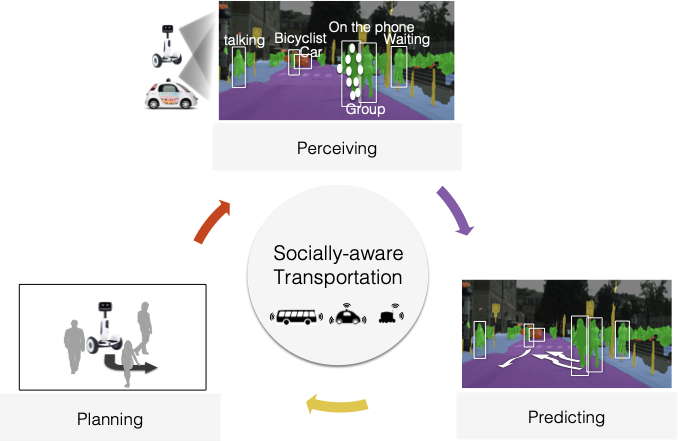
VITA projects 2021

<https://vita.epfl.ch/>

The VITA lab is interested in democratizing robots that will co-exist with humans in a safe, efficient, and socially-aware way. Self-driving cars, delivery robots, or social robots are examples of such robots. To realize this future, we propose empowering robots with a type of cognition we call ***socially-aware AI***, i.e., robots that can not only perceive human behavior but reason with social intelligence - the ability to make safe and consistent decisions in unconstrained crowded social scenes.

In the next pages, a list of Semester/Master projects is proposed to tackle the 3 **P**illars of a *socially-aware AI* system tailored for transportation applications: **Perception**, **Prediction**, and **Planning**.



The ideal student:

- is a strong coder and a fast learner

- is experienced in python

- has strong expertise in **Deep Learning** / ML and Computer Vision

- is interesting in reshaping transportation systems with AI

List of projects:

[Perception](#_c5c9pr7zeb6c)

[Generalize 3D Object Detection: Datasets and Metrics](#_c7mxpcm0yxmj)

[Action predictions for self-driving cars](#_brq71ampw3qf)

[Awesome Image Synthesis](#_800438ptj2xp)

[Long term pose tracking](#_f20kjzsr7ro)

[Envelope feature extraction and segmentation of buildings](#_suprdsgnil04)

[Multi-Task Learning with Auxiliary Tasks](#_kii7rwxthexh)

[One Network to Rule Them All](#_kii7rwxthexh)

[Sizing at first Sight](#_vgwnkevduiag)

[Dense Reconstruction From Sparse Measurements](#_1cgrmztxlqu9)

[Prediction](#_eq62wy8y2igp)

[Pedestrian Crossing Intention Prediction](#_kypp9ovwpd9q)

[Pedestrian Behavior prediction,](#_bid08knp8hjg)

[Vehicle behavior prediction](#_f7kktpdyjp7m)

[Pedestrian Stop & Go Forecasting](#_868y6d7f9qj5)

[Pose / Bounding box / Intention prediction](#_4nbf4girhcnv)

[Planning](#_smsq6haqipdv)

[Offline Reinforcement Learning for Autonomous Driving](#_25oa3x2okmx4)

[A self-navigating robot](#_iuu9ec1cr24)

[Ethical transportation](#_hwezp0po6ktq)

[Theoretical](#_2scx2humbior)

[Out-of-distribution Generalization via Self-supervis](#_cb3l2k1wnhrt)ed Learning

[Neural Network enhanced Discrete choice models](#_cb3l2k1wnhrt)

[Audio super-resolution](#_w0uuepw1f9cl)

[Rethinking Operational Research, Reinforcement Learning](#_4p3o045xrorc)

[Interdisciplinary Projects](#_6gg4bbwcpe73)

[A Path-planning Mars Rover](#_om757ittwxh8)

[Control Panel parts Recognition using RGB camera](#_mxm74yp1fd5z)

