UMR IRISA

Activity Report 2022

Team INTUIDOC

Intuitive user interaction for documents

D6 – Signal, Image, Language

Université de Rennes | rennes | INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES RENNES | cnrs | Inria | Inserm | CentraleSupélec | IMT Atlantique Bretagne-Pays de la Loire École Mines-Télécom | Université Bretagne Sud ubs

1 Team composition

Researchers and faculty

Eric Anquetil, Professor, Insa, head of the team
Bertrand Coüasnon, Associate Professor, Insa, HDR
Nathalie Girard, Associate Professor, Univ. Rennes 1
Aurélie Lemaitre, Associate Professor, Univ. Rennes 2, HDR
Ivan Leplumey, Associate Professor, Insa
Yann Ricquebourg, Associate Professor, Insa

Associate members

Jean Camillerapp, Retired Professor
Yann Soullard, External Colloborator, Associate Professor, Univ. Rennes 2

Research engineers, technical staff

Hugo Hazard, Insa Research Engineer from 1st July
Bruno Hortollary, Insa Research Engineer, from 1st October
Iwan Le Floch, Insa Research Engineer, until 1st September
Pauline Nerdeux, Insa Research Engineer, until 1st July
Ali Yesilkanat, Insa Research Engineer from 1st September

PhD students

Islam Barchouch, Insa PhD student, from 1st October
Killian Barrère, Insa PhD student
Simon Corbillé, Univ. Rennes 1 PhD student
Martin Dornier, CIFRE/Insa PhD student
Camille Guerry, Insa PhD student, until 16th December
Florent Imbert, Insa PhD student
Clément Leroy, Insa PhD student, until 24th June
Arnaud Lods, CIFRE/Insa PhD Student
William Mocaër, Insa PhD student
Timothée Neitthoffer, CIFRE/Insa PhD Student
Solène Tarride, CIFRE/Insa PhD Student, until 31st January

PostDoc

Omar Krichen, Insa Post Doc
Wassim Swaileh, Insa Post Doc

Administrative assistant

Nadia Derouault

2 Overall objectives

2.1 Overview

The IntuiDoc team focuses its work on handwriting, gesture (2D and 3D), and documents under various aspects: analysis, recognition, composition, interpretation. We are also interested in human-document interaction and graphical/gestural man-document interaction. This research relates to the handwriting and the documents under different forms: manuscript, printed paper form, pen-based and touch-based interaction, graph, images, heterogeneous documents, etc.

The roadmap of the IntuiDoc team is on the frontier of several research axes: Pattern recognition, Machine-Learning, Artificial Intelligence, Human-Machine Interaction, Uses and Digital Learning. The aim is to explore new scientific challenges of the domain of the Human-Document Interaction with a specific focus on interactive, incremental and evolving learning based on the integration of the user in all the processes of analysis and decision making.

Today, four major emerging scientific axes are investigated with strong partnerships with national and international laboratories and companies:

- "On-line" evolving cross-learning of 2D (touch and pen–based) and 3D gestures (Kinect and Leap Motion);
- "On-line" analysis of drawing, sketching and handwriting with pen-based tablet for digital learning (e-education);
- Interactive learning of document structure without ground-truth;
- Document collection analysis for big-data.

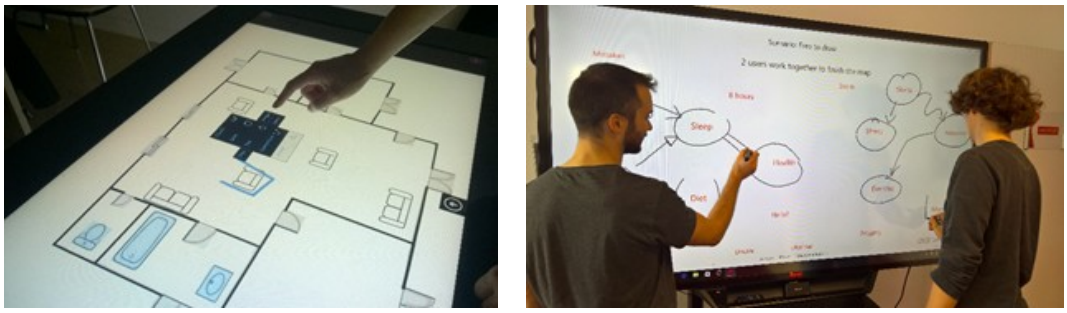
2.2 Scientific foundations

2.2.1 On-line evolving cross-learning of 2D and 3D gestures.

2D evolving recognizer for gesture commands With the increasing use of touch and pen-based sensitive screens, human-computer interactions are evolving. New interaction methods have been designed to take advantage of the new potential of interaction offered by these interfaces. Among them, a new concept has recently appeared: to associate commands to gestures. Those gesture commands enable users to execute various actions simply by drawing symbols. This new man-machine interaction can be used for on-line composition of complex documents such as electrical sketches or floor plan. In order to use such gesture commands, a recognition system is required. For users to easily memorize more than a dozen of gesture commands, it is important to enable gesture set customization. The classifier used to recognize drawn symbols must hence be customizable, able to learn from very few data, and evolving to learn new classes on-the-fly and improve during its use. The objective of this work is to obtain a gesture command system that cooperates as best as possible with the user, learning from its mistakes without soliciting the user too often. Gesture commands lead to a cross-learning

situation where the user has to learn and memorize the gestures, and the classifier has to learn and recognize drawn gestures. We study the impact of different strategies to supervise the online training of an evolving recognizer for gesture commands, and how to optimize this cooperation between the user and the recognition system. In particular, we design an inner confidence measure to solicit the user when some data samples don't fit the classifier model, and that it will be very gainful to learn from it.

Multi-touch gesture recognition Due to the recent prevalence of multi-touch devices, multi-touch gesture recognition has gained a large interest in the last decade. Unlike mono-touch gesture recognition which tracks the movement of a single point of input, multi-touch gesture often tracks many points of contact in parallel as they appear, move and disappear. The recognition for multi-touch gestures is challenging because of the complex chronological relation between the fingers' trajectories. We explore new methods for modelling the shape, relative temporal and motion information in multi-touch gesture by a model of graph and graph embedding approach. In our future work we aim at developing a strategy to detect the pattern of multi-touch gesture at runtime, to be able to address direct manipulation by command gesture.



(a) (b)

Figure 1: Multi-Touch and Multi-User Interaction

Multiple user freely-drawn sketch recognition and 3D action gesture recognition Another scientific challenge is also to address large multi-touch display that allows multiple users to simultaneously interact in the same context and work together. Indeed, many researches and commercial products propose tangible interfaces which support simultaneous participation of multiple users. This is a really new research topic to automatically recognise and interpret in real time the freely-drawn sketch of multiple users.

Finally, in this axe, we investigate the validity of transferring the expertise on hand-drawn symbol representation [DA13] to recognise 3D action gesture. This new research topic will be conducted in collaboration with MIMETIC project team of Inria through the co-supervision of 2 successive PhD theses on the subject of early recognition of 2D/3D gestures (see section 3.10). We base this proposition on the observation that

[DA13] A. Delaye, E. Anquetil, "HBF49 feature set: A first unified baseline for online symbol recognition", *Pattern Recognition 46*, 1, 2013, p. 117 – 130.

patterns produced by a human motion, in particular 2D hand-drawn symbols and 3D actions, share several important properties. They are both governed by kinematic constraints that must be considered while modelling such human motions. We hypothesise that both recognition problems could be addressed in similar ways.

