes** aimed at investigating personalization of robot capabilities at different levels of possible Human-Robot Interaction (HRI), namely "**Physical**, "**Cognitive**", and "**Social**". This requires a set of research skills ranging from computer science and AI to automation, ethics, and psychology, and the use of a collaborative design approach. **Integration Milestones**, will provide co-working societal problems where collaborative design/implementation is fostered and supported by partner organizations.

2ND RESIDENTIAL WORKSHOP

PAL Robotics in Barcelona,

21-22-23 June 2022

PERSEO
European Training Network on
PErsonalized Robotics as SErvice Oriented applications

The workshop organized a public special focus webinar on the robots addressing problems connected to the Integration Scenarios for Market & Society on a platform working in groups

During the 2nd Residential Meeting, the ESRs were engaged in a workshop in which the individual projects were presented in the form of position papers, linking state of the art and preliminary results. Position papers were peer-reviewed and orally discussed by all workshop participants, including an external advisory group that provided fruitful feedback to each ESR on the directions of research and potential methodological and theoretical aspects to consider.

Workshop and Seminars

Designing for End-Users
by Cristian Leorin

Presentation of the scenario and hands-on session
by Severin Lemaignan

ARI tutorial - Entrepreneurship in the market of personal robots
BY HRI TEAM, PAL ROBOTICS

