IRISA
UMR

Activity Report 2018

Team LINKMEDIA

Creating and Exploiting Explicit Links between Multimedia Fragments

Joint team with Inria Rennes – Bretagne Atlantique

D6 – Media and Interactions

Table of contents

<div align="center">

Project-Team LINKMEDIA

</div>

Creation of the Project-Team: 2014 July 01

> *Linkmedia went through the regular Inria evaluation process in October 2018, after 4 years of existence. Linkmedia switched leaders on November 01, 2018.*

Keywords:

Computer Science and Digital Science:

       A3.3.2. - Data mining
       A3.3.3. - Big data analysis
       A5.3.3. - Pattern recognition
       A5.4.1. - Object recognition
       A5.4.3. - Content retrieval
       A5.7. - Audio modeling and processing
       A5.7.1. - Sound
       A5.7.3. - Speech
       A5.8. - Natural language processing
       A9.2. - Machine learning
       A9.3. - Signal analysis
       A9.4. - Natural language processing

Other Research Topics and Application Domains:

       B9. - Society and Knowledge
       B9.3. - Medias
       B9.6.10. - Digital humanities
       B9.10. - Privacy

1. Team, visitors, external collaborators

Research Scientists
    Laurent Amsaleg [Team leader since Nov 2018, CNRS, Researcher, HDR]
    Yannis Avrithis [Inria, Advanced Research Position]
    Vincent Claveau [CNRS, Researcher]
    Teddy Furon [Inria, Researcher, HDR]
    Guillaume Gravier [Team leader until Oct 2018, CNRS, Senior Researcher, HDR]
    Miaojing Shi [Inria, Starting Research Position]

Faculty Members
    Ewa Kijak [Univ de Rennes I, Associate Professor]
    Simon Malinowski [Univ de Rennes I, Associate Professor]
    Christian Raymond [INSA Rennes, Associate Professor, until Nov 2018]
    Pascale Sébillot [INSA Rennes, Professor, HDR]

Post-Doctoral Fellows
    Mateusz Budnik [CNRS]
    Anne-Lyse Minard [CNRS, until Aug 2018]

PhD Students
    Ricardo Carlini Sperandio [Univ Federal de Minas Gerais]

Clément Dalloux [CNRS]
Mikail Demirdelen [Univ de Rennes I]
Cheikh Brahim El Vaigh [Inria]
Marzieh Gheisari Khorasgani [Inria]
Yann Lifchitz [Safran, from Feb 2018]
Cédric Maigrot [Univ de Rennes I]
Cyrielle Mallart [Ouest France, from Dec 2018]
Oriane Siméoni [Inria]
Hanwei Zhang [Chinese Scolarship Council]
Mathieu Laroze [Wipsea]
François Torregrossa [SoLocal, from Dec 2018]

Technical staff
Firas Hmida [Univ de Rennes I, until Jan 2018]

Interns
Dominique Barbe [Ecole Normale Supérieure Rennes, from Feb 2018 until Jun 2018]
Alan Turpin [Univ de Rennes I, from May 2018 until Aug 2018]

Administrative Assistant
Aurélie Patier [Univ de Rennes I]

Visiting Scientists
Yohan Gumiel [PUC Parana, Brazil, from Jun 2018 until Aug 2018]
Lucas Oliveira [PUC Parana, Brazil, from Jun 2018 until Aug 2018]
Yanlin Qian [Tampere Univ of Technology, Finland, from Jun 2018 until Aug 2018]

2. Overall Objectives

2.1. Context

Linked media appears today as a major challenge, with numerous potential applications in all areas of multimedia. The strong increase of ubiquitous access to the Internet and the resulting convergence of media on the network open countless opportunities for linked media and reinforce the key role of such a challenge. New applications centered on the notion of linked media are emerging today, such as second screen applications and recommendation services. However, because of the lack of adequate technology, linking related content is mostly deferred to human operators in current applications or to user behavior analysis, e.g., via collaborative filtering, thus indirectly considering the content. This fact severely limits the opportunities offered by a web of media, in terms of creativity, scalability, representativeness and completeness, thus negatively impacting the spread of linked media and the development of innovative services in the Internet of media.

Most of the research effort in automatic multimedia content analysis has been devoted so far to describing and indexing content on which core tasks around information retrieval and recommendation are built to develop multimedia applications. This general philosophy mostly reposes on a vision where documents are considered as isolated entities, i.e., as a basic unit which is indexed or analyzed regardless of other content items and of context. Considering documents in isolation has enabled key progress in content-based analysis and retrieval on a large scale: e.g., design of generic descriptors, efficient techniques for content-based analysis as well as fast retrieval methodology. But ignoring the links, implicit or explicit, between content items also appears as a rather strong limitation with direct consequences on algorithms and applications, both in terms of performance and in terms of possibilities.

2.2. Scientific objectives

LINKMEDIA investigates a number of key issues related to multimedia collections structured with explicit links: Can we discover what characterizes a collection and makes its coherence? Are there repeating motifs that create natural links and which deserve characterization and semantic interpretation? How to explicitly create links from pairwise distances? What structure should a linked collection have? How do we explain the semantic of a link? How explicit links can be used to improve information retrieval? To improve user experience?

In this general framework, the global objective of LINKMEDIA is to develop the scientific, methodological and technological foundations facilitating or automating the creation, the description and the exploitation of multimedia collections structured with explicit links. In particular, we target a number of key contributions in the following areas:

- designing efficient methods dedicated to multimedia indexing and unsupervised motif discovery: efficiently comparing content items at large scale and finding repeating motifs in an unsupervised manner are two key ingredients of multimedia linking based on a low-level representation of the content;

- improving techniques for structuring and semantic description: better description of multimedia content at a semantic—i.e., human interpretable—level, making explicit the implicit structure when it exists, is still required to make the most of multimedia data and to facilitate the creation of links to a precise target at a semantic level;

- designing and experimenting approaches to multimedia content linking and collection structuring: exploiting low-level and semantic content-based proximity to create explicit links within a collection requires specific methodology departing from pairwise comparison and must be confronted with real data;

- studying new paradigms for the exploitation of linked multimedia content as well as new usages: explicit links within media content collections change how such data is processed by machines and ultimately consumed by humans in ways that have yet to be invented and studied.

3. Research Program

3.1. Scientific background

LINKMEDIA is a multidisciplinary research team, with multimedia data as the main object of study. We are guided by the data and their specificity—semantically interpretable, heterogeneous and multimodal, available in large amounts, unstructured and disconnected—, as well as by the related problems and applications.

With multimedia data at the center, orienting our choices of methods and algorithms and serving as a basis for experimental validation, the team is directly contributing to the following scientific fields:

- multimedia: content-based analysis; multimodal processing and fusion; multimedia applications;

- computer vision: compact description of images; object and event detection;

- natural language processing: topic segmentation; information extraction;

- information retrieval: high-dimensional indexing; approximate k-nn search; efficient set comparison.

LINKMEDIA also takes advantage of advances in the following fields, adapting recent developments to the multimedia area:

- signal processing: image processing; compression;

- machine learning: deep architectures; structured learning; adversarial learning;

- security: data encryption; differential privacy;

- data mining: time series mining and alignment; pattern discovery; knowledge extraction.

3.2. Workplan

Research activities in LINKMEDIA are organized along three major lines of research which build upon the scientific domains already mentioned.

3.2.1. Unsupervised motif discovery

As an alternative to supervised learning techniques, unsupervised approaches have emerged recently in multimedia with the goal of discovering directly patterns and events of interest from the data, in a totally unsupervised manner. In the absence of prior knowledge on what we are interested in, meaningfulness can be judged based on one of three main criteria: unexpectedness, saliency and recurrence. This last case posits that repeating patterns, known as motifs, are potentially meaningful, leading to recent work on the unsupervised discovery of motifs in multimedia data [50], [48], [49].

LINKMEDIA seeks to *develop unsupervised motif discovery approaches which are both accurate and scalable*. In particular, we consider the discovery of repeating objects in image collections and the discovery of repeated sequences in video and audio streams. Research activities are organized along the following lines:

- developing the scientific basis for scalable motif discovery: sparse histogram representations; efficient co-occurrence counting; geometry and time aware indexing schemes;
- designing and evaluating accurate and scalable motif discovery algorithms applied to a variety of multimedia content: exploiting efficient geometry or time aware matching functions; fast approximate dynamic time warping; symbolic representations of multimedia data, in conjunction with existing symbolic data mining approaches;
- developing methodology for the interpretation, exploitation and evaluation of motif discovery algorithms in various use-cases: image classification; video stream monitoring; transcript-free natural language processing (NLP) for spoken document.

3.2.2. Description and structuring

Content-based analysis has received a lot of attention from the early days of multimedia, with an extensive use of supervised machine learning for all modalities [51], [45]. Progress in large scale entity and event recognition in multimedia content has made available general purpose approaches able to learn from very large data sets and performing fairly decently in a large number of cases. Current solutions are however limited to simple, homogeneous, information and can hardly handle structured information such as hierarchical descriptions, tree-structured or nested concepts.

LINKMEDIA aims at *expanding techniques for multimedia content modeling, event detection and structure analysis*. The main transverse research lines that LINKMEDIA will develop are as follows:

- context-aware content description targeting (homogeneous) collections of multimedia data: latent variable discovery; deep feature learning; motif discovery;
- secure description to enable privacy and security aware multimedia content processing: leveraging encryption and obfuscation; exploring adversarial machine learning in a multimedia context; privacy-oriented image processing;
- multilevel modeling with a focus on probabilistic modeling of structured multimodal data: multiple kernels; structured machine learning; conditional random fields.

3.2.3. Linking and collection data model

Creating explicit links between media content items has been considered on different occasions, with the goal of seeking and discovering information by browsing, as opposed to information retrieval via ranked lists of relevant documents. Content-based link creation has been initially addressed in the hypertext community for well-structured texts [44] and was recently extended to multimedia content [52], [47], [46]. The problem of organizing collections with links remains mainly unsolved for large heterogeneous collections of unstructured documents, with many issues deserving attention: linking at a fine semantic grain; selecting relevant links; characterizing links; evaluating links; etc.

LINKMEDIA targets pioneering research on media linking by *developing scientific ground, methodology and technology for content-based media linking* directed to applications exploiting rich linked content such as navigation or recommendation. Contributions are concentrated along the following lines:

- algorithmic of linked media for content-based link authoring in multimedia collections: time-aware graph construction; multimodal hypergraphs; large scale k-nn graphs;

- link interpretation and characterization to provide links semantics for interpretability: text alignment; entity linking; intention vs. extension;

- linked media usage and evaluation: information retrieval; summarization; data models for navigation; link prediction.

4. Application Domains

4.1. Asset management in the entertainement business

Regardless of the ingestion and storage issues, media asset management—archiving, describing and retrieving multimedia content—has turned into a key factor and a huge business for content and service providers. Most content providers, with television channels at the forefront, rely on multimedia asset management systems to annotate, describe, archive and search for content. So do archivists such as the Institut National de l'Audiovisuel, the Nederlands Instituut voor Beeld en Geluid or the British Broadcast Corporation, as well as media monitoring companies, such as Yacast in France. Protecting copyrighted content is another aspect of media asset management.