[Converting 3D depth map and camera view to 2D elevation map](#_prjk8d5c0slc)

| image | | Title | Description |
| --- | --- | --- | --- |
| Perception | | | |
|  | | Generalize 3D Object Detection: Datasets and Metrics | Many interesting datasets for 3D object detection have been recently launched: KITTI, NuScenes, Waymo, Argo, Lyft, A2D2, and more. Each one has different characteristics ([here](https://arxiv.org/abs/2005.08139) an overview), so there is no guarantee that an algorithm trained on a dataset will perform well in the real world.  In addition, the performances of state-of-the-art algorithms are evaluated using standard metrics, averaging over all the results of the val/test set. How well existing algorithms will perform in difficult cases? For example, an algorithm is able to predict the 3D distance of people walking in the street. But what happens if a person is lying down in the street?  In this project, S(he) will study cross-dataset generalization and / or new metrics to assess the long tail of perception for humans. (S)he will also evaluate common algorithms using the newly proposed metrics to see how they perform when tested on difficult instances.  Related [Paper](https://arxiv.org/abs/1906.06059) and [GitHub](https://github.com/vita-epfl/monoloco):  L. Bertoni, et al., MonoLoco: Monocular 3D Pedestrian Localization and Uncertainty Estimation, ICCV 2019  Project requirement: experience with Python, interest in Data Science or Computer Vision  Contact: [lorenzo.bertoni@epfl.ch](mailto:lorenzo.bertoni@epfl.ch) to customize this project based on your interest |
|  | | Action predictions for self-driving cars | Can autonomous vehicles detect human actions?  Is the person crossing the street distracted?  Is (s)he looking at the phone?  Has (s)he established an eye contact?  Perceiving these actions is crucial to make self-driving cars safe. This project aims to recognize many different actions of people in the context of autonomous driving, by leveraging the great advances of 2D pose estimation algorithms (For example: [GitHub](https://github.com/vita-epfl/openpifpaf))  Related [Paper](https://arxiv.org/abs/1906.06059) and [GitHub](https://github.com/vita-epfl/monoloco):  L. Bertoni, et al., MonoLoco: Monocular 3D Pedestrian Localization and Uncertainty Estimation, ICCV 2019  Project requirement: previous experience with Python and PyTorch  Contact: [lorenzo.bertoni@epfl.ch](mailto:lorenzo.bertoni@epfl.ch) |
|  | | Awesome Image Synthesis | Deep generative models have made stunning progress over the past years (for instance take a look at this website: <http://thispersondoesnotexist.com>). It’s now possible to artificially generate photo-realistic images, arts, audios, videos, and more broadly a large family of high dimensional signals. It paves the way to rethinking data generation, augmentation, and manipulation. There are multiple variations to work on:   * One possibility is to dive into state-of-the-art deep generative models (e.g., Generative Adversarial Networks) to study their performance across various domains and applications (e.g., generate arts, photo-realistic images of humans, music, depth signals). * The conditional image synthesis does not work the same as unconditional ones. In this project, the goal is to improve the image quality generated by the conditional Generative Adversarial Network (cGAN). We focus on the class of problems where images are generated given semantic inputs, such as scene segmentation masks or human body poses. We want to synthesize images which are realistic and at the same time reasonable in the context that can be useful for virtual tryon, augmented reality, etc. * Learning a representation for other tasks: Person re-identification is among the most challenging tasks in self-driving cars. This is a hard task because of nuisances such as body poses, illuminations, or changes in camera view-point. Thanks to the advancements in generative models, noticeable improvement in realistic image generation was made, recently. But it still shows some weaknesses, especially where large geometric transformations are appeared and also in low-resolution images. In this work, we will go beyond the classification techniques in person re-identification and study how deep generative models can achieve a hidden view of a pedestrian which is useful for other tasks. We also study how capable they are in learning a disentangled representation. * ...   Related papers:   * O. Gafni, and L. Wolf. "Wish You Were Here: Context-Aware Human Generation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. * S. Saadatnejad, A. Alahi. “Next Steps for Image Synthesis using Semantic Segmentation”, strc 2020, [link](https://infoscience.epfl.ch/record/281018?&ln=en) * S. Saadatnejad, A. Alahi. “Pedestrian image generation for self-driving cars”. Strc 2019.   Keywords: Conditional GAN, representation learning, context-aware image synthesis, Deep Generative models, Style transfer  Contact: [saeed.saadatnejad@epfl.ch](mailto:saeed.saadatnejad@epfl.ch) |
