scientific inserts

I — XL

II

Speech and Audio Processing

Overview

Heads: Prof. Hervé Bourlard (MS and PhD, Polytechnic University, Mons, Belgium, 1982 and 1992), Dr. Philip N. Garner (MEng, University of Southampton, UK, 1991; PhD, University of East Anglia, UK, 2011), Dr. Mathew Magimai-Doss (MS by Research, Indian Institute of Technology Madras, India, 1999; PhD, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2005), Dr. Petr Motlicek (MS and PhD, Brno University of Technology, Czech Republic, 1999 and 2003).

Group overview

Speech processing has been one of the mainstays of Idiap's research portfolio for many years, covering most of the aspects of speech processing such as multilingual automatic speech recognition (ASR), speech synthesis, speech coding (including very low bit-rate), automatic speech intelligibility evaluation, or speech processing for analysis of motor speech disorders (e.g. pathological speech). The expertise and activities of the group encompass statistical automatic speech recognition (based on hidden Markov models (HMMs), or hybrid systems exploiting deep neural networks (DNN) and new deep learning architectures), text-to-speech (TTS), speaker recognition (with extensions towards text-dependent and forensics scenarios) and generic audio processing (covering sound source localization, microphone arrays, speaker diarization, audio indexing, perceptual background noise analysis for telecommunication systems) and, more recently, compressive sensing, and sparse recovery theories applied to ASR.

The Speech and Audio Processing group in 2020 was composed of 1 head of group, 3 principal investigators, 1 sabbatical academic visitor, 2 research associates, 9 postdocs, 12 PhD students, and 7 interns.

Key scientific outputs

Our primary research directions have traditionally been HMMs and DNN based approaches applied in acoustic modelling for various speech processing tasks. Use of techniques from HMM and HMM-DNN based ASR in HMM and HMM-DNN based speech synthesis resulted in a unified approach to speech recognition and synthesis. The group was well placed to take full advantage of recent advances in new architectures of deep learning, studied in particular through PyTorch and other open source frameworks. Advances in ASR are usually researched through Kaldi toolkit, now used by most of the international speech community, or its combination with other deep learning tools (particularly Pytorch).

In 2020, several key research contributions were achieved by the group, including: (1) multilingual automatic speech recognition, especially in cross-lingual adaptation, and automatic speech recognition in lowresourced language conditions, (2) speaker recognition, through both text-independent and particularly text-dependent (i.e. particularly for speaker verification) scenarios and information fusion for large-scale speaker identification, (3) large scale media processing, including multilingual broadcast news recognition, and spoken query for spoken term detection, (4) new Compressive Sensing and Sparse Recovery theories to ASR, and dualities with sparse DNN auto-encoders, (5) detection of impairments in speech signal to uncover motor speech disorders, and (5) paralinguistic speech processing with minimal prior knowledge.

Beside that, the group is also involved in the deployment of speech and speaker recognition algorithms for industrial applications, and is regularly involved in international evaluation campaigns.

Additional information and a list of projects are available from www.idiap.ch/speech.

Automatic speech recognition (ASR)

In recent years, our ASR research activities have been expanded from mono-lingual to cross-/multi-lingual processing. More specifically, in addition to focusing on "majority" languages other than English, French, or German, Idiap is actively carrying research in several ASR directions, including:

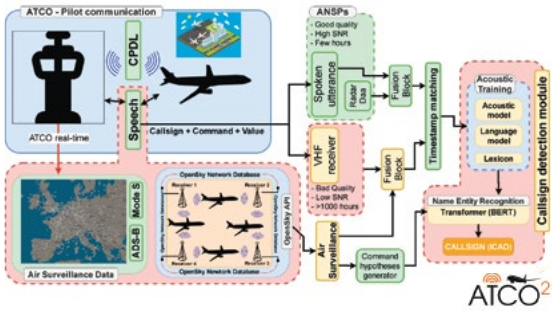


Figure 1

ATCO2 platform collecting audio data and combining them with contextual information.

→ **Robust parametrisation of ASR models**
We are investigating new features (e.g., posterior-based features) and new acoustic models (new forms of HMMs, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech, or dialect). In the context of the recently started EC H2020 projects (ATCO2 and HAAWAII)[1], we are developing semi-supervised learning methods for rapid adaptation of speech recognition models to new (unseen) domains using unlabelled data.

→ **Cross-lingual and multi-lingual speech recognition (specifically for low-resource scenarios)**
From 2017, Idiap collaborates on the US IARPA SARAL project[2]. As illustrated in Figure 2, the project aims at developing cross-lingual retrieval and summarization techniques that will work for any language in the world, given minimal resources to work with. In those contexts, we focus on investigating and exploiting fast acoustic model adaptation techniques in cross-lingual and multi-lingual scenarios. The concept is also partially exploited in an industrial collaboration with Uniphore[3], a world leader in Conversational AI, to develop ASR engines for specific Asian languages.

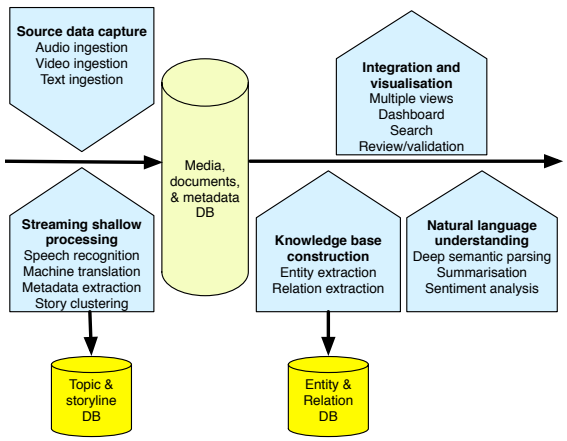


Figure 2

An overview of a typical multilingual multimedia processing stream, as used in ongoing US-DARPA SARAL project in the context of multilingual speech recognition at Idiap.

→ **Swiss languages**
We continuously improve our multilingual speech recognisers for Swiss German and Swiss French and also apply the most recent advances in speech technology employing Deep Neural Networks (DNN). Since 2015, we collaborate with recapp IT AG on a wider range of Swiss dialects towards the first commercial product that performs Swiss German (dialect) speech recognition. Idiap also works on an ongoing Innosuisse project SM2[4] with aim to develop a customisable technology for ASR followed by "semantic keyword and concept detection and spoken document summarization" applied to e-learning domain. In late 2019, the group initiated a new collaboration with Swisscom, enabling us to investigate "lexicon-free" advances in the field, which are particularly suitable for dialects with no standardised orthography.

→ **Exploiting compressive sensing and sparse recovering theories for ASR**
Through SNSF funded projects PHASER, PHASER-QUAD and SHISSM[5], Idiap is still continuing development of new theoretical links between compressive sensing, sparse auto-encoders, and statistical/HMM-DNN approaches towards improving ASR performance and noise robustness.

Text-to-speech synthesis (TTS)

Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS has been central to several past projects. The group has tracked the recent developments in deep learning which will dominate future research. Work under MASS (Multilingual Affective Speech Synthesis) brought the concept of emotion into the speech synthesis, particularly via modelling of prosody. As the MASS project is replaced by NAST (neural architectures for speech technology), the research focusses on how to integrate emotional indicators based on formant position into state of the art deep learning solutions. Current capability is reflected in the open-source package IdiapTTS[6], originally released in 2019.

Speaker recognition & verification

Idiap is actively carrying R&D on significantly improving capabilities of voice technologies in suspect identification applicable to very large scale data. Two activities were pursued in 2020: (1) combining speaker recognition with other tasks such as natural language understanding and network analysis to combat organized crime through ROXANNE H2020 project[7], and (2) improving the state-of-the-art speaker technologies by integrating recent advances in machine learning (especially through the participation on NIST speaker recognition evaluations[8]).

Pathological speech processing

Speech and language impairments can occur due to various reasons such as, due to neurological disorders, oral cancer, hearing loss. In recent years, Idiap has been actively involved in such impaired or pathological speech processing in collaboration with clinical researchers. The SNSF Sinergia project MoSpeeDi[9] (Motor Speech Disorder) focuses on (i) developing accurate models to characterize non-impaired and impaired phonetic speech planning and motor speech programming, and ii) developing automatic techniques to detect and classify several speech impairments as well as to assess the intelligibility of patients.

1 https://www.atco2.org,https://www.haawaii.de
2 https://www.idiap.ch/en/scientific-research/projects/SARAL
3 https://www.uniphore.com
4 https://www.idiap.ch/en/scientific-research/projects/SM2
5 https://www.idiap.ch/en/scientific-research/projects/SHISSM
6 https://github.com/idiap/IdiapTTS
7 https://www.roxanne-euproject.org
8 https://www.nist.gov/itl/iad/mig/speaker-recognition
9 https://www.idiap.ch/en/scientific-research/projects/MOSPEEDI

The EU H2020 MSCA-ITN-ETN project TAPAS[10] which is targeting three key research problems, (1) detection, (2) therapy, and (3) assisted living so that it works well for people with speech impairments and also helps in making informed clinical choices.

Other directions

- **Physiological influences**
  Under the NAST project and the new NCCR Evolving Language, we are investigating both how physiological processes of perception and production can influence speech technology, and how the quite mature technology can say something about our understanding of physiology. In 2020, two distinct threads began: in a first, stemming from the TTS work, we investigate how to combine conventional DNNs with the spiking neurons thought to be closer to the physiological function. In a second thread we investigate the latent understanding of the ear and cochlea, including the "efferent" path.

- **Sign language processing**
  In the context of SNSF Sinergia project SMILE,[11] Idiap is developing a sign language assessment system that can assist Swiss German sign language learners in standardizing a vocabulary production test to be aligned with levels A1 and A2 of the Common European Framework of Reference for Languages. Demonstration available at-vimeo.com/297803984.

- **Sound localization and microphone array**
  Idiap continues to work on distant speech processing by contributing to the Perception and Activity Understanding group through EC H2020 MuMMER project[12], focusing on audio source localization, speech detection and speaker re-identification applied in robotics.

- **Effective processing of speech using embedded devices**
  Integrating voice technologies in low powered devices with limited computing capabilities, Idiap is collaborating with Logitech through the SHAPED Innosuisse project[13].

- **Joint acquisition and modeling of speech and physiological signals**
  In recently funded SNSF project TIPS[14], Idiap is collaborating with CSEM[15] to develop a platform where speech and physiological signals are collected in a synchronous manner through a wearable cooperative sensor and processed to develop novel speech- and physiology-based applications.

Key publications

[1] P. N. Garner and S. Tong, "A Bayesian approach to recurrence in neural networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020

[2] N. Antonello and P. N. Garner, "A t-distribution based operator for enhancing out of distribution robustness of neural network classifiers," IEEE Signal Processing Letters, vol. 27, pp. 1070–1074, Jun. 2020

[3] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Automatic pathological speech intelligibility assessment exploiting subspace-based analyses," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1717–1728, May 2020

[4] I. Kodrasi and H. Bourlard, "Spectro-temporal sparsity characterization for dysarthric speech detection," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1210–1222, Apr. 2020

[5] R. Prasad, G. Yilmaz, O. Chetelat, et al., "Detection of s1 and s2 locations in phonocardiogram signals using zero frequency filter," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2020

10 https://www.tapas-etn-eu.org
11 https://www.idiap.ch/scientific-research/projects/SMILE
12 http://mummer-project.eu
13 https://www.idiap.ch/en/scientific-research/projects/SHAPED
14 https://www.idiap.ch/en/scientific-research/projects/TIPS
15 www.csem.ch

Social Computing

Overview

Head: Prof. Daniel Gatica-Perez (PhD, University of Washington, USA, 2001; EPFL Adjunct Professor)

Group overview

Social computing is an interdisciplinary domain that integrates theory and models from ubiquitous computing, social media, machine learning, and social sciences to analyze human behavior in everyday life for social good.

The Social Computing group in 2020 was composed of one group head, two postdoctoral researchers, three PhD students, one scientific collaborator, and two EPFL master students. The main research lines investigated in 2020 included mobile crowdsourcing for health and cities; social media analytics; and ubiquitous interaction analysis.

Key scientific outputs

Publications on (1) mobile crowdsensing to understand youth nightlife and eating habits; and (2) multimodal analysis of social interaction. 15 EPFL PhD students have graduated from the group since 2002.

Additional information and a list of projects are available from www.idiap.ch/socialcomputing.

Mobile crowd-sensing for health and cities

First, in the context of the SNSF Dusk2Dawn project[16] (Characterizing Youth Nightlife Spaces, Activities, and Drinks, in collaboration with La Trobe University and the University of Zurich), we investigated the use of mobile crowdsensing to characterize urban phenomena related to nightlife (Figure 3). This included the automatic recognition of night drinking activity from smartphone sensor data (location, motion, bluetooth, wiki, and app logs) in Switzerland [1]. Furthermore, we investigated the methodological possibilities of analyzing visual crowdsourced data to understand phenomena of interest to alcohol research, thus bringing novelty to work in public health [2]. This research was complemented with qualitative methods used to understand how youth relate to their personal spaces in the context of nightlife [3].

Second, the European H2020 WeNet project[17], is building diversity-aware algorithms for mobile sensing to support the well-being of young adults, and is developing a series of large-scale experiments in several European universities, as well as universities in Latin America and Asia. One key motivation of this work is the advocacy for diversity in data and algorithms to improve the representation of non-western citizens. In the context of eating habits, we conducted a comparative study across two of these countries, focused on the recognition of both social context and food categories in eating episodes from smartphone sensing [4].

Regarding mobile crowdsourcing for social innovation, we continued our work using the Civique platform that allows to collect mobile data for local causes)[18]. The platform has been used in a variety of applications, ranging from supporting cities to collect information related to urban issues like street harassment, to teaching students about humanitarian technologies. In 2020, the platform was used in the context of the Corona Citizen Science project (in collaboration with EPFL and the University of Lausanne). This project explored the use of smartphones to capture some aspects of the experience of Swiss residents during the COVID-19 lockdown[19].

Ubiquitous interaction analytics

In the context of the SNSF HealthVlogging project (in collaboration with the University of Lausanne), we are investigating how human-centered, audio-visual analysis of health-related videos shared in platforms like YouTube can complement the work of health psychologists studying new practices of health promotion. This line of work complements other research in our group, including the SNSF Digital Lives project (in collaboration with the University of Neuchâtel and the University of Lausanne) [5].

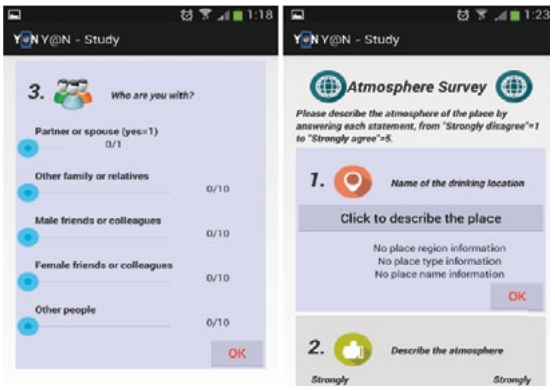


Figure 3

Mobile crowd-sensing app for nightlife.

16 http://www.idiap.ch/project/dusk2dawn
17 https://www.internetofus.eu/
18 https://www.civique.org
19 https://www.idiap.ch/en/allnews/citizen-science-explores-the-aftermaths-of-the-covid-19-lockdown

Key publications

[1] T.T. Phan, F. Labhart, S. Muralidhar, and D. Gatica-Perez, Understanding Heavy Drinking at Night Using Smartphone Sensing and Active Human Engagement, in Proc. Int. Conf. on Pervasive Computing Technologies for Healthcare (Pervasive Health), Atlanta, Oct. 2020.

[2] F. Labhart, T. T. Phan, D. Gatica-Perez, and E. Kunstche, Shooting shots: Estimating alcoholic drink sizes in real life using event-level reports and annotations of close-up pictures, Drug and Alcohol Review, published online Oct. 2020.

[3] K. Pelzelmayer, S. Landolt, J. Truong, F. Labhart, D. Santani, E. Kuntsche, and D. Gatica Perez, Youth Nightlife at Home: Towards a Feminist Conceptualisation of Home, Children's Geographies, Vol. 19, Issue 1, 2021.

[4] L. Meegahapola, S. Ruiz-Correa, and D. Gatica-Perez, Alone or With Others? Understanding Eating Episodes of College Students with Mobile Sensing, in Proc. Int. Conf. on Mobile and Ubiquitous Multimedia (MUM), Essen, Nov. 2020.

[5] S. Muralidhar, E. Kleinlogel, E. Mayor, M. Schmid Mast, A. Bangerter, and D. Gatica-Perez, Understanding Applicants' Reactions to Asynchronous Video Interviews though Self-Reports and Nonverbal Cues, in Proc. ACM Int. Conf. on Multimodal Interaction (ICMI), Utrecht, Oct. 2020.

Machine Learning

Overview

Head: Prof. François Fleuret (MS École Normale Supérieure de Paris and University of Paris VI, 1995; PhD, University of Paris VI, France, 2000; EPFL Adjunct Professor)

Group overview

Machine learning encompasses computer techniques that modulate their behavior according to exemplar data. It has resulted in technologies at the core of many modern every-day dataprocessing software and apparatus. The objective of the Machine Learning group is to develop novel machine-learning techniques of general use, with a particular interest in algorithmic efficiency and training from small data-sets. The research we conduct can be motivated by a general and fundamental problem, or by a concrete industrial application or use case.