4.2. Multimedia Internet

One of the most visible application domains of linked multimedia content is that of multimedia portals on the Internet. Search engines now offer many features for image and video search. Video sharing sites also feature search engines as well as recommendation capabilities. All news sites provide multimedia content with links between related items. News sites also implement content aggregation, enriching proprietary content with user-generated content and reactions from social networks. Most public search engines and Internet service providers offer news aggregation portals.

4.3. Multiscreen TV

The convergence between television and the Internet has accelerated significantly over the past few years, with the democratization of TV on-demand and replay services and the emergence of social TV services and multiscreen applications. These evolutions and the consequently ever growing number of innovative applications offer a unique playground for multimedia technologies. Recommendation plays a major role in connected TV. Enriching multimedia content, with explicit links targeting either multimedia material or knowledge databases, appears as a key feature in this context, at the core of rich TV and second screen applications.

4.4. E-learning

On-line courses are rapidly gaining interest with the recent movement for massive open on-line courses (MOOCs). Such courses usually aggregate multimedia material, such as a video of the course with handouts and potentially text books, exercises and other related resources. This setting is very similar to that of the media aggregation sites though in a different domain. Automatically analyzing and describing video and textual content, synchronizing all material available across modalities, creating and characterizing links between related material or between different courses are all necessary features for on-line courses authoring.

5. Highlights of the Year

5.1. Highlights of the Year

Our activities in relation with fake news was highlighted in a number of occasions such an audition at the French National Assembly by the parliamentary group *Économie numérique de la donnée, de la connaissance et de l'intelligence artificielle*, at the Forum des sciences cognitives in Nancy, in a press conference organized by the CNRS in May 2018, through a Fake News hackathon organized in Rennes by Irisa and Inria, through many articles in the press as well as during the European cyber-week conference held in November 2018.

5.1.1. Awards

- Miaojing Shi was selected as the Forty under Forty Talents by French Tech hubs China.
- Miaojing Shi was in the final pitch (in progress) of the 2018 Innovation Awards by Comité France Chine.

6. New Results

6.1. Low-level content description and indexing

6.1.1. Scalability of the NV-tree: Three Experiments

Participants: Laurent Amsaleg, Björn Þór Jónsson [Univ. Copenhagen], Herwig Lejsek [Videntifier Tech.].

The NV-tree is a scalable approximate high-dimensional indexing method specifically designed for large-scale visual instance search. We report in [10] on three experiments designed to evaluate the performance of the NV-tree. Two of these experiments embed standard benchmarks within collections of up to 28.5 billion features, representing the largest single-server collection ever reported in the literature. The results show that indeed the NV-tree performs very well for visual instance search applications over large-scale collections.

6.1.2. Prototyping a Web-Scale Multimedia Retrieval Service Using Spark

Participants: Laurent Amsaleg, Gylfi Þór Gudmundsson [School of Computer Science, Reykjavik], Björn Þór Jónsson [Univ. Copenhagen], Michael Franklin [Computer Science Division, Berkeley].

The world has experienced phenomenal growth in data production and storage in recent years, much of which has taken the form of media files. At the same time, computing power has become abundant with multi-core machines, grids, and clouds. Yet it remains a challenge to harness the available power and move toward gracefully searching and retrieving from web-scale media collections. Several researchers have experimented with using automatically distributed computing frameworks, notably Hadoop and Spark, for processing multimedia material, but mostly using small collections on small computing clusters. In [3] we describe a prototype of a (near) web-scale throughput-oriented MM retrieval service using the Spark framework running on the AWS cloud service. We present retrieval results using up to 43 billion SIFT feature vectors from the public YFCC 100M collection, making this the largest high-dimensional feature vector collection reported in the literature. We also present a publicly available demonstration retrieval system, running on our own servers, where the implementation of the Spark pipelines can be observed in practice using standard image benchmarks, and downloaded for research purposes. Finally, we describe a method to evaluate retrieval quality of the ever-growing high-dimensional index of the prototype, without actually indexing a web-scale media collection.

6.1.3. Extreme-value-theoretic estimation of local intrinsic dimensionality

Participants: Laurent Amsaleg, Teddy Furon, Oussama Chelly [National Institute of Informatics], Stéphane Girard [MISTIS, Inria Grenoble], Michael Houle [National Institute of Informatics], Ken-Ichi Kawarabayashi [National Institute of Informatics], Michael Nett [Google].

This work is concerned with the estimation of a local measure of intrinsic dimensionality (ID) recently proposed by Houle. The local model can be regarded as an extension of Karger and Ruhl's expansion dimension to a statistical setting in which the distribution of distances to a query point is modeled in terms of a continuous random variable. This form of intrinsic dimensionality can be particularly useful in search, classification, outlier detection, and other contexts in machine learning, databases, and data mining, as it has been shown to be equivalent to a measure of the discriminative power of similarity functions. Several estimators of local ID are proposed and analyzed based on extreme value theory, using maximum likelihood estimation, the method of moments, probability weighted moments, and regularly varying functions, see [2]. An experimental evaluation is also provided, using both real and artificial data.

6.1.4. *Intrinsic dimensionality for Information Retrieval*

Participant: Vincent Claveau.

Examining the properties of representation spaces for documents or words in Information Retrieval (IR) brings precious insights to help the retrieval process. Following the work presented in the previous paragraph, it has been shown that intrinsic dimensionality is chiefly tied with the notion of indiscriminateness among neighbors of a query point in the vector space. In this work [13], we revisit this notion in the specific case of IR. More precisely, we show how to estimate indiscriminateness from IR similarities in order to use it in representation spaces used for documents and words. We show that indiscriminateness may be used to characterize difficult queries; moreover we show that this notion, applied to word embeddings, can help to choose terms to use for query expansion.

6.1.5. *Heat Map Based Feature Ranker*

Participants: Christian Raymond, Carlos Huertas [Autonomous University of Baja California, Mexico], Reyes Uarez-Ramirez [Autonomous University of Baja California, Mexico].

In [6], we present Heat Map Based Feature Ranker, an algorithm to estimate feature importance purely based on its interaction with other variables. A compression mechanism reduces evaluation space up to 66% without compromising efficacy. Our experiments show that our proposal is very competitive against popular algorithms, producing stable results across different types of data. We also show how noise reduction through feature selection aids data visualization using emergent self-organizing maps.

6.1.6. *Time series retrieval and indexing using DTW-preserving shapelets*

Participants: Laurent Amsaleg, Ricardo Carlini Sperandio, Simon Malinowski, Romain Tavenard [Univ. Rennes 2].

Dynamic Time Warping (DTW) is a very popular similarity measure used for time series classification, retrieval or clustering. DTW is, however, a costly measure, and its application on numerous and/or very long time series is difficult in practice. We have proposed a new approach for time series retrieval: time series are embedded into another space where the search procedure is less computationally demanding, while still accurate. This approach is based on transforming time series into high-dimensional vectors using DTW-preserving shapelets. That transform is such that the relative distance between the vectors in the Euclidean transformed space well reflects the corresponding DTW measurements in the original space. We have also proposed in [12] strategies for selecting a subset of shapelets in the transformed space, resulting in a trade-off between the complexity of the transformation and the accuracy of the retrieval. Experimental results using the well known time series datasets demonstrate the importance of this trade-off. This transformation can then be used to build efficient time series indexing schemes.

6.1.7. *Fast Spectral Ranking for Similarity Search*

Participants: Yannis Avrithis, Teddy Furon, Ahmet Iscen [Univ. Prague], Giorgos Tolias [Univ. Prague], Ondra Chum [Univ. Prague].

Despite the success of deep learning on representing images for particular object retrieval, recent studies show that the learned representations still lie on manifolds in a high dimensional space. This makes the Euclidean nearest neighbor search biased for this task. Exploring the manifolds online remains expensive even if a nearest neighbor graph has been computed offline. This work introduces an explicit embedding reducing manifold search to Euclidean search followed by dot product similarity search. This is equivalent to linear graph filtering of a sparse signal in the frequency domain. To speed up online search, we compute an approximate Fourier basis of the graph offline. We improve the state of art on particular object retrieval datasets including the challenging Instre dataset containing small objects. At a scale of 10^5 images, the offline cost is only a few hours, while query time is comparable to standard similarity search [15].

6.1.8. *Mining on Manifolds: Metric Learning without Labels*

Participants: Yannis Avrithis, Ahmet Iscen [Univ. Prague], Giorgos Tolias [Univ. Prague], Ondra Chum [Univ. Prague].

In this work we present a novel unsupervised framework for hard training example mining [17]. The only input to the method is a collection of images relevant to the target application and a meaningful initial representation, provided e.g. by pre-trained CNN. Positive examples are distant points on a single manifold, while negative examples are nearby points on different manifolds. Both types of examples are revealed by disagreements between Euclidean and manifold similarities. The discovered examples can be used in training with any discriminative loss. The method is applied to unsupervised fine-tuning of pre-trained networks for fine-grained classification and particular object retrieval. Our models are on par or are outperforming prior models that are fully or partially supervised.

6.1.9. *Hybrid Diffusion: Spectral-Temporal Graph Filtering for Manifold Ranking*

Participants: Yannis Avrithis, Teddy Furon, Ahmet Iscen [Univ. Prague], Giorgos Tolias [Univ. Prague], Ondra Chum [Univ. Prague].

State of the art image retrieval performance is achieved with CNN features and manifold ranking using a k-NN similarity graph that is pre-computed off-line. The two most successful existing approaches are temporal filtering, where manifold ranking amounts to solving a sparse linear system online, and spectral filtering, where eigen-decomposition of the adjacency matrix is performed off-line and then manifold ranking amounts to dot-product search online. The former suffers from expensive queries and the latter from significant space overhead. Here we introduce a novel, theoretically well-founded hybrid filtering approach allowing full control of the space-time trade-off between these two extremes. Experimentally, we verify that our hybrid method delivers results on par with the state of the art, with lower memory demands compared to spectral filtering approaches and faster compared to temporal filtering [16].

6.1.10. *Transactional Support for Visual Instance Search*

Participants: Laurent Amsaleg, Björn Þór Jónsson [Univ. Copenhagen], Herwig Lejsek [Videntifier Tech.].

This work addresses the issue of dynamicity and durability for scalable indexing of very large and rapidly growing collections of local features for visual instance retrieval. By extending the NV-tree, a scalable disk-based high-dimensional index, we show how to implement the ACID properties of transactions which ensure both dynamicity and durability. We present a detailed performance evaluation of the transactional NV-tree, showing that the insertion throughput is excellent despite the effort to enforce the ACID properties [20].

6.1.11. *Time-series prediction for capacity planning*

Participants: Simon Malinowski, Colin Leverger [Orange Labs], Thomas Guyet [AgroCampus Ouest], Vincent Lemaire [Orange Labs].

In a collaboration with Orange Labs, we have worked on KPI time series prediction in order to improve capacity planning. A software has been develloped. This software is detailed in [32]. It aims at visualizing and comparing different time series prediction techniques on user-defined input data. We have also developed a novel prediction algorithm that focuses on time series for with a seasonality [21]. It uses the combination of a clustering algorithm and Markov Models to produce day-ahead forecasts. Our experiments on real datasets show that in the case study, our method outperforms classical approaches (AR, Holt-Winters).

6.1.12. Scale-adaptive CNN for Crowd counting

Participants: Miaojing Shi, Lu Zhang [Fudan Univ.], Qiaobo Chen [Shanghai Jiaotong Univ.].