|  | | Long term pose tracking | Design an auxiliary task and extend an existing tracklet decoder for longer duration pose tracking in OpenPifPaf.  Strong Python skills required.  Related papers:  [1] PifPaf: Composite Fields for Human Pose Estimation  **Contact**: [sven.kreiss@epfl.ch](mailto:sven.kreiss@epfl.ch) |
|  | | Envelope feature extraction and segmentation of buildings | Using image recognition and deep learning techniques, we plan to detect essential features of building envelopes such as building type, the material of the façade and etc. only from images.  The prospective results of the project can be used to develop existing risk or energy models.  In the first step, a small dataset of Switzerland buildings should be annotated. In the second step, a multi-task model is designed and trained on this dataset with some modifications. Finally, we study how a generative model is useful in removing obstacles.  Related paper:   * Kang et al, “Building Instance Classification Using Street View Images”, Arxiv 2018, [link](https://arxiv.org/pdf/1802.09026.pdf)   Keywords: CNN, Semantic Segmentation, Multi-task Learning, Envelope feature extraction  Contact: [saeed.saadatnejad@epfl.ch](mailto:saeed.saadatnejad@epfl.ch) |
|  | | Multi-Task Learning with Auxiliary Tasks | Multi-Task Learning aims to transfer learned knowledge between multiple tasks to yield better overall performance. In cases with only one task of interest, it is still possible to leverage all other annotations available, that we call auxiliary supervision, to optimize the results on that one main task. In this project, the student will study how to combine auxiliary information in order to maximize the transfer from auxiliary tasks to the main one. It will be applied on several Computer Vision tasks, in the context of scene understanding.  Related paper:   * T. Mordan; N. Thome; G. Henaff; M. Cord: Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection. NeurIPS 2018. ([link](http://papers.neurips.cc/paper/7406-revisiting-multi-task-learning-with-rock-a-deep-residual-auxiliary-block-for-visual-detection), [github](https://github.com/vita-epfl/rock-pytorch))   Keywords: Multi-Task Learning, Scene understanding, Convolutional Neural Networks  Contact: [taylor.mordan@epfl.ch](mailto:taylor.mordan@epfl.ch) |
|  | | One Network to Rule Them All | Fields are an emerging representation in Deep Learning, particularly suited to describe full images and reason on their contents. At VITA lab, we have developed a unified framework using fields for various tasks useful for autonomous driving: object detection, human description (behaviors, visual attributes, body and hand poses, tracking), animal and vehicle pose estimation, scene graphs (i.e., predicting relations between objects in a scene).  The goal of the project will be to learn a single field-based network to solve all these tasks together, using the resources already available for all of them.  In a second time, if time allows it, additional tasks will be added, in order to ultimately get as close as possible to a thorough scene understanding model.  Strong coding skills and experience with PyTorch are required.  Related paper:   * S. Kreiss; L. Bertoni; A. Alahi: OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association. ArXiv 2021 ([link](https://arxiv.org/abs/2103.02440), [github](https://github.com/openpifpaf/openpifpaf)) * T. Mordan; M. Cord; P. Pérez; A. Alahi: Detecting 32 Pedestrian Attributes for Autonomous Vehicles. ArXiv 2020 ([link](https://arxiv.org/abs/2012.02647), [github](https://github.com/vita-epfl/detection-attributes-fields)) * G. Adaimi; S. Kreiss; A. Alahi: Perceiving Traffic from Aerial Images. ArXiv 2020 ([link](https://arxiv.org/abs/2009.07611), [github](https://github.com/vita-epfl/butterflydetector))   Keywords: Fields, Multi-Task Learning, Scene Understanding  Contact: [taylor.mordan@epfl.ch](mailto:taylor.mordan@epfl.ch) |
|  | | Sizing at first Sight | How far can we push the limits of 3D reconstruction? Can a deep learning method estimate the ring size of a finger given a single image? In this project, the student will study 3D reconstruction of deformable objects like hands, human body… While affordable depth sensor can reconstruct objects with few mm accuracy, (S)He will investigate how to increase the accuracy to less than 0.5 mm given multimodal data (RGB-D).  Contact: [alexandre.alahi@epfl.ch](mailto:alexandre.alahi@epfl.ch) |
|  | | Dense Reconstruction From Sparse Measurements | The past years have witnessed a race for more affordable depth sensors by compromising resolution. Yet, the reduced resolution is limiting the range of applications. In this work, we will study how Deep Generative models (e.g., GAN) can be used to reconstruct high resolution depth map given the combination of a low resolution sensor augmented with an RGB image.  Contact: [alexandre.alahi@epfl.ch](mailto:alexandre.alahi@epfl.ch) |
| Prediction | | | |