REMINDER
PROJECT SUMMARY

PERSEO is a Marie Sklodowska Curie Innovative Training Network

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955778

www.perseo.eu

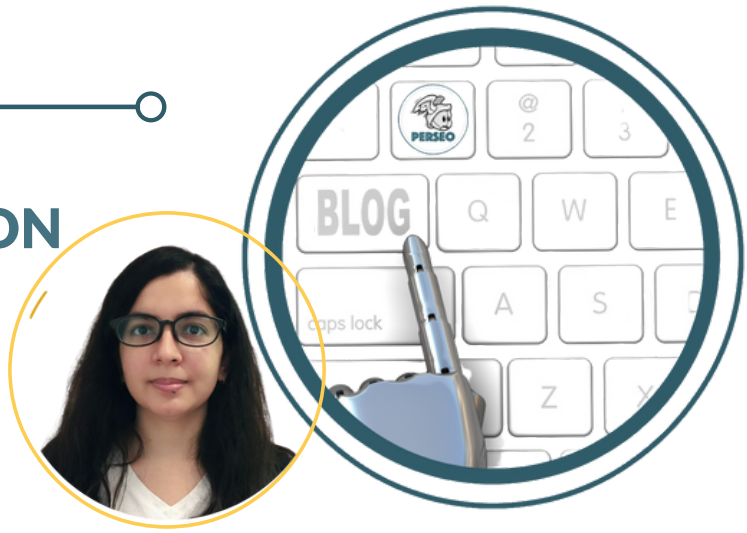
2ND RESIDENTIAL WORKSHOP

PAL Robotics in Barcelona,

21-22-23 June 2022

ESR POST BLOG

BRING ME THERE: AN INSTRUCTION GENERATION APPROACH FOR VISION AND LANGUAGE NAVIGATION



NIYATI RAWAL/ ESR7

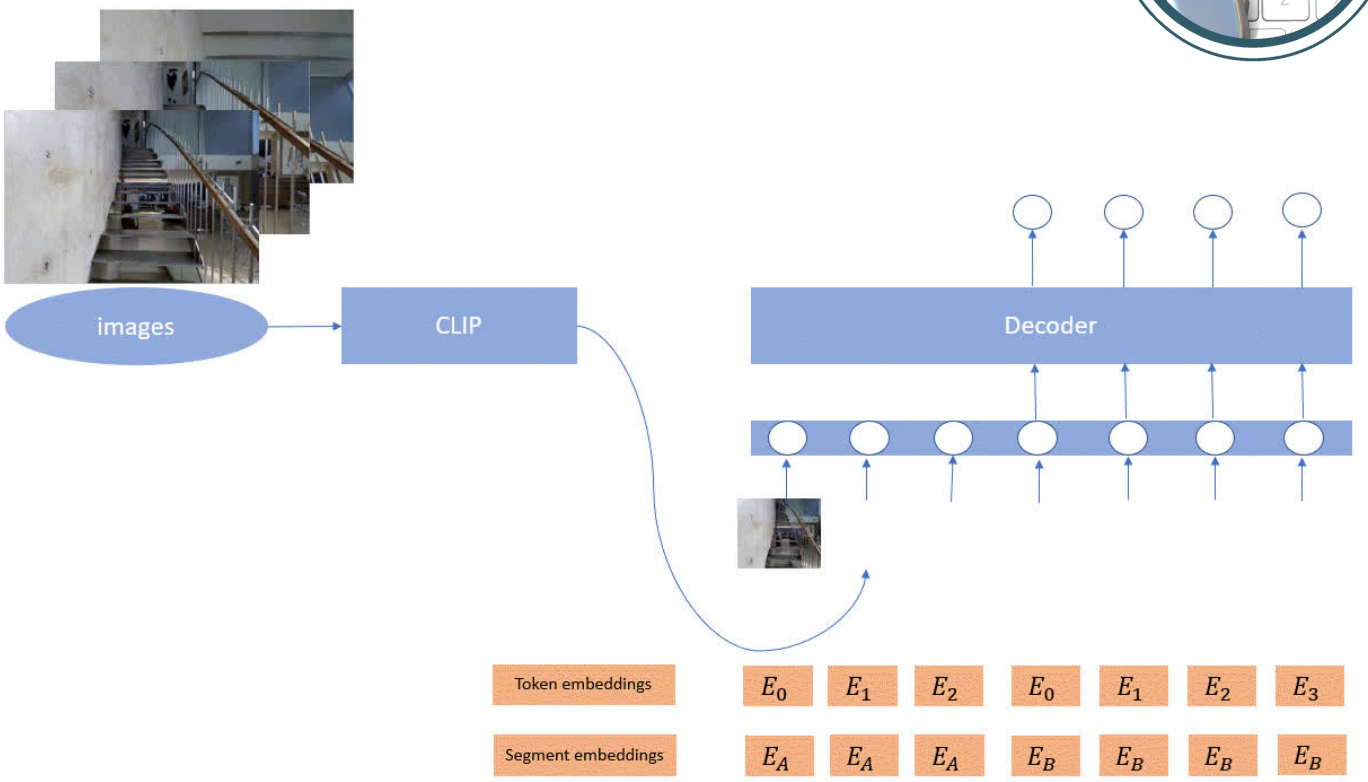
While many researchers study computer vision, natural language processing, or robotics, we are mainly interested in the intersection of these three domains. Specifically, we are interested in Vision and Language Navigation (VLN). For an agent or a robot to perform VLN, it can perceive the 360-degree view of the environment and is given human instructions in the form “Take a right, going past the kitchen into the hallway”. The goal for the agent is to follow human instruction and navigate in a previously unknown environment.

Recently, the research in VLN has been growing. Typically, VLN involves an agent that follows human instructions and navigates in a previously unknown environment. In some of the recent works, there is a focus on generating better quality instructions that could lead to better conversational agents. Moreover, using a combination of synthetic instructions along with human instructions could also improve navigation performance. In my first project, we propose an architecture inspired by a Generative Pre-Trained Transformer (GPT). The model generates synthetic instructions of the path the agent has traversed and consists of a transformer decoder that generates sentences for a sequence of images from the environment describing the agent’s path to a particular point in time.

While the original GPT model has been implemented for NLP applications of text generation or text summarization, the model that we propose describes the actions the agent has to take in an environment until it reaches the target location, given the sequence of images of the traversed path. The overall approach is as follows. First, we take the images from the environment the agent has traversed. These images are fed into a trained vision encoder CLIP to extract the features from the images which are then fed into a GPT-2 decoder model along with the first Begin of String (BoS) token. The GPT-2 decoder predicts the subsequent language tokens as an output and is thereby, able to generate a complete language instruction to describe the actions the agent has taken to reach the final location in the last image. Unlike in GPT, where the input of the model is only the text up to the previous time step, in our case the input of the model consists of both images and text. Inspired by BERT, we use segment embeddings and position embeddings in addition to token embeddings to effectively segregate image and textual information.

MEET NIYATI

CLICK HERE



The proposed model generates high quality synthetic instructions that could potentially improve the ability of an agent to navigate in a previously unknown environment. The model could also be extended to other applications. For example, when the human asks an agent a question where an object is, it can provide the human with instructions on how they could reach it. This would enable humans and robots to collaborate and interact with each other in the same environment. The robot could also help a human navigate in an indoor environment when the human is unable to figure out the way. Or the other way round, the human could guide the robot in case the robot gets stuck in a particular location through continuous interaction between the human and the robot.