The task of crowd counting is to automatically estimate the pedestrian number in crowd images. To cope with the scale and perspective changes that commonly exist in crowd images, this work proposes a scale-adaptive CNN (SaCNN) architecture with a backbone of fixed small receptive fields. We extract feature maps from multiple layers and adapt them to have the same output size; we combine them to produce the final density map. The number of people is computed by integrating the density map. We also introduce a relative count loss along with the density map loss to improve the network generalization on crowd scenes with few pedestrians, where most representative approaches perform poorly on. We conduct extensive experiments and demonstrate significant improvements of SaCNN over the state-of-the-art [31].

6.1.13. Revisiting Perspective information for Efficient Crowd counting

Participants: Miaojing Shi, Zhaohui Yang [Peking Univ.], Chao Xu [Peking Univ.], Qijun Chen [Tongji Univ.].

A major challenge of crowd counting lies in the perspective distortion, which results in drastic person scale change in an image. Density regression on the small person area is in general very hard. In this work, we propose a perspective-aware convolutional neural network (PACNN) for efficient crowd counting, which integrates the perspective information into density regression to provide additional knowledge of the person scale change in an image. Ground truth perspective maps are firstly generated for training; PACNN is then specifically designed to predict multi-scale perspective maps, and encode them as perspective-aware weighting layers in the network to adaptively combine the outputs of multi-scale density maps. The weights are learned at every pixel of the maps such that the final density combination is robust to the perspective distortion. We conduct extensive experiments to demonstrate the effectiveness and efficiency of PACNN over the state-of-the-art [42].

6.1.14. Phone-Level Embeddings for Unit Selection Speech Synthesis

Participants: Laurent Amsaleg, Antoine Perquin [EXPRESSION team, IRISA], Gwénolé Lecorvé [EXPRESSION team, IRISA], Damien Lolive [EXPRESSION team, IRISA].

Deep neural networks have become the state of the art in speech synthesis. They have been used to directly predict signal parameters or provide unsupervised speech segment descriptions through embeddings. In [25] we present four models with two of them enabling us to extract phone-level embeddings for unit selection speech synthesis. Three of the models rely on a feed-forward DNN, the last one on an LSTM. The resulting embeddings enable replacing usual expert-based target costs by an euclidean distance in the embedding space. This work is conducted on a French corpus of an 11 hours audiobook. Perceptual tests show the produced speech is preferred over a unit selection method where the target cost is defined by an expert. They also show that the embeddings are general enough to be used for different speech styles without quality loss. Furthermore, objective measures and a perceptual test on statistical parametric speech synthesis show that our models perform comparably to state-of-the-art models for parametric signal generation, in spite of necessary simplifications, namely late time integration and information compression.

6.1.15. Disfluency Insertion for Spontaneous TTS: Formalization and Proof of Concept

Participants: Pascale Sébillot, Raheel Qader [EXPRESSION team, IRISA], Gwénolé Lecorvé [EXPRESSION team, IRISA], Damien Lolive [EXPRESSION team, IRISA].

This is an exploratory work to automatically insert disfluencies in text-to-speech (TTS) systems. The objective is to make TTS more spontaneous and expressive. To achieve this, we propose to focus on the linguistic level of speech through the insertion of pauses, repetitions and revisions. We formalize the problem as a theoretical process, where transformations are iteratively composed. This is a novel contribution since most of the previous work either focus on the detection or cleaning of linguistic disfluencies in speech transcripts, or solely concentrate on acoustic phenomena in TTS, especially pauses. We present a first implementation of the proposed process using conditional random fields and language models. The objective and perceptual evaluation conducted on an English corpus of spontaneous speech show that our proposition is effective to generate disfluencies, and highlights perspectives for future improvements [26]

6.1.16. *Bi-directional Recurrent End-to-End Neural Network Classifier for Spoken Arab Digit Recognition*

Participants: Christian Raymond, Naima Zerari [University of Batna 2, Algeria], Hassen Bouzgou [University of Batna 2, Algeria].

In [30], we propose a general end-to-end approach to sequence learning that uses Long Short-Term Memory (LSTM) to deal with the non-uniform sequence length of the speech utterances. The neural architecture can recognize the Arabic spoken digit spelling of an isolated Arabic word using a classification methodology, with the aim to enable natural human-machine interaction. The proposed system consists to, first, extract the relevant features from the input speech signal using Mel Frequency Cepstral Coefficients (MFCC) and then these features are processed by a deep neural network able to deal with the non uniformity of the sequences length. A recurrent LSTM or GRU architecture is used to encode sequences of MFCC features as a fixed size.

6.1.17. *Are Deep Neural Networks good for blind image watermarking?*

Participants: Teddy Furon, Vedran Vukotić [Lamark, France], Vivien Chappelier [Lamark, France].

Image watermarking is usually decomposed into three steps: i) some features are extracted from an image, ii) they are modified to embed the watermark, iii) and they are projected back into the image space while avoiding the creation of visual artefacts. The feature extraction is usually based on a classical image representation given by the Discrete Wavelet Transform or the Discrete Cosine Transform for instance. These transformations need a very accurate synchronisation and usually rely on various registration mechanisms for that purpose. This paper investigates a new family of transformation based on Deep Learning networks. Motivations come from the Computer Vision literature which has demonstrated the robustness of these features against light geometric distortions. Also, adversarial sample literature provides means to implement the inverse transform needed in the third step. This work [29] shows that this approach is feasible as it yields a good quality of the watermarked images and an intrinsic robustness.

6.2. Description and structuring

6.2.1. *Automatic classification of radiological reports for clinical care*

Participants: Anne-Lyse Minard, Alfonso Gerevini [Università degli Studi di Brescia, Italy], Alberto Lavelli [Fondazione Bruno Kessler, Italy], Alessandro Maffi [Università degli Studi di Brescia, Italy], Roberto Maroldi [Università degli Studi di Brescia, Italy, Azienda Socio Sanitaria Territoriale Spedali Civili di Brescia, Italy], Ivan Serina [Università degli Studi di Brescia, Italy], Guido Squassina [Azienda Socio Sanitaria Territoriale Spedali Civili di Brescia, Italy].

Radiological reporting generates a large amount of free-text clinical narratives, a potentially valuable source of information for improving clinical care and supporting research. The use of automatic techniques to analyze such reports is necessary to make their content effectively available to radiologists in an aggregated form. In this paper we focus on the classification of chest computed tomography reports according to a classification schema proposed for this task by radiologists of the Italian hospital ASST Spedali Civili di Brescia. The proposed system is built exploiting a training data set containing reports annotated by radiologists. Each report is classified according to the schema developed by radiologists and textual evidences are marked in the report. The annotations are then used to train different machine learning based classifiers. We present in this paper a method based on a cascade of classifiers which make use of a set of syntactic and semantic features. The resulting system is a novel hierarchical classification system for the given task, that we have experimentally evaluated [5].

6.2.2. *Revisiting the medial axis for planar shape decomposition*

Participants: Yannis Avrithis, N. Papanelopoulos [NTU Athens], S. Kollias [Univ. Lincoln].

We introduce a simple computational model for planar shape decomposition that naturally captures most of the rules and salience measures suggested by psychophysical studies, including the minima and short-cut rules, convexity, and symmetry [7]. It is based on a medial axis representation in ways that have not been explored before and sheds more light into the connection between existing rules like minima and convexity. In particular, vertices of the exterior medial axis directly provide the position and extent of negative minima of curvature, while a traversal of the interior medial axis directly provides a small set of candidate endpoints for part-cuts. The final selection follows a prioritized processing of candidate part-cuts according to a local convexity rule that can incorporate arbitrary salience measures. Neither global optimization nor differentiation is involved. We provide qualitative and quantitative evaluation and comparisons on ground-truth data from psychophysical experiments. With our single computational model, we outperform even an ensemble method on several other competing models.

6.2.3. Is ATIS too shallow to go deeper for benchmarking Spoken Language Understanding models?

Participants: Christian Raymond, Frédéric Béchet [Aix Marseille University].

We started a collaboration about benchmarking scientific benchmarks. We started in [11] by the ATIS (Air Travel Information Service) corpus, that will be soon celebrating its 30th birthday. Designed originally to benchmark spoken language systems, it still represents the most well-known corpus for benchmarking Spoken Language Understanding (SLU) systems. In 2010, in a paper titled "What is left to be understood in ATIS?", Tur et al. discussed the relevance of this corpus after more than 10 years of research on statistical models for performing SLU tasks. Nowadays, in the Deep Neural Network (DNN) era, ATIS is still used as the main benchmark corpus for evaluating all kinds of DNN models, leading to further improvements, although rather limited, in SLU accuracy compared to previous state-of-the-art models. We propose in this paper to investigate these results obtained on ATIS from a qualitative point of view rather than just a quantitative point of view and answer the two following questions: what kind of qualitative improvement brought DNN models to SLU on the ATIS corpus? Is there anything left, from a qualitative point of view, in the remaining 5% of errors made by current state-of-the-art models?

6.2.4. KRAUTS: A German Temporally Annotated News Corpus

Participants: Anne-Lyse Minard, Strötgen Jannik [Max Planck Institute for Informatics, Germany], Lukas Lange [Max Planck Institute for Informatics, Germany], Manuela Speranza [Fondazione Bruno Kessler, Italy], Bernardo Magnini [Fondazione Bruno Kessler, Italy].

In recent years, temporal tagging, i.e., the extraction and normalization of temporal expressions, has become a vibrant research area. Several tools have been made available, and new strategies have been developed. Due to domain-specific challenges, evaluations of new methods should be performed on diverse text types. Despite significant efforts towards multilinguality in the context of temporal tagging, for all languages except English, annotated corpora exist only for a single domain. In the case of German, for example, only a narrative style corpus has been manually annotated so far, thus no evaluations of German temporal tagging performance on news articles can be made. In this paper, we present KRAUTS, a new German temporally annotated corpus containing two subsets of news documents: articles from the daily newspaper DOLOMITEN and from the weekly newspaper DIE ZEIT. Overall, the corpus contains 192 documents with 1,140 annotated temporal expressions, and has been made publicly available to further boost research in temporal tagging citejannik:hal-01844834.

6.2.5. Active learning to assist annotation of aerial Images in environmental surveys

Participants: Ewa Kijak, Mathieu Laroze [OBELIX team, IRISA], Romain Dambreville [OBELIX team, IRISA], Chloe Friguet [OBELIX team, IRISA], Sébastien Lefèvre [OBELIX team, IRISA].

Remote sensing technologies greatly ease environmental assessment over large study areas using aerial images, e.g. for monitoring and counting animals or ships. Such data are most often analyzed by a manual operator, leading to costly and non scalable solutions. If object detection algorithms are used to fasten and automate the counting processes, these algorithms need to have prior ground truth available, which is a time-consuming and tedious process for field experts or engineers. We introduced a method to assist the annotation process in aerial images by introducing an active learning algorithm, allowing interaction with the expert such as class confirmation or correction at the labeling stage, and querying the expert with groups of samples taken from the same image to ease user annotation. Usual active learning algorithms perform instance selection from the whole set of input data. In this work, the selection of the queried instances is constrained by requiring that they belong to a group, (a part of) an image in our case, to ease the annotator task as the queried instances are proposed in their comprehensive context. We defined a score to rank the images and identify the one that should be annotated at each iteration, based on both uncertainty and true positives. The main objective is to reduce the number of human interactions on the overall process, starting from a first annotated image, rather than reaching the maximum final accuracy. Therefore, the annotation cost is measured through the gain in interactions (corrections of the classifier decisions by the annotator) with respect to a labeling task from scratch. At each iteration, the classifier is retrained according to a specific subset of data. Several strategies have been compared and their performances regarding the interaction gain have been discussed [19], [36].