| Are They Going to Cross? A Benchmark Dataset and Baseline for Pedestrian  Crosswalk Behavior | | Pedestrian Crossing Intention Prediction | The ability to predict pedestrian behavior at crossroads is crucial for road safety, traffic management systems, Advanced Driver Assistance Systems (ADAS), and more broadly autonomous vehicles. In particular, knowing in advance whether pedestrians will cross the road in front of the vehicles is a major concern.  In this project, the student will study and implement (in PyTorch) several approaches from the literature to complete this task.  Related papers:   * T. Mordan; M. Cord; P. Pérez; A. Alahi: Detecting 32 Pedestrian Attributes for Autonomous Vehicles. ArXiv 2020 ([link](https://arxiv.org/abs/2012.02647), [github](https://github.com/vita-epfl/detection-attributes-fields)) * H. Razali; T. Mordan; A. Alahi: Pedestrian Intention Prediction: A Convolutional Bottom-Up Multi-Task Approach. 2020 ([link](https://drive.google.com/file/d/1bgQwPH6YX1EqEUOPAkxO3n8QZM5IWOG1/view), [github](https://github.com/HaziqRazali/Pedestrian-Intention-Prediction))   Keywords: Advanced Driver Assistance Systems, Autonomous Vehicles, Pedestrian Intention Prediction, Human Behavior Analysis  Contact: [taylor.mordan@epfl.ch](mailto:taylor.mordan@epfl.ch) |
|  | | Pedestrian Behavior prediction, Focus: **Sequence modelling,**  **multimodality,Generalization** | Self-driving or assisted-driving vehicles should not only have a thorough understanding of the current situation, but also forecast the future states. The forecasting step is essential for planning the right actions, avoiding dangerous situations, and raising early enough warnings.  In this project, the student will benchmark new state-of-the-art designs of forecasting algorithms that learn to predict the motion behavior of humans. The student will focus on building on top of the already well-established [TrajNet++](https://github.com/vita-epfl/trajnetplusplusbaselines) framework on the following areas:   1. **Interaction and Sequence Modelling**: The student will explore new sequence modelling architectures to better model human motion and social interactions. The student can also extend current models to better improve generalization / transferability of models. (Possibility of a **publication.)** 2. **Multimodality**: Human motion is inherently multimodal: given a past history of human motion paths, multiple socially plausible futures are possible. Training neural networks for forecasting based solely on minimization of *L2* loss leads to the model outputting the mean of all the possible outcomes. In this project, students will explore methods to model complex multimodal distributions. (Possibility of a **publication.)** 3. **Extending TrajNet++ with Cues and Datasets:** Students will focus on extending TrajNet++ with additional cues like scene information, head orientation and newly released datasets. Strong python/Pytorch coding is a strong requirement.     Reference papers:  Interaction and Sequence Modelling:   * Alahi, Alexandre et al. **“Social LSTM: Human Trajectory Prediction in Crowded Spaces.”** 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) * Vemula, Anirudh et al. “**Social Attention: Modeling Attention in Human Crowds**.” 2018 IEEE International Conference on Robotics and Automation (ICRA) (2018) * Giuliari, Francesco et al. “**Transformer Networks for Trajectory Forecasting**.” ArXiv abs/2003.08111 (2020)   Multimodality   * A. Gupta; j. Johnson; l. Fei-Fei; S. Savarese; A. Alahi : **Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks**. CVPR 2018. * Javad Amirian and Jean-Bernard Hayet and Julien Pettre : **Social Ways: Learning Multi-Modal Distributions of Pedestrian Trajectories with GANs**, CVPR Workshop 2019 * Liang et al. **The Garden of Forking Paths: Towards Multi-Future Trajectory Prediction**   TrajNet++   * P. Kothari; S. Kriess, A. Alahi: **Human Trajectory Forecasting in Crowds: A Deep Learning Perspective,** arxiv preprint   Related [GitHub](https://github.com/vita-epfl/trajnetplusplusbaselines) link  Project requirement: previous experience with PyTorch is a must.  KeyWords: Trajectory forecasting, social interactions, sequence modelling, multimodality.  Contact: [parth.kothari@epfl.ch](mailto:parth.kothari@epfl.ch) to customize this project based on your interest |
|  | | Vehicle behavior prediction | Can self-driving cars “drive” the same as humans? Is this dream going to come true?  Learning how to drive/predict from sensors’ inputs is one of the main building blocks of self-driving systems. In this project we use existing datasets to learn human driving behavior. However, it requires addressing multiple exciting challenges. The three main challenges are defined as three sub-projects as follows :   1. ***Causal learning****.* Causal learning has become a hot topic in recent years. This is due to the fact that neural networks tend to use “any” clue that helps them decrease the loss, and sometimes the leveraged clues are totally irrelevant to the task from a human point of view. The goal of this project is to first analyze existing models in learning correct causal relations, then, to study a causal model structure. Here are two related works: [1](https://sites.google.com/view/causal-confusion), [2](https://arxiv.org/abs/2009.12547). 