Over the recent years, the group has been composed on average of four PhD students, and one or two developers working on industrial applications. We also maintained a sustained collaboration with EPFL's eSpace Center, CVLab and MLO lab.

Key scientific outputs

In 2019, our work has resulted in contributions that improved surface reconstruction from large signal and event-based sensors, interpretability of deep architectures neural networks, more stable training procedures for generative adversarial models, and importance-sampling methods for dealing with very large images.

Additional information and a list of projects are available from www.idiap.ch/ml.

Jacobian matching and regularization

Transfer learning aims at taking advantage of pre-existing models to facilitate the training of new models, either by speeding it up, or by allowing it with very small amount of training data. The key notion is to "transfer" structures learned by the existing network. Our algorithm consists of a novel penalty that not only forces the new model to mimic the response of the existing one, but to also mimic the dynamic of change of the output, given changes of the input. We have put the same tools to use for interpretability, relying on the amplitude of the derivatives of a neural network's output with respect to internal activation to provide a consistent estimate of the importance of different parts of the signal in driving the network's decision.

Stable Adversarial Optimization

"Generative Adversarial Networks" rely on training jointly two models, one synthesizing realistic signals (images, sound, text) and another trying to discriminate synthetic from genuine examples. Such techniques have demonstrated striking performance in many application domains, but involve a complex and unstable optimization problem. We have developed a new method that consists of training several such pairs in parallel, and maintaining carefully their statistical independence. This insures that their joint behavior has a good "covering" property, and we show experimentally that the resulting synthesis is less likely to miss sub-families of samples.

Re-sampling for deep models

The computational requirement for deep neural networks is one of their most problematic characteristics, even though most of the computation is spent on samples that are properly handled, and could be ignored. For training, we have derived a tractable upper bound of the per-sample gradient norm that allow to prioritize re-sampling examples and reduce the variance of the stochastic gradient estimates. During inference, we have designed a new model to handle megapixel images, composed of a first network that computes an attention map on a downscaled input, and a second network that processes the full resolution input at locations sampled according to the attention scores.

Depth estimation with deep models

End-to-end deep-learning networks are now the most efficient method for stereo matching. However, existing networks are memory-hungry and unable to process even modest-size images, and they have to be trained for a specific disparity range. The Deep Stereo network that we developed addresses both issues: First, its architecture relies on novel bottleneck modules that drastically reduce the memory footprint in inference, and additional design choices allow to handle greater image size during training. We have extended this approach to a new class of event-based sensors which provides extremely reactive visual information, at the cost of the point-wise intensity estimation and spatial resolution. Our algorithm relies on a new integration module that combines evidences through time to allow a classical deep architecture to be put to use.

Key publications

[1] T. Chavdarova, G. Gidel, F. Fleuret, and S. Lacoste-Julien, "Reducing Noise in GAN Training with Variance Reduced Extragradient", in Proceedings of the international conference on Neural Information Processing Systems (NeurIPS), pages 391–401, 2019.

[2] S. Srinivas and F. Fleuret, "Full-Gradient Representation for Neural Network Visualization," in Proceedings of the international conference on Neural Information Processing Systems (NeurIPS), pages 4126–4135, 2019.

[3] S. Tulyakov, F. Fleuret, M. Kiefel, P. Gehler, and M. Hirsch, "Learning an event sequence embedding for event-based deep stereo," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 1527–1537, 2019.

[4] A. Katharopoulos and F. Fleuret, "Processing Megapixel Images with Deep Attention-Sampling Models," in Proceedings of the International Conference on Machine Learning (ICML), pages 3282–3291, 2019.

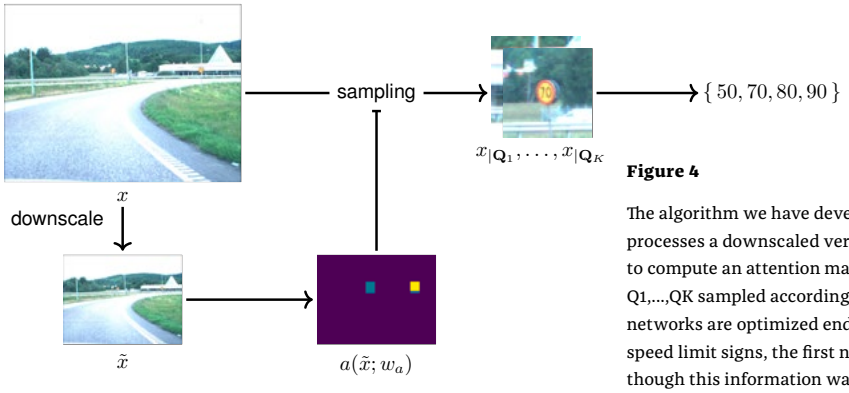


Figure 4

The algorithm we have developed is composed of two models: One processes a downscaled version x˜ of the image in a regular manner to compute an attention map, the second processes only locations Q1,...,QK sampled according to the attention of the first. The two networks are optimized end-to-end jointly. On this data-set of speed limit signs, the first network "learned" to look at signs even though this information was never provided during training, and the second network learned to recognize the different speed numbers to make the final prediction.

Perception and Activity Understanding

Overview

Head: Dr. Jean-Marc Odobez (PhD, INRIA/Rennes University, France, 1994; EPFL MER)

Group overview

The group conducts research in human activity analysis from multi-modal data, investigating fundamental tasks like person tracking, pose estimation, non-verbal behavior detection, or the temporal interpretation of this information in forms of gestures, activities, or social relationships. These tasks are addressed through the design of principled algorithms extending models from computer vision, multimodal signal processing, and machine learning, in particular probabilistic graphical models and deep learning. Behavior, human and human-robot interaction analysis, surveillance, traffic monitoring or multimedia content analysis are the main application domains.

Over the last 5 years, the group was composed on average of two post-doctoral researchers, four PhD students, one research engineer from the development team.

Key scientific outputs

The group is known for its work on probabilistic multi-object tracking, non-verbal behavior extraction, and temporal motif discovery. In 2015 and 2016, the PAU team ranked first at the MediaEval Person discovery challenge, related to its work on multi-modal person face diarization in the

EU EUMSSI project, Its patended work on 3D face and gaze tracking has led to the creation of the Eyeware SA company in 2016. Also, the team ranked second at the ICCV 2019 Facebook Synthetic Eye Generation Challenge. In recent years, the group has been investigating deep learning methods for several tasks like gesture recognition, audio-visual speaking activity modeling, gaze, audio localization and speech/nonspeech detection, body landmark detection, and multimedia processing (cross-modal transfer learning, shape recognition, text localization and semantic categorization). The group also integrates its algorithms into real-time perceptual systems used in collaborative projects (HRI system for the Pepper robotic platform in the EU MuMMER project, see video), or by companies such as an anti-tailgating detection system, or our emotion recognition method for job interviews developed in the context of the innoswiss ADVANCE project and featured in the national TV[20]. During the period 2016–2020, the group published 15 journal papers and 37 conference papers, and filled 3 patents.

✳ **Additional information and a list of projects are available from** www.idiap.ch/perception.

Human activity analysis, non-nerbal behavior (NVB) extraction, human-robot interactions

→ **Head and body pose inference from RGB-D (color and depth) data**
The team has developed a robust and accurate head pose tracking framework from RGB-D data combining the benefits of the online fitting of a 3D face morphable model with the online 3D reconstruction of the full head (Figure 5b), making head tracking a commodity for situations up to 1.5m. Research on 2D and 3D body pose estimation is also conducted. Leveraging our previous work on lightweight and efficient CNN structures for 2D body landmark estimation we proposed a residual approach to infer the 3D pose of people [1] (see Figure 5a).

→ **Gaze analytics**
Due to visually unobservable gaze variabilities across people, we have investigated several methods for building user-specific models from some user samples. This year, we pushed our initial models further and demonstrated that it was possible to design a fully unsupervised gaze representation learning approach, allowing the possibility to leverage internet-scale data. This was achieved by jointly learning a representation network and a gaze redirection network only from pairs of eye images, using as task the redirection synthesis of the second image from the first image and the difference between the two gaze representations derived from the image pairs [3]. Finally, in the context of the ROSALIS SNSF project[21], we investigated the use of gaze coordination priors involved in object manipulations and speaker interaction for obtaining weakly labeled gaze samples [4].

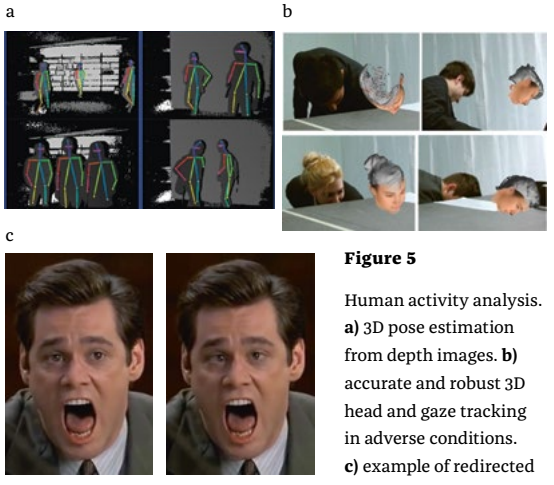


Figure 5

Human activity analysis. **a)** 3D pose estimation from depth images. **b)** accurate and robust 3D head and gaze tracking in adverse conditions. **c)** example of redirected gaze (left image: original; right one: redirected gaze towards the nose).

20 Mise au point, TSR, 18/10/2020.

21 https://www.idiap.ch/en/scientific-research/projects/ROSALIS

Multimedia and multimodal analysis

→ **Semantic text recognition (OCR)**
In the Innosuisse VIEW project, we investigated deep learning methods for the detection, segmentation, categorization and recognition of text content in slides, allowing further semantic tagging (see Figure 6).

→ **Audio analysis**
Within the EU MuMMER project on social robotics, we investigated different DNN architectures for sound processing. We proposed an efficient multi-task approach for the joint localization and categorization (speech vs non-speech) of multiple sources from a microphone array, a frequent situation for robots placed in public spaces. This is illustrated in Fig. 6, and was nominated for the best student paper award at Interspeech. More recently, we proposed methods that drastically reduce the amount of annotated data needed to train our architectures for each new microphone array: use of simulated data; domain adaptation, and weakly supervision (knowing the number of sources, not their location). achieving performance close to the full supervised case but with much less annotated data.

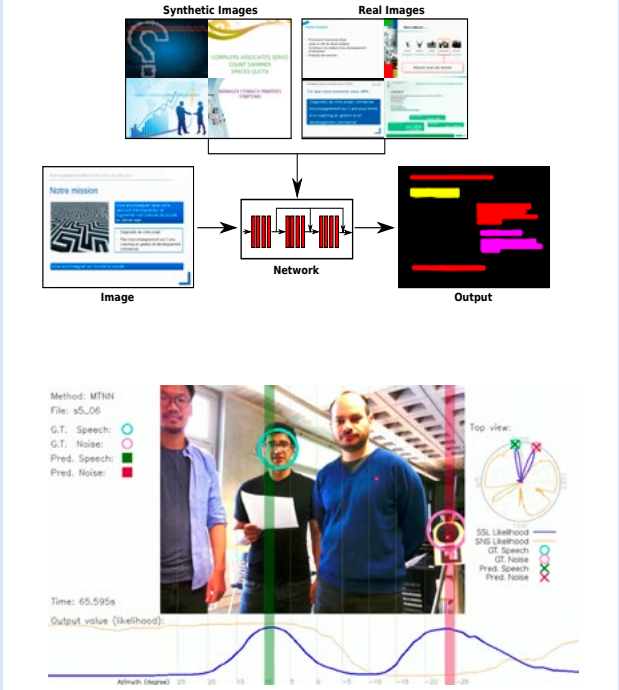


Figure 6

Up: Semantic Text Segmentation Network, localizing text in images while also localizing and recognizing text regions. **Down:** Multiple sound source localization and discrimination with a multitask network learning to predict for each direction (circle on the right) the sound source likelihood. Sound sources are then mapped to the visual data.

Key publications

[1] A. Martínez, M. Villamizar, O. Canévet, J.-M. Odobez, "Residual Pose: A Decoupled Approach for Depth-based 3D Human Pose Estimation," in Int. Conf. on Intelligen Robots and Systems (IROS), 2020.

[2] G. Liu, Y. Yu, K. Funes and J.-M Odobez. "A Differential Approach for Gaze Estimation," in IEEE Trans. Pattern Anal. Machine Intelligence (PAMI) , Vol 43(3), 2021.

[3] Y. Yu, and J.-M Odobez. "Unsupervised Representation Learning for Gaze Estimation," in Int Conf. on Vision and Pattern Recognition (CVPR), June 2020.

[4] R. Siegfried, B. Aminian and J.-M. Odobez. "ManiGaze: a Dataset for Evaluating Remote Gaze Estimator in Object Manipulation Situations," ACM Symp. on Eye Tracking Research and Applications (ETRA), 2020.

Uncertainty Quantification and Optimal Design

Overview

Head: Prof. David Ginsbourger (Ph.D. Mines Saint-Etienne 2009, Habilitation Univ. Bern 2014, Titularprofessor Univ. Bern 2018)

Group overview

The Uncertainty Quantification and Optimal Design group focuses on quantifying and reducing uncertainties in the context of natural and artificial complex systems. Application domains notably include energy and geosciences, with collaborations ranging from safety engineering to hydrology and climate sciences. In all these fields the study of complex systems often relies on expensive data acquisition and model runs, calling for adapted experimental design strategies. UQOD members are coming from and keeping strong academic ties to the Institute of Mathematical Statistics and Actuarial Science (IMSV) of the University of Bern (UniBE). During the year 2020, the UQOD group has been composed of a permanent senior researcher, two PhD students, an intern, and three visiting students (one PhD student from the University of Neuchâtel and two UniBE master students).

Key scientific outputs

Current contributions include efficient algorithms for Bayesian optimization and for estimating and quantifying uncertainties on implicitly defined parameter regions using Gaussian Process (GP) models. Other recent results deal with the interplay between the choice of covariance kernels and properties of GPs, with implications in high-dimensional GP modelling, in function prediction under structural constraints, and in problems with set-valued inputs. Ongoing work encompasses uncertainty quantification and efficient design of experiments for data- and simulation-driven inverse problem solving with applications in computational cosmology and geophysics (Current SNSF project).

Bayesian optimization and emulation with Gaussian Process models

Bayesian global optimization relying on GPs has become a standard for optimizing prohibitively expensive to evaluate systems, e.g. with response(s) stemming from heavy numerical simulations. More generally, sequential design of (computer) experiments based on GPs have flourished to efficiently address a variety of goals. This constitutes one of the core domains of expertise of the UQOD group, with recent contributions ranging from theoretical to methodological questions (such as parallelization, handling large data sets, coping with high-dimensional inputs, with set-valued inputs, etc.) and applications. Notably, the group has been involved in a collaboration with researchers in hydrogeology with the aim to investigate Bayesian optimization for contaminant source localization relying on flow simulations.

The UQOD group also investigates GP emulation per se, with a particular focus on the incorporation of expert knowledge and the identification of structural properties of objective functions through the specification of covariance kernels and the estimation of their parameters. Recent work directions include investigations on positive definite kernels over sets of finite sets using RKHS embeddings, with application to Bayesian (combinatorial) optimization [2]. Figure 7 represents two instances of a score discrepancy landscape corresponding to an optimal and to an arbitrary subset of five monitoring wells among twenty-five, respectively, in the framework of the aforementioned contaminant localization problem. It is found that Bayesian Optimization algorithms based on the considered RKHS embeddings most of the time successfully locate the optimal well combination (out of a total of 53 130 candidates) in a few dozens of iterations. Furthermore, UQOD has been involved in collaborations around the adaptation of GP models and Bayesian Optimization algorithms to challenging frameworks, be it for instance in terms of high-dimensionality or of heteroscedastic multivariate trajectories in the context of robot learning. The former topic has been at the center of a research work with M. Binois (now with INRIA Sophia-Antipolis) and O. Roustant (now with INSA Toulouse) on the choice of the low-dimensional domain in high-dimensional global optimization via random embeddings, which as resulted in the article [1]. On a different note, a collaboration with Noémie Jaquier and Sylvain Calinon from Idiap's Robot Learning and Interaction (RLI) group has resulted in [5], where it is shown how to design in a multi-output GP framework a probabilistic model that encapsulates variability information inherited from on a preliminary Gaussian Mixture model learnt from demonstrations while being generative and naturally accommodating via-points and related constraints.

Figure 7

Contaminant localization: Score discrepancy map: location of selected wells (input S), score discrepancy landscape, and the spatial sum of score discrepancy objective function value f(S) from [2].