2.2.2 Artificial Intelligence for e-education

This research axis is more recent in the team. It focuses on the design of artificial intelligence engines for e-education. It is a very active line of research since it is associated with 7 major research projects over the past 5 years with a funding budget of 1.456 Million euros.

Handwriting analysis for digital learning at school The scientific problem we tackle here is to quantitatively evaluate a cursive handwriting with respect to a reference model and recommendations of a teacher. In order to be able to teach children how to write, we must be able to analyse their handwriting, to evaluate if the letters, words, sentences are correctly written, and to detail which aspects of the child's handwriting do not correspond to the teacher models (corrective feedback). This problem is completely different from the classical task of character recognition, where the challenge is to determine to which class the data samples belongs.

Our objective is to be able to analyse, qualify and evaluate handwriting, with regards to reference models, and for multiple distinct aspects like: shape (for legibility), drawing direction and order (for ductus), speed and fluidity for instance. We use an analysis system based on an evolving fuzzy classifier. It allows to easily define reference models from few data samples to customise "on the fly" the writing exercises to the children. Then, the analysis system can be used to evaluate drawn gestures, regarding a specific feature set, and finally give a confidence score.

The ANR Joint laboratory (LabCom: Script&Labs) between IntuiDoc and Learn&Go The axe of Artificial Intelligence for e-education has been launched with the IntuiScript project founded by the French government as part of innovative national projects (BPI-PIA2). IntuiScript targets towards offering an advanced digital writing experience at school by using tablets and tactile digital devices (with finger touch and stylus). This project was structured around the conception of a digital workbook to help teachers and children from three to seven years old during the handwriting learning process:

- it allows children to work in autonomy with an on-line and real time feedback;
- it proposes automatically pedagogical exercises that are adapted to children difficulties based on the automatic analysis of children writing;
- it provides a precise off-line analysis of children writing (i.e. order, direction, shape) to help teachers to understand children writing skills and difficulties.

This project was based on a user-centred design approach that includes several cycles of conception followed by experiments. Therefore, feedback of children and teachers

related to these experiments have been used to improve the education scenario. More than 1,000 primary school students from Brittany have taken part to the experiments in the project.

This four years project was a real success. It resulted in the launch of the product "Kaligo" today distributed in schools by the company Learn & Go.

Figure 2: Handwriting analysis for digital learning

With the success of IntuiScript project, the IntuiDoc team and Learn&Go company created the "Script&Labs" LabCom to innovate on Digital Learning (2017-2021).

The scientific principles of the joint laboratory belong to artificial intelligence (pattern recognition, Machine Learning) and man-agent interaction. Interpretation, adaptation and learning are the heart of its researches, aiming to conceive automated interpretation systems for children productions: writing, arithmetic operations, geometric figures. This scientific know-how form the foundation of new solutions in Digital Learning, leading to more independence and customisation for each student learning process.

The results of the joint laboratory will directly be used in innovating educational modules focusing in active learning, handwritten input, analysis and personalised help via immediate feedback on student production [SAB17].

Three lines have been currently chosen :

- writing learning for young children
- learning numbers and first arithmetic calculations
- geometry learning by “generative drawing”

Beyond these three lines, LabCom Script&Labs helps working on “active and collaborative learning from secondary school to higher education” through e-Fran « Actif » et Cominlabs « e-Fil » projects.

The “ScriptAndLabs” joint laboratory is founded by ANR for four years (n° ANR-16-LVC2-0008-01).

In 2019, we won with Learn&Go company a new project from the Ministry of Education (DNE) as part of the Innovation - Artificial Intelligence Project (P2IA). The

[SAB17] D. Simonnet, E. Anquetil, M. Bouillon, “Multi-Criteria Handwriting Quality Analysis with Online Fuzzy Models”, *Pattern Recognition*, 2017, `https://hal.archives-ouvertes.fr/hal-01515397`.

objective is to design a software environment based on artificial intelligence for helping to learn French (writing / spelling) for teachers and students in cycle 2. In this project, we are responsible for designing and developing the automated analysis engine of handwriting to identify the spelling/graphemes mistakes of pupils in CP, CE1 and CE2.

Moreover, as a result of the "Script&Labs" joint laboratory, we have signed several service contracts in 2021 and 2022 with the company Learn And Go in order to adapt the Kaligo application to help students learn to write for Germany, Luxembourg and Switzerland for traditional schools and children with special needs.

This has had a significant impact on the team. IntuiDoc is now recognized by the French Ministry but also internationally in the design of intelligent tutors to help learning of children at school: writing, arithmetic, geometry ... This has opened the way to new collaborative projects as the ANR bilateral, (KIHT 2021-2024) with STABILO International GmbH (see section 5.1.1).

"On-line" analysis of drawing for new learning strategies based on "generative drawing" One of the key topics in this axe is how to encourage new learning strategy based on "generative drawing" using pen-based numerical devices. The goal is to improve the learning of students by considering "learning" as a generative activity. In this scope, the potential induced by pen-based tablet is really interesting. The goal is to investigate how we can automatically generate intelligent "corrective" or "predictive" feedbacks to a user during his drawing process: summarizing, mapping, drawing, sketching... We base this work on the visual grammar CD-CMG [MA09] (Context Driven Multi-set Grammar), to model the domain knowledge and interpret the hand-drawn sketches on the fly. We adapted this grammar to the Geometry domain to cover the concepts taught in middle-school (cf. Figure 3).

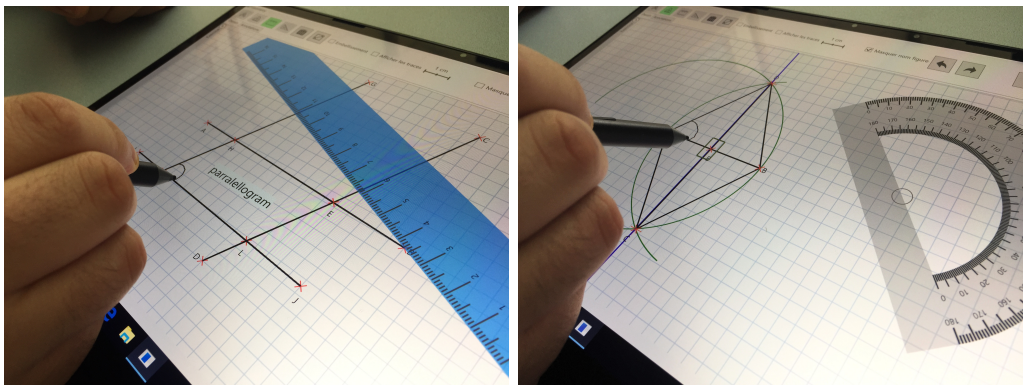


Figure 3: Pen-based and Gesture-based software for geometry learning.

We explore this new research area in collaboration with researchers in psychology of the LP3C/LOUSTIC Laboratory of Rennes. To support this multi-disciplinary challenge, we have frirstly developed the innovative four year national project "ACTIF"

[MA09] S. Macé, E. Anquetil, "Eager interpretation of on-line hand-drawn structured documents: The DALI methodology", *Pattern Recognition 42*, 12, 2009, p. 3202 – 3214, New Frontiers in Handwriting Recognition.