2. ***Generalizability***. is a key goal in data-driven models. However, neural networks highly depend on the training data and perform relatively poorly on out-of-distribution samples. In this project we try to improve the model's generalizability by different approaches, such as [this work.](https://yueatsprograms.github.io/ttt/home.html) 3. ***Knowledge injection****.* Neural networks learn policies from data, however, there exist many types of knowledge which can be injected to the network to improve performance and generalizability. For instance, how can we teach networks to respect road boundaries, or to avoid collisions. This line of work is adding knowledge to the neural networks. [More info](https://towardsdatascience.com/augmenting-neural-networks-with-constraints-optimization-ac747408432f).   **Keywords**: Prediction / planning,, generalization, causal learning, knowledge injection  **Contact**: [mohammadhossein.bahari@epfl.ch](mailto:mohammadhossein.bahari@epfl.ch) |
|  | | Pedestrian Stop & Go Forecasting | Forecasting pedestrians' behaviors around self-driving cars is essential for safety, but is one of the most challenging tasks. We suggest that predicting when a pedestrian will walk or stand still using visual cues is at the right trade-off level between annotation requirement for learning and behavior analysis information. In this project, the student will formally define the task of Pedestrian Stop & Go Forecasting, and setup the learning problem from existing datasets. The student will then evaluate simple baselines and adapt existing approaches to this problem.  Keywords: Pedestrian Behavior, Forecasting, Convolutional Neural Network, Recurrent Neural Network,  Contact: [taylor.mordan@epfl.ch](mailto:taylor.mordan@epfl.ch) |
|  | | Pose / Bounding box / Intention prediction | One of the most challenging tasks for a self-driving vehicle when it sees a pedestrian is to predict not only the intention but also his location as early as possible. In this project, the student should study deep learning methods (e.g., CNN with attention mechanisms, LSTM) to predict the intention of the person. They should also investigate proper metrics for evaluation. The method can be used for early warning systems.  In the first step, we try to improve the intention prediction by a side-task which is a bounding box prediction, then we use other side-tasks such as pose prediction.  Predicting the pose can also be useful in decision making of disable people.  Related paper:   * S. A. Bouhsain, S. Saadatnejad, A. Alahi: “Pedestrian Intention Prediction: A Multi-task Perspective”, Arxiv 2020   Keywords: Forecasting, CNN/RNN  Contact: [saeed.saadatnejad@epfl.ch](mailto:saeed.saadatnejad@epfl.ch) |
| Planning | | | |
|  | | Offline Reinforcement Learning for Autonomous Driving | Imitation learning (IL) and reinforcement learning (RL) have been two dominant paradigms for data-driven planning / control for decades, yet they both have fundamental shortcomings: IL policies are inherently limited by imperfect human demonstrations; RL trial-and-errors often raise significant cost and safety concerns. Given a set of human driving experience, can machines learn to drive even better than human teachers? Offline reinforcement learning is a reviving framework that holds tremendous promise for making it possible [1]. In this project, we will explore this exciting area of research in the context of autonomous driving [2].  Reference: [1] [Offline Reinforcement Learning: From Algorithms to Practical Challenges, 2020](https://sites.google.com/view/offlinerltutorial-neurips2020), [2] [CrowdNav, 2019](https://arxiv.org/abs/1809.08835)  Prerequisite: fluent in Pytorch / Tensorflow, passionate about research / open problems, familiar with RL basics  Contact: [yuejiang.liu@epfl.ch](mailto:yuejiang.liu@epfl.ch) |
|  | | A self-navigating robot | In this project, the student will study existing path planning algorithms (e.g., RL based methods) that enable social robots navigate complex environments such as crowded corridors or even race with humans. (S)He will work with our Segway robotic platform named loomo:  <https://actu.epfl.ch/news/our-robot-interacting-with-vip-visitors-in-autonom/>  <https://www.youtube.com/watch?v=3AnXPqoIfvU&feature=emb_title>  Contact: [yuejiang.liu@epfl.ch](mailto:yuejiang.liu@epfl.ch) |
|  | | Ethical transportation | Self-driving cars, self navigating trams, or any intelligent Transportation Systems (ITS) will have to make decisions that can affect humanity. Which actions should a self-driving vehicle take in critical scenarios such as crashing against a wall to avoid collision with a group of kids? What are the ethical impacts of the actions that we (researchers) are automating in the design of intelligent transportation systems?  In this project, the student needs to identify this pandora’s box of research questions regarding ethical impacts of developing ITS. (S)He will develop a platform similar to the MIT moral machines by 1) building a crowd-sourced picture of human opinion on how machines should make decisions when faced with ethical dilemmas in ITS, and 2) crowd-sourcing assembly and discussion of potential scenarios of moral consequence.  **Contact**: [brian.sifringer@epfl.ch](mailto:brian.sifringer@epfl.ch) |