6.3. Search, linking and navigation

6.3.1. *Detecting fake news and tampered images in social networks*
Participants: Cédric Maigrot, Ewa Kijak, Vincent Claveau.

Social networks make it possible to share information rapidly and massively. Yet, one of their major drawback comes from the absence of verification of the piece of information, especially with viral messages. This is the issue addressed by the participants to the Verification Multimedia Use task of Mediaeval 2016. They used several approaches and clues from different modalities (text, image, social information).

One promising approach is to examine if the image (if any) has been doctored. In recent work [23], we study context-aware methods to localize tamperings in images from social media. The problem is defined as a comparison between image pairs: an near-duplicate image retrieved from the network and a tampered version. We propose a method based on local features matching, followed by a kernel density estimation, that we compare to recent similar approaches. The proposed approaches are evaluated on two dedicated datasets containing a variety of representative tamperings in images from social media, with difficult examples. Context-aware methods are proven to be better than blind image forensics approach. However, the evaluation allows to analyze the strengths and weaknesses of the contextual-based methods on realistic datasets.

In further work [9], [22], we explore the interest of combining and merging these approaches in order to evaluate the predictive power of each modality and to make the most of their potential complementarity.

6.3.2. *A Crossmodal Approach to Multimodal Fusion in Video Hyperlinking*
Participants: Christian Raymond, Guillaume Gravier, Vedran Vukotić.

With the recent resurgence of neural networks and the proliferation of massive amounts of unlabeled data, unsupervised learning algorithms became very popular for organizing and retrieving large video collections in a task defined as video hyperlinking. Information stored as videos typically contain two modalities, namely an audio and a visual one, that are used conjointly in multimodal systems by undergoing fusion. Multimodal autoencoders have been long used for performing multimodal fusion. In this work, we start by evaluating different initial, single-modal representations for automatic speech transcripts and for video keyframes. We progress to evaluating different autoencoding methods of performing multimodal fusion in an offline setup. The best performing setup is then evaluated in a live setup at TRECVID's 2016 video hyperlinking task. As in offline evaluations, we show that focusing on crossmodal translations as a way of performing multimodal fusion yields improved multimodal representations and that our simple system, trained in an unsupervised manner, with no external information information, defines the new state of the art in a live video hyperlinking setup. We conclude by performing an analysis on data gathered after the live evaluations at TRECVID 2016 and express our thoughts on the overall performance of our proposed system [8].

6.3.3. *A study on multimodal video hyperlinking with visual aggregation*

Participants: Mateusz Budnik, Mikail Demirdelen, Guillaume Gravier.

Video hyperlinking offers a way to explore a video collection, making use of links that connect segments having related content. Hyperlinking systems thus seek to automatically create links by connecting given anchor segments to relevant targets within the collection. In 2018, we pursued our long-term research effort towards multimodal representations of video segments in a hyperlinking system based on bidirectional deep neural networks, which achieved state-of-the-art results in the TRECVid 2016 evaluation. A systematic study of different input representations was done with a focus on the aggregation of the representation of multiple keyframes. This includes, in particular, the use of memory vectors as a novel aggregation technique, which provides a significant improvement over other aggregation methods on the final hyperlinking task. Additionally, the use of metadata was investigated leading to increased performance and lower computational requirements for the system [35].

6.3.4. *Opinion mining in social networks*

Participants: Anne-Lyse Minard, Christian Raymond, Vincent Claveau.

As part of the DeFT text-mining challenge, we participated in the elaboration of a task on fine-grained opinion mining in tweets [34] and to the analysis of the participants' results. We have also proposed systems [33] for each sub-task: (i) tweet classification according to the topic of the tweet, (ii) tweet classification according to their polarity, (iii) detection of the polarity markers and target of opinion in tweets. For the two first tasks, the approaches we proposed rely on a combination of boosting, decision trees and Recurrent Neural Networks. For the last task, we experimented with RNN coupled with a CRF layer. All of these systems performed very well and ranked in the best performing systems for each of the task.

6.3.5. *Biomedical Information Extraction in social networks*

Participants: Anne-Lyse Minard, Christian Raymond, Vincent Claveau.

This year, we participated in SMM4H challenge about extracting medical information from social networks. Faour tasks were proposed: (i) detection of posts mentioning a drug name, (ii) classification of posts describing medication intake, (iii) classification of adverse drug reaction mentioning posts, (iv) Automatic detection of posts mentioning vaccination behavior. In [24], we presented the systems developed by IRISA to participate to these four tasks. For these tweet classification tasks, we adopt a common approach based on recurrent neural networks (BiLSTM). Our main contributions are the use of certain features, the use of Bagging in order to deal with unbalanced datasets, and on the automatic selection of difficult examples. These techniques allow us to reach 91.4, 46.5, 47.8, 85.0 as F1-scores for Tasks 1 to 4, ranking us among the 3 first participants for each task.

6.3.6. *Information Extraction in the biomedical domain*

Participants: Clément Dalloux, Vincent Claveau, N. Grabar [STL-CNRS].

Automatic detection of negated content is often a pre-requisite in information extraction systems, especially in the biomedical domain. Following last year work, we propose two main contributions in this field [43]. We first introduced a new corpora built with excerpts from clinical trial protocols in French and Brazilian Portuguese, describing the inclusion criteria for patient recruitment. The corpora are manually annotated for marking up the negation cues and their scope. Secondly, two supervised learning approaches are been proposed for the automatic detection of negation. Besides, one of the approaches is validated on English data from the state of the art: the approach shows very good results and outperforms existing approaches, and it also yields comparable results on the French data.

We also have developed other data-sets (annotated corpora). Indeed, textual corpora are extremely important for various NLP applications as they provide information necessary for creating, setting and testing these applications and the corresponding tools. They are also crucial for designing reliable methods and reproducible results. Yet, in some areas, such as the medical area, due to confidentiality or to ethical reasons, it is complicated and even impossible to access textual data representative of those produced in these areas. We propose the CAS corpus [14] built with clinical cases, such as they are reported in the published scientific literature in French. We describe this corpus, containing over 397,000 word occurrences, and its current annotations (PoS, lemmas, negation, uncertainty).

As part of this work, we also developed software available as web-services on http://allgo.inria.fr (see the Software section).

6.3.7. *Revisiting Oxford and Paris: Large-Scale Image Retrieval Benchmarking*

Participants: Yannis Avrithis, F. Radenovic [Univ. Prague], Ahmet Iscen [Univ. Prague], Giorgos Tolias [Univ. Prague], Ondra Chum [Univ. Prague].

In this work [27] we address issues with image retrieval benchmarking on standard and popular Oxford 5k and Paris 6k datasets. In particular, annotation errors, the size of the dataset, and the level of challenge are addressed: new annotation for both datasets is created with an extra attention to the reliability of the ground truth. Three new protocols of varying difficulty are introduced. The protocols allow fair comparison between different methods, including those using a dataset pre-processing stage. For each dataset, 15 new challenging queries are introduced. Finally, a new set of 1M hard, semi-automatically cleaned distractors is selected. An extensive comparison of the state-of-the-art methods is performed on the new benchmark. Different types of methods are evaluated, ranging from local-feature-based to modern CNN based methods. The best results are achieved by taking the best of the two worlds. Most importantly, image retrieval appears far from being solved.

6.3.8. *Unsupervised object discovery for instance recognition*

Participants: Oriane Siméoni, Yannis Avrithis, Ahmet Iscen [Univ. Prague], Giorgos Tolias [Univ. Prague], Ondra Chum [Univ. Prague].

Severe background clutter is challenging in many computer vision tasks, including large-scale image retrieval. Global descriptors, that are popular due to their memory and search efficiency, are especially prone to corruption by such a clutter. Eliminating the impact of the clutter on the image descriptor increases the chance of retrieving relevant images and prevents topic drift due to actually retrieving the clutter in the case of query expansion. In this work, we propose a novel salient region detection method. It captures, in an unsupervised manner, patterns that are both discriminative and common in the dataset. Saliency is based on a centrality measure of a nearest neighbor graph constructed from regional CNN representations of dataset images. The descriptors derived from the salient regions improve particular object retrieval, most noticeably in a large collections containing small objects [28].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. *Few shot learning for object recognition in aerial images (CIFRE PhD)*

Participants: Yannis Avrithis, Yann Lifchitz.

Duration: 3 years, started in March 2018
Partner: Safran

This is a CIFRE PhD thesis project aiming to study architectures and learning techniques most suitable for object recognition from few samples and to validate these approaches on multiple recognition tasks and use-cases related to aerial images.

7.1.2. *Incremental dynamic construction of knowledge bases from text mining (CIFRE PhD)*

Participants: Guillaume Gravier, Pascale Sébillot, Cyrielle Mallart.

Duration: 3 years, started in Dec. 2018
Partner: Ouest France

In the context of a newspaper, the thesis explores the combination of text mining and knowledge representation techniques to assist the extraction, interpretation and validation of valuable pieces of information from the journal's content so as to incrementally build a full-scale knowledge base. This thesis is in close relation with the iCODA Inria Project Lab, with direct contribution to the project's results.

7.1.3. *Embedding heterogeneous data for directory search (CIFRE PhD)*

Participants: Guillaume Gravier, Vincent Claveau, François Torregrossa.

Duration: 3 years, started in Dec. 2018
Partner: SoLocal

The thesis aims at learning how to jointly exploit heterogeneous sources of information (e.g., names, activity sector, user profiles, queries, etc.) in the design of neural network embeddings for information retrieval and language understanding. Applications cover natural language query analysis and personalized information retrieval in Pagesjaunes' directory.

7.1.4. *Active learning on adaptive representations for object detection in high-resolution imaging (CIFRE PhD)*

Participants: Ewa Kijak, Mathieu Laroze.

Duration: 3 years, started in Jun. 2016
Partner: Wipsea

Wipsea is a Rennaise startup, created in 2013 whose core business is image analysis for ecology. WIPSEA helps biologists, Unmanned Aerial Vehicle (UAV) companies and environmental consultants to automatically process automatically captured pictures by proposing image-processing algorithms that detect and characterize all kinds of animals such as dugongs, whales, elephants, and orangutans... The main objective of the thesis is to develop a detection method to optimize the interaction between the expert and the system in order to converge as quickly as possible to an object detection solution adapted to a given task.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *CominLabs Project Linking Media in Acceptable Hypergraphs (LIMAH)*

Participants: Vincent Claveau, Guillaume Gravier, Pascale Sébillot.

Duration: 4.5 years, started in April 2014
Partners: Telecom Bretagne (IODE), Univ. Rennes II (CRPCC, PREFics), Univ. Nantes (LINA/TAL)
URL: http://limah.irisa.fr

LIMAH aims at exploring hypergraph structures for multimedia collections, instantiating actual links reflecting particular content-based proximity—similar content, thematic proximity, opinion expressed, answer to a question, etc. Exploiting and developing further techniques targeting pairwise comparison of multimedia contents from an NLP perspective, LIMAH addresses two key issues: How to automatically build from a collection of documents an hypergraph, i.e., a graph combining edges of different natures, which provides exploitable links in selected use cases? How collections with explicit links modify usage of multimedia data in all aspects, from a technology point of view as well as from a user point of view? LIMAH studies hypergraph authoring and acceptability taking a multidisciplinary approach mixing ICT, law, information and communication science as well as cognitive and ergonomy psychology.