Computer experiments for the quantification and the reduction of uncertainties

Besides global optimization, UQOD has indeed also been focusing on sequential strategies dedicated to other goals such as locating parameter regions leading to a response exceeding a given threshold, corresponding e.g. to an abnormal behaviour of the considered system. More generally, the combination of GP modelling and stochastic simulation techniques have been investigated for quantifying and reducing uncertainties on sets, with example application in safety engineering with goals such as identifying sets of dangerous and/or safe configurations of a complex system and if possible providing some measures of confidence along with the estimate(s). Previously developed "asymmetric nested Monte Carlo" algorithms and their use for efficiently estimating orthant probabilities of high-dimensional Gaussian vectors have been instrumental to derive conservative set estimates on a neutronic criticality safety test case (from IRSN, the French Institut de Radioprotection et de Sûreté Nucléaire) and also to derive sequential design strategies dedicated to this class of conservative set estimation problem. The underlying long-standing collaboration with colleagues from Neuchâtel, CentraleSupélec and IRSN, has resulted in the article [3]. Current perspectives include the adaptation of such (conservative) set estimation strategies to larger scale inverse problem from geophysics (collaboration with UNIL), e.g., when uncertainties on the mass density field inside a volcano are to be reduced via well-chosen gravimetric measurements.

Key publications

[1] M. Binois, D. Ginsbourger, and O. Roustant, "On the choice of the low-dimensional domain for global optimization via random embeddings," in J. Glob. Optim. 76, 69–90 (2020).

[2] D. Ginsbourger and T. Krityakierne, "Kernels over sets of finite sets using RKHS embeddings, with application to Bayesian (combinatorial) optimization," Accepted to AISTATS 2020.

[3] D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet, "Adaptive Design of Experiments for Conservative Estimation of Excursion Sets," in Technometrics (Published online: 23 Dec 2019).

[4] D. Armentano, J.-M. Azaïs, D. Ginsbourger, and J.R. León, "Conditions for the finiteness of the moments of the volume of level sets," in Electron. Commun. Probab., 24:17 (2019).

[5] N. Jaquier, D. Ginsbourger, and S. Calinon, "Learning from demonstration with model-based Gaussian process," Conference on Robot Learning (CoRL), 2019.

Genomics and Health Informatics

Overview

Head: Dr. Raphaëlle Luisier (Master of Science in Bioengineering and Biotechnology, EPFL, 2009 & PhD in Bioinformatics from the Basel University, 2013)

Group overview

The Genomics & Health Informatics Group was created in 2019 and is developing statistical and machine learning methods, including deep learning, to extract key information from multimodal and longitudinal biological data of various types such as genomic, clinical, and imaging data. Specifically, the group aims to address salient biological questions related to poorly understood human disorders by interpreting and integrating complex high-content data sets. The group works in close collaboration with biologists and clinicians (Patani laboratory, Francis Crick Institute, London; Serio lab, Kings College, London), thereby maintaining a fertile ground for innovation, learning and discoveries with real therapeutic prospects. The group has expertise in genomics, bioinformatics, RNA biology, neuroscience, data science, and data visualisation. The group has access to unique high quality data including longitudinal RNA-sequencing (Patani laboratory, Francis Crick Institute, London) and time-lapse cellular imaging (Serio lab, Kings College, London).

Key scientific outputs

In 2020 the Genomics & Health Informatics group has developed methods enabling the following discoveries published in leading peered-review journals in neuroscience and pathology, as briefly described below.

Additional information and a list of projects are available from www.idiap.ch/genomics.

Advance in histopathological tissue sections analysis

We previously reported novel ALS disease phenotypes in tissue sections from mouse transgenic models and sporadic ALS post-mortem tissue (Luisier et al., 2018; Tyzack et al., 2019). These studies demonstrate that histopathological analysis of tissue sections is an invaluable resource in neurodegeneration research. However, cell-to-cell variation in both the presence and severity of a given phenotype is however a key limitation of this approach. The Genomics & Health Informatics group directly addressed these issues by combining automated image processing with machine learning methods to substantially improve the speed and reliability of identifying phenotypically diverse Motor Neurons (MNs) populations in collaboration with Patani's laboratory

and Serio's laboratory. The developed method enabled the following key advances which were published in Brain Pathology (Hagemann et al., 2021):

- Establishment of an image processing pipeline for automated identification and profiling of Motor Neurons in ALS pathological tissue sections: This approach enabled unbiased analysis of hundreds of cells, from which hundreds of features were readily extracted.

- Harnessing machine learning methods to robustly phenotype MNs at single cell resolution: We automated the identification of phenotypically distinct MNs subpopulations, and revealed common aberrant phenotypes related to cellular shape.

These findings showcase the potential of combining histopathology with automated image processing and machine learning and might prove transformational in our understanding of ALS and neurodegenerative diseases more broadly.

Discovery of a novel hallmark of Amyotrophic Lateral Sclerosis disease

We recently uncovered, in collaboration with Rickie Patani's laboratory (Francis Crick Institute), aberrant SFPQ intron-retaining transcripts (IRTs) in the cytoplasm and SFPQ protein mislocalization as a new hallmark of ALS (Luisier et al., 2018).

Following up on these findings, the Genomics & Health Informatics group developed bioinformatics methods to a new transcriptomic data-set, leading to key biological findings published in Brain (Tyzack et al, 2021). First, ALS-causing VCP gene mutations were showed to lead to accumulation of > 100 aberrant IRTs specifically within the cytoplasm, therefore revealing that cytoplasmic IRT is a widespread molecular phenomenon that has been somewhat overlooked. Second, we demonstrated that the RBPs known to be mislocalized in ALS (namely TDP-43, SFPQ and FUS), avidly bind specifically to the aberrant cytoplasmic IRT pool, as opposed to any individual IRT, which somewhat shifts the focus in ALS pathogenesis from exclusively aberrant protein localisation to include the aberrant localization of IRTs.

These findings provide fundamental insight into the current knowledge of post-transcriptional mechanisms underlying initial disease events in ALS and bear potential therapeutic significance.

Key publications

[1] M Petric-Howe, H Crerar, J Neeves, GE Tyzack, R Patani* and R Luisier*. "Diminished miRNA activity is associated with aberrant cytoplasmic intron retention in ALS pathogenesis," in BioaRxiv. *These authors contributed equally.

[2] C Andreassi*, R Luisier*, H Crerar, S Franke, NM Luscombe, G Cuda, M Gaspari and A Riccio. "3′ UTR Remodelling of Localized Transcripts in Sympathetic Neurons Axons," in Cell Reports (2021). *These authors contributed equally.

[3] GE Tyzack, J Neeves, P Klein, H Crerar, O Ziff, DM Taha, R Luisier*, NM Luscombe* and R Patani*. "An aberrant cytoplasmic intron retention programme is a blueprint for ALS-related RBP mislocalization," in Brain (2021). *These authors contributed equally.

[4] C Hagemann, GE Tyzack, DM Taha, H Devine, L Greensmith, J Newcombe, R Patani*, A Serio* and R Luisier*. "Automated and unbiased classification of motor neuron phenotypes with single cell resolution in ALS tissue," in Brain Pathology (2021). *These authors contributed equally.

[5] R Luisier, M Girgin, MP Lutolf, A Ranga. "Mammary epithelial morphogenesis in 3D combinatorial microenvironments," in Scientific Reports (2020).

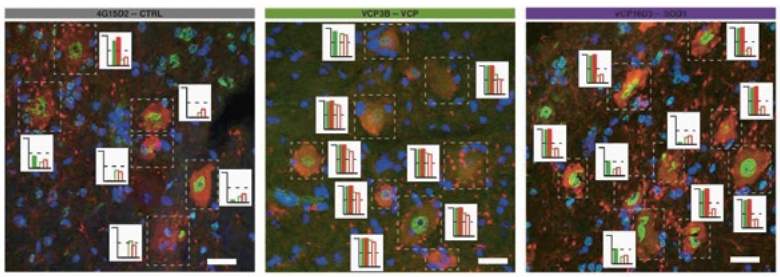


Figure 8

Examples of Motor Neurons in the ventral spinal cord of wild-type, ALS-related VCP and SOD1 mutant mice. MN cytoplasm is identified by ChAT immunofluorescence, nuclei are counterstained with DAPI, and SFPQ proteins are localised with immunofluorescence (green channel). For each Motor Neurons, the disease probability is obtained with different classifiers. Scale bar = 26 µm. (fig. from C.H. et al. Brain Pathology, 2021)

Robot Learning and Interaction

Overview

Head: Dr. Sylvain Calinon (MS and PhD, EPFL, 2003 and 2007)

Group overview

The Robot Learning and Interaction group, created in 2014, focuses on human-centered robotics applications in which the robots can acquire new skills by interactions, with strong generalization demands. It requires the development of intuitive interfaces to acquire meaningful demonstrations, the development of models that can exploit the structure and geometry of the acquired data in an efficient way, and the development of adaptive control techniques that can exploit the learned task variations and coordination patterns.

The Robot Learning & Interaction group in 2020 was composed of 3 postdoctoral fellows, 7 PhD students and 1 visiting PhD student.

Key scientific outputs

Development of learning techniques that only need a small number of demonstrations, by exploiting structures that can be found in a wide range of robotic tasks, and by exploiting bidirectional human-robot interaction as a way to collect better data.

Additional information and a list of projects are available from www.idiap.ch/rli.

Combination of controllers as a product of experts

We formulate the problem of combining controllers as a product of experts, where each expert can takes care of a specific aspect of the task to achieve. This approach allows the orchestration of different controllers, which can be learned separately or altogether (by variational inference). With this probabilistic formulation, the robot counteracts perturbations that have an impact on the fulfillment of the task, while ignoring other perturbations. This approach creates bridges with research in biomechanics and motor control, described under various terminologies, including minimal intervention principle, uncontrolled manifolds or optimal feedback control.

Geometry-aware learning and control

Data encountered in robotics are characterized by simple but varied geometries, which are often underexploited when developing learning and control algorithms. We exploit Riemannian geometry to extend algorithms initially developed for standard Euclidean data, by taking into account the structures of these manifolds in a unified manner.

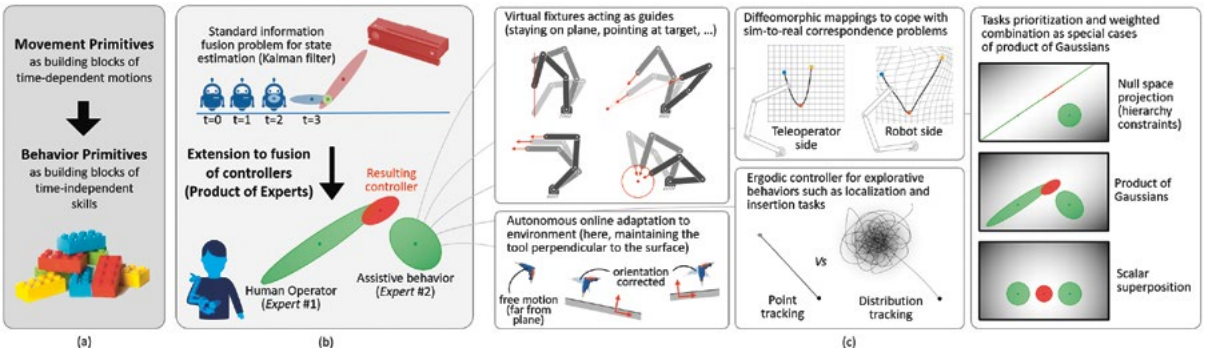


Figure 9

(a) We propose to extend the notion of movement primitives (superposition of basis functions in the form of time-dependent trajectories) to the notion of behavior primitives (superposition of time-independent controllers). **(b)** State estimation in robotics is typically formulated as an information fusion problem (Kalman filter combining a motion model with perception information, both taking into account uncertainties). The resulting combination of the two sources of information corresponds to a product of Gaussians. We propose to combine controllers in the same principled way, by using a product of experts (PoE) formulation in which the user is part of the shared control problem, with the other controllers assisting the user to achieve the task. **(c)** We developed behavior primitives covering a wide range of control strategies in both position and force domains, including the autonomous adaptation to object locations, the consideration of tasks variations and tasks prioritization, or the use of ergodic controllers for exploration behaviors.

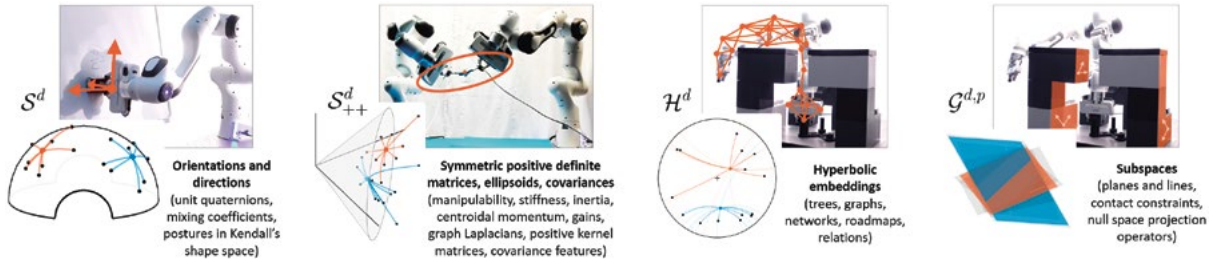


Figure 10

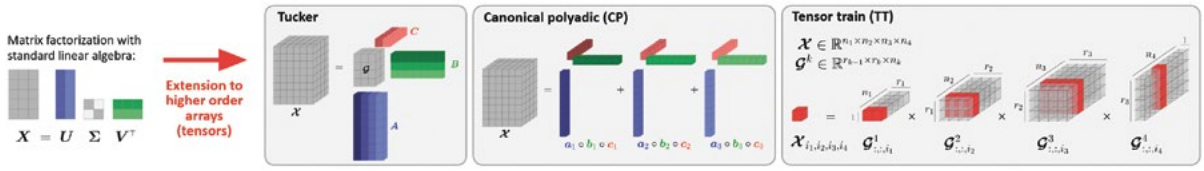


Figure 11

Tensor factorization in robotics applications

Another type of structure that we exploit relates to the organization of data as multidimensional arrays (also called tensors). These data appear in various robotic tasks, either as the natural organization of sensory/motor data (tactile arrays, images, kinematic chains), or as the result of standardized preprocessing steps (moving time windows, covariance features). Developed in the fields of multilinear algebra and tensor methods, these approaches extend linear factorization techniques such as singular value decomposition to multilinear decomposition, without requiring the transformation of the tensors into matrices or vectors. We exploit these techniques to provide robots with the capability to learn tasks from only few tensor datapoints, by relying on the multidimensional nature of the data.

Key publications

[1] Calinon, S., "Gaussians on Riemannian Manifolds: Applications for Robot Learning and Adaptive Control," IEEE Robotics and Automation Magazine (RAM), 27:2, pp. 33–45, 2020.

[2] Chatzilygeroudis, K., Vassiliades, V., Stulp, F., Calinon, S. and Mouret, J.-B., "A Survey on Policy Search Algorithms for Learning Robot Controllers in a Handful of Trials," IEEE Trans. on Robotics, 32:2, pp. 328–347, 2020.

[3] Jaquier, N., Rozo, L., Caldwell, D.G. and Calinon, S., "Geometry-aware Manipulability Learning, Tracking and Transfer," International Journal of Robotics Research (IJRR), 2020 (in press).

[4] Lembono, T.S., Paolillo, A., Pignat, E. and Calinon, S., "Memory of Motion for Warm-starting Trajectory Optimization," IEEE Robotics and Automation Letters (RA-L), 5:2, pp. 2594–2601, 2020.

Natural Language Understanding

Overview

Head: Dr. James Henderson (BSc, Massachusetts Inst. Technology, USA, 1987; MSE & PhD, Univ. Pennsylvania, USA, 1991,1994; MER & Chargé de Cours, Univ. Geneva, 2008–2012,2012–2018)

Group overview

The Natural Language Understanding group (NLU) works at the intersection of machine learning and natural language processing, with an emphasis on representation learning for the meaning of language, attention-based deep learning models, and structured prediction. We model summarization, abstraction (textual entailment), machine translation, knowledge extraction, syntactic and semantic structure, and lexical semantics, among other natural language processing problems (NLP). We develop deep learning models of the discovery and prediction of entities and their relations at multiple levels of representation for multiple tasks.

During 2020, the NLU group had the following members: the head of the group, and 7 PhD students. One PhD student became a postdoctoral researchers during the year.

Key scientific outputs

During 2020, the work of the NLU group has produced several key publications. A novel version of the Transformer deep learning architecture was developed which maps graphs to graphs, including a version which refines the output graph over several iterations. This is now the most accurate model of syntactic parsing on standard benchmarks. Two methods for improving portability and transfer for models of natural language inference were developed, and a novel embedding-based alternative to sequence-to-sequence models was proposed. Also, the head-of-group published a paper mapping the history and future of deep learning in NLU, which received a lot of attention.

Additional information and a list of projects are available from www.idiap.ch/nlp.