(BPI – e-FRAN) with the support of the Brittany Region. Today, this research work is extended through the new ANR "Triangle" project (see section 5.2.5) led by the University of Rennes 2 (LP3C laboratory). It focuses on the finalization of the Intelligent Tutor for geometry learning support on a touch tablet/stylus in middle school (IntuiGeo). It will lead to experimentations and a deployment in the academies of Rennes and Poitiers.

2.2.3 Interactive learning of document structure without ground-truth

Interactive Rule Inference We work on the interactive learning of document structure, in the context of a thesis that has just ended. This work enables to combine statistical methods with syntactical approaches (grammars). Indeed, statistical methods are not able to convey two-dimension hierarchical structures that are common in document analysis. On the opposite, rule-based syntactical methods often require a fastidious manual step for the specification of the various organisations of the document physical layouts. The objective is to model the logical structure with rules and to learn the physical structure. This learning is based on databases of documents with ground-truth that are really costly to label. The current and future work aims at learning physical properties without ground-truth. The scientific context is to lean on large amount of documents and on generic document system analysis. We want to show that some knowledge can stand out from the repetition of physical structures, thanks to non-supervised learning methods. The challenge is to define strategies to make this learning possible thanks to an interaction with the user, which brings a semantic knowledge to the physical detected elements.

Combination Deep Learning / Syntactical Analysis In collaboration with Richard Zanibbi from the Rochester Institute of Technology (RIT), Rochester, New York, USA, we will continue to work on interactive learning by combining deep learning technics, syntactical analysis and user interaction to introduce learning of segmentation. Deep learning methods like convolutional neural networks or recurrent neural networks have shown very interesting results in recognition by being able to make a common segmentation and recognition, with a good introduction of local context. But they are limited to a local context, which is interesting for the recognition of letters and words in a handwritten text line, but is not enough for a modelling and an understanding of a complex structure like the one we can find in a complex structured document. We propose to study the strong combination of deep learning and syntactical methods to build a document structure recognition system able to deal with segmentation problems by learning them. The syntactical part models the structure and brings complex context to the deep learning recognition. The objective is to introduce in the architecture of the neural networks the large contextual information and to make the neural networks able to give not only a recognition but also information of localization of the recognized element. Indeed this localization information is important for the syntactical part to continue and explore different solutions in the global recognition of the document. To train the neural networks, we will have to focus also on a semi–automatic generation of datasets and ground truth, made by the grammatical description of the document, in combination with unsupervised clustering and a user interaction to generate ground

truth with a minimum of manual work.

Spread Applications These combinations could open large perspectives by simplifying the grammatical description as much as possible by learning the document structure, including regions of interest (segmentation), region types (classification) and their relationships (parsing/structure). Many applications could be studied on domains where it is important to combine deep learning and strong a priori knowledge. We will also make this combination able to deal with born digital documents (pdf, XML...) to address the huge quantities of documents, which need a real understanding for information extraction.

2.2.4 Document collection analysis for historical big-data

Strategies for Sequential Collection The DMOS-PI method proposes a framework for the analysis of collections of documents. It enables to share information from the collection between the pages, thanks to an iterative mechanism of analysis. This mechanism also makes it possible to integrate an asynchronous interaction between automatic analysis and human operators. We propose to work on modelling strategies of analysis for the analysis of collections of documents. The strategies could sequence the various iterative treatments of documents pages, the global treatments and the interactions. The interest is to exploit as much knowledge as possible on the collection in order to make the extraction of information in each analyzed pages more reliable, and to make the understanding between the various data at the collection level easier. In this context, the ANR HBDEX project has been selected. It is led by the PSE "Paris School of Economics" ("Ecole d'économie de Paris"), with the LITIS lab in Rouen and the CAMS-EHESS. This project focuses on the extraction of historical big-data for digital humanities, applied to financial data. The objective is to analyse masses of tabular data: daily listing on the Stock Exchange from the 19th and 20th centuries. The analysis will be based on the redundancy between the successive days of listing and the consistency between the global sequences of data. This modelling will enable a fast adaptation to other kinds of historical tabular data that only exist on a paper form (economic, demographic, meteorological), but that is necessary to constitute historical big-data databases. This opens a large possibility of applications on documents found in all statistical institutes.

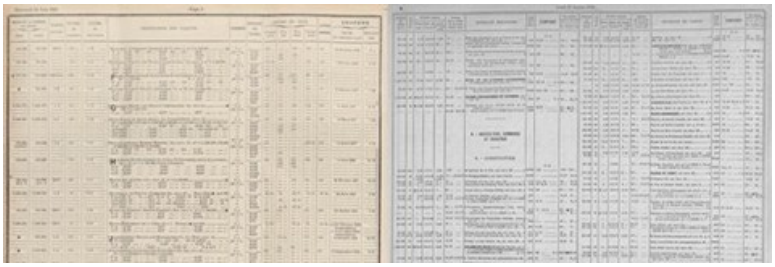


Figure 4: Examples of daily listing on the Paris Stock Exchange

Adaptive Document Layout Analysis We propose to integrate the interactive document structure learning without ground truth and the collection modelling to generate an adaptive document layout analysis system where a user, with few interactions, could make the recognition system learn new layouts to adapt itself and improve the global recognition quality. We will build this adaptive document layout system on the European project EURHISFIRM (InfraDev). EURHISFIRM designs a world-class research infrastructure (RI) to connect, collect, collate, align, and share detailed, reliable, and standardized long-term financial, governance, and geographical data on European companies. This project is leaded by the PSE "Paris School of Economics" ("Ecole d'économie de Paris"), with seven partners working on quantitative economics and finance, economic and social history, and the LITIS Lab in Rouen working with us on document images analysis. We will work on a system to extract high-quality data from historical serial printed sources, to address three issues: (i) lowering the costs of data extraction from the same source; (ii) lowering the cost of adaptation of the system from one source to the other; (iii) developing effective data validation process. Interactions between the system and experts on the sources lay at the heart of the conception.

3 Scientific achievements

3.1 Online Children Handwriting Recognition and Segmentation

Participants: Simon Corbillé, Éric Anquetil, Élisa Fromont (LACODAM Team).

This section presents works done in the thesis entitled: Hybridization of AI approach "Transparent" and "Deep Learning" for the automated analysis of graphic productions of students in the context of education. We consider the task of analyzing children handwriting in the context of a dictation task. The objective is to detect orthographic and phonological errors. To achieve this goal, we use two approaches.

In the first approach, we extend an existing handwriting analysis engine, based on an explicit segmentation of the handwritten input, originally developed for children copying exercises. We present a new approach, based on the combination of this analysis engine with a deep learning word recognition approach in order to improve both the recognition and segmentation performance. Explicit segmentation needs prior knowledge, and the deep network recognition predictions are a reliable approximation of the ground truth which can guide the analysis process. We propose to combine multiple prior knowledge strategies to further improve the analysis performance. Furthermore, we exploit the deep network approximate implicit segmentation to optimize the existing analysis process in terms of complexity. This work has been published in IJDAR (International Journal on Document Analysis and Recognition) [4].