8.1.2. *CominLabs Project BigCLIN*

Participants: Vincent Claveau, Ewa Kijak, Clément Dalloux.

Duration: 3 years, started in September 2016
Partners: STL-CNRS, Inserm/CHU Rennes, Inria
URL: https://bigclin.cominlabs.u-bretagneloire.fr/fr

Data collected or produced during clinical care process can be exploited at different levels and across different domains. Yet, a well-known challenge for secondary use of health big data is that much of detailed patient information is embedded in narrative text, mostly stored as unstructured data. The project proposes to address the essential needs when reusing unstructured clinical data at a large scale. We propose to develop new clinical records representation relying on fine-grained semantic annotation thanks to new NLP tools dedicated to French clinical narratives. To efficiently map this added semantic information to existing structured data for further analysis at big scale, the project also addresses distributed systems issues: scalability, management of uncertain data and privacy, stream processing at runtime, etc.

8.1.3. *Computer vision for smart phones (MobilAI)*

Participants: Yannis Avrithis, Mateusz Budnik.

Duration: 2 years, started in September 2018
Partners: Lamark, Quai des Apps, AriadNext

The ability of our mobile devices to process visual information is currently not limited by their camera or computing power but by the network. Many mobile apps suffer from long latency due to data transmitted over the network for visual search. MobilAI aims to provide fast visual recognition on mobile devices, offering quality user experience whatever the network conditions. The idea is to transfer efficient deep learning solutions for image classification and retrieval onto embedded platforms such as smart phones. The intention is to use such solutions in B2B and B2C application contexts, for instance recognizing products and ordering online, accessing information about artifacts in exhibitions, or identifying identity documents. In all cases, visual recognition is performed on the device, with minimal or no access to the network.

8.2. National Initiatives

8.2.1. *ANR Project IDFRAud*

Participant: Teddy Furon.

Duration: 3 years, started in Feb. 2015
Partners: AriadNext, IRCGN, École Nationale Supérieure de Police

The IDFRAud project consists in proposing an automatic solution for ID analysis and integrity verification. Our ID analysis goes through three processes: classification, text extraction and ID verification. The three processes rely on a set of rules that are externalized in formal manner in order to allow easy management and evolving capabilities. This leads us to the ID knowledge management module. Finally, IDFRAud addresses the forensic link detection problem and to propose an automatic analysis engine that can be continuously applied on the detected fraud ID database. Cluster analysis methods are used to discover relations between false IDs in their multidimensional feature space. This pattern extraction module will be coupled with a suitable visualization mechanism in order to facilitate the comprehension and the analysis of extracted groups of inter-linked fraud cases.

8.2.2. *FUI 19 NexGenTV*

Participants: Vincent Claveau, Guillaume Gravier, Ewa Kijak, Anne-Lyse Minard.

Duration: 2.5 years, started in May 2015
Partners: Eurecom, Avisto Telecom, Wildmoka, Envivio-Ericsson

Television is undergoing a revolution, moving from the TV screen to multiple screens. Today's user watches TV and, at the same time, browses the web on a tablet, sends SMS, posts comments on social networks, searches for complementary information on the program, etc. Facing this situation, NexGen-TV aims at developing a generic solution for the enrichment, the linking and the retrieval of video content targeting the cost-cutting edition of second screen and multiscreen applications for broadcast TV. The main outcome of the project will be a software platform to aggregate and distribute video content via a second-screen edition interface connected to social media. The curation interface will primarily make use of multimedia and social media content segmentation, description, linking and retrieval. Multiscreen applications will be developed on various domaine, e.g., sports, news.

8.2.3. Inria Project Lab Knowledge-driven data and content collaborative analytics (iCODA)

Participants: Laurent Amsaleg, Vincent Claveau, Cheikh Brahim El Vaigh, Guillaume Gravier, Pascale Sébillot.

Duration: 4.5 years, started in April 2017
Partners: Inria project-teams Linkmedia, CEDAR, GraphIK and ILDA, with Ouest-France, Le Monde and AFP

One of today's major issues in data science is the design of algorithms that allow analysts to efficiently infer useful information and knowledge by collaboratively inspecting heterogeneous information sources, from structured data to unstructured content. Taking data journalism as an emblematic use-case, the goal of the project is to develop the scientific and technological foundations for knowledge- mediated user-in-the-loop collaborative data analytics on heterogenous information sources, and to demonstrate the effectiveness of the approach in realistic, high-visibility use-cases. The project stands at the crossroad of multiple research fields—content analysis, data management, knowledge representation, visualization—that span multiple Inria themes, and counts on a club of major press partners to define usage scenarios, provide data and demonstrate achievements.

8.2.4. Inria-BNF: Classification d'images patrimoniales (CIP)

Participants: Laurent Amsaleg, Guillaume Gravier, Ewa Kijak, Yannis Avrithis.

Duration: 1 year, started in Dec 2018

This project is within the context of the collaborations between Inria and the French Ministry of Culture. In that context, we have started a collaboration with the French National Library (BNF) which collects, preserves and makes known the national documentary heritage. This collaboration aims at facilitating the automatic classification of heritage images through the use of recent deep-learning techniques. Such images are quite specific: they are not at all similar with what deep-learning techniques are used to work with, that is, the classification of heritage images does not target modern categories such as planes, cars, cats and dogs because this is irrelevant and because heritage collections do not include images of contemporary objects. Furthermore, heritage images come in vast quantities, but they are little annotated and deep-learning techniques can hardly rely on massive annotations to easily learn. Last, the learning has to be continuous as curators may need to add or modify existing classes, without re-learning everything from scratch.

The techniques of choice to reach that goal include the semi-supervised learning, low-shot learning techniques, knowledge transfer, fine tuning existing models, etc.

8.3. European Initiatives

8.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: CHIST ERA
- Project acronym: ID_IOT
- Project title: Identification for the Internet of things
- Duration: 3 years, started in Oct 2016.

- Coordinator: Boris Skoric (Eindhoven Univ. of Technology (NL))
- Other partners: Inria-RBA (Teddy Furon, Marzieh Gheisari), Univ. of Geneva (CH)
- Abstract: The IoT will contain a huge number of devices and objects that have very low or nonexistent processing and communication resources, coupled to a small number of high-power devices. The weakest devices, which are most ubiquitous, will not be able to authenticate themselves using cryptographic methods. This project addresses these issues using physical unclonable functions (PUFs). PUFs, and especially quantum readout PUFs, are ideally suited to the IoT setting because they allow for the authentication and identification of physical objects without requiring any crypto or storage of secret information.

  Furthermore, we foresee that back-end systems will not be able to provide security and privacy via cryptographic primitives due to the sheer number of IoT devices. Our plan is to address these problems using privacy preserving database structures and algorithms with good scaling behaviour. Approximate nearest neighbour (ANN) search algorithms, which have remarkably good scaling behaviour, have recently become highly efficient, but do not yet have the right security properties and have not yet been applied to PUF data. Summarised in a nutshell, the project aims to improve the theory and practice of technologies such as PUFs and ANN search in the context of generic IoT authentication and identification scenarios.

Program: Joint Programming Initiative Cultural Heritage
- Project acronym: READ-IT
- Project title: Reading Europe: Advanced Data Investigation Tools
- Duration: June 2018 - May 2021
- Coordinator: Le Mans Université, Institute of Human and Social Sciences-Digital Humanities (FR)
- Other partners: CNRS, Institut de Recherche en Informatique et Systèmes Aléatoires (FR); The Open University (UK); Utrecht University, Digital Humanities Lab (NL); Czech Litterature Institute (CZ)
- Abstract: READ-IT will build a unique large-scale, user-friendly, open access, semantically-enriched investigation tool to identify and share groundbreaking evidence about 18th-21st century Cultural Heritage of reading in Europe. The interdisciplinary collaboration between digital humanists, human & social sciences scholars and computer researchers will investigate innovative ways of gathering new resources through crowdsourcing and web-crawling as well as linking and reusing preexisting datasets.

8.4. International Initiatives

8.4.1. *NSFC Project: Using transfer learning to understand visual objects and their relationships*

Participant: Miaojing Shi.

Duration: 2 years, start in Jan 2019
Partners: Tongji University, China

This project is supported by China National Joint Research Fund for Overseas Chinese Scholars. Machine Perception tasks have flourished since the advent of deep learning techniques. Next key problem lies on visual scene understanding. To make sense of visual scenes, we need to rely on the visual object relationships inside. The challenge for this task is that 1) the training data is limited, on particular those unusual seen objects/object relationships; 2) visual relationships become complicated and diverse with an increase of object numbers. This research shall employ the transfer learning methods to transfer available knowledge of visual relationships to new objects with unknown relationships. The significance of this research is not just to enhance the machine perception ability; it allows us to leverage a relatively small amount of expensively annotated images to detect new objects and their relationships in a much larger dataset without or with only cheap image-level labels.

8.4.2. Inria International Partners

8.4.2.1. Informal International Partners

- Peking University, China
- Sichuan University, China
- Czech Technical University, Czech Republic
- Computer Science Department, IT University of Copenhagen, Denmark
- Tampere University of Technology, Finland
- National Institute of Informatics, Tokyo, Japan

8.4.3. Participation in Other International Programs

- STIC-AMSUD Project : TRANSFORM
  Participants: Simon Malinowski, Guillaume Gravier, Laurent Amsaleg.

  TRANSFORM is a research project that involves Linkmedia Team, PUC Minas, Brazil and Univ. of Chile. It aims at studying complex transformations of multimedia data in order to facilitate its manipulation. TRANSFORM focuses on transforming multimedia data into compact representations that are suited for indexing and retrieval purposes.

- ConFAP-CNRS Project: FIGTEM
  Participants: Vincent Claveau, Ewa Kijak, Clément Dalloux.

  FIGTEM is a research project that involves STL-CNRS, CHU Rennes, PUC Parana, Curitiba and led by LinkMedia. This project aims at developing natural language processing methods, including information extraction and indexing, dedicated to the clinical trial domain. The goal is to populate a formal representation of patients (via their electronic patient records) and clinical trial data in different languages (French, English, Portuguese).

- NSFC Project : Perception and self-learning of service robot in dynamic scenarios
  Participant: Miaojing Shi.