Deep Learning Architectures for Graphs

Deep learning models based on self-attention, in particular Transformer, have revolutionised the state-ofthe-art in many NLP tasks, particularly when combined with pretraining (e.g. BERT). We have developed a version of Transformer which maps graphs to graphs, instead of sequences to sequences. Given the nodes of a graph, such as the words of a sentence, Graph2Graph Transformer can input arbitrary graphs and output arbitrary graphs over these nodes, using the self-attention mechanism. We also developed Recursive Non-autoregressive Graph2Graph Transformer (Figure 13(a)) for the iterative refinement of predicted graphs. When combined with BERT pretraining, these models are now the most accurate models for syntactic parsing, a core benchmark for structured prediction and NLU. State-of-the-art results are also achieved in coreference resolution. We are currently investigating other structure prediction tasks and other NLP tasks which can benefit from structured inputs. An important current

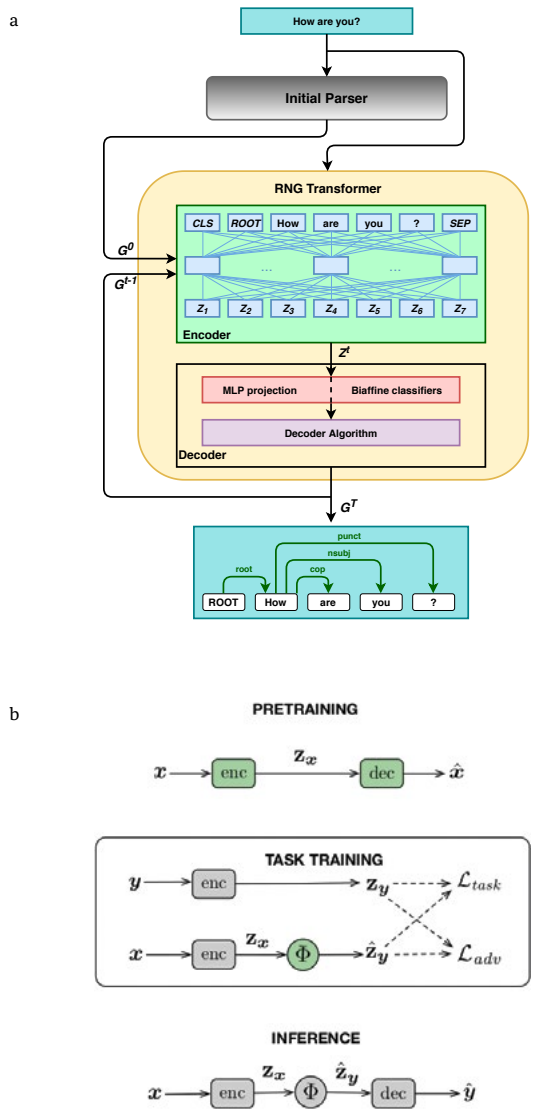


Figure 12

(a) Recursive Non-autoregressive Graph2Graph Transformer for syntactic parsing. **(b)** Embedding-to-Embedding training of sequence-to-sequence models.

topic of research is developing effective architectures for inducing the nodes of the graph, instead of having them pre-specified.

Representation Learning for NLP Tasks

Deep learning models of natural language induce hidden representations which can be effectively transferred across tasks. One topic we have addressed is how to improve this transfer by reducing the models' reliance on idiosyncrasies of the training data which don't generalise to the real task, called annotation biases. We train a bias-only model and use it to identify the biased datapoints, thereby improving performance on out-of-domain datasets which do not have the same biases. Another topic we have addressed is how to learn new tasks in the embedding space itself. After training a sequence-to-sequence auto-encoder, another sequence-to-sequence task can be learned by learning to map the encoding of the task's input to the encoding of the task's output. From this output embedding the output sequence can be generated with the decoder of the auto-encoder. This Embedding-to-Embedding method (Figure 13(b)) works well with small training datasets and on unsupervised tasks. We are currently working on alternative methods to improve model transfer and reduce bias, and on learning embedding-to-embedding mappings in other types of auto-encoder hidden representations.

Key publications

[1] Alireza Mohammadshahi and James Henderson. "Recursive Non-Autoregressive Graph-to-Graph Transformer for Dependency Parsing with Iterative Refinement". Accepted to Transactions of the Association for Computational Linguistics (TACL), arXiv:2003.13118 [cs.CL].

[2] Florian Mai, Nikolaos Pappas, Ivan Montero, Noah A. Smith, James Henderson. "Plug and Play Autoencoders for Conditional Text Generation". In Proc. of EMNLP 2020, Online, 2020.

[3] Alireza Mohammadshahi and James Henderson. "Graph-to-graph Transformer for Transition-Based Dependency Parsing". In Findings of ACL: EMNLP 2020, Online, 2020.

[4] James Henderson. "The Unstoppable Rise of Computational Linguistics in Deep Learning". In Proc. of ACL 2020 (theme track), Online, 2020.

[5] Rabeeh Karimi Mahabadi, Yonatan Belinkov and James Henderson. "End-to-End Bias Mitigation by Modelling Biases in Corpora". In Proc. ACL 2020, Online, 2020.

Computational Bioimaging

Overview

Head: Prof. Michael Liebling (MS, EPFL, 2000; PhD, EPFL 2004; postdoc, Caltech, 2004–2007; Assistant Prof (2007–2013), Associate Prof (2013–2017), Associate Adjunct Prof (2017–) UC Santa Barbara (UCSB)

Group overview

Research in the Computational Bioimaging Group focuses on developing image acquisition, reconstruction and analysis algorithms to study live biological systems. Practical tools aim at (i) extending the physical limits of imaging hardware via techniques including super-resolution and deconvolution methods, and (ii) quantitative analysis of complex biological systems: motion-based image analysis, cell tracking, microscopic fluid flow estimation, and integration of multi-modality images.

In 2020, the Computational Bioimaging Group was composed of the head of the group, four PhD students, one postdoc, and two interns.

Key scientific outputs

Recent milestones include temporal super-resolution methods for sensitive fluorescence cameras and approaches for quantitative heart development imaging in animal models.

Additional information and a list of projects are available from www.idiap.ch/cbi.

Temporal resolution doubling, aliasing mitigation in microscopy with modulated color illumination

As part of the SNSF project COMP-Bio "Computational Biomicroscopy: Advanced Image Processing Methods to Quantify Live Biological Systems," we developed image capture techniques specifically adapted for optical microscopy of biological samples.

Despite recent developments in microscopy, temporal aliasing remains an issue when imaging dynamic samples. Modern sampling frameworks, such as generalized sampling, mitigate aliasing but require the measurement of temporally-overlapping and potentially negative-valued inner-products. Conventional cameras cannot collect these directly, as they operate sequentially and are only sensitive to light intensity. We proposed to mitigate aliasing for monochrome samples by implementing a generalized sampling approach that uses a plain color camera with a custom modulated color illumination. Color solves the temporal overlap problem (spectral

multiplexing) and the use of B-spline segments ensures positivity for the projection kernels. Reconstruction requires spectral unmixing and inverse filtering. We implemented this method using a color LED illuminator and evaluated its performance by imaging a rotating grid and its applicability by imaging the beating zebrafish embryo heart in transmission and light-sheet microscopes. Compared to stroboscopic imaging, our method mitigates aliasing with performance improving as the projection order increases.

Still using modulated illumination, we further proposed a temporal resolution doubling method that is applicable for fluorescence light-sheet microscopy. This method is relevant for studying dynamic biological processes, such as heart development and function in zebrafish embryos, which often relies on multichannel fluorescence labeling to distinguish multiple anatomical features, yet also demands high frame rates to capture rapid cell motions. Specifically, we have extend a method we proposed previously, hueencoded shutter method (HESM), by using samples labeled with multiple fluorophores, whose emission signal can either be used to distinguish multiple anatomical features when imaged in multi-channel mode or, if the fluorophores are co-localized in a dynamic tissue, to increase the frame rate via HESM. We experimentally showed, in the case of a beating heart in a zebrafish embryo, that we can increase the frame rate by a factor two while preserving the ability to image static features labeled in distinct channels. With a suitable illumination setup and fluorescent labeling, the method could generalize to other applications where flexibility between multiple channel and high-speed fluorescence imaging is desirable. For fluorophores that are not co-localized, the imaging system is similar to a conventional light sheet microscope.

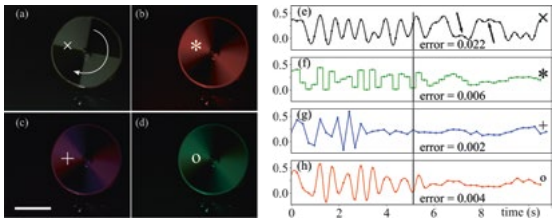


Figure 13

Evaluation of generalized sampling via modulated illumination framework. Our experimental setup consisted of a rotating grid that we imaged with different sampling modes in our framework. (a) Still frame of the grid as it rotates (arrow indicates direction), the rotation speed was linearly increased over time) under stroboscopic light. (b)-(d) Images of the same grid as in (a), but acquired using B-splines as pre-filters, with degrees (b) 0, (c) 1, and (d) 2. (e)-(h) time profile extracted from the reconstructed image series corresponding to the locations indicated by the symbols x, *, +, and o in (a), (b), (c), and (d), respectively. Sampling with higher degree pre-filters limits the effects of aliasing. Scale bars: 2 cm. (fig. adapted from C.J. et al. J. Biomed Opt., 2020

Our lab is further interested in developing effective volumetric microscopy methods, along with efficient 3D reconstruction techniques. In a study that involved Optical Projection Tomography (OPT), a technique that provides isotropic resolution for samples up to a few millimeters in size, we studied the relative merits of complex image acquisition protocols versus computationally more demanding deconvolution approaches. High resolution OPT can be achieved by deconvolving Focal-Plane-Scanning (FPS-OPT) data but it requires to accurately know the system's Point Spread Function (PSF). While the presence of noise and inaccuracies in the PSF model or parameters affects reconstruction quality, their effect is difficult to assess quantitatively in practice and the computational cost of naive simulations is prohibitively expensive. We developed an efficient approach to carry out FPS-OPT simulations for a wide range of illumination geometries, including Focal-Sheet-Scanning OPT (FSS-OPT), a method using a lateral light-sheet illumination to perform FPS-OPT. Our simulation framework can accomodate large size 3D data by dividing the forward model into elements that can be efficiently processed by GPUs. We compared the performance of FPS-OPT and FSS-OPT on simulated data. In the presence of Poisson noise, we show that FSS-OPT outperforms FPS-OPT with deconvolution even if all model parameters are accurately known.

Key publications

[1] Ch. Jaques and M. Liebling, "Aliasing mitigation in optical microscopy of dynamic biological samples by use of temporally modulated color illumination and a standard RGB camera," J. Biomed. Opt., 25(10), 106505, 2020

[2] Ch. Jaques, A. Ernst, N. Mercader and M. Liebling, "Temporal resolution doubling in fluorescence light-sheet microscopy via a hue-encoded shutter and regularization," OSA Continuum, 3(8), 2195–2209, 2020

[3] A. Shajkofci and M. Liebling, "Spatially-Variant CNN-Based Point Spread Function Estimation for Blind Deconvolution and Depth Estimation in Optical Microscopy," IEEE Trans. Image Proces., 29, pp. 5848–5861, 2020

[4] O. Mariani, A. Ernst, N. Mercader, and M. Liebling, "Reconstruction of image sequences from ungated and scanning-aberrated laser scanning microscopy images of the beating heart," IEEE Trans. Comput. Imag., 6, pp. 385–395, 2020.

[5] F. Marelli, and M. Liebling, "Optics versus computation: influence of illumination and reconstruction model accuracy in Focal-plane-scanning optical projection tomography," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), in press.

Biometrics Security and Privacy

Overview

Head: Dr. Sébastien Marcel (PhD, University of Rennes, France, 2000; Visiting Professor, University of Cagliari, 2010; Lecturer, EPFL since 2013; Lecturer, UNIL since 2018; Lecturer, Master AI since 2019)

Group overview

Biometrics refers to the automatic recognition of individuals based on their physiological and/or behavioural characteristics. The Biometrics Security and Privacy (BSP) group at Idiap focuses on four main areas of research:

- Biometric recognition: We investigate the development of accurate and fair (unbiased) recognition algorithms, notably for face, voice, and vein biometric modalities.
- Presentation attack detection (PAD): We look for new and better ways of detecting presentation attacks (direct attacks) on face, voice, and vein biometric recognition systems.
- Morphing attack detection and Deepfakes detection: We investigate the threat of morphing attacks (indirect attacks) and more generally Deepfakes to develop effective morphing and Deepfake detection techniques.
- Biometric template protection: We research effective methods of preserving both the security of biometric recognition systems and the privacy of their users by protecting the biometric models ("templates") that are employed by the system for recognition purposes.

The BSP group prioritises reproducibility in research. This is important for ensuring that our work can be both verified and built upon by the wider research community. To enable reproducibility, we mainly make use of our Python-based signal-processing and machine-learning toolbox, Bob (http://www.idiap. ch/software/bob/), which we make freely available for academic purposes. The group also develops and maintains the BEAT platform (https://www.beat-eu.org/platform/), an MLaaS platform compliant with Swiss and European data-security norms.

The group participates in several large-scale biometrics projects at Swiss (SNSF), European (H2020) or world-wide levels (eg. IARPA/DARPA) but also conducts projects directly with companies.

The BSP group provides also expertise to the Swiss Center for Biometrics Research and Testing, which, among other things, carries out tests and evaluations on commercial products related to biometrics.

The BSP group in 2020 was composed of 1 head of group, 4 research associates, 4 postdocs, 2 PhD students, 3 interns and 1 engineer.

Key scientific outputs

The BSP group has been pioneering the work on mobile biometrics (face and speaker recognition) and on PAD in face and speaker recognition by sharing the first open databases, organising the first International competitions and producing the first reproducible research studies in the domain. Regarding face PAD, the group confirmed that the current trend using discriminant classifiers is prone to over-fitting hence resulting in a lack of generalisation on unseen presentation attacks. These results question the efficiency and practicality of the existing PAD systems, as well as, call for creation of databases with larger variety of realistic presentation attacks. The BSP group also investigated approaches for heterogeneous face recognition, vein recognition and more recently Deepfakes to assess its threat to face recognition.

Additional information and a list of projects are available from www.idiap.ch/biometric.

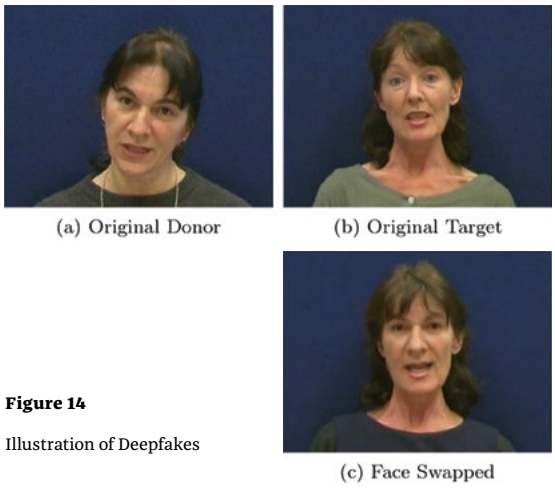


Figure 14

Illustration of Deepfakes

Deepfakes: a New Threat to Face Recognition?

It is becoming increasingly easy to automatically replace a face of one person in a video with the face of another person by using a pre-trained generative adversarial network (GAN). Recent public scandals call for automated ways to detect these Deepfake videos. To help developing such methods, we produced and analysed Deepfake videos. We showed that the state of the art face recognition systems based on Neural Networks are vulnerable to Deepfake videos, with about 90% false acceptance rates , which means methods

for detecting Deepfake videos are necessary. By considering several baseline approaches, we found that audio-visual approach based on lip-sync inconsistency detection was not able to distinguish Deepfake videos. The best performing method, which is based on visual quality metrics and is often used in presentation attack detection domain, resulted in 8.97% equal error rate on high quality Deepfakes. Our experiments demonstrate that GAN-generated Deepfake videos are challenging for both face recognition systems and existing detection methods, and the further development of face swapping technology will make it even more so.

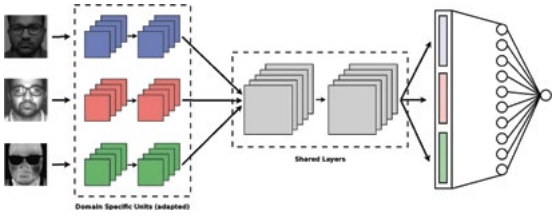
Presentation attack detection

One important aspect of biometric systems is their reliability not only when assaulted by impostors, but also under different types of attacks. One possible security treat is presentation attacks (aka spoofing attacks): an action of outwitting a biometric sensor by presenting a counterfeit biometric evidence of a valid user. It is a direct attack to the sensory input of the biometric system and the attacker does not need previous knowledge about the recognition algorithm. Most of the biometric modalities are not resistant to presentation attacks: a biometric system is usually designed to only recognise identities without concern whether the sample comes from a live person or not. Despite the existence of very sophisticated biometric systems nowadays, the task of implementing presentation attack detection (PAD aka anti-spoofing) schemes for them has attracted much less attention.

We developed a multi-spectral face device (Figure 16) to capture synchronised, high speed and high resolution image sequences under different image domains (VIS, NIR, SWIR, Thermal and 3D). The main hypothesis is that bona fide samples are easier to discriminate from presentation attacks with the appropriate combination of image domains. We proposed a novel Deep Convolutional Neural Network architecture (Figure 15) to learn multi-spectral complementary information.

Figure 15

Illustration of multi-spectral face Presentation Attack Detection (PAD) device **(right)** and architecture **(down)**.



Heterogeneous face recognition

The task of Heterogeneous Face Recognition (Figure 16) consists in to match face images that were sensed in different modalities, such as sketches to photographs, thermal images to photographs or near infrared to photographs. We demonstrated that high level features of Deep Convolutional Neural Networks trained on visual spectra images are domain independent and can be used to encode faces sensed in different image domains.