Ground truth instruction:

Exit bedroom and head straight toward living area, and wait at entrance.

Generated instruction:

exit the bathroom and walk through the bedroom . then stand near the entrance of the front door .

MEET NIYATI

CLICK HERE

ESR POST BLOG

WHAT DO YOU REALLY WANT? THE ROBOT CHALLENGE OF RECOGNIZING HUMAN INTENTION.

ESTEVE VALLS MASCARO'
ESR 6

Imagine that your friends are visiting you at home for dinner today, and you need to re-organize the space to make room for everyone. However, you are alone and the furniture is too heavy. Lucky for you, you just purchased a robot whose strength you could use for moving tables, sofa, and so on; and you actually see that it knows how to perform those tasks. You then command the robot to corner the sofa on the wall, but you see that the trajectory executed is not the one desired, and neither is the end goal. Therefore, you decide to help carry the sofa to guide the robot to perform the task. “This robot does not take me into account”, you think, while observing that it is still moving the sofa without following your guidance. “Would it be great if the robot could not only execute tasks but understand what I want?”

In this previous example, a typical Human-Robot Collaboration (HRC) situation has emerged that did not satisfy the expectations of the user. Even though the robot knew the task and how to execute the steps, it lacked the understanding of what the human really wanted. To enhance Human-Robot Collaboration, robots need to recognize and predict human intention to adapt accordingly and assist better in the task at hand.

Currently researchers tend to address human intention by answering (i) ‘What are we doing?’, (ii) ‘Where are we?, (iii) ‘What are we interested in?’, (iv) ‘What are we interacting with?’ or (v) ‘What are we saying?’ . The answers to these questions are not trivial at all, and each is alone a long-standing challenge in the research community. However, we claim that human intention is not about dealing with each topic separately, but to understand the underlying dependencies that coexists among them.

Our work then consists of not only answering these different questions from a human-like intuitive perspective but also understanding the relationship between each question and answer to construct a higher-level system, which will lead us to recognize and predict human intention.

Then, the robot will be able to understand what you really want and to match its behavior to your intention. You will be able to corner the sofa by guiding the robot based on your gaze, your movement and your language and finally your friends will be able to come for dinner!

MEET ESTEVE

CLICK HERE

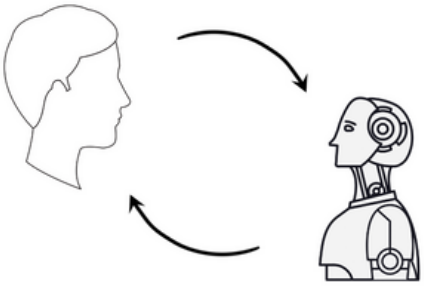
ESR ABSTRACT

IN-DEPTH

ERROR 001: ROBOT MISBEHAVED

FRANCESCO VIGNI
ESR12

Research is the art of improving mankind's knowledge.



As soon as research joins forces with technological progress we are able to envision new tools and solutions for the problems of today and of the future. It has happened to all of us to see the **404 Page not found** while browsing the internet.

It's annoying.

This type of error is returned when the resource (link) we are requesting (by clicking) cannot be found. Errors are of utmost importance! The codes (like 404) are **well defined** to let the user clearly understand the origin of the problem.

When building robots we have to handcraft the errors in a similar fashion that is used for standard HTTP protocol. However, the problem increases in complexity as soon as we build **robots that interact with humans.** With these, we have to carefully design the **appearance, functionality, and behaviors.**

The type of errors that we aim to capture in these robots is not limited to technical issues! The **interaction itself** can progress in an unexpected way, and it is important to keep track of it (for instance with error-like diagnostics).

LET'S ENVISION THE FUTURE

Imagine a social robot deployed in a mall for customer service. Its task is to navigate among the crowd and engage in conversations with customers of the mall and respond to simple requests like:

- Where can I buy flowers here?
- What time does the hairdresser close?

This robot is not just a tool in a factory. It has to employ **social skills** to achieve its task! Its success can depend on two separate points:

1. an excellent technical platform (hardware and software are working as expected)
2. excellent employment of social skills

While the first one holds true for all products (We buy a car and expect it to work properly), the second point is relevant due to the social context the robot is expected to operate.

Social skills are difficult even for humans, they vary according to the local culture and demographic factors. The goal of deploying robots that can control their social skills is ambitious but doable!

We as engineers have to keep iterating with human scientists and eventually **define a standard protocol for social errors** that will arise during interactions with this type of robot. Clearly a misfunctioning of a motor of the robot is a type of error different from using inappropriate social behavior.