  This project is the Key Program of National Natural Science Foundation of China, which involves Miaojing Shi from Linkmedia and is led by Tongji University. It aims at improving the perception of service robot in dynamic scenarios through self-learning.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Zenilton Kleber (PUC Minas, Brazil), November 2018, 2 days
Silvio Guimaraes (PUC Minas, Brazil), December 2018, 10 days
Michael Houle (National Institute of Informatics), July 2018, 7 days

8.5.1.1. Internships

Yohann Gumiel and Lucas Oliveira, PhD students at PUC Parana, Curitiba, Brazil, stayed 3 months (June-August 2018), in Rennes in the framework of the FIGTEM project.
Yanlin Qian, PhD student at Tampere University of Technology, 3 months (Jun-August 2018)

8.5.2. *Visits to International Teams*

8.5.2.1. Research Stays Abroad

- Guillaume Gravier & Simon Malinowski, PUC Minas, Brazil, March 2018, 1 week.
- Vincent Claveau, PUC Parana, Curitiba, Brazil, November 2018, 1 week, in the framework of the FIGTEM project.
- Miaojing Shi, Tongji University, China, March 2018, 1 week.
- Miaojing Shi, Xian Jiaotong University, China, April 2018, 1 week.
- Yannis Avrithis, National and Kapodistrian University of Athens, 3 visits on February, May and October 2018, 1 week each.
- Laurent Amsaleg & Teddy Furon, National Institute of Informatics, March 2018, 1 week

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. *Scientific Events Organisation*

- LinkMedia organized the joint conferences CORIA - TALN in Rennes, May 2018, including several workshops and a hackathon on Fake News detection. Full program: https://project.inria.fr/coriataln2018/fr/.
- Guillaume Gravier and Vincent Claveau organized a special conference about "IA and media" in the framework of the European Cyber Week in Nov 2018, Rennes. Full program: https://www.european-cyber-week.eu/medias-ia-et-cyber.html

9.1.1.1. General Chair, Scientific Chair

- Laurent Amsaleg is general co-chair of ACM Intl. Conf. on Multimedia 2019.
- Guillaume Gravier is technical program co-chair for ACM Intl. Conf. on Multimedia 2019.
- Vincent Claveau was general chair of CORIA-TALN 2018.

9.1.1.2. Member of the Organizing Committees

- Guillaume Gravier was special session co-chair for IEEE Intl. Conf. on Content-Based Multimedia Indexing 2018
- Simon Malinowski was member of the Organizing Committee of the AALTD 18 Workshop (co-located with ECML-PKDD 18)
- Ewa Kijak was publication chair for IEEE Intl. Conf. on Content-Based Multimedia Indexing 2018

9.1.2. *Scientific Events Selection*

9.1.2.1. Chair of Conference Program Committees

- Vincent Claveau and Pascale Sébillot were program chairs of TALN 2018.
- Anne-Lyse Minard was demo chair of CORIA-TALN 2018.

9.1.2.2. Member of the Conference Program Committees

- Laurent Amsaleg was a PC member of: Extraction et Gestion de Connaissances, European Signal Processing Conference, ACM International Conference on Multimedia Retrieval, Multimedia Modeling, Content-Based Multimedia Indexing, IEEE International Conference on Multimedia & Expo, International Conference on Similarity Search and Applications, Multimedia Information Retrieval and Applications.

- Guillaume Gravier was a PC member of: ACM International Conference on Multimedia, ACM International Conference on Multimedia Retrieval, IEEE International Conference on Multimedia & Expo, European Conference on Information Retrieval, International Symposium on Multimedia, International Conference on Multimedia Modeling

- Simon Malinowski was a PC member of: Advanced analytics and Learning on Temporal Data 2018

- Vincent Claveau was a PC member of: European Conference on Information Retrieval (ECIR), Computational Linguistics (CoLing), International Conference in Artificial Intelligence (IJCAI), International Conference on MultiMedia Modeling (MMM), EMNLP (demo), ACL (demo), HumaNum workshop with EGC, TexMine workshop with EGC, DeFT workshop with TALN.

- Miaojing Shi was a PC member of: International Conference on Computer Vision and Pattern Recognition (CVPR), International Conference on Image Processing, Applications and Systems (IPAS), Asian Conference on Computer Vision (ACCV).

- Ewa Kijak was a PC member of: ACM International Conference on Multimedia, IEEE Intl. Conf. on Content-Based Multimedia Indexing

- Yannis Avrithis was a PC member of: IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR), European Conference on Computer Vision (ECCV).

- Teddy Furon was a PC member of: IEEE Workshop on Information Forensics and Security, IEEE International Conference on Image Processing, International Conference on Similarity Search and Applications.

- Pascale Sébillot was a PC member of: AAAI Conf. on Artificial Intelligence, European Conference on Information Retrieval (ECIR), International Joint Conference on Artificial Intelligence and European Conference on Artificial Intelligence (IJCAI-ECAI), Language Resources and Evaluation Conference (LREC), Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT).

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Laurent Amsaleg was guest editor for the special issue of the Information Systems Journal, best papers from the SISAP 2016 conference.

- Laurent Amsaleg was guest editor for the special issue of the Information Systems Journal, best papers from the SISAP 2017 conference.

- Guillaume Gravier is associate editor of IEEE Trans. on Multimedia

- Vincent Claveau is Chief Editor of the ISTE journal RIDoWS

- Vincent Claveau is member of the editorial board of the journal Traitement Automatique des Langues (TAL)

- Christian Raymond is member of the Editorial board of the journal Discours

- Pascale Sébillot is editor of the Journal Traitement Automatique des Langues.

- Pascale Sébillot is member of the editorial board of the Journal Traitement Automatique des Langues.

9.1.3.2. Reviewer - Reviewing Activities

- Laurent Amsaleg was a reviewer for: IEEE Transactions on Information Forensics and Security, ACM Transactions on Information Systems.

- Christian Raymond was a reviewer for: Multimedia Tools and Applications

- Vincent Claveau was a reviewer for: Multimedia Tools and Applications, Technique et Sciences Informatiques, Traitement Automatique des Langues

- Miaojing Shi was a reviewer for: IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Image Processing, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia
- Yannis Avrithis was a reviewer for: IEEE Transactions on Pattern Analysis and Machine Intelligence
- Teddy Furon was a reviewer for: IEEE Transactions on Information Forensics and Security, IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, MDPI Entropy Journal.
- Pascale Sébillot reviewed for Traitement Automatique des Langues.

9.1.4. Invited Talks

- Miaojing Shi gave an invited talk at Shanghai Jiaotong University, Shanghai.
- Miaojing Shi gave an invited talk at Tongji University, Shanghai.
- Miaojing Shi gave an invited keynote at FranceIsAI 2018, Oct 17-19, Paris.

9.1.5. Leadership within the Scientific Community

- Laurent Amsaleg is a member of the Steering Committee of SISAP for the 2016-2020 term
- Guillaume Gravier is member of the Community Council of the Mediaeval Multimedia Evaluation series
- Guillaume Gravier coordinated the answer to the call "Institut inter disciplinaire de l'IA" for Rennes.
- Guillaume Gravier was president of the Scientific Evaluation Committee of the National Research Agency for the theme 'Knowledge, data, content, big data - AI'
- Guillaume Gravier is a member of the scientific committee of the preGDR Traitement Automatique des Langues
- Vincent Claveau is a deputy head of the GdR CNRS MaDICS www.madics.fr
- Vincent Claveau is the finance head of ARIA www.asso-aria.org
- Pascale Sébillot is a member of the permanent steering committee of Conf. Francophone en Traitement Automatique des Langues Naturelles.
- Pascale Sébillot is a member of the committee of leaders of the preGDR Traitement Automatique des Langues reflection axes (in charge of the Intermodality and Multimodality axis (with B. Favre, LIS Marseille)).

9.1.6. Scientific Expertise

- Miaojing Shi was expert evaluator for the French National Research Agency (ANR).

9.1.7. Research Administration

- Guillaume Gravier is deputy director of Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA, UMR 6074)
- Guillaume Gravier is a member of the Board of the technology cluster Images & Réseaux.
- Guillaume Gravier is a member of the Board of the Comité des Projets of Inria - Rennes Bretagne Atlantique.
- Pascale Sébillot is a member of the Conseil National des Universités 27th section (computer science).
- Pascale Sébillot is the director of the Computer Science Laboratory, INSA Rennes.
- Pascale Sébillot is the deputy director of the Scientific Advisory Committee of IRISA UMR 6074.
- Pascale Sébillot is a member of the theses advisory committee of the MathSTIC doctoral school.
- Pascale Sébillot is a member of the board of the MathSTIC doctoral school.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence: Laurent Amsaleg & Teddy Furon, Indexation multimédia, 10h, L3-cours invité, ENS Rennes, France

Licence: Laurent Amsaleg, Bases de données avancées, 2h, L3-option génie mathématique, INSA Rennes, France

Licence: Guillaume Gravier, Probability and statistics, 16h, L3, INSA Rennes, France

Licence: Guillaume Gravier, Natural Language Processing, 12h, L3 & M1, INSA Rennes, France

Engineering school: Vincent Claveau, Machine Learning, 16h, 3rd year, INSA Rennes, France

Engineering school: Vincent Claveau, Information Retrieval, 10h, 3rd year, ENSSAT Lannion, France

Master: Vincent Claveau, Information Retrieval, 10h, M2 MIAGE, Univ. Rennes, France

Master: Laurent Amsaleg, Bases de données avancées, 25h, M2, INSA Rennes, France

Master: Guillaume Gravier, Data analysis and probabilistic modeling, 30h, M2, University Rennes 1, France

Master: Simon Malinowski, Méthodes de prédiction, 32h, M2 Miage, Univ. Rennes

Master: Simon Malinowski, Apprentissage automatique, 24h, M2 Miage, Univ. Rennes

Master: Simon Malinowski, Fouille de données symboliques, 12h, M2 Miage, Univ. Rennes

Master: Christian Raymond, Dialogue, 4H, M2 SIF, Univ. Rennes

Master: Ewa Kijak, Image processing, 67h, M1, ESIR, France

Master: Ewa Kijak, Supervised machine learning, 15h, M2R, University Rennes 1, France

Master: Ewa Kijak, Supervised machine learning, 15h, M2, University Rennes 1, France

Master: Ewa Kijak, Image classification, 45h, M1, ESIR, France

Master: Ewa Kijak, Image indexing, 17h, M2, University Rennes 1, France

Master: Ewa Kijak, Indexing and multimedia databases, 15h, M2, ENSSAT, France

Master: Ewa Kijak, Computer vision, 22h, M2, ESIR, France

Master: Yannis Avrithis, Deep learning for vision, 20h, M2 SIF, France

Master: Teddy Furon, Rare event simulations, 20h, M2 option génie mathématique, INSA, France

Master: Pascale Sébillot, Natural Language Programming, 6h, M1, INSA Rennes, France

9.2.2. Supervision

PhD in progress: Hanwei Zhang, Deep Learning in Adversarial Contexts, October 2017, Laurent Amsaleg, Yannis Avrithis, Teddy Furon & Ewa Kijak

PhD in progress: Marzieh Gheisari-Khorasgani, Secure identification in the Internet of Things, January 2018, Laurent Amsaleg & Teddy Furon

PhD in progress: Cédric Maigrot, Détection de fausses informations dans les réseaux sociaux, October 2015, Laurent Amsaleg, Vincent Claveau & Ewa Kijak

PhD in progress: Ricardo Carlini-Sperandio, Unsupervised motif mining in multimedia time series, August 2015, Laurent Amsaleg & Guillaume Gravier

PhD in progress: Antoine Perquin, Universal speech synthesis through embeddings of massive heterogeneous data, October 2017, Laurent Amsaleg, Gwénolé Lecorvé & Damien Lolive (with Expression, IRISA team)

PhD in progress: Tong Xue, Visualization and collaborative analysis of document collections and extracted knowledge for data journalism, October 2018, Laurent Amsaleg & Anastasia Bezerianos

PhD in progress: Mathieu Laroze, Active learning on adaptive representations for object detection in high-resolution imaging, started June 2016, Romain Dambreville, Chloe Friguet, Ewa Kijak and Sebastien Lefevre (with OBELIX, IRISA team)

PhD in progress: Mikail Demirdelen, User-adapted multi-document multimedia synthesis, started Oct. 2016, Guillaume Gravier & Pascale Sébillot

PhD in progress: Cheikh Brahim El Vaigh, Incremental content to data linking leveraging ontological knowledge in data journalism, started Oct. 2017, Guillaume Gravier, Pascale Sébillot and François Goasdoué (with CEDAR, Inria team)

PhD in progress: Cyrielle Mallart, Incremental dynamic construction of knowledge bases from text mining, started Dec. 2018, Guillaume Gravier & Pascale Sébillot (with Ouest-France)

PhD in progress: Oriane Siméoni, Invariance and supervision in visual learning, started Oct. 2016, Yannis Avrithis & Guillaume Gravier

PhD in progress: Zhaohui Yang, learning visual models with minimal human supervision, Jan. 2018, Miaojing Shi & Yannis Avrithis & Chao Xu (Peking University, Beijing)

PhD in progress: Feng Xiong, Perception and self-learning of service robot in dynamic scenarios. Nov. 2019, Miaojing Shi & Qijun Chen (Tongji University, Shanghai)

PhD in progress: Yann Lifchitz, Few shot learning for object recognition in aerial images. Started Mar. 2018, Yannis Avrithis, Sylvaine Picard (Safran) and Andrei Bursuc (Valeo).