| Theoretical | | |
| --- | --- | --- |
|  | OoD Generalization via Self-supervised Learning | Recent advances in machine learning have brought tremendous successes in the i.i.d setting, where test samples come from the same distribution as the training data. However, to deploy the learned models to safety-critical systems like self-driving vehicles, we also need them to function robustly in the real-world under a wide variety of conditions, some of which may be underrepresented by, or even absent from, the collected training data. When this discrepancy occurs, conventional learning techniques often fail drastically [1]. In this project, we aim to address this fundamental challenge from the self-supervision perspective [2]. In particular, we will explore methods that can incorporate prior knowledge to reliably adapt models at test time.  Reference: [1] [Wilds: A Benchmark of in-the-Wild Distribution Shift, 2020](https://arxiv.org/pdf/2012.07421.pdf) [2] [Social NCE, 2020](https://arxiv.org/abs/2012.11717)  Prerequisite: fluent in Pytorch / Tensorflow, passionate about research / open problems, familiar with ML fundamentals  Contact: [yuejiang.liu@epfl.ch](mailto:yuejiang.liu@epfl.ch) |
|  | Neural Network enhanced Discrete choice models | On one hand, Discrete Choice Models have emerged as a convenient and powerful theoretical framework for analyzing individual travel behavior, yet often lack predictive power. On the other hand, Neural Networks have shown stunning predictability performance, yet with poor interpretability. In the project, the student will connect both worlds and study how extract the best of each, aiming for explainable A.I.  Applications to modeling human behavior: Transport Mode Choice, Intention Prediction, Trajectory Forecasting, … .  Reference papers:   * Sifringer, Brian, Virginie Lurkin, and Alexandre Alahi. "Enhancing discrete choice models with representation learning." *Transportation Research Part B: Methodological* 140 (2020): 236-261.   **Keywords**: Discrete Choice Models, Statistical Models, Neural Networks, Representation Learning, Novel Architectures, Interpretability  **Contact**: [brian.sifringer@epfl.ch](mailto:brian.sifringer@epfl.ch) |
|  | Audio super-resolution | Signal re-construction and even super-resolution can be advantageous in many applications. In this use-case, the goal of the project is to inspire onseleves of recent advances in Image enhancing and apply it to a different type of data: Audio. Mainly, Generative Adversarial Models (GAN) and associated new methods are investigated.  Reference Papers:       **Keywords**: GAN, Audio, Signal Enhancing, Embedding, Super-Resolution, Fourier Transforms, Mel Spectrogram, Convolutional Neural Networks  **Contact**: [brian.sifringer@epfl.ch](mailto:brian.sifringer@epfl.ch) |
|  | Rethinking Operational Research, Reinforcement Learning | Recent success stories of playing atari games or Alpha Go have shown that it is possible to rethink traditional Operational Research (OR) problems with Deep Reinforcement Learning (RL) techniques. The goal of this project is to revisit popular OR problems with Deep RL and compare them.  Contact: [alexandre.alahi@epfl.ch](mailto:alexandre.alahi@epfl.ch) |