In the second approach, we propose to combine the predictions of an accurate Seq2Seq model with the predictions of a R-CNN object detector and refine the bounding box predictions provided by the detector with a segmentation lattice compute from the online signal. Preliminary results were presented in [17]. Figure 5 illustrates the architecture of the method.

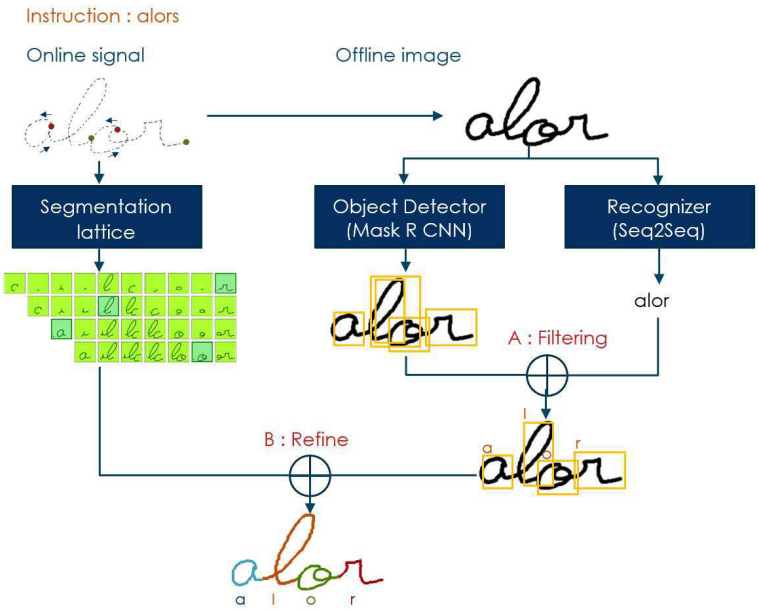


Figure 5: Overview of the architecture of the method.

3.2 Analysis of French spelling in handwriting

Participants: Éric Anquetil, Nathalie Girard, Simon Corbillé, Omar Krichen, Pauline Nerdeux.

Keywords: Online handwriting analysis, spelling analysis.

This work is a continuation of previous works done in the context of the Innovation - Artificial Intelligence Project (P2IA). The objective is to detect orthographic and phonological errors. To achieve this goal, we extend an existing handwriting analysis engine, based on an explicit segmentation of the handwritten input, originally developed for children copying exercises. In this context, the writing to be analyzed is still degraded, since the target population is primary school children, and we are also faced with spelling mistakes, which makes the writing analysis task more complex. The analysis process, in a copying context, was guided by the instruction, which provided a reliable approximation of the ground truth. In a dictation context, this approach becomes obsolete, since the child does not see the instruction he/she has to reproduce. To overcome this problem, we present a new approach, based on the combination of this analysis engine with a deep learning word recognition approach, namely a Seq2Seq architecture, in order to improve both the recognition and segmentation performance. Explicit segmentation needs prior knowledge, and the deep network recognition predictions are a reliable approximation of the ground truth which can guide the analysis process. We propose to combine multiple prior knowledge strategies to further improve the analysis performance. Furthermore, we exploit the deep network approximate implicit segmentation to optimise the existing analysis process in terms of complexity. This resulted in the definition of new strategies for prior knowledge generation:

- Deep prediction strategy: consists in using the seq2seq predicted word as prior knowledge for the analysis system;
- Fusion and competition: consists in combining multiple prior knowledge strategies, such as the instruction, the deep network prediction, and phonetic hypotheses similar to the dictated word.

These two strategies improved the analysis system performance in terms of recognition and segmentation, such as shown in table 1.

Table 1: Segmentation (IOU) and recognition performance of each strategy

Strategy	**Segmentation rate (IoU)**	**Recognition rate**
Instruction strategy	88.66 %	64%
Deep prediction strategy	90.4% ± 0.54%	72.18± 0.73 %
Fusion and competition	92.82% ± 0.28 %	**83.28%** ± 0.51 %

This work has been published in IJDAR [4].

3.3 Kaligo-based Intelligent Handwriting Teacher (KIHT)

Participants: Éric Anquetil, Yann Soullard, Wassim Swaileh, Florent Imbert, Romain Tavenard (LETG lab).

This work takes place within the KIHT French-German bilateral ANR project. This project is composed of two academic partners, IRISA (France) and KIT (Germany) and two industrial partners, Learn & Go (France) and Stabilo (Germany). The Stabilo company has developed a specific digital pen composed of IMU sensors, called Digipen, on which we will base our work.
Our work is to get a handwriting trajectory reconstruction from the Digipen. The Stabilo Digipen captures time series from IMU's sensors (accelerometers, gyroscope, magnetometer, force sensors) during writing, and our goal is to reconstruct the online handwriting trace. The fisrt step has been to create a database will serve as a benchmark to help advanced research, on this new scientific challenge. Then, to reconstruct the handwriting trajectory from IMU signals, we present a complete pipeline which consists of :

- preprocessing including an alignment strategy between input and ground-truth time series for training as well as a learning pipeline based on touching strokes,
- a neural network architecture inspired by a Temporal Convolutional Network (TCN) to cope with noisy signals coming from the inertial sensors,
- an evaluation protocol based on the Frechet distance to assess the quality of the reconstructed trajectories.

This work was presented at SIFED 2023 [11], and submitted to the IJDAR 2023 journal track.

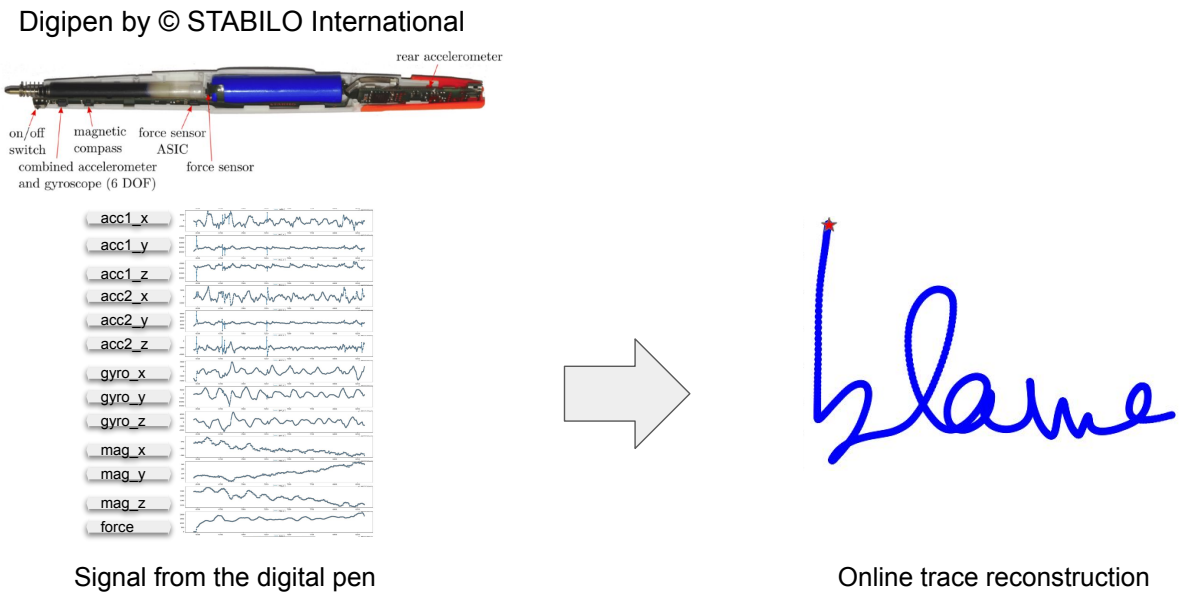


Figure 6: The goal of this project is to design solutions based on deep learning to generate the on-line handwriting trajectory from kinematic sensors coming from the digital pen.