9.2.3. *Juries*

Laurent Amsaleg

- PhD, reviewer, André Belchior Mourão, *Towards an Architecture for Efficient Distributed Search of Multimodal Information*, Universidade Nova de Lisboa, April 2018.

Ewa Kijak

- PhD, member, Himalaya Jain, *Learning compact representations for large scale image search*, University Rennes 1, June 2018.

Guillaume Gravier

- HDR, reviewer, Anthony Larcher, Le Mans Université, Déc. 2018.
- PhD, reviewer, Sébastien Delecraz, Aix-Marseille Université, Dec. 2018.

Teddy Furon

- PhD, member, Dimche Kostadinov, *Nonlinear Transform Learning: Model, Applications, and Algorithms*

Pascale Sébillot

- PhD, member, Alexandre Bérard, Université Grenoble Alpes, June 2018.
- PhD, president, Caroline Langlet, Télécom ParisTech, Sept. 2018.

Vincent Claveau was reviewer for the mid-term PhD auditions of Anna Koroleva, LIMSI, Christopher Norman, LIMSI, Zheng Zhang, LIMSI.

9.3. Popularization

9.3.1. *Articles and contents*

- Interviews in order to popularize: Guillaume Gravier was interviewed on AI technology for Sciences Ouest
- Vincent Claveau gave interviews about Fake News detection in the following journals: Stratégies 06/06/2018 and 12/10/2018, LCI.fr 8/06/2018, Le Monde.fr 24/05/2018, Sciences Ouest n363, Sciences et Avenir 30/05/2018 and xx/12/2018, 01net n894
- Ewa Kijak was interviewed by LeTemps.ch

9.3.2. Interventions

- Public exhibitions: Vincent Claveau and Guillaume Gravier presented at Futur.e.s
- Vincent Claveau participated in a press conference organized by the CNRS in May 2018
- Vincent Claveau gave an invited talk about NLP and fake news at the Forum des sciences cognitives in Nancy
- Vincent Claveau was heard at the French National Assembly by the parliamentary group *Économie numérique de la donnée, de la connaissance et de l'intelligence artificielle*
- Miaojing Shi was invited for a panel discussion about *China is AI: perspective on the booming Chinese AI ecosystem* at FranceIsAI, Oct 2018, Paris.
- Miaojing Shi was invited to the 2nd Edition of the Purple Night organized by the France China Committee in order to value the Franco-Chinese cooperation.

9.3.3. Internal action

- Regional events: Guillaume Gravier was invited conferencer on AI for Café des Sciences (general public audience, $\sim$100 pers.)

9.3.4. Creation of media or tools for science outreach

- Cédric Maigrot and Vincent Claveau participated to the educational video from Cité des Sciences about Fake News detection https://bit.ly/2Rz7MAp

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] T. FURON. *Binary Tardos Codes and Zero-bit Watermarking*, Université de Rennes 1, October 2018, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01932766

Articles in International Peer-Reviewed Journals

[2] L. AMSALEG, O. CHELLY, T. FURON, S. GIRARD, M. E. HOULE, K.-I. KAWARABAYASHI, M. NETT. *Extreme-value-theoretic estimation of local intrinsic dimensionality*, in "Data Mining and Knowledge Discovery", November 2018, vol. 32, n⁰ 6, pp. 1768–1805 [*DOI :* 10.1007/S10618-018-0578-6], https://hal.archives-ouvertes.fr/hal-01864580

[3] L. AMSALEG, G. T. GUDMUNDSSON, B. Þ. JÓNSSON, M. J. FRANKLIN. *Prototyping a Web-Scale Multimedia Retrieval Service Using Spark*, in "ACM Transactions on Multimedia Computing, Communications and Applications", July 2018, vol. 14, n⁰ 3s, pp. 1 - 24 [*DOI :* 10.1145/3209662], https://hal.archives-ouvertes.fr/hal-01853379

[4] P. BOSILJ, E. KIJAK, S. LEFÈVRE. *Partition and Inclusion Hierarchies of Images: A Comprehensive Survey*, in "Journal of Imaging", February 2018, vol. 4, n⁰ 2, pp. 1-31 [*DOI :* 10.3390/JIMAGING4020033], https://hal.archives-ouvertes.fr/hal-01848445

[5] A. GEREVINI, A. LAVELLI, A. MAFFI, R. MAROLDI, A.-L. MINARD, I. SERINA, G. SQUASSINA. *Automatic classification of radiological reports for clinical care*, in "Artificial Intelligence in Medicine", June 2018, pp. 1-10 [*DOI :* 10.1016/J.ARTMED.2018.05.006], https://hal.archives-ouvertes.fr/hal-01844840

[6] C. HUERTAS, R. JUÁREZ-RAMÍREZ, C. RAYMOND. *Heat Map based Feature Ranker: In Depth Comparison with Popular Methods*, in "Intelligent Data Analysis", 2018, vol. 22, n⁰ 5, pp. 1009-1037, forthcoming [*DOI :* 10.3233/IDA-173481], https://hal.inria.fr/hal-01848544

[7] N. PAPANELOPOULOS, Y. AVRITHIS, S. KOLLIAS. *Revisiting the medial axis for planar shape decomposition*, in "Computer Vision and Image Understanding", November 2018, pp. 1-20, forthcoming [*DOI :* 10.1016/J.CVIU.2018.10.007], https://hal.inria.fr/hal-01930939

[8] V. VUKOTIĆ, C. RAYMOND, G. GRAVIER. *A Crossmodal Approach to Multimodal Fusion in Video Hyperlinking*, in "IEEE MultiMedia", 2018, vol. 25, n⁰ 2, pp. 11-23, forthcoming [*DOI :* 10.1109/MMUL.2018.023121161], https://hal.inria.fr/hal-01848539

Articles in National Peer-Reviewed Journals

[9] C. MAIGROT, E. KIJAK, V. CLAVEAU. *Fusion par apprentissage pour la détection de fausses informations dans les réseaux sociaux*, in "Revue des Sciences et Technologies de l'Information - Série Document Numérique", 2018, pp. 1-26, forthcoming [*DOI :* 10.3166/DN.1.2-3.1-26], https://hal.inria.fr/hal-01843607

International Conferences with Proceedings

[10] L. AMSALEG, B. Þ. JÓNSSON, H. LEJSEK. *Scalability of the NV-tree: Three Experiments*, in "SISAP 2018 - 11th International Conference on Similarity Search and Applications", Lima, Peru, October 2018, pp. 1-14, https://hal.inria.fr/hal-01843046

[11] F. BÉCHET, C. RAYMOND. *Is ATIS too shallow to go deeper for benchmarking Spoken Language Understanding models?*, in "InterSpeech 2018", Hyderabad, India, September 2018, pp. 1-5, https://hal.inria.fr/hal-01835425

[12] R. CARLINI SPERANDIO, S. MALINOWSKI, L. AMSALEG, R. TAVENARD. *Time Series Retrieval using DTW-Preserving Shapelets*, in "SISAP 2018 – 11th International Conference on Similarity Search and Applications", Lima, Peru, Springer, October 2018, pp. 257-270 [*DOI :* 10.1007/978-3-030-02224-2_20], https://hal.inria.fr/hal-01841995

[13] V. CLAVEAU. *Indiscriminateness in representation spaces of terms and documents*, in "ECIR 2018 - 40th European Conference in Information Retrieval", Grenoble, France, LNCS, Springer, March 2018, vol. 10772, pp. 251-262 [*DOI :* 10.1007/978-3-319-76941-7_19], https://hal.archives-ouvertes.fr/hal-01859568

[14] N. GRABAR, V. CLAVEAU, C. DALLOUX. *CAS: French Corpus with Clinical Cases*, in "LOUHI 2018 - The Ninth International Workshop on Health Text Mining and Information Analysis", Bruxelles, France, Ninth International Workshop on Health Text Mining and Information Analysis (LOUHI) Proceedings of the Workshop, October 2018, pp. 1-7, https://hal.archives-ouvertes.fr/hal-01937096

[15] A. ISCEN, Y. AVRITHIS, G. TOLIAS, T. FURON, O. CHUM. *Fast Spectral Ranking for Similarity Search*, in "CVPR 2018 - IEEE Computer Vision and Pattern Recognition Conference", Salt Lake City, United States, IEEE, June 2018, pp. 1-10, https://hal.inria.fr/hal-01839744

[16] A. ISCEN, Y. AVRITHIS, G. TOLIAS, T. FURON, O. CHUM. *Hybrid Diffusion: Spectral-Temporal Graph Filtering for Manifold Ranking*, in "ACCV 2018 - Asian Conference on Computer Vision", Perth, Australia, December 2018, pp. 1-16, https://hal.inria.fr/hal-01909348

[17] A. ISCEN, G. TOLIAS, Y. AVRITHIS, O. CHUM. *Mining on Manifolds: Metric Learning without Labels*, in "CVPR 2018 - IEEE Computer Vision and Pattern Recognition Conference", Salt Lake City, United States, IEEE, June 2018, pp. 1-10, https://hal.inria.fr/hal-01843085

[18] S. JANNIK, A.-L. MINARD, L. LANGE, M. SPERANZA, B. MAGNINI. *KRAUTS: A German Temporally Annotated News Corpus*, in "LREC 2018 - 11th International Conference on Language Resources and Evaluation", Miyazaki, Japan, May 2018, https://hal.archives-ouvertes.fr/hal-01844834

[19] M. LAROZE, R. DAMBREVILLE, C. FRIGUET, E. KIJAK, S. LEFÈVRE. *Active Learning to Assist Annotation of Aerial Images in Environmental Surveys*, in "CBMI 2018 - International Conference on Content-Based Multimedia Indexing", La Rochelle, France, 2018 International Conference on Content-Based Multimedia Indexing (CBMI), September 2018, pp. 1-6, https://hal.archives-ouvertes.fr/hal-01845487

[20] H. LEJSEK, F. H. ÁSMUNDSSON, B. Þ. JÓNSSON, L. AMSALEG. *Transactional Support for Visual Instance Search*, in "SISAP 2018 - 11th International Conference on Similarity Search and Applications", Lima, Peru, October 2018, pp. 1-14, https://hal.inria.fr/hal-01843043