| Interdisciplinary Projects | | |
| --- | --- | --- |
|  | Deep learning approach for 360 Stereo Camera Odometry | For the Xplore Mars Rover project, knowing the accurate position ( not using GPS ) of the rover is one of the most important and challenging aspect of Xplore Navigation Subsystem. This step is crucial for our path planner , the mapping of the environment and our motor controller. So far we use a IMU and encoder, but those solutions cannot detect slip and can output wrong positions estimations.  Unfortunately 3D LiDar are extremely expensive and it is a known fact that stereo cameras can with some degree replace LiDars ( less precise and more prone to perturbations ). The goal of this project is to use the stream of stereo cameras image pair to estimate the position of the rover.  In this project you will design a model to estimate the position of the rover given a sequence of 360 stereo image pairs.  **Reference documents:**  [1] VINET : <https://arxiv.org/abs/1701.08376>  **Keywords:** Deep Learning, Odometry, Stereo Cameras  **Contact**: [george.adaimi@epfl.ch](mailto:george.adaimi@epfl.ch) |
|  | Deep learning approach to 3D Lidar Odometry | For the Xplore Mars Rover project, knowing the accurate position ( not using GPS ) of the rover is one of the most important and challenging aspect of Xplore Navigation Subsystem. This step is crucial for our path planner , the mapping of the environment and our motor controller. So far we use a IMU and encoder, but those solutions cannot detect slips and can output wrong position estimations.  The goal of this project is to use only the 3d point cloud generated from our 3d Lidar (Ouster OS1 64) to estimate the rover position. The current state of the art method is called LOAM (see ref.) but it is hard to compute and to calibrate. Recently Deep Learning approaches (CNN taking as input 2 consecutive scans , LSTM on the all sequence , etc … ) have showed promising results, but often lacked of an appropriate dataset.  In this project you will design a supervised learning model to estimate the position of the rover given a sequence of 3d LiDar scans.  **Reference documents:**  [1] **LOAM**:<https://www.youtube.com/watch?v=8ezyhTAEyHs> <https://www.researchgate.net/publication/282704722_LOAM_Lidar_Odometry_and_Mapping_in_Real-time>  [2] **LodoNet** : https://arxiv.org/abs/2009.00164  **Keywords:** Deep Learning, Odometry, 3D LiDAR  **Contact**: [george.adaimi@epfl.ch](mailto:george.adaimi@epfl.ch) |
| **Source:** nationalgeographic.com | A curious rover | For the Xplore Mars Rover, the rover has to move in an unknown environment. Because the sensors we use cant yield a complete elevation map of this environment, many parts of this maps are yet to be discovered, which becomes a immense challenge for the path planner.  In this project you will optimize our current path planner to take into consideration such unknown areas in the map.  Note : You should not assume that the coordinates of the target positions are known ( AR-tag in the competition ). This is the interesting part of the project and where the appropriate ‘curiosity behavior’ matters a lot.  **Keywords:** Navigation, Curiosity  **Contact**: [george.adaimi@epfl.ch](mailto:george.adaimi@epfl.ch) |
|  | Blind local planner with reinforcement learning | In the Xplore Mars Rover, so far our local and global planner depend on the accuracy our of 3D LiDar sensor to detect obstacles. Unfortunately our LiDar has a min-range of 1m, and close obstacles are a terrifying threat. Even worst we do not take into consideration the types of soils the rover could be on ( sand, wet sand, rocks, leafs, etc.. ), which risk to over-use the resources of the rover.  The aim of this project is to improve our local planner not to depend solely on the LiDar but to use the current speed, weight distribution, angles of the rover joints, etc ….  In this project you will train a reinforcement learning model to adapt the rover local planner to handle obstacles (ex: rocks, holes , …) as well as different types of soil (ex: sand , gravel …)  **Reference documents:**  [1] https://leggedrobotics.github.io/rl-blindloco/  **Keywords:** Reinforcement learning, Navigation  **Contact**: [george.adaimi@epfl.ch](mailto:george.adaimi@epfl.ch) |