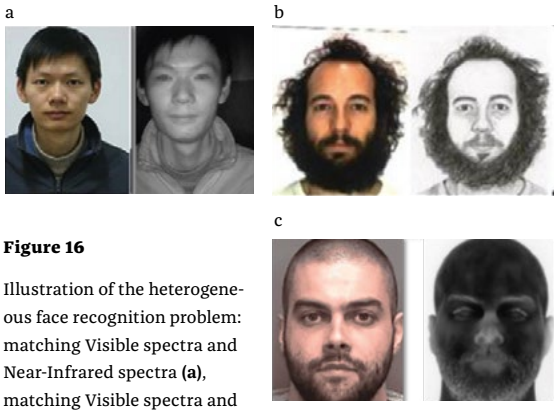


Figure 16

Illustration of the heterogeneous face recognition problem: matching Visible spectra and Near-Infrared spectra **(a)**, matching Visible spectra and sketch **(b)**, matching Visible spectra and Thermal spectra **(c)**.

Remote photoplethysmography

Photoplethysmography (PPG) consists in measuring the variation in volume inside a tissue, using a light source. The aim of remote photoplethysmography (rPPG) is to measure the same variations, but using ambient light instead of structured light and widely available sensors such as a simple webcam (Figure 17).

We presented a new, publicly available database containing a relatively large number of subjects recorded under two different lighting conditions. Also, three state-of-the-art rPPG algorithms from the literature were selected, implemented and released as open source free software.

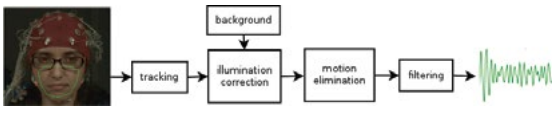


Figure 17

Illustration of remote photoplethysmography: colors from the video signal are filtered to produce an estimation of the heart beat signal.

Biosignal Processing

Overview

Swiss Centre for Biometrics Research and Testing

In 2014, the Idiap Research Institute launched the "Swiss Centre for Biometrics Research and Testing" (www.biometrics-center.ch), a competence centre within the Institute following recent successes in coordinating International research projects in Biometrics (MOBIO, TABULA RASA and BEAT). The aim of this centre is to serve as a legacy for these projects and to push for industry-driven research and testing in biometrics.

The centre attracted the attention of large companies (license, research and testing agreements) and led to many new projects (DARPA, IARPA, InnoSuisse). In 2020, the centre has developed over three directions:

- Maintaining and evolving the BEAT platform: The platform is now used in multiple research projects such as ALLIES, LEARN-REAL and the major H2020 AI4EU project. In parallel we are working towards improving the platform with the aim to create the Idiap AI platform.
- Evaluation and Testing: we engaged with the FIDO Alliance (https://fidoalliance.org), focused on providing open and free authentication standards to help reduce the world's reliance on passwords, and became an accredited FIDO lab able to perform certification of biometrics products. In 2020 we also became an Android accredited lab and we conducted several evaluations of biometric products.
- Joining the CITeR cooperative research center: we joined the US CITeR (Center for Identification Technology Research) National Science Foundation (NSF) Industry/University Cooperative Research Center (I/UCRC) as the first non-US academic site with two affiliates (IDEMIA and SCIPA) to fund our research activities.

Key publications

[1] G. Heusch, A. George, D. Geissenbuhler, Z. Mostaani and S. Marcel, "Deep Models and Shortwave Infrared Information to Detect Face Presentation Attacks", IEEE Transactions on Biometrics, Behavior, and Identity Science, 2020.

[2] T.d.F. Pereira and S. Marcel, "Fairness in Biometrics: a figure of merit to assess biometric verification systems", arXiv pre-print, 2020.

[3] A. George and S. Marcel, "Learning One Class Representations for Face Presentation Attack Detection using Multi-channel Convolutional Neural Networks", IEEE Transactions on Information Forensics and Security, 2020.

[4] P. Korshunov and S. Marcel, "Vulnerability assessment and detection of Deepfake videos", IAPR International Conference on Biometrics, pp 258–264, 2019.

[5] T. de Freitas Pereira, A. Anjos and S. Marcel, "Heterogeneous Face Recognition Using Domain Specific Units", IEEE Transactions on Information Forensics and Security, Vol 14(7), pp 1803–1816, 2019.

Head: Dr. André Anjos (BSc, MSc & PhD, Federal University of Rio de Janeiro, Brazil, 1999, 2001, 2006; Visiting Professor, State University of Bauru, 2015; Lecturer, EPFL, 2013–; Lecturer, Master of AI, 2019–)

Group overview

Biosignals are signals from living beings and their analysis to support clinical practice and research. This group currently focuses on image (e.g. retinography, X-ray, computed tomography), vital signs, clinical and multimodal data analysis for healthcare and related applications. Current trends in the field show refreshed interest on the use of machine learning techniques, complementing basic signal and sequence processing, all of which are key domains of research at Idiap. It leverages on Idiap's expertise on human subject handling, data acquisition, open science and data processing.

In 2020, the Biosignal Processing Group was composed of the head of the group, and 2 master students.

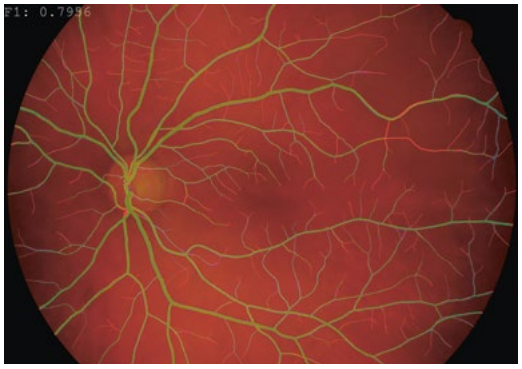
Key scientific outputs

The group currently develops 4 thematic areas of research: computer-aided diagnosis (CAD) from medical images (e.g. retinographies, chest X-ray, computed tomography), computer models for Tuberculosis (TB) detection and follow-up treatment, the analysis of time sequences with health data (e.g. vital signs or wearables data), and, finally, reproducibility in AI research. We are particularly interested in the clinical interpretability of AI models we build, as well as challenges associated to data scarcity in the medical domain. The milestones for 2021 include continued work in above cited areas, and the development of new projects on computed tomography (CT) scans for the detection of rare vessel diseases, our reproducible research platform, and finally continued participation in EPFL courses and the Idiap's AI master program. Additional funds have been request for the Swiss National Science Foundation for hiring two doctoral students.

Additional information and a list of projects are available from www.idiap.ch/bio-sig.

Semantic Segmentation for Medical Imaging

Since the introduction of U-Nets in 2015, the field of medical image segmentation has seen renewed interest bringing in a variety of fully convolutional (deep) neural network (FCN) architectures for binary and multi-class segmentation problems promising very attractive results, with applications in computed tomography, retinography, and histopathology to cite a few. Despite the incredible progress, the lack of annotated images (due to cost), and rigor in the comparison of trained models has led the community to believe larger and more dense networks provide better results. We addressed these



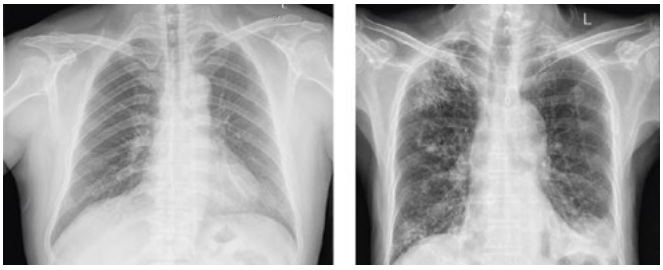


Figure 16

(left) Vessel segmentation for Retinography – predicted vessel maps vs. ground truths for a DL model evaluated on the high-resolution HRF test-set. True positives, false positives and false negatives are displayed in green, blue and red respectively. **(right)** Radiological signs on healthy (left) and PTB-affected lungs (right).

gaps in different ways. The first was to conduct and publish rigorous (open source, reproducible) benchmarks with popular retinography datasets and state-of-the-art FCN models in which we throw a new light over some published figures, opening space for new developments. In addition, we proposed a simple extension, of a compact U-Net architecture, dubbed W-Net, by concatenating two U-Nets together, which reaches outstanding performance on several popular datasets, still using orders of magnitude less learnable weights than any previously published approach.

Computer-aided Diagnosis for Tuberculosis

Tuberculosis (TB) is one of the leading causes of death from a single infectious agent. In many highburden regions around the world, which often lack specialized healthcare professionals, Chest X-Ray (CXR) exams continue to be of vital importance in the diagnosis and follow-up of the various presentations of the disease. In this context, we investigate the benefits of automatic Pulmonary Tuberculosis (PTB) detection methods based on radiological signs found on CXR. Contrary to direct scoring from images, implemented in most related work, indirect detection offers natural interpretability of automated reasoning. We identify generalization difficulties for direct detection models trained exclusively on the modest amount of publicly available CXR images from PTB patients. We subsequently show that a model, pre-trained on tens of thousands of CXR images using automatically annotated radiological signs, offers a more adequate base for development. By relaying radiological signs through a simple linear classifier, one is able to obtain state-of-the-art results on all three publicly available datasets. We further discuss limitations imposed by the limited number of PTB-specific radiological signs available on public datasets, and evaluate possible performance gains that could be obtained if more were available. This work is fully reproducible.

Reproducible Research

Since the last decade, we have been actively looking at the reproducibility of published work and how to lower the entrance barrier of publication readers. We argue it is insufficient, in most cases, to only publish software leading to results if original data remains inaccessible. Reproducibility should imply in the following characteristics: repeatability, share-ability, extensibility and stability, which is not guaranteed by most published material to date. We propose a software suite called Bob that possesses such characteristics, demonstrating its flexibility to various tasks. From another perspective, there are legitimate cases in which raw data leading to research conclusions cannot be published. Furthermore, in a growing number of use-cases, the availability of both software does not translate to an accessible reproducibility scenario. To bridge this gap, we built an open platform for research in computational sciences related to pattern recognition and machine learning, to help on the development, reproducibility and certification of results obtained in the field.

Key publications

[1] A. Galdran, A. Anjos, J. Dolz, H. Chakor, H. Lombaert, and I. Ben Ayed, "The Little W-Net That Could: State-of-the-Art Retinal Vessel Segmentation with Minimalistic Models", Submitted to Computerized Medical Imaging and Graphics, 2021.

[2] G. Raposo, A. Trajman, A. Anjos, "Towards Deep Interpretable Tuberculosis Detection from Chest X-Ray Images", Submitted to the 24th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2021.

[3] A. C. B. H. Ferreira, D. D. Ferreira, H. C. Oliveira, I. C. de Resende, A. Anjos, and M. H. B. de M. Lopes, "Competitive neural layer-based method to identify people with high risk for diabetic foot", Computers in Biology and Medicine, vol. 120, p. 103744, May 2020.

[4] A. Anjos, M. Günther, T. de Freitas Pereira, P. Korshunov, A. Mohammadi, and S. Marcel, "Continuously Reproducing Toolchains in Pattern Recognition and Machine Learning Experiments", at Thirty-fourth International Conference on Machine Learning, 2017.

22 https://www.idiap.ch/software/bob

23 https://www.idiap.ch/software/beat

Energy Informatics

Overview

Head: Dr. Jérôme Kämpf (BSc, University of Kent at Canterbury; MSc, University of Lausanne, 2001 and 2003; PhD, Ecole Polytechnique Fédérale de Lausanne, 2009)

Group overview

The Energy Informatics concepts are to exploit state-of-the-art Information and Communication Technologies to tackle global warming and climate change challenges. The aim is to increase the integration of renewable and distributed energy sources by making energy systems smarter, and to increase energy efficiency beyond what improvements at component level can achieve. In that vein, the Energy Informatics Group at Idiap researches into ways of simulating energy transition pathways with intelligent control and adjustment mechanisms of evolving shelters with retrofitting and use, renewable energy production and energy storage in a changing climate. Two key application areas are more deeply studied: the building automation with its adaptation to human behavior and the energy management with its indirect modeling of human activity to anticipate energy needs.

In 2020, the Energy Informatics Group was composed of 1 head of group, 2 postdocs, 2 exchange PhD students, 1 exchange MSc student and 3 interns.

Key scientific outputs

In building automation, surrogate models were developed to evaluate glare in the eye of the occupants and work plane illuminance. The proposed machine learning models being slightly less precise than physical models yet much faster, allow for their implementation in embarked computers to control in quasi real-time blinds and electric lighting in office buildings.

In energy management, a virtual model of the whole city of Fribourg was realized in 2.5D to study the Urban Heat Island effect in future climates and mitigation scenarios including vegetation, water bodies and albedo change. The methodology was based on the open-source Urban Energy Simulation tool CitySim. CitySim is now further developed within the group to provide such specific answers to climate change challenges.

⁂ **Additional information and a list of projects are available from** www.idiap.ch/energy.

Building automation

Building occupants tend to negatively perceive building automation as it may alter their comfort for the sake of energy savings. The main challenge identified is therefore to control (or suggest actions on) the building infrastructure in order to minimise the energy intake while maintaining a sufficient comfort for the occupants. Noteworthy, human comfort has multiple facets and depends on each individual's history and preferences. The use of smart sensors is a unique opportunity to learn and adapt the automation to the users.

Through the thesis of an exchange Master student from Politecnico di Milano, we completed the equipment of 3 rooms in the Idiap building with CO2, temperature, relative humidity sensors and smart thermostats for radiators. Pipeline Machine Learning models were developed to evaluate occupants' presence, window and door states through an analysis of the monitoring data feed from the sensors. A further thermal model was developed to optimise the continuous valve position of the smart thermostats for radiators as a function of occupants' presence, window and door states. On a longer term, the work is the premises for prediction algorithms of the air quality and energy consumption aiming at an optimal control for both low-tech and high-tech buildings.

Through the project LUCIDELES, we study the trade-off between energy and visual comfort in buildings. Machine learning based surrogate models were developed to simulate the daylight penetration through Venetian blinds and its effect on work-plane illuminance and glare perceived by the occupants. These predictive models were then used to control optimally the blinds' state to avoid glare while maximising daylight in buildings, thus minimising the need for electric lighting. A first laboratory environment named CELLS for testing the algorithms was set-up at the BlueFactory in Fribourg within the framework of the IEA SHC Task 61 (task61.iea-shc.org) on Integrated Solutions for Daylighting and Electric Lighting: From component to user centered system efficiency.

Energy management

While urban areas cover about 2% of the planet surface, they consume 3/4 of the total resources. According to the Pareto rule, major energy-related efforts should target cities and their inhabitants. The efforts comprises the integration of renewable and distributed energy sources, which in dense urban centers is a challenge. In particular the mutual shading between buildings and vegetation affect the active and passive solar performance, while improving the urban comfort and heat island.

Through the projet ICU, we evaluated using physically-based simulations the Urban Heat Island (UHI) effect in the city of Fribourg. Diverse evolution scenarios of the city under potential future climates were investigated, leading to a quantified evaluation of the temperature increase in the city due to climate change. A database was set-up with the information available on the built and natural environements in Fribourg, and linked to the urban energy simulator CitySim to predict the hourly dynamics of the surface temperatures. The results, available in false-color maps representing the surface temperatures, can be visualised using any GIS interface for their dissemination.

Retrofitting buildings in city districts is often a complex trade-off between the financial investment, grey energy and associated CO_2 emissions for a specific reduction of the annual energy demand. Our multiobjective optimisation approach based on the CMA-ES/HDE hybrid Evolutionary Algorithm was used by an exchange PhD student from the Free University of Bolzano. He linked it to the tool CitySim to compute the annual energy demand of buildings for space heating, while evaluating the financial, grey energy and CO_2 emissions associated to the envisaged renovations using its life cycle analysis module. The proposed results can be adopted to support the decision makers in the development of efficient strategies to apply at city or district scale.

Through the project EnerMaps, we prepared a unique database schema to store 50 energy-related datasets that will be displayed in a web-based graphical user interface named the EnerMaps Data Management Tool. Meanwhile we started the project Eguzki with the evaluation of physically-based simulation tools for District Heating Networks, comparing their performance with real monitoring provided by partner energy providers. AI-based surrogate models will be developed based on the simulation results.

CREM – Centre de Recherches Énergétiques et Municipales

We maintain a close scientific collaboration with CREM (www.crem.ch), an actor specialized in the field of energy sustainability in urban areas. CREM has a large network of communes, public utilities and companies making it an essential partner for case-studies and dissemination. We share a Master AI student on applied artificial intelligence for District Heating Network's substations data feed.

Key publications

[1] K. H. Poon, J. Kämpf, S. E. R. Tay, N. H. Wong and T. G. Reindl, Parametric study of URBAN morphology on building solar energy potential in Singapore context, in: Urban Climate, 33(100624), 2020

[2] C. Basurto and J. Kämpf, An Integrated and strategic evaluation of automatic blind controls to achieve energy and occupant's comfort objectives, in: Proceedings of the 5th IBPSA-England Conference on Building Simulation and Optimization (Virtual), Loughborough, UK, 2020

Active and Granted Projects in 2020

An overview of the projects that have been active during the year 2020 is presented in Section 4.1. The projects are grouped in three categories, namely National Research Projects, European and International Research Projects and Industry-oriented Projects. Finally the list of projects accepted during 2020 but starting in 2021 is presented.