[21] C. LEVERGER, V. LEMAIRE, S. MALINOWSKI, T. GUYET, L. ROZÉ. *Day-ahead time series forecasting: application to capacity planning*, in "AALTD'18 at ECML 2018", Dublin, Ireland, September 2018, https://arxiv.org/abs/1811.02215 , https://hal.inria.fr/hal-01912002

[22] C. MAIGROT, V. CLAVEAU, E. KIJAK. *Fusion-based multimodal detection of hoaxes in social networks*, in "WI 2018 - Web Intelligence Chile IEEE/WIC/ACM International Conference on Web Intelligence", Santiago, Chile, Proceedings of WI 2018 - IEEE/WIC/ACM International Conference on Web Intelligence, December 2018, pp. 1-8, https://hal.archives-ouvertes.fr/hal-01936720

[23] C. MAIGROT, E. KIJAK, V. CLAVEAU. *Context-aware forgery localization in social-media images: a feature-based approach evaluation*, in "ICIP 2018 - IEEE International Conference of Image Processing", Athènes, Greece, IEEE, October 2018, pp. 545-549 [*DOI : 10.1109/ICIP.2018.8451726*], https://hal.inria.fr/hal-01843611

[24] A.-L. MINARD, C. RAYMOND, V. CLAVEAU. *IRISA at SMM4H 2018: Neural Network and Bagging for Tweet Classification*, in "SMM4H 2018 - Social Media Mining for Health Applications, Workshop of EMNLP", Brussels, Belgium, Proceedings of the Social Media Mining for Health Applications Workshop, associated with EMNLP 2018, October 2018, pp. 1-2, https://hal.archives-ouvertes.fr/hal-01937019

[25] A. PERQUIN, G. LECORVÉ, D. LOLIVE, L. AMSALEG. *Phone-Level Embeddings for Unit Selection Speech Synthesis*, in "SLSP 2018 - 6th International Conference on Statistical Language and Speech Processing", Mons, Belgium, Springer, October 2018, pp. 21-31 [*DOI : 10.1007/978-3-030-00810-9_3*], https://hal.archives-ouvertes.fr/hal-01840812

[26] R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Disfluency Insertion for Spontaneous TTS: Formalization and Proof of Concept*, in "SLSP 2018 - 6th International Conference on Statistical Language and Speech Processing", Mons, Belgium, Springer, October 2018, pp. 1-12 [*DOI : 10.1007/978-3-030-00810-9_4*], https://hal.inria.fr/hal-01840798

[27] F. RADENOVIĆ, A. ISCEN, G. TOLIAS, Y. AVRITHIS, O. CHUM. *Revisiting Oxford and Paris: Large-Scale Image Retrieval Benchmarking*, in "CVPR 2018 - IEEE Computer Vision and Pattern Recognition

Conference", Salt Lake City, United States, IEEE, June 2018, pp. 1-10, https://arxiv.org/abs/1803.11285 , https://hal.inria.fr/hal-01843075

[28] O. SIMÉONI, A. ISCEN, G. TOLIAS, Y. AVRITHIS, O. CHUM. *Unsupervised Object Discovery for Instance Recognition*, in "WACV 2018 - IEEE Winter Conference on Applications of Computer Vision", Lake Tahoe, United States, IEEE, March 2018, pp. 1-10 [*DOI :* 10.1109/WACV.2018.00194], https://hal.inria.fr/hal-01842143

[29] V. VUKOTIC, V. CHAPPELIER, T. FURON. *Are Deep Neural Networks good for blind image watermarking?*, in "WIFS 2018 - International Workshop on Information Forensics and Security", Hong-Kong, China, December 2018, pp. 1-7, https://hal.inria.fr/hal-01936750

[30] N. ZERARI, S. ABDELHAMID, H. BOUZGOU, C. RAYMOND. *Bi-directional Recurrent End-to-End Neural Network Classifier for Spoken Arab Digit Recognition*, in "ICNSLP 2018 - 2nd International Conference on Natural Language and Speech Processing", Algier, Algeria, IEEE, April 2018, pp. 1-6 [*DOI :* 10.1109/ICNLSP.2018.8374374], https://hal.inria.fr/hal-01835440

[31] L. ZHANG, M. SHI, Q. CHEN. *Crowd Counting via Scale-Adaptive Convolutional Neural Network*, in "WACV 2018 - IEEE Winter Conference on Applications of Computer Vision", Lake Tahoe, United States, IEEE, March 2018, pp. 1-9 [*DOI :* 10.1109/WACV.2018.00127], https://hal.inria.fr/hal-01830946

National Conferences with Proceedings

[32] C. LEVERGER, R. MARGUERIE, V. LEMAIRE, T. GUYET, S. MALINOWSKI. *PerForecast : un outil de prévision de l'évolution de séries temporelles pour le planning capacitaire*, in "EGC 2018 - Conférence Extraction et Gestion des Connaissances", Paris, France, vol. RNTI, January 2018, vol. E, n⁰ 34, pp. pp.455-458, https://hal.inria.fr/hal-01911243

[33] A.-L. MINARD, C. RAYMOND, V. CLAVEAU. *Participation de l'IRISA à DeFT 2018 : classification et annotation d'opinion dans des tweets*, in "DEFT 2018 - Défi Fouille de Textes", Rennes, France, ateliers DEFT 2018, May 2018, pp. 1-13, https://hal.archives-ouvertes.fr/hal-01844261

[34] P. PAROUBEK, C. GROUIN, P. BELLOT, V. CLAVEAU, I. ESHKOL-TARAVELLA, A. FRAISSE, A. JACKIEWICZ, J. KAROUI, L. MONCEAUX, T.-M. JUAN-MANUEL. *DEFT2018 : recherche d'information et analyse de sentiments dans des tweets concernant les transportsen Île de France.*, in "DEFT 2018 - 14ème atelier Défi Fouille de Texte", Rennes, France, Actes de la conférence Traitement Automatique des Langues, TALN 2018, May 2018, vol. 2, pp. 1-11, https://hal.archives-ouvertes.fr/hal-01839407

Conferences without Proceedings

[35] M. BUDNIK, M. DEMIRDELEN, G. GRAVIER. *A study on multimodal video hyperlinking with visual aggregation*, in "ICME 2018 - IEEE International Conference on Multimedia and Expo", San Diego, United States, IEEE, July 2018, pp. 1-6 [*DOI :* 10.1109/ICME.2018.8486549], https://hal.archives-ouvertes.fr/hal-01862199

[36] M. LAROZE, R. DAMBREVILLE, C. FRIGUET, E. KIJAK, S. LEFÈVRE. *Apprentissage actif pour l'annotation d'images aériennes appliqué aux suivis environnementaux*, in "RFIAP 2018 - Reconnaissance des Formes, Image, Apprentissage et Perception", Marne-la-Vallée, France, June 2018, pp. 1-3, https://hal.archives-ouvertes.fr/hal-01828591

Scientific Books (or Scientific Book chapters)

[37] J.-L. MINEL, E. MORIN, P. SÉBILLOT, I. TELLIER. *Préface*, in "Varia - Traitement Automatique des Langues", April 2018, vol. 58, n⁰ 1, pp. 7-11, https://hal.archives-ouvertes.fr/hal-01789046

Books or Proceedings Editing

[38] V. CLAVEAU, P. SÉBILLOT (editors). *Actes de la conférence Traitement Automatique de la Langue Naturelle, TALN 2018 : Volume 1 : Articles longs, articles courts de TALN*, May 2018, https://hal.archives-ouvertes.fr/hal-01843560

[39] A.-L. LIGOZAT, P. CELLIER, A.-L. MINARD, V. CLAVEAU, C. GROUIN, P. PAROUBEK (editors). *Actes de la conférence Traitement Automatique de la Langue Naturelle, TALN 2018 : Volume 2 : Démonstrations, articles des Rencontres Jeunes Chercheurs, ateliers DeFT*, May 2018, https://hal.archives-ouvertes.fr/hal-01843585

Research Reports

[40] T. FURON. *The illusion of group testing*, Inria Rennes Bretagne Atlantique, March 2018, n⁰ RR-9164, pp. 1-19, https://hal.inria.fr/hal-01744252

[41] A. ISCEN, Y. AVRITHIS, G. TOLIAS, T. FURON, O. CHUM. *Hybrid Diffusion: Spectral-Temporal Graph Filtering for Manifold Ranking*, Inria Rennes - Bretagne Atlantique, July 2018, pp. 1-16, https://hal.inria.fr/hal-01846002

[42] M. SHI, Z. YANG, C. XU, Q. CHEN. *Revisiting Perspective Information for Efficient Crowd Counting*, Inria Rennes - Bretagne Atlantique, December 2018, pp. 1-10, https://hal.inria.fr/hal-01831109

Other Publications

[43] C. DALLOUX, N. GRABAR, V. CLAVEAU, C. MORO. *Portée de la négation : détection par apprentissage supervisé en français et portugais brésilien*, Actes de la conférence TALN 2018 - Traitement Automatique de la Langue Naturelle, May 2018, vol. 1, TALN 2018 - 25e conférence sur le Traitement Automatique des Langues Naturelles, Poster, https://hal.archives-ouvertes.fr/hal-01937100

References in notes

[44] M. AGOSTI, J. ALLAN. *Special issue on methods and tools for the automatic construction of hypertext*, in "Information Processing and Retrieval", 1997, vol. 33, n⁰ 2

[45] C. A. BHATT, M. S. KANKANHALLI. *Multimedia data mining: State of the art and challenges*, in "Multimedia Tools and Applications", 2011, vol. 51, n⁰ 1, pp. 35–76

[46] O. CHUM, J. MATAS. *Large-Scale Discovery of Spatially Related Images*, in "Pattern Analysis and Machine Intelligence", 2010, vol. 32, n⁰ 2, pp. 371-377

[47] M. ESKEVICH, G. J. F. JONES, S. CHEN, R. ALY, R. ORDELMAN, D. NADEEM, C. GUINAUDEAU, G. GRAVIER, P. SÉBILLOT, T. DE NIES, P. DEBEVERE, R. V. DE WALLE, P. GALUŠCKOVÁ, P. PECINA, M. LARSON. *Multimedia Information Seeking through Search And Hyperlinking*, in "Intl. Conf. on Multimedia Retrieval", 2013

[48] C. HERLEY. *ARGOS: Automatically extracting Repeating Objects from multimedia Streams*, in "IEEE Trans. on Multimedia", Feb. 2006, vol. 8, n° 1, pp. 115–129

[49] A. PARK, J. R. GLASS. *Unsupervised pattern discovery in speech*, in "IEEE Transaction on Acoustic, Speech and Language Processing", Jan. 2008, vol. 16, n° 1, pp. 186–197

[50] K. M. PUA, J. GAUCH, S. E. GUCH, J. Z. MIADOWICZ. *Real time repeated video sequence identification*, in "Computer Vision and Image Understanding", 2004, vol. 93, n° 2004, pp. 310–327

[51] C. G. M. SNOEK, M. WORRING. *Multimodal Video Indexing: A Review of the State-of-the-Art*, in "Multimedia Tools and Applications", 2005, vol. 25, n° 1, pp. 5–35

[52] O. DE ROOIJ, M. WORRING. *Browsing Video Along Multiple Threads*, in "IEEE Transactions on Multimedia", 2010, vol. 12, n° 2, pp. 121–130