3.4 Kaligo Switzerland

Participants: Eric Anquetil, Nathalie Girard, Pauline Nerdeux, Omar Krichen.

Keywords: Online handwriting analysis, character analysis, German handwriting.

After having developed the handwriting analysis engines for the French and English languages for the company Learn&Go, within the scope of the Labcom Script&Labs (see section 5.3.4), our work consisted in developing analysis engines for the interpretation of German script letters to be used in an educational context in Swiss classrooms. The goal, as for previous engines, is to qualitatively analyse the child written letter in order to produce precise feedback on the shape, order, and direction of his/her handwriting. Extensive work has been done to deal with the difference between the English/French script letters and the German ones, such as the Eszett (ß), or to deal with different types of accented letters (ä, ö, ü, Ä, Ö, Ü). The engine has also been adapted to deal with the confusion that these new letters bring to the analysis process (*e.g.* how to distinguish between "ß" and "B").

The final delivery, composed of the digits engine, lower and upper case letters engines for German language, as well as the databases used for the learning process, was provided to the Learn&Go company on June 24Th, 2022.

3.5 Interactive and iterative analysis of price-lists

Participants: Camille Guerry, Iwan Le Floch, Bertrand Coüasnon, Aurélie Lemaitre.

This work is done in the context of the ANR project HBDEX (see part 6.2.5). The aim of the ISICA method (Interactive Strategy for Iterative Collection Analysis) proposed in the PhD of Camille Guerry [1] is to model a strategy that automatically exploits the sequential aspect of a collection to improve document recognition.

The strategy that we propose in ISICA is composed of different iterations. The aim of each iteration is to recognize and validate one kind of structural element in the documents at different levels of granularity. An iteration consists of:

1. a first individual structural analysis of each document, based on a combination of deep-learning and syntactical approach;
2. a transverse validation phase, using knowledge coming from the context of the collection;
3. a second individual structural analysis. For this step, the recognition system exploits the information validated by step 2 to specify the grammatical description used in step 1.

For the validation phase (step 2.), we first construct a sequence composed of data that comes from the structural analysis phase (step 1). We then apply a validation process to this sequence. We proposed two different validation processes, depending on the knowledge that we have on the collection.

- Rupture detection: we apply a bottom-up rupture detection method used in signal processing [TOV20] to detect changes in our data. We then post process the obtained slots to detect recognition errors and correct them.
- Sequence alignment: we proposed a method inspired by sequence alignment methods used in bioinformatics to identify elements coming from different images.

After this validation phase, we generate questions for the user if there are ambiguities or if the information needed is not available in the original images.

In a first instance, we focus on daily price lists coming from Paris stock exchange markets of the end of the 19th and beginning of the 20th century. This sequential collection is composed of 40 years of documents (more than 40 000 images).

We have evaluated the ISICA method for each of the granularity levels: columns, sections, stocks and other values. On the columns, we show that ISICA allows to decrease the error rate from 7.89 % to 0.44 %. For section recognition, we go from an F-measure of 0.9937 to 1 using our strategy. Thanks to ISICA, we improve the

[TOV20] C. Truong, L. Oudre, N. Vayatis, "Selective review of offline change point detection methods.", *in : Signal Processing, 167, 107299*, IEEE, p. 770–778, 2020.

F-measure from 0.914 to 0.988 on stock identification and we significantly reduce the number of questions needed to obtain satisfactory results (48,067 questions without ISICA to 1,996 with ISICA). On the values associated to the stocks, we decrease the error rate from 8.12 % to 2.93 % in the most favourable cases. Moreover only 1.88 % of the modifications made by the strategy in ISICA required user interaction.

We also proposed a processing chain that allows this strategy to be easily parallelized to produce a large amount of data. We are currently producing data for daily price lists coming from Paris stock exchange markets of the end of the 19th and beginning of the 20th century ("La Coulisse"). In the context of the project H2020 Eurhisfirm, we have also tested this strategy on other European price list collections. This strategy is designed to be generic and could be applied to other types of collections.

3.6 Transformer-Based Neural Network for Handwritten Text Recognition

Participants: Killian Barrère, Bertrand Coüasnon, Aurélie Lemaitre, Yann Soullard.

Keywords: Transformer, Handwritten text recognition, Neural networks.

This section describes the works done by Killian Barrère during the second year of his thesis, that aims to study recent advances in the field of neural networks. We improved the light Transformer architecture we proposed to perform well on few data. Improvements we tried were mostly related to convolutional layers and normalization techniques. We obtain good results without using additional data thanks to the light design of the architecture, with a limited number of layers and number of parameters per layer. Moreover our architecturereaches results at the state of the art by using well-designed synthetic data. Our architecture also has many advantages such as a faster convergence. We published 3 papers [7] [8] [16].

We then proposed an even lighter architecture for the context of historical documents in which only a few annotated data are available. To train efficiently our architecture, we designed synthetic data adapted and improved in various ways (such as a selection of handwritten fonts, augmentation and backgrounds) for historical documents. Examples of synthetic data generated by our pipeline are available inside Figure 7. We also proposed a training strategy to train efficiently a Transformer architecture with few available data. We performed experiments on the ICFHR READ 2018 dataset that proposes interesting scenarios to evaluate how well a model perform with a few annotated data from a given collection. We obtain results at the level of the best-performing architectures from the state of the art, demonstrating that our architecture combined with our proposed strategy is able to train on a few annotated data while obtaining excellent results.

3.7 Analysis of musical scores

Participants: Ali Yesilkanat, Hugo Hazard, Aurélie Lemaitre, Yann Soullard, Bertrand Coüasnon, Jean Camillerapp, Nathalie Girard..

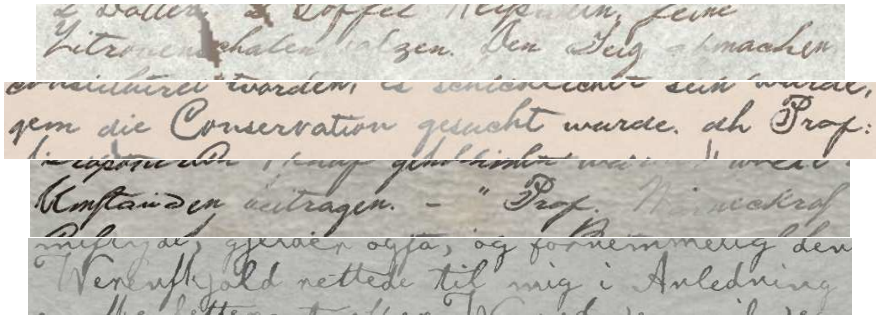


Figure 7: Examples of generated synthetic text-lines, trying to match the style of the ICFHR 2018 READ dataset.