Projects in Progress during 2020

National Research Projects

Name	**ADEL** (Automatic Detection of Leadership from Voice and Body)
Funding	UNIL
Coordinator	University of Lausanne
Duration	2020.06.01 – 2021.05.31
Partner(s)	Idiap Research Institute, IMD Switzerland, École Polytechnique Fédérale de Lausanne

Name	**ADIVA** (Automatic analysis of verbal and non-verbal behavior and provision of feedback in video selection interviews)
Funding	SNF – Division I
Coordinator	University of Neuchâtel
Duration	2018.10.01 – 2020.03.31
Partner(s)	Idiap Research Institute, University of Lausanne

Name	**ALLIES** (Autonomous Lifelong learnIng intelLigent Systems)
Funding	SNF -ERA-NET
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2021.09.30
Partner(s)	Laboratoire national de métrologie et d'essais, Université du Maine, Universitat Politecnica de Catalunya

Name	**CODIMAN** (A future that works: Cobotics, digital skills and the re-humanization of the workplace)
Funding	SNF – NRP77
Coordinator	Berner Fachhochschule
Duration	2020.05.01 – 2024.04.30
Partner(s)	Idiap Research Institute

Name	**COMPBIO** (Computational biomicroscopy: advanced image processing methods to quantify live biological systems)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2018.04.01 – 2022.03.31

Name	**DOMAT** (On-demand Knowledge for Document-level Machine Translation)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2022.09.30
Partner(s)	HES-SO Vaud

Name	**DUSK2DAWN** (Characterizing Youth Nightlife Spaces, Activities, and Drinks)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2017.07.01 – 2021.06.30
Partner(s)	University of Zurich, La Trobe University

Name	**EVOLANG** (The Origins and Future of Language)
Funding	SNF – NCCR
Coordinator	University of Zurich
Duration	2020.06.01 – 2024.05.31
Partner(s)	Idiap Research Institute, Ecole Polytechnique Fédérale de Lausanne, Eidgenoessische Technische Hochschule Zuerich, University of Basel, University of Fribourg, University of Geneva, University of Lausanne, University of Neuchâtel, Zurich University of the Arts

Name	**FLOSS** (Flexible Linguistically-guided Objective Speech aSessment)
Funding	Hasler Foundation
Coordinator	Idiap Research Institute
Duration	2017.03.01 – 2020.03.31

Name	**HEALTHVLOGGING** (Social media culture and the (re)shaping of healthrelated practices by Youtubers)
Funding	SNF Spark
Coordinator	University of Lausanne
Duration	2019.12.01 – 2021.03.31
Partner(s)	Idiap Research Institute

Name	**HEAP** (Human-Guided Learning and Benchmarking of Robotic Heap Sorting)
Funding	SNF – ERA NET
Coordinator	University of Lincoln
Duration	2019.04.01 – 2022.03.31
Partner(s)	Idiap Research Institute, Istituto Italiano di Tecnologia, Institut de Recherche en Informatique et en Automatique, Technische Universitaet Wien

Name **INTREPID** (Automated interpretation of political and economic policy documents: Machine learning using semantic and syntactic information)
Funding SNF – Sinergia
Coordinator Graduate Institute of International and Development Studies
Duration 2019.01.01 – 2022.12.31
Partner(s) Idiap Research Institute

Name **IPEQ** (Uncertainty quantification and efficient design of experiments for data- and simulation-driven inverse problem solving)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2018.11.01 – 2022.08.31
Partner(s) University of California at Davis

Name **ISUL** (Importance sampling for large-scale unsupervised learning)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2020.02.28

Name **LAB_FIDO** (Création d'un laboratoire de test biométrique agréé)
Funding Loterie Romande
Coordinator Idiap Research Institute
Duration 2018.12.01 – 2020.01.31

Name **LAOS** (Learning Representations of Abstraction in Text)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2018.11.01 – 2022.09.30

Name **LEARN-REAL** (LEARNing physical manipulation skills with simulators using REAListic variations)
Funding SNF – ERA NET
Coordinator Idiap Research Institute
Duration 2019.04.01 – 2022.03.31
Partner(s) École Centrale de Lyon, Istituto Italiano di Tecnologia

---

Name **MASS** (Multilingual Affective Speech Synthesis)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2017.05.01 – 2020.09.30

Name **MOSPEEDI** (Motor Speech Disorders)
Funding SNF – Sinergia
Coordinator University of Geneva
Duration 2017.10.01 – 2020.09.30
Partner(s) Idiap Research Institute, University Hospitals of Geneva, Université Paris 3

Name **NAST** (Neural Architectures for Speech Technology)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2020.02.01 – 2024.01.31

Name **NATAI** (The Nature of Artificial Intelligence)
Funding SNSF – Agora
Coordinator Idiap Research Institute
Duration 2020.10.01 – 2023.09.30
Partner(s) Musée de la Main UNIL/CHUV

Name **NKBP** (Deep Learning Models for Continual Extraction of Knowledge from Text)
Funding SNF – Division II Lead Agency
Coordinator Idiap Research Institute
Duration 2020.06.01 – 2024.05.31
Partner(s) Katholieke Universiteit Leuven

Name **ROSALIS** (Robot skills acquisition through active learning and social interaction strategies)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2018.04.01 – 2022.03.31

Name **SHISSM** (Sparse and hierarchical Structures for Speech Modeling)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2018.01.01 – 2021.12.31

Name **SMILE** (Scalable Multimodal sign language Technology for sIgn language Learning and assessmEnt)
Funding SNF – Sinergia
Coordinator Idiap Research Institute
Duration 2016.03.01 – 2020.10.31
Partner(s) University of Surrey, University of Applied Sciences of Special Needs Education

---

Name **TACT-HAND** (Improving control of prosthetic hands using tactile sensors and realistic machine learning)
Funding SNF – DACH
Coordinator Deutsches Zentrum Fuer Luft – Und Raumfahrt Ev
Duration 2016.04.01 – 2020.05.31
Partner(s) Idiap Research Institute, Universitaet Bielefeld

Name **TIPS** (Towards Integrated processing of Physiological and Speech signals)
Funding SNF Division II
Coordinator Idiap Research Institute
Duration 2019.11.01 – 2023.10.31
Partner(s) Centre Suisse d'Électronique et de Microtechnique, Coaching & Moderation

---

Name **TOT** (Trust Over Time)
Funding EPFL
Coordinator École Polytechnique Fédérale de Lausanne
Duration 2019.10.01 – 2020.09.30
Partner(s) Idiap Research Institute, Radio Television Suisse

Name **VERIFAKE** (Vulnerability assessment and detection of Deepfake videos)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2019.06.01 – 2020.05.31

European and International Research Projects

Name	**4DHEART** (4D analysis of heart development and regeneration using advanced light microscopy)
Funding	H2020-MSCA
Coordinator	Fundacion Centro Nacional de Investigaciones Cardiovasculares Carlos Iii
Duration	2016.10.01 – 2020.09.30
Partner(s)	Idiap Research Institute, Centre National de La Recherche Scientifique, Universität Bern, Acquifer AG, Bitplane AG, Leica Microsystems Cms Gmbh, 4D-Nature ImAGing Consulting, S. L., Centre Europeen de Recherche En Biologie et Medecine

Name	**AI4EU** (A European AI On Demand Platform and Ecosystem)
Funding	H2020-RIA-ICT
Coordinator	Thales Services SA
Duration	2019.01.01 – 2021.12.31
Partner(s)	Idiap Research Institute, ABB AS, AGI Research SRO, Allianz SE, Atos Spain SA, Aristotle University of Thessaloniki, Blumorpho SAS, Budapest University of Technology and Economics, Bureau de Recherches Géologiques et Minières, Barcelona Supercomputing Center, CARTIF Foundation, Commissariat à l'énergie atomique et aux énergies alternatives, CINECA - Consorzio Interuniversitario, Consiglio Nationale delle Ricerche, Centre National de la Recherche Scientifique, Agencia Estatal Consejo Superior De Investigaciones Cientifícas, National Center for Sicentific Research "Demokritos", German Research Center for Artificial Intelligence, German Aerospace Center, EIT Digital, Eötvös Loránd University, European Organisation for Security, FundingBox Research, Fondazione Bruno Kessler, Fraunhofer Gesellschaft, France Digitale, Foundation for Research and Technology – Hellas, Forum Virium Helsinki, Grassroots Arts and Research, France Hub IA, Industrial Data Space e. V., Institut Mines-Télécom-IMT, Institut National de Recherche en Informatique et Automatique, Associacao Do Instituto Superior Tecnico Para A Investigacao E Desenvolvimento, Centre for Research and Technology Hellas, Jožef Stefan Institute, Karlsruhe Institute of Technology, Know-Center Gmbh Research Center for Datadriven Business & Big data Analytics, University of Leeds, Loupe 16, Università degli studi di Roma "La Sapienza", Norwegian University of Science and Technology, National University of Ireland Galway, Office National d'Etudes et Recherches Aérospatiales, Orange SA, Örebro University, PG WConsulting, Université Paris I Panthéon-Sorbonne, QWANT, Siemens AG, SAP SE, Smile, Smart Rural, Simula Research Laboratory, Thales Alenia Space, Thomson Licensing, Telenor ASA, Tilde SIA, Thales Research & Technology France, Technische Universität Berlin, Delft University of Technology, Centre for Intelligent Technologies, Technical University of Munich, Technical University of Vienna, Alma Mater Studiorum – University of Bologna, University College Cork, University of Coimbra, Université Grenoble Alpes, Unilever UK, National and Kapodistrian University of Athens, Universitat Politecnica de Catalunya, Technical University of Madrid, Universita Degli Studi Di Siena, Sorbonne University, Ca' Foscari University of Venice, Vrije Universiteit Brussel, Women in AI, Wavestone

Name	**AI4MEDIA** (A European Excellence Centre for Media, Society and Democracy)
Funding	H2020-ICT
Coordinator	Centre for Research and Technology Hellas
Duration	2020.09.01 – 2024.08.31
Partner(s)	Idiap Research Institute, Athens Technology Center, Aristotelio Panepistimio Thessalonikis, Commissariat a l'Energie Atomique et aux Energies Alternatives, Communaute d'Universites et Etablissements Universite Cote D'Azur, Deutsche Welle, Institute of Information Science and Technologies, F6S Network Limited, Flemish Radio and Television Broadcasting Organization, Fraunhofer-Gesellschaft, Globalz SA, Grassroots Arts And Research UG, HES-SO Valais, IBM Ireland Limited, Imagga Technologies Ltd, Interdigital R&D France, Institut de Recherche et de Coordination Acoustique Musique, Joanneum Research Forschungsgesellschaft, Katholieke Universiteit Leuven, Modl.ai ApS, Polytechnic University of Bucharest, Queen Mary University of London, RAI-Radiotelevisione Italiana SPA, The Netherlands Institute for Sound and Vision, University of Amsterdam, University of Florence, University of Malta, University of Trento

Name	**ATCO2** (Automatic collection and processing of voice data from air-traffic communications)
Funding	H2020-CSJU
Coordinator	Idiap Research Institute
Duration	2019.11.01 – 2022.02.28
Partner(s)	Brno University of Technology, OpenSky Network, ReplayWell, Romagna Tech, Evaluations and Language Resources Distribution Agency, University of Saarland

Name	**BATL** (Biometric Authentification with Timeless Learner)
Funding	USA IARPA
Coordinator	University of Southern California
Duration	2017.03.01 – 2021.05.31
Partner(s)	Idiap Research Institute

Name **COLLABORATE** (Co-production CeLL performing Human-Robot Collaborative AssEmbly)
Funding H2020-RIA-DT
Coordinator Aristotle University Of Thessaloniki
Duration 2018.10.01 – 2022.03.31
Partner(s) Idiap Research Institute, Arcelik A.S., Association pour la Recherche et le Développement des méthodes et processes, ASTI Mobile Robotics, Blue Ocean Robotics APS, Centre for Research and Technology Hellas, Centro Ricerche Fiat SCPA, Jozef Stefan Institute, Katholieke Universiteit Leuven, Kolektor, University of Patras, Pratt & Whitney Rzeszów, Universita Degli Studi di Genova

---

Name **ENERMAPS** (Open Source Tools to Share, Compare, and Reuse Low-Carbon Energy Data)
Funding H2020
Coordinator Centre de recherches énergetiques et municipales
Duration 2020.04.01 – 2022.03.31
Partner(s) Idiap Research Institute, Zentrum Für Energiewirtschaft Und Umwelt (E-THINK), Accademia Europea di Bolzano, OpenAire Make, Revolve Media, Technische Universitaet Wien

---

Name **HAAWAII** (Highly Automated Air Traffic Controller Workstations with Artificial Intelligence Integration)
Funding H2020
Coordinator Deutsches Zentrum Fuer Luft und Raumfahrt Ev
Duration 2020.06.01 – 2022.11.30
Partner(s) Idiap Research Institute, Austro Control, Croatia Control, Isavia OHF, NATS, Brno University of Technology

---

Name **ICARUS** (Innovative AppRoach to Urban Security)
Funding H2020
Coordinator Forum Européen Pour La Sécurité Urbaine
Duration 2020.06.01 – 2024.05.31
Partner(s) Idiap Research Institute, Ethical and Legal Plus S.L., Erasmus Universiteit Rotterdam, Eurocircle Association, Fachhochschule Salzburg GmbH, Globalz SA, KENTRO MELETON ASFALEIAS, University of Leeds, Lisbon Municipal Police, Makesense, Commune de Nice, Panteion University Of Social And Political Sciences, Riga Municipal Police, City of Rotterdam, Landeshauptstadt Stuttgart, City of Torino, University of Salford

---

Name **MEMMO** (Memory of Motion)
Funding H2020-RIA-ICT
Coordinator Centre national de la recherche scientifique
Duration 2018.01.01 – 2022.06.30
Partner(s) Idiap Research Institute, University of Edinburgh, Max Planck Society for the Advancement of Sciences, University of Oxford, PAL ROBOTICS SL, AIRBUS SAS, Wandercraft, Centre de médecine physique et de réadaptation, Costain Group PLC

Name **MUMMER** (MultiModal Mall Entertainment Robot)
Funding H2020-ICT
Coordinator University of Glasgow
Duration 2016.03.01 – 2020.02.28
Partner(s) Idiap Research Institute, Centre National de La Recherche Scientifique, Aldebaran Robotics, Teknologian Tutkimuskeskus Vtt, Kiinteistö Oy Ideapark Ab

---

Name **ROXANNE** (Real time network, text, and speaker analytics for combating organized crime)
Funding H2020-SU-SEC
Coordinator Idiap Research Institute
Duration 2019.09.01 – 2022.08.31
Partner(s) Trilateral Research LTD, Brno University of Technology, Phonexia s.r.o., SAIL LABS Technology GmbH, Capgemini Technology Services, The International Criminal Police Organization, Saarland University, KENTRO MELETON ASFALEIAS, Gottfried Wilhelm Leibniz Universitaet Hannover, Università Cattolica del Sacro Cuore – Transcrime, AEGIS IT RESEARCH UG, AIRBUS Defence and Space SAS (Innovation Coordinator), Police of Czech Republic, Romanian Minister of Interior, Lithuanian Forensic Science Centre, Police Service of Northern Ireland, ADITESS Advanced Integrated Technology Solutions & Services LTD, Ministry of Interior Croatia, Netherlands Forensic Institute, Internet of Things applications and Multi- Layer development, Ministry Of Public Security – Israel National Police, Hellenic Police, An Garda Síochána

---

Name **SARAL** (Summarization and domain-Adaptive Retrieval of Information Across Languages)
Funding USA IARPA
Coordinator University of Southern California
Duration 2017.10.01 – 2021.08.31
Partner(s) Idiap Research Institute, Massachusetts Institute of Technology, Raytheon Company, Reenselaer Polytechnic Institute, University of Massachusetts Amherst, Northeastern University

---

Name **SAVI** (Spotting Audio-Visual Inconsistencies)
Funding USA DARPA
Coordinator Sri International
Duration 2016.05.19 – 2020.05.18
Partner(s) Idiap Research Institute

Name **SWAN** (Secure Access Control over Wide Area Network)
Funding Research Council of Norway
Coordinator Hogskolen I Gjovik
Duration 2016.01.01 – 2020.01.31
Partner(s) Idiap Research Institute, Morpho, Bankenverband, Universitetet I Oslo, Zwipe As

Name	**TAPAS** (Training Network on Automatic Processing of PAthological Speech)
Funding	H2020-MSCA
Coordinator	Idiap Research Institute
Duration	2017.11.01 – 2021.10.31
Partner(s)	University of Sheffield, Philips, Radboud University Nijmegen – Stichting Katholieke Universiteit, Ludwig-Maximilians-Universität München, Institut de Recherche En Informatique de Toulouse, Antwerpen University Hospital, Friedrich-Alexander-Universität Erlangen Nuernberg, Instituto de Engenharia de Sistemas E Computadores, Investigacao E desenvolvimento Em Lisboa, Interuniversitair Micro-Electronica Centrum Imec Vzw, Stichting Het Nederlands Kanker Instituut – Antoni Van Leeuwenhoek Ziekenhuis, Universitaet Augsburg