One rightfully might ask: “what is an appropriate behavior for this type of robot?”

The development of social robotics has borrowed theories from human-human interactions to increase comfort and trust in the interactions.
Examples of these behaviors are:

- An appropriate interpersonal distance while interacting
- A proper volume of the robot speaker if it has to speak to the user
- A policy for turn-taking in conversations that is acceptable by users

Keeping this in mind we can see that diagnostics of social skills can improve the development of robots. This error (due to its human nature) can hardly be detected by the robot. It needs feedback from the user. The feedback from the user can be implicit or explicit. Explicit feedback can be the user saying “Sorry I don’t like how you are behaving! I’m leaving!”. Implicit feedback can exploit cutting-edge emotion recognition to estimate whether the user is stressed or bored during the interaction. This latter tool can use models from Artificial Intelligence to detect events like disengagement that can be due to dissatisfaction with the interaction.

Given these simple examples of appropriate robot behaviors, it will be extremely important to define a standardized social errors protocol for robotics that can help the future development of truly social robots.

In this sense, the 001 Error could be mapped to the social misbehavior of the robot, providing the users and the developers with information on the state of the interaction (exteroceptor sensors) when the error was detected.

MEET FRANCESCO

CLICK HERE

TRAININGS
EVENTS

PERSEO SUMMER SCHOOL

on service-based and cloud robotic

20- 22 SEPTEMBER2022,
PALMA DE MAJORCA, SPAIN

The Summer School 2022 on service-based and cloud-based robotics was organized by the PERSEO ETN European Doctoral Training Network on Personalized Robotics.

Focus: technologies based on cloud services and robotics to support customization in human-robot interaction.

A series of lectures by international experts in the field (*see list of speakers below) and practical seminars on AI cloud and robotics systems such as IBM Watson AI Services and the Noosware Cloud Robotics platform. The school was attended by PhD students and researchers.

During the 2nd Residential Meeting, the ESRs will be requested to organize a workshop where the individual projects will be submitted in the form of position papers, linking state of the art and preliminary results. Position papers will be peer-reviewed and orally discussed by all workshop participants, including an external advisory group who will provide fruitful feedback to each ESR on the directions of research and potential methodological and theoretical aspects to consider.
ESR will have hands-on experience on the robots addressing problems connected to the Integration Scenarios for Market & Society on a platform working in groups.

PERSE

@PerseoProject

TRAININGS
EVENTS

PERSEO SUMMER SCHOOL

on service-based and cloud robotic

20- 22 SEPTEMBER2022,
PALMA DE MAJORCA, SPAIN

PERSEO.EU

PERSE
@PerseoProject

SPEAKERS

Prof. Michael Beetz (Bremen University, Germany)
Paving the way for a Cambrian explosion in robotic: open knowledge services for robotics applications.

Prof. Filippo Cavallo (Firenze University, Italy)
Internet of Robotics Things technologies for active and healthy aging: scenarios and challenges.

Prof. Tony Prescott (University of Sheffield, UK)
Designing Social Robots that people can relate to

Dr. Ben Kehoe (iRobot)
Severless Computing: a force moltiplier for cloud robotics

Prof. Elisa Rubegni (Lancaster University, UK)
Facing AI socio-technical challenges in vulnerable groups.

Prof. Gabriel Oliver (Universitat de les Illes Balears, Spain)
On the Autonomous Observation of Marine Environments

Dr. Ioanna Giorgi (Plymouth/Kent University, UK)
Multilingual Cognition of Robot Companions

Prof. Pramida Caleb-Solly (University of Nottingham, UK)
Rethinking Assistive Robotics

Prof. Irene Lopatovska (Pratt Institute, USA)
Alexa personalisation

Prof. Clare Dixon (University of Manchester, UK)
Verification and validation for robotics

Prof. Lucas Cordeiro (University of Manchester, UK)
Security of Software Systems with Applications on the Internet of Things

ORGANISED BY:

UNIVERSITY OF MANCHESTER, SHEFFIELD HALLAM UNIVERSITY,
UNIVERSITY OF NAPLES AND UNIVERSITY OF BALEARIC ISLANDS.