In the context of Collabscore project (section 5.2.2), we work in musical score recognition. The work is divided in two aspects: the construction of recognizers for musical symbols, based on deep learning, and the grammatical description of musical contents, based on the syntactic DMOS method.

3.7.1 Recognizers for musical symbols

In this section, the objective is the recognition of historical printed scores using the Common Western Music Notation. We apply such a detector to detect a large class set of music symbols on the printed scores to achieve this goal. But using them out-of-the-box on music sheets with large resolution spaces can lead to suboptimal performance due to the dense nature of music scores with many tiny objects; therefore, we applied the detector on crops rather than running on the entire score. Later, we employed a high-resolution backbone for the detector and applied it to the full-page printed musical scores. Figure 8 illustrates the detected music symbols on a music sheet.

3.7.2 Grammatical description of musical scores

The grammatical description is based on DMOS method. The terminal elements are composed of vertical segments and connected components from our Music image processing library (Musique). Then the resulting connected components of the object detector from the previous part are added to the list of terminal elements. Using these terminals and after having detected the stafflines (which is another role of the Musique library), the grammatical rules allow us to detect each bar of the score, by detecting the barlines. In each bar, we detect a potential header and then all the voice elements (notes, rests, chords...) as well as their associated characteristics (accidentals, augmentation dots, articulations...) For the moment, only monophonic scores for a single instrument are well detected, and we are currently working on detection of polyphonic scores for orchestras.

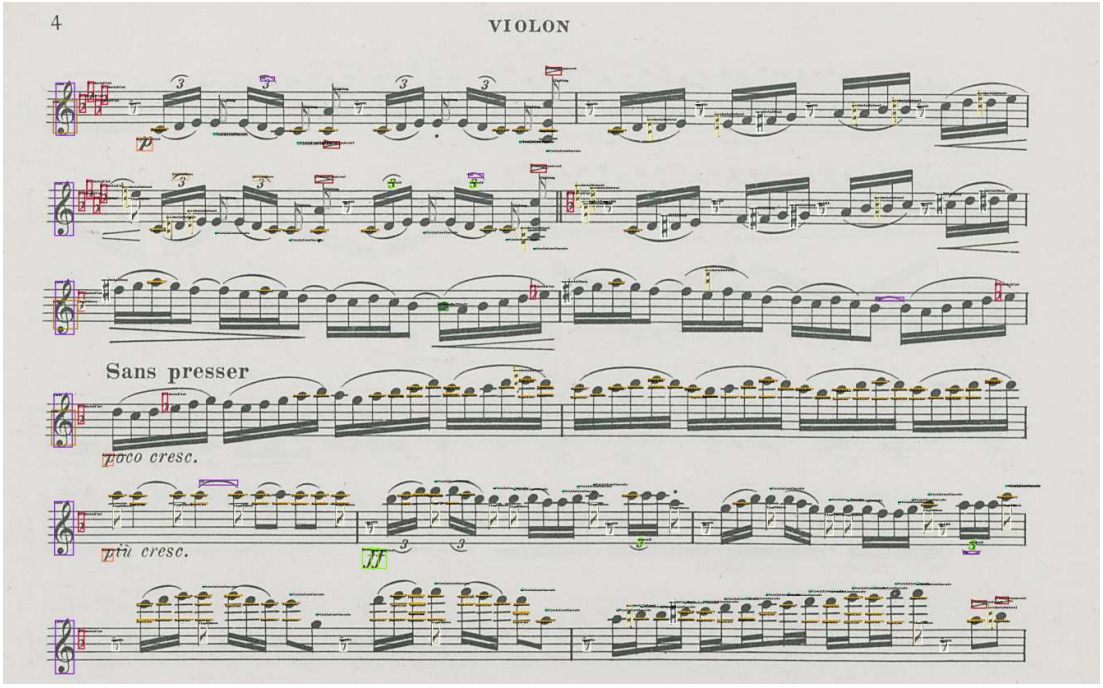


Figure 8: Results of the musical score recognizer on a monophonic score from Barcarolle, Camille Saint-Saëns.

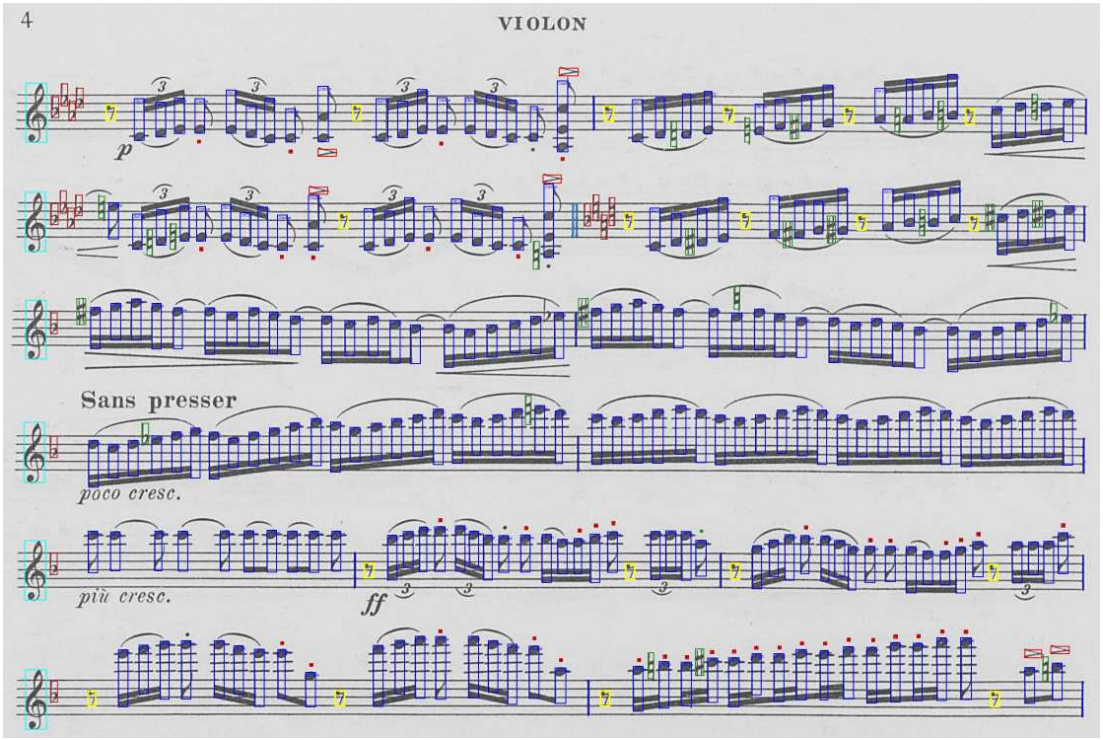


Figure 9: Results on a monophonic score from Barcarolle, Camille Saint-Saëns.

3.8 Knowledge Integration inside Multitask Network for Unseen Identity Document Recognition

Participants: Timothée Neitthoffer, Bertrand Coüasnon, Aurélie Lemaitre, Yann Soullard.

Keywords: Knowledge Integration, Multitask Learning, OCR, Text Localisation, Identity Documents, Deep Network.

This work describes the works done by Timothée Neitthoffer during the end of its second year of his thesis that aims to perform joint text localisation and recognition in identity documents (IDs).