Name	**TRESPASS-ETN** (TRaining in Secure and PrivAcy-preserving biometricS)
Funding	H2020-MSCA
Coordinator	Eurecom
Duration	2020.01.01 – 2023.12.31
Partner(s)	Idiap Research Institute, Hochschule Darmstadt, Chalmers Tekniska Hoegskola AB, Katholieke Universiteit Leuven, Rijksuniversiteit Groningen, Universidad Autónoma de Madrid

Name	**WENET** (The Internet of US)
Funding	H2020-RIA-FETPROACT
Coordinator	University of Trento
Duration	2019.01.01 – 2022.12.31
Partner(s)	Idiap Research Institute, Aalborg University, Amrita Vishwa Vidyapeetham, Ben-Gurion University of the Negev, University of Tübingen, Instituto Potosino de Investigacion Cientifica y Tecnologica, Jilin University, London School of Economics and Political Science, Martel GmbH, National University of Mongolia, Open University of Cyprus, U-Hopper SRL, Universidad Catolica Nuestra Senora de La Ascuncion

Industry-oriented Projects

Name	**AADES** (Adaptive and Asynchronous Detection and Segmentation)
Funding	Armasuisse
Coordinator	Idiap Research Institute
Duration	2018.10.01 – 2021.09.30

Name	**ADVANCE** (Augmented dialogue tool based on verbal and non-verbal behavior computing)
Funding	Innosuisse
Coordinator	Idiap Research Institute
Duration	2019.03.01 – 2020.08.31
Partner(s)	CM Profiling, HES-SO Fribourg

Name	**AMICLEAR** (Détection de fibres d'AMIante par CLassification d'images de microscopie Électronique pour l'Analyse de Résidus)
Funding	Fondation The Ark
Coordinator	Idiap Research Institute
Duration	2019.08.01 – 2020.04.01
Partner(s)	Amiscan

Name	**AMS_SPONSORSHIP** (Sensor Fusion and Active Sensing for World-View Understanding)
Funding	Industrial
Coordinator	Idiap Research Institute
Duration	2019.09.01 – 2024.08.31
Partner(s)	ams

Name	**BOAT** (Automated Braces generation for Orthopaedic Anatomical Treatment of fractures)
Funding	Innosuisse
Coordinator	Idiap Research Institute
Duration	2019.04.05 – 2020.08.04
Partner(s)	HES-SO Fribourg, Swibrace SA

Name	**COB'HOOK**
Funding	Innosuisse
Coordinator	Idiap Research Institute
Duration	2019.05.01 – 2020.10.31
Partner(s)	Richemont

Name	**COMINT** (Speech Technology for COMINT)
Funding	Armasuisse
Coordinator	Idiap Research Institute
Duration	2020.09.01 – 2023.08.30

Name	**DAHL** (Domain Adaptation via Hierarchical Lexicons)
Funding	Industrial
Coordinator	Idiap Research Institute
Duration	2019.11.01 – 2020.10.31
Partner(s)	Swisscom

Name	**EGUZKI** (Programme de simulation de réseaux de chauffage à distance basé sur l'intelligence artificielle pour la résolution rapide et prédictive de réseaux complexes bouclés.)
Funding	OFEN
Coordinator	RWB Valais SA
Duration	2020.09.01 – 2023.08.30
Partner(s)	Idiap Research Institute, Altis Groupe SA, Oiken SA, SATOM SA

Name **GAZESENSESCREEN** (GazeSense Screen)
Funding Fondation The Ark
Coordinator Eyeware
Duration 2020.01.01 – 2020.11.30
Partner(s) Idiap Research Institute

Name **HAPPYGLOVE** (Electronic goniometer hAPPy Glove feasibility study)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2020.02.03 – 2020.09.30
Partner(s) Kinntek Sarl

Name **ICU** (Ilots de chaleur en ville de Fribourg : identification, anticipation et stratégie d'adaptation et de valorisation)
Funding OFEN
Coordinator HES-So Fribourg
Duration 2019.01.01 – 2020.12.31
Partner(s) Idiap Research Institute

Name **INNO-MOBILET** (Smart Battery Network)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2020.09.01 – 2021.02.28
Partner(s) Mobi-Let Sarl

Name **INNO-SIMPLYHOME** (Improving indoor environment, energy consupmtion and comfort in existing buildings)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2019.10.01 – 2020.03.31
Partner(s) Cleveron AG

Name **IVECT** (Impact of greening on the energy balance and thermal comfort of buildings and districts)
Funding OFEN
Coordinator HES-So Valais
Duration 2020.12.01 – 2023.11.30
Partner(s) Idiap Research Institute, Centre de recherches énergetiques et municipales, État du Valais, City of Zurich, Zurich University of Applied Sciences

---

Name **LUCIDELES** (Leveraging User-Centric Intelligent Daylight and Electric Lighting for Energy Saving)
Funding OFEN
Coordinator University of Fribourg
Duration 2020.02.01 – 2021.12.31
Partner(s) Idiap Research Institute, Regent lighting AG

Name **MALAT** (Machine Learning for Air Traffic)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2020.03.01 – 2023.02.28
Partner(s) SkySoft ATM

Name **MARGIN** (Multi-modal federated age verification)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2020.07.01 – 2021.12.31
Partner(s) Privately SA, EPFL

Name **RISE** (Rich Interpersonal Skill analytics for rEcruitment)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2018.05.01 – 2020.04.30
Partner(s) University of Lausanne

Name **SEWS2** (Smart Early Waning Score System for in and out-hospital care via anomaly detection)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2019.11.01 – 2020.05.31
Partner(s) Vtuls Sàrl

Name **SHAPED** (Speech Hybrid Analytics Platform for consumer and Enterprise Devices)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2018.03.01 – 2020.02.29
Partner(s) Logitech S.A.

Name **SM2** (Extracting Semantic Meaning from Spoken Material)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2018.11.01 – 2020.04.30
Partner(s) Recapp

Name **SPORTPROFILING** (Sport Profiling)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 2019.08.01 – 2020.06.01
Partner(s) Action Types Swiss Sàrl, ProKey Coach

Name **SRML** (Super-resolution through Machine Learning)
Funding Armasuisse
Coordinator Idiap Research Institute
Duration 2020.07.01 – 2021.01.31

Name **USP** (Unique Stability Plates: Advanced Aluminium Solution for High Precision Milling)
Funding Innosuisse
Coordinator Idiap Research Institute
Duration 2019.05.01 – 2020.11.01
Partner(s) Allega GmbH, Constellium Valais SA, Hes-so Valais

Projects Awarded in 2020 and Starting in the following year

Name	**EMIL** (Emotion in the loop – a step towards a comprehensive closed-loop deep brain stimulation in Parkinson's disease)
Funding	SNF – Bridge Discovery
Coordinator	University of Bern
Duration	2021.06.01 – 2025.05.31
Partner(s)	Idiap Research Institute, Centre Suisse d'Electronique et de Microtechnique

Name	**HARDENING** (Heterogeneous face recognition for unified identity management)
Funding	Innosuisse
Coordinator	Idiap Research Institute
Duration	2021.02.01 – 2022.07.31
Partner(s)	Facedapter Sarl

Name	**SMILE-II** (Scalable Multimodal sign language technology for sIgn language Learning and assessmEnt Phase-II)
Coordinator	Idiap Research Institute
Duration	2021.01.01 – 2024.12.31
Partner(s)	University of Applied Sciences of Special Needs Education, University of Surrey, University of Zurich

Name	**STEADI** (Storytelling Algorithm for Digital Interviews)
Funding	SNF – Division I
Coordinator	University of Neuchâtel
Duration	2021.02.01 – 2025.01.31
Partner(s)	Idiap Research Institute, University of Lausanne

Name	**SWITCH** (Learning by Switching Roles in Physical Human-Robot Collaboration)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2021.03.01 – 2024.02.28
Partner(s)	Jozef Stefan Institute

Articles in scientific journals in 2020

[1] M. Á. Álvarez-Carmona, E. Villatoro-Tello, M. Montes-y Gómez, and L. V. Pineda, "Author profiling in social media with multimodal information," Journal of Computación y Sistemas, vol. 24, no. 3, Apr. 2020.

[2] N. Antonello and P. N. Garner, "A t-distribution based operator for enhancing out of distribution robustness of neural network classifiers," IEEE Signal Processing Letters, vol. 27, pp. 1070–1074, Jun. 2020.

[3] S. Calinon, "Gaussians on Riemannian manifolds for robot learning and adaptive control," IEEE Robotics and Automation Magazine, 2020.

[4] K. Chatzilygeroudis, A. Vassiliades, F. Stulp, S. Calinon, and J. B. Mouret, "A survey on policy search algorithms for learning robot controllers in a handful of trials," IEEE Transactions on Robotics, vol. 32, no. 2, pp. 328–347, Apr. 2020.

[5] M. Cook, S. Kuntsche, F. Labhart, and E. Kuntsche, "Do different drinks make you feel different emotions? Examination of young adolescents' beverage-specific alcohol expectancies using the alcohol expectancy task," Addictive Behaviors, 2020.

[6] N. Dawalatabad, S. Madikeri, C. C. Sekhar, and H. A. Murthy, "Novel architectures for unsupervised information bottleneck based speaker diarization of meetings," IEEE/ACM Transactions on Audio Speech and Language Processing, 2020.

[7] P. Dighe, A. Asaei, and H. Bourlard, "On quantifying the quality of acoustic models in hybrid DNNHMM ASR," Speech Communication, vol. 119, pp. 24–35, May 2020.

[8] M. Ewerton, O. Arenz, and J. Peters, "Assisted teleoperation in changing environments with a mixture of virtual guides," Advanced Robotics, vol. 34, no. 18, pp. 1157–1170, Jul. 2020.

[9] A. C. B. H. Ferreira, D. D. Ferreira, H. C. Oliveira, I. C. de Resende, A. Anjos, and M. H. B. de Moraes Lopes, "Competitive neural layer-based method to identify people with high risk for diabetic foot," Computers in Biology and Medicine, vol. 120, May 2020.

[10] P. N. Garner and S. Tong, "A Bayesian approach to recurrence in neural networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[11] A. George and S. Marcel, "Learning one class representations for face presentation attack detection using multi-channel convolutional neural networks," IEEE Transactions on Information Forensics and Security, 2020.

[12] G. Heusch, A. George, D. Geissenbuhler, Z. Mostaani, and S. Marcel, "Deep models and shortwave infrared information to detect face presentation attacks," IEEE Transactions on Biometrics, Behavior, and Identity Science, 2020.

[13] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Automatic pathological speech intelligibility assessment exploiting subspace-based analyses," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1717–1728, May 2020.

[14] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Subspace-based learning for automatic dysarthric speech detection," IEEE Signal Processing Letters, vol. 28, pp. 96–100, Dec. 2020.

[15] C. Jaques, A. Ernst, N. Mercader, and M. Liebling, "Temporal resolution doubling in fluorescence light-sheet microscopy via a hue-encoded shutter and regularization," OSA Continuum, vol. 3, no. 8, Aug. 2020.

[16] C. Jaques and M. Liebling, "Aliasing mitigation in optical microscopy of dynamic biological samples by use of temporally modulated color illumination and a standard RGB camera," Journal of Biomedical Optics, vol. 25, no. 10, p. 106 505, Oct. 2020.

[17] N. Jaquier, L. Rozo, D. G. Caldwell, and S. Calinon, "Geometry-aware manipulability learning, tracking and transfer," International Journal of Robotic Research, 2020.

[18] I. Kodrasi and H. Bourlard, "Spectro-temporal sparsity characterization for dysarthric speech detection," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1210–1222, Apr. 2020.

[19] F. Labhart, T.-T. Phan, D. Gatica-Perez, and E. Kuntsche, "Shooting shots: Estimating alcoholic drink sizes in real life using event-level reports and annotations of close-up pictures," Drug and Alcohol Review, 2020.

[20] F. Labhart, F. Tarsetti, O. Bornet, D. Santani, J. Truong, S. Landolt, D. Gatica-Perez, and E. Kuntsche, "Capturing drinking and nightlife behaviours and their social and physical context with a smartphone application – investigation of users' experience and reactivity," Addiction Research and Theory, vol. 28, no. 1, pp. 62–75, Jan. 2020.

[21] T. S. Lembono, A. Paolillo, E. Pignat, and S. Calinon, "Memory of motion for warm-starting trajectory optimization," IEEE Robotics and Automation Letters, vol. 5, no. 2, L. Dongheui, Ed., pp. 2594–2601, 2020.

[22] G. Liu, Y. Yu, and J.-M. Odobez, "A differential approach for gaze estimation," IEEE Transaction on Pattern Analysis and Machine Intelligence, 2020.

[23] R. Luisier, M. Girgin, M. P. Lutolf, and A. Ranga, "Mammary epithelial morphogenesis in 3d combinatorial microenvironments," Scientific Reports, vol. 10, no. 1, 2020.

[24] O. Mariani, A. Ernst, N. Mercader, and M. Liebling, "Reconstruction of image sequences from ungated and scanning-aberrated laser scanning microscopy images of the beating heart," IEEE Transactions on Computational Imaging, vol. 6, pp. 385–395, 2020.

[25] A. Martinez-Gonzalez, M. Villamizar, O. Canévet, and J.-M. Odobez, "Efficient convolutional neural networks for depth-based multi-person pose estimation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 11, pp. 4207–4221, Nov. 2020.

[26] L. B. Meegahapola and D. Gatica-Perez, "Smartphone sensing for the well-being of young adults: A review," IEEE Access, Dec. 2020.

[27] R. Moyen and T. Gentilhomme, "Adaptive ensemble-based optimisation for petrophysical inversion," Mathematical Geosciences, 2020.

[28] B. Pais, P. Buluschek, G. DuPasquier, T. Nef, N. Schütz, H. Saner, D. Gatica-Perez, and V. Santschi, "Evaluation of 1-year in-home monitoring technology by home-dwelling older adults, family caregivers, and nurses," Frontiers in Public Health, vol. 8, p. 9, Oct. 2020.

[29] D. Panda, S. R. Dash, R. Ray, and S. Parida, "Predicting the causal effect relationship between COPD and cardio vascular diseases," Informatica, vol. 44, no. 4, Dec. 2020.

[30] K. Pelzelmayer, S. Landolt, J. Truong, F. Labhart, D. Santani, E. Kuntsche, and D. Gatica-Perez, "Youth nightlife at home: Towards a feminist conceptualisation of home," Children's Geographies, 2020.

[31] K. H. Poon, J. Kämpf, S. E. R. Tay, N. H. Wong, and T. G. Reindl, "Parametric study of urban morphology on building solar energy potential in singapore context," Urban Climate, vol. 33, no. 100624, 2020.

[32] D. Ram, L. Miculicich, and H. Bourlard, "Neural network based end-to-end query by example spoken term detection," IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020.

[33] R. Sahu, S. R. Dash, L. A. Cacha, R. R. Poznanski, and S. Parida, "Epileptic seizure detection: A comparative study between deep and traditional machine learning techniques," Journal of Integrative Neuroscience, vol. 19, no. 1, pp. 1–9, 2020.

[34] A. Shajkofci and M. Liebling, "Free annotated data for deep learning in microscopy? A hitchhiker's guide," Photoniques, no. 104, pp. 30–33, 2020.

[35] A. Shajkofci and M. Liebling, "Spatially-variant CNN-based point spread function estimation for blind deconvolution and depth estimation in optical microscopy," IEEE Transactions on Image Processing, vol. 29, pp. 5848–5861, Apr. 2020.

[36] K. Smit, E. Kuntsche, D. Anderson-Luxford, and F. Labhart, "Fun/intoxication pre-drinking motives lead indirectly to more alcohol-related consequences via increased alcohol consumption on a given night," Addictive Behaviors, 2020.

[37] J. Truong, F. Labhart, D. Santani, D. Gatica-Perez, E. Kuntsche, and S. Landolt, "The emotional entanglements of smartphones in the field: On emotional discomfort, power relations, and research ethics," Area, vol. 52, no. 1, 2020.

[38] G. E. Tyzack, G. Manferrari, J. Newcombe, N. M. Luscombe, R. Luisier, and R. Patani, "Paraspeckle components NONO and PSPC1 are not mislocalized from motor neuron nuclei in sporadic ALS," Brain, 2020.

[39] M. Villamizar, O. Canévet, and J.-M. Odobez, "Multi-scale sequential network for semantic text segmentation and localization," Pattern Recognition Letters, vol. 129, pp. 63–69, 2020.

[40] M. Villamizar, A. Martínez-González, O. Canévet, and J.-M. Odobez, "Watchnet++: Efficient and accurate depth-based network for detecting people attacks and intrusion," Machine Vision and Applications, 2020.

[41] E. Villatoro-Tello, S. Parida, P. Motlicek, and O. Bojar, "Inferring highly-dense representations for clustering broadcast media content," The Prague Bulletin of Mathematical Linguistics, Oct. 2020.

PhD Theses in 2020

[1] T. Chavdarova and F. Fleuret, "Deep generative models and applications," PhD thesis, EPFL, Jul. 2020.

[2] C. Jaques, "Active illumination and computational methods for temporal and spectral super-resolution microscopy," PhD thesis, EPFL, 2020.

[3] N. Jaquier, "Robot skills learning with riemannian manifolds : Leveraging geometry-awareness in robot learning, optimization and control," PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Jul. 2020.