ESR PUBLICATIONS

PERSEO.EU

Familiar Acoustic Cues for Legible Service Robots

Angelopoulos Georgios, Vigni Francesco, Rossi Alessandra, Russo Giuseppina, Turco Mario, Rossi Silvia

RO- MAN 2022

Neuromorphic Computing for Interactive Robotics: A Systematic Review

Muhammad Aitsam, Sergio Davies, Alessandro Di Nuovo | **IEEE Access**

You Are In My Way: Non-verbal Social Cues for Legible Robot Navigation Behaviors

Angelopoulos Georgios, Rossi Alessandra, Di Napoli Claudia, Rossi Silvia | **IROS 2022**

Robust Human Motion Forecasting using Transformer-based Model

Esteve Valls Mascaro, Shuo Ma, Hyemin Ahn, Dongheui Lee | **IROS 2022**

Regulatory frameworks and standards for agricultural robotics in the European Union

Andrea Bertolini, Rocco limongelli

BDS Publishing

Anthropomorphism and Human-Robot Interaction (HRI)

Ricarda Wullenkord, Dimitri Lacroix, Friederike Eyssel | **Cambridge University Press**

Regulatory frameworks and standards for agricultural robotics in the European Union

Andrea Bertolini, Rocco limongelli | **BDS Publishing**

#NEWS

THE PAPER WINNER OF

EGO4D LONG-TERM ANTICIPATION CHALLENGE IN CVPR 2022 AND ECCV 2022

IS "INTENTION-CONDITIONED LONG-TERM HUMAN EGOCENTRIC ACTION FORECASTING"

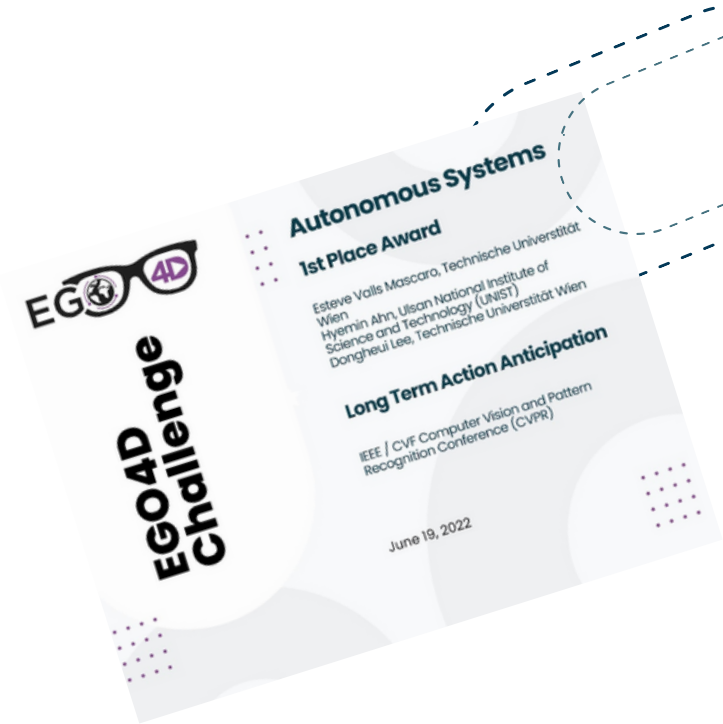
by Esteve Valls Mascarò

Abstract

To anticipate how a person would act in the future, it is essential to understand the human intention since it guides the subject towards a certain action. In this paper, we propose a hierarchical architecture which assumes a sequence of human action (low-level) can be driven from the human intention (high-level). Based on this, we deal with long-term action anticipation task in egocentric videos.

Our framework first extracts this low- and high-level human information over the observed human actions in a video through a Hierarchical Multi-task Multi-Layer Perceptrons Mixer (H3M). Then, we constrain the uncertainty of the future through an Intention-Conditioned Variational Auto-Encoder (I-CVAE) that generates multiple stable predictions of the next actions that the observed human might perform. By leveraging human intention as high-level information, we claim that our model is able to anticipate more time-consistent actions in the long-term, thus improving the results over the baseline in Ego4D dataset.

This work results in the state-of-the-art for Long-Term Anticipation (LTA) task in Ego4D by providing more plausible anticipated sequences, improving the anticipation scores of nouns and actions. Our work ranked first in both CVPR@2022 and ECCV@2022 Ego4D LTA Challenge.



MEET ESTEVE

CLICK HERE

CONTACTS

Prof. *SILVIA ROSSI*
Project Coordinator

Interdepartmental Center for Advances in Robotic Surgery (ICAROS)

University of Naples Federico II

silvia.rossi@unina.it

PERSEO

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 955778