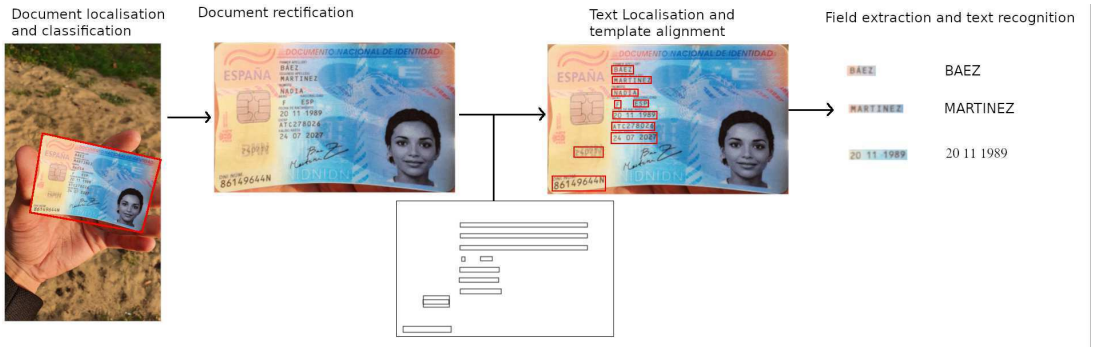


Figure 10: Traditional pipeline for IDs recognition. First the document is localised in the captured image and classified before being rectified. Text localisation is performed on the rectified document. Based on its type, the corresponding template is obtained from a knowledge base. This template is aligned based on the text localisation step and the fields are extracted. They are then recognized by the OCR.

Identity Document recognition is a key step in Know Your Customer (KYC) applications where a service provider is required to verify information from IDs. IDs belonging to the same type (example Spanish ID card version 2021) share the same field structure (template). Traditional identity document pipelines leverage this template to guide the localisation of the fields and then recognize them. However, such systems have to be tuned to the different document templates to correctly perform on these. Thus, they can not be directly used on new types of documents from new countries or on new versions. We address the task of performing text localisation and recognition in the context of new document types, i.e. where only the template is available while there is no labeled document samples from the new type of document. To that end, we adapt a multitask network to the identity document domain and we propose to take advantage of the template as an additional input of the network in order to guide the localisation of the fields. We explore various ways to integrate the template inside the network architecture. To evaluate our approach, we design a new task for the MIDV2020 database from the rectified in-the-wild photos. Our method achieves better results on the new task as well as on a private industrial database composed of real world examples.

A paper is submitted for ICDAR 2023.

3.9 Unsupervised learning of face image representations

Participants: Martin Dornier, Bertrand Coüasnon, Christian Raymond, Yann Ricquebourg.

Keywords: Neural networks, Unsupervised learning, facial landmark detection.

This CIFRE PhD (n°2020/0133) in partnership this InterDigital R&D France focuses on training neural networks with minimal annotated data in several InterDigital's fields of area.

First works focused on face alignment with limited annotated data. We developed an architecture based on a auto-encoder enhanced with the addition of "skip-connections" This work [10] has been published at the International Conference of Image Analysis and Processing 2022.

We also developed another architecture, again for face alignment with an extension to 3D landmarks, based on StyleGAN. This new architecture improves over our previous work. An article about this new model has been submitted to the journal Pattern Recognition

3.10 Early recognition of 2D gestures

Participants: William Mocaër, Eric Anquetil, Richard Kulpa.

Keywords: Online Handwritten Gesture Recognition, Convolutional Neural Network, Untrimmed.

Funded by the IntuiDoc team and by the EUR Digisport, this thesis has the objective to combine the knowledge of two teams, IntuiDoc and MimeTIC, to solve a common challenge: early gesture recognition.

It can be 2D gestures; symbols drawn on tablets, or 3D gestures: gestures done by an human (hands or full body). The final goal of early recognition is to be able to elaborate very reactive applications. First we focused on early recognition of 2D gestures.

Inspired by 3D gesture recognition approach, we built a new CNN architecture called OLT-C3D (for Online Long-Term Convolutional 3D) based on a spatio-temporal CNN. The originality of the network is it's ability to take a skeleton representation along the time, and apply spatio-temporal convolutions which does not consider the future. Also, thanks to dilated convolutions, the network is able to see far in the past and we can avoid the use of recurrent cells, the network is shown in figure 11.

The representation of the 2D gesture given in input of the network is a sequence of images, which fixed-length increments between each frame.

To make the network make stable decisions, the CTC Loss is added to train the network.

This works well for untrimmed 2D gesture recognition, and this lead to two new pub-

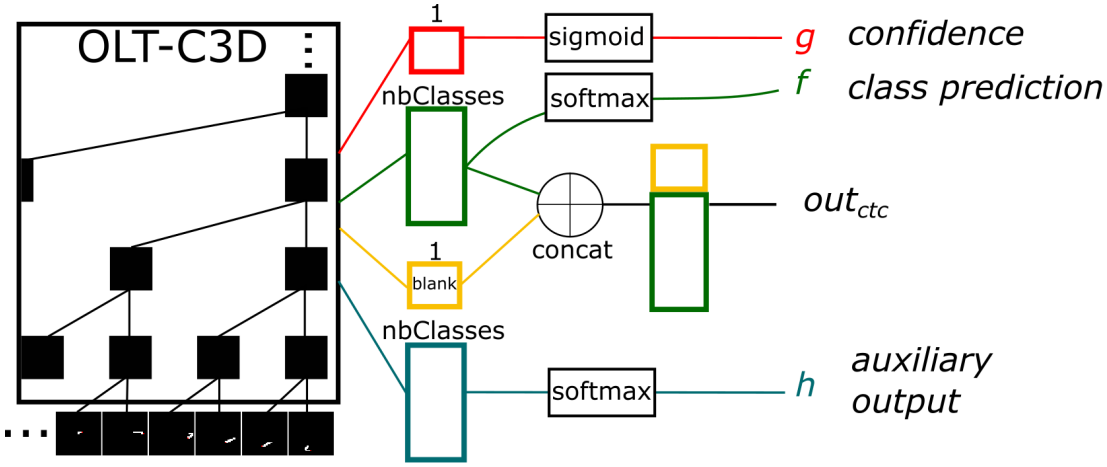


Figure 11: Network OLT-C3D with multiple outputs.

lications at RFIAP2022 [13] (oral presentation in June in Vannes), and ICPR2022 [12] (poster presentation in August 2022, in Montréal, Canada).

4 Software development

4.1 Software Deposit

All the presented softwares have been deposited in APP. More details on those softwares can be found on Intuidoc web site (http://www.irisa.fr/intuidoc).

4.2 IA for Digital learning: Handwriting analysis software

Contact: Eric Anquetil

Keywords: Handwriting Analysis, digital learning, fuzzy logic.

In the area of digital learning associated to the IntuiScript project and the ANR LabCom ScriptAndLabs (see section 5.3.4 and 3.3), we have developed four software:

- The ISF (Isolated Symbol Feature) software library allows the characterization of the meaning and direction of on-line manuscript tracing.
- The ISA (Isolated Symbol Analysis) software library allows the analysis of isolated symbols: capital letters, cursive letters, numbers
- The IWA (Isolated Word Analysis) software library enables the analysis and segmentation of handwritten words.
- The DAP (Drawing Precision Analysis) software library allows the analysis of the accuracy of a graphical plot against a guidance.