[4] F. Labhart, "Context is everything: Using a smartphone app to capture young people's drinking behaviours, cognitions, environments, and consequences," PhD thesis, La Trobe University, Melbourne, Australia, Oct. 2020.

[5] L. Miculicich, "Discourse phenomena in machine translation," PhD thesis, École polytechnique fédérale de Lausanne, 2020.

[6] A. Mohammadi, "Trustworthy face recognition: Improving generalization of deep face presentation attack detection," PhD thesis, École polytechnique fédérale de Lausanne, 2020.

[7] M. Pettinato, "Detection of disguised speech in forensic science by humans and automatic systems," Master Thesis, PhD thesis, Université de Lausanne Ecole des Sciences Criminelles, Jul. 2020.

[8] T.-T. Phan, "Understanding eating and drinking in context from crowdsourced data," PhD thesis, EPFL, May 2020.

[9] E. Pignat, "Product of experts for robot learning from demonstration," PhD thesis, EPFL, 2020.

[10] A. Shajkofci, "Weakly supervised deep learning methods for biomicroscopy," PhD thesis, EPFL, 2020.

[11] S. Tong, "Multilingual training and adaptation in speech recognition," PhD thesis, EPFL, 2020.

[12] Y. Yu, "Accurate nod and 3D gaze estimation for social interaction analysis," PhD thesis, EDEE, EPFL, Mar. 2020.

Articles in conference proceedings in 2020

[1] C. Basurto and J. Kämpf, "An integrated and strategic evaluation of automatic blind controls to achieve energy and occupant's comfort objectives," in Proceedings of the 5th IBPSA-England Conference on Building Simulation and Optimization (Virtual), Loughborough, UK, Sep. 2020.

[2] H. Bredin, R. Yin, J. M. Coria, P. Korshunov, M. Lavechin, D. Fustes, H. Titeux, W. Bouaziz, and M.-P. Gill, "Pyannote. audio: Neural building blocks for speaker diarization," in IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2020.

[3] O. Canévet, W. He, P. Motlicek, and J.-M. Odobez, "The mummer data set for robot perception in multi-party HRI scenarios," in Proceedings of the 29th IEEE International Conference on Robot & Human Interactive Communication, Sep. 2020.

[4] E. Courdier and F. Fleuret, "Real-time segmentation networks should be latency aware," in Asian Conference on Computer Vision, 2020.

[5] N. Cummins, Y. Pan, Z. Ren, J. Fritsch, V. S. Nallanthighal, H. Christensen, D. Blackburn, B. Schuller, M. Magimai.-Doss, H. Strik, and A. Härmä, "A comparison of acoustic and linguistics methodologies for Alzheimer's dementia recognition," in Proceedings of Interspeech, 2020, pp. 2182–2186.

[6] P. Das, J. McGrath, Z. Fang, A. Boyd, G. Jang, A. Mohammadi, S. Purnapatra, D. Yambay, S. Marcel, M. Trokielewicz, P. Maciejewicz, K. Bowyer, A. Czajka, and S. Schuckers, "Iris liveness detection competition (livdet-iris) – the 2020 edition," in International Joint Conference on Biometrics, 2020.

[7] M. Fabien, E. VILLATORO-TELLO, P. Motlicek, and S. Parida, "Bertaa: Bert fine-tuning for authorship attribution," in Proceedings of the 17th International Conference on Natural Language Processing, 2020.

[8] J. Fritsch, S. P. Dubagunta, and M. Magimai.-Doss, "Estimating the degree of sleepiness by integrating articulatory feature knowledge in raw waveform based CNNs," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Barcelona, Spain, May 2020.

[9] H. Girgin, E. Pignat, N. Jaquier, and S. Calinon, "Active improvement of control policies with bayesian gaussian mixture model," in Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2020.

[10] W. He, L. Lu, B. Zhang, J. Mahadeokar, K. Kalgaonkar, and C. Fuegen, "Spatial attention for far-field speech recognition with deep beamforming neural networks," in 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 7499–7503.

[11] J. Henderson, "The unstoppable rise of computational linguistics in deep learning," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online: Association for Computational Linguistics, Jul. 2020, pp. 6294–6306.

[12] E. Hermann and M. Magimai.-Doss, "Dysarthric speech recognition with lattice-free mmi," in International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2020, pp. 6109–6113.

[13] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Synthetic speech references for automatic pathological speech intelligibility assessment," in International Conference on Acoustics, Speech, and Signal Processing, Barcelona, Spain, 2020, pp. 6099–6103.

[14] N. Jaquier, L. Rozo, and S. Calinon, "Analysis and transfer of human movement manipulability in industry-like activities," in Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2020.

[15] Z.-G. Juan, P. Motlicek, Q. Zhan, R. Braun, and K. Vesely, "Automatic speech recognition benchmark for air-traffic communications," in Proc. Interspeech 2020, Oct. 2020, pp. 2297–2301.

[16] R. Karimi Mahabadi, Y. Belinkov, and J. Henderson, "End-to-end bias mitigation by modelling biases in corpora," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020.

[17] A. Katharopoulos, A. Vyas, N. Pappas, and F. Fleuret, "Transformers are rnns: Fast autoregressive transformers with linear attention," in Proceedings of International Conference on Machine Learning, 2020.

[18] B. Khonglah, S. Madikeri, S. Dey, H. Bourlard, P. Motlicek, and J. Billa, "Incremental semi-supervised learning for multi-genre speech recognition," in Proceedings of ICASSP 2020, 2020.

[19] I. Kodrasi, M. Pernon, M. Laganaro, and H. Bourlard, "Automatic discrimination of apraxia of speech and dysarthria using a minimalistic set of handcrafted features," in Proc. of Annual Conference of the International Speech Communication Association, Oct. 2020, pp. 4991–4995.

[20] K. Kotwal and S. Marcel, "Cnn patch pooling for detecting 3d mask presentation attacks in nir," in IEEE International Conference on Image Processing, Oct. 2020.

[21] T. Kulak, J. Silverio, and S. Calinon, "Fourier movement primitives: An approach for learning rhythmic robot skills from demonstrations," in Robotics: Science and Systems, 2020.

[22] T. S. Lembono, C. Mastalli, P. Fernbach, N. Mansard, and S. Calinon, "Learning how to walk: Warmstarting optimal control solver with memory of motion," in International Conference on Robotics and Automation, 2020.

[23] S. Madikeri, B. Khonglah, S. Tong, P. Motlicek, H. Bourlard, and D. Povey, "Lattice-free maximum mutual information training of multilingual speech recognition system," in In Proceedings of Interspeech 2020, ISCA, 2020, pp. 4746–4750.

[24] F. Mai, N. Pappas, I. Montero, N. A. Smith, and J. Henderson, "Plug and play autoencoders for conditional text generation," in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, Online, 2020.

[25] A. Martínez-González, M. Villamizar, O. Canévet, and J.-M. Odobez, "Residual pose: A decoupled approach for depth-based 3d human pose estimation," in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020.

[26] K. Matoba and F. Fleuret, "Exact preimages of neural network aircraft collision avoidance systems," in Machine Learning for Engineering Modeling, Simulation, and Design Workshop at Neural Information Processing Systems 2020, Nov. 2020.

[27] L. B. Meegahapola, S. Ruiz-Correa, and D. Gatica-Perez, "Alone or with others? understanding eating episodes of college students with mobile sensing," in 19th International Conference on Mobile and Ubiquitous Multimedia, ser. MUM 2020, ACM, Essen, Germany: Association for Computing Machinery, Nov. 2020, 162–166.

[28] L. B. Meegahapola, S. Ruiz-Correa, and D. Gatica-Perez, "Protecting mobile food diaries from getting too personal," in 19th International Conference on Mobile and Ubiquitous Multimedia, ser. MUM 2020, Essen, Germany: Association for Computing Machinery, Nov. 2020, 212–222.

[29] L. Miculicich and J. Henderson, "Partially-supervised mention detection," in Proceedings of the Third Workshop on Computational Models of Reference, Anaphora and Coreference, 2020.

[30] A. Mohammadi, S. Bhattacharjee, and S. Marcel, "Domain adaptation for generalization of face presentation attack detection in mobile settings with minimal information," in 45th International Conference on Acoustics, Speech, and Signal Processing, Barcelona, Spain: IEEE, 2020.

[31] A. Mohammadi, S. Bhattacharjee, and S. Marcel, "Improving cross-dataset performance of face presentation attack detection systems using face recognition datasets," in 45th International Conference on Acoustics, Speech, and Signal Processing, IEEE, 2020.

[32] A. Mohammadshahi and J. Henderson, "Graph-to-graph transformer for transition-based dependency parsing," in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, ACL, Online: Association for Computational Linguistics, Nov. 2020, 3278–3289.

[33] K. Motokura, M. Takahashi, M. Ewerton, and J. Peters, "Plucking motions for tea harvesting robots using probabilistic movement primitives," in IEEE International Conference on Robotics and Automation, 2020.

[34] S. Muralidhar, E. P. Kleinlogel, E. Mayor, A. Bangerter, M. Schmid Mast, and D. Gatica-Perez, "Understanding applicants' reactions to asynchronous video interviews through self-reports and nonverbal cues," in Proc. ACM Int. Conf. on Multimodal Interaction (ICMI), Utrecht, Oct. 2020.

[35] T. Nakazawa, H. Nakayama, C. Ding, R. Dabre, S. Higashiyama, H. Mino, I. Goto, W. Pa Pa, A. Kunchukuttan, S. Parida, O. Bojar, and S. Kurohashi, "Overview of the 7th workshop on asian translation," in Proceedings of the 7th Workshop on Asian Translation, Association for Computational Linguistics, 2020.

[36] H. H. Nguyen, J. Yamagishi, I. Echizen, and S. Marcel, "Generating master faces for use in performingwolf attacks on face recognition systems," in International Join Conference on Biometrics, Sep. 2020.

[37] M. Panteris, S. Manschitz, and S. Calinon, "Learning, generating and adapting wave gestures for expressive human-robot interaction," in Proc. ACM/IEEE Intl Conf. on Human-Robot Interaction (HRI), 2020, pp. 386–388.

[38] A. Paolillo, T. S. Lembono, and S. Calinon, "A memory of motion for visual predictive control tasks," in International Conference on Robotics and Automation, 2020.

[39] S. Parida, P. Motlicek, A. R. Dash, S. R. Dash, D. K. Mallick, S. P. Biswal, P. Pattnaik, B. N. Nayak, and O. Bojar, "Odianlp's participation in wat2020," in Proceedings of the 7th Workshop on Asian Translation, ACL Anthology, 2020.

[40] S. Parida, E. VILLATORO-TELLO, S. Kumar, M. Fabien, and P. Motlicek, "Detection of similar languages and dialects using deep supervised autoencoders," in Proceedings of the 17th International Conference on Natural Language Processing, 2020.

[41] S. Parida, E. VILLATORO-TELLO, S. Kumar, P. Motlicek, and Q. Zhan, "Idiap submission to swissgerman language detection shared task," in Proceedings of the 5th Swiss Text Analytics Conference (SwissText) & 16th Conference on Natural Language Processing (KONVENS), CEUR Workshop Proceedings, Jun. 2020.

[42] T.-T. Phan, F. Labhart, S. Muralidhar, and D. Gatica-Perez, "Understanding heavy drinking at night through smartphone sensing and active human engagement," in Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare, Oct. 2020.

[43] E. Pignat, H. Girgin, and S. Calinon, "Generative adversarial training of product of policies for robust and adaptive movement primitives," in In Proc. Conference on Robot Learning (CoRL), 2020.

[44] E. Pignat, T. S. Lembono, and S. Calinon, "Variational inference with mixture model approximation for applications in robotics," in International Conference on Robotics and Automation, 2020.

[45] A. Pozo, T.-T. Phan, and D. Gatica-Perez, "Learning urban nightlife routines from mobile data," in Proc. Int. Conf. on Mobile and Ubiquitous Multimedia, Essen, Germany, Nov. 2020.

[46] R. Prasad, G. Yilmaz, O. Chetelat, and M. Magimai.-Doss, "Detection of s1 and s2 locations in phonocardiogram signals using zero frequency filter," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2020.

[47] S. S. Sarfjoo, S. Madikeri, P. Motlicek, and S. Marcel, "Supervised domain adaptation for textindependent speaker verification using limited data," in Interspeech, 2020.

[48] A. Shajkofci and M. Liebling, "Deepfocus: A few-shot microscope slide auto-focus using a sample invariant cnn-based sharpness function," in International Symposium on Biomedical Imaging, Accepted to ISBI 2020, 2020.

[49] R. Siegfried, B. Aminian, and J.-M. Odobez, "Manigaze: A dataset for evaluating remote gaze estimator in object manipulation situations," in Symposium on Eye Tracking Research and Applications, Stuttgart, Germany: ACM, Jun. 2020, p. 5.

[50] P. T. Sivaprasad, F. Mai, T. Vogels, M. Jaggi, and F. Fleuret, "Optimizer benchmarking needs to account for hyperparameter tuning," in Proceedings of the 37th International Conference on Machine Learning, Vienna, Austria, 2020.

[51] S. Tornay, O. Aran, and M. Magimai.-Doss, "An hmm approach with inherent model selection for sign language and gesture recognition," in Proceedings of the International Conference on Language Resources and Evaluation LREC 2020, 2020.

[52] S. Tornay, N. C. Camgoz, R. Bowden, and M. Magimai.-Doss, "A phonology-based approach for isolated sign production assessment in sign language," in Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion), Oct. 2020.

[53] S. Tornay, M. Razavi, and M. Magimai.-Doss, "Towards multilingual sign language recognition," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Barcelona, Spain, May 2020.

[54] E. VILLATORO-TELLO, S. Parida, S. Kumar, P. Motlicek, and Q. Zhan, "Idiap & uam participation at germeval 2020: Classification and regression of cognitive and motivational style from text," in Proceedings of the GermEval 2020 Shared Task on the Classification and Regression of Cognitive and Motivational style from Text, Jun. 2020.

[55] E. VILLATORO-TELLO, G. Ramírez-de-la Rosa, S. Parida, S. Kumar, and P. Motlicek, "Idiap and uam participation at mex-a3t evaluation campaign," in Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020) co-located with 36th Conference of the Spanish Society for Natural Language Processing (SEPLN 2020), vol. 2664, CEUR Workshop Proceedings, Sep. 2020, p. 6.

[56] A. Vyas, A. Katharopoulos, and F. Fleuret, "Fast transformers with clustered attention," in Proceedings of the International Conference on Neural Information Processing Systems, Dec. 2020.

[57] Y. Yu and J.-M. Odobez, "Unsupervised representation learning for gaze estimation," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2020.

List of Patents in 2020

While evaluating the success of technology transfer is a difficult task, invention disclosures and patents are usually key metrics in the field. As a necessary first step to any patent filing, an invention disclosure aims to identify a technology with a certain level of maturity and some promising economic potential.

In 2020, 25 invention and software disclosures have been filled. A patent committee examines each invention disclosure and decides to move forward with a patent filing or not. The committee is composed of the head of technology transfer, two representatives of the direction and one representative of the researchers. Since 2012, 11 patents have been filed: 3 patents are granted, 7 applications are published and pending (1 of them has been sold) and 1 application was abandoned.

Granted patents

- **IDIAP-1** [US 9,689,959 B2] A. Asaei, H. Bourlard, V. Cevher, "Method, apparatus and computer program product for determining the location of a plurality of speech sources"
- **IDIAP-2** [US 9,058,541 B2] C. Dubout, F. Fleuret, "Object detection method, object detector and object detection computer program"
- **IDIAP-8** [US 9,973,503 B2] S. Marcel, A. Anjos, P. Abbet, "Method and internet-connected server for reviewing a computer-executable experiment"

Pending applications

- **SAMSUNG-1** [US 2014/0149104 A1, EP 2736042 A1] N-H. Kim, P. Motlicek, P. N. Garner, D. Imseng, J-W. Lee, J-M. Cho, "Apparatus and method for constructing multilingual acoustic model and computer readable recording medium for storing program for performing the method"
- **IDIAP-5** [WO 2015/192879 A1] K. A. Funes Mora, J-M. Odobez, "A gaze estimation method and apparatus"
- **IDIAP-9** [WO 2017/221049 A1, US 2019/0180040 A1] A. Anjos, S. Marcel, "A data-network connected server, a device, a platform and a method for conducting computer-executable experiments"
- **globID-1** [WO 2019/150254 A1] L. Sonna Momo, L. Cerqueira Torres, S. Marcel, A. Anjos, M. Liebling, A. Shajkofci, S. Amoos, A. Woeffray, A. Sierro, "Method and device for biometric vascular recognition and/or identification"
- **IDIAP-10** [EP 3 691 258 A1] M. Liebling, C. Jacques, "System and method for acquiring images"
- **IDIAP-11** [EP 3 719 679 A1] S. Marcel, V. Krivokuca, "A method for protecting biometric templates, and a system and method for verifying a speaker's identity"

Sold applications

- **IDIAP-4** [WO 2016/023582 A1] S. Marcel, "A method of detecting a falsified presentation to a vascular recognition system". This patent has been sold.

Abandonned applications

- **IDIAP-6** [US 2017/0069306 A1] A. Asaei, M. Cernak, H. Bourlard, "Signal processing method and apparatus based on structured sparsity of phonological features"

encarté scientifique

I — XL