Through industrial collaboration with Learn&Go company, these software have been successfully integrated in the pen-based tablet solution: Kaligo. This solution is distributed by Learn&Go company. It is focused on learning writing at school from children aged 3 to 7. In 2021, at the LabCom Script&Labs end, a delivery of the source code related to these libraries has been done to the company Learn&Go.

4.3 DALI: a framework for the design of pen-based document sketching systems

Contact: Eric Anquetil

Keywords: Sketch recognition, pen-based interaction, visual language theory, industrial transfer.

DALI is a framework for the interpretation of hand-drawn sketches drawn on tablet PCs. The first property of the *DALI* method is its genericity, which means that it can be used to design pen-based software to sketch various natures of documents. It is based on the visual language and grammar theory that makes it possible to model bidimensional symbols and documents *DALI* interprets the user strokes *on-the-fly*, directly during the design of the document; it means that each time the user draws a stroke, the system analyses it and produces a visual feedback, showing how it is interpreted.

This way, the user is an actor of the interpretation process, because he can progressively correct the errors of the system. Thus, the interpretation process can rely on the information given by the user to better interpret the following strokes. The coupling of these two properties increases significantly the efficiency and the robustness of the sketch interpretation process.

The *DALI* method has been used to design several pen-based prototypes, for instance for the sketching of musical scores, electrical sketches, UML class diagrams, architectural floor plans, etc.

It has been transferred to the Script&Go society, which led to the design of *Script&Go Electrical Sketches*. These softwares are today commercialized and used daily by hundreds of technicians in France. *Script&Go Electrical Sketches* has been rewarded with the "Trophées de l'innovation" 2008 for uses, applications and communicating solutions for enterprises", in the category named "Solutions Métiers".

In 2018, DALI framework has been extended to design Dplan (Dali plan) library. The DPlan library allows the analysis and interpretation in real time of pen-based plan sketching on numeric tablet (walls, rooms, doors, windows...). Dplan library has been integrated in 2018 in the "IntuiDiag" software transferred to the Innax company. This transfer has been supported by a development fund from *SATT Ouest Valorisation*.

In 2021, DAlI framework has been extended to design the DGeo (Dali Geometry) library. The DGeo library allows on the fly analysis of geometric sketches on pen-based tablets. This library has been integrated into the "IntuiGeo" software.

4.4 DocRead : an automatic generator of recognition systems on structured documents

Contact: Bertrand Coüasnon

Keywords: Recognition, structured document, musical scores, mathematical formulae, table structures, forms, archives.

DocRead is an automatic generator of recognition systems on structured documents. It has been developed thanks to the DMOS-P method [Coü06]. It is made of a compiler of the EPF language (with which it is possible to describe a document), a parser associated to this language, an early vision module (binarization and line segments detection) and a classifier having also a reject option.

This generator allows us a fast adaptation to a new type of document. Indeed, it is only necessary to define a new grammar in EPF, which describes the new type of document. Then, if necessary, a new learning of the classifier is done to make it able to recognize new symbols. The new recognition system adapted to a new structured document is produced by compilation.

With this generator, we already have been able to produce recognition systems of structured documents. The recent, years, we mainly focused on the following ones:

- TabRead: a prototype for table structures recognition;
- Mexicanread: a software for field localisation in ancient mexican marriage records [LCCC18].
- MapRead: a software for the segmentation of historical maps [LC21].

4.5 DocRead development platform

Participants: Jean Camillerapp, Aurélie Lemaitre.

As presented in the previous section, DocRead is a method to develop new recognition systems. We set up a complete development environment for the creation of recognition system. This environment is based on:

- a virtual machine containing all the appropriate libraries,

[Coü06] B. Coüasnon, "DMOS, a Generic Document Recognition Method: Application to Table Structure Analysis in a General and in a Specific Way", *International Journal on Document Analysis and Recognition, IJDAR 8*, 2-3, June 2006, p. 111–122.

[LCCC18] A. Lemaitre, J. Camillerapp, C. Carton, B. B. Coüasnon, "A combined strategy of analysis for the localization of heterogeneous form fields in ancient pre-printed records", *International Journal on Document Analysis and Recognition 21(4)*, 269-282, July 2018, `https://hal.inria.fr/hal-01858192`.

[LC21] A. Lemaitre, J. Camillerapp, "Segmentation of historical maps without annotated data", *in : 6th International Workshop on Historical Document Imaging and Processing (HIP'21)*, *The 6th International Workshop on Historical Document Imaging and Processing*, Association for Computing Machinery, p. 19–24, Lausanne, France, September 2021, `https://hal.inria.fr/hal-03374571`.

- a kernel of lambda Prolog, PM-MALI,
- a complete environment with Eclipse, with a dedicated plugin for both lambda Prolog and EPF development.

Until 2020, we were using a virtual machine in a 32-bit environment. Since 2021, Jean Camillerapp has been porting the PM-MALI kernel to a 64-bit environment. This has required a deep study of the architecture of PM-MALI, with many adaptations.

Thanks to this work on PM-MALI, we updated the DocRead development platform with a new virtual machine in 64-bit environment. Aurélie Lemaitre has been working in 2022 on the finalization of this new virtual environment.

5 Contracts and collaborations

5.1 International Initiatives

5.1.1 French-German bilateral ANR project in artificial intelligence (KIHT - Kaligo-based Intelligent Handwriting Teacher)

Participant: Éric Anquetil, Yann Soullard, Wassim Swaileh, Florent Imbert, Romain Tavenard (LETG lab).

- Role : Project leader toward the French ANR (Stabilo for the German part)
- Partners: STABILO International GmbH, Karlsruher Institut für Technologie Institut für Technik der Informations¬verarbeitung, LearnAndGo company, Institut National des Sciences Appliquées de Rennes (IntuiDoc, IRISA Laboratory).
- 36 months (2021-2024)
- Contract: INSA, KIHT

In this project, we will design a new intelligent device to help learning handwriting in classrooms. The originality of the project consists in designing a new handwriting capture device developed by the company STABILO: a digital pen equipped with kinematic sensors (inertia measurement units (IMUs)) that allows writing on any surface (screen and paper).

The Stabilo company, supported by the German laboratory KIT, has the task of designing the hardware of the digital pen as well as embedding the AI algorithms developed. On our side, we are working, through a thesis and a post-doc, on the design of an original and powerful deep neural network architecture to automatically synthesise the online handwriting from the kinematic signals produced by the digital pen sensors.

5.2 National Initiatives

5.2.1 ANR HBDEX : Exploitation of Historical Big Data for Digital Humanities

Participant: Bertrand Coüasnon, Aurélie Lemaitre, Camille Guerry, Iwan Le Floch.

- Partners: Paris School of Economics (PSE), Université de Rouen Normandie (LITIS), CAMS-EHESS
- 48 months (2017-2022)
- Contract: INSA, ANR

This project focuses on massive historical data extraction for digital humanities, applied to financial data. The objective is to analyze masses of printed tabular data: daily stock exchange quotation lists for the "La Coulisse" Stock Exchange in Paris during the 19th and 20th century.

We work on modeling strategies of analysis for sequential data for using the redundancy between the successive days of listing and the consistency between the global sequences of data.

Final results are described in section 3.5.

5.2.2 ANR CollabScore: Shared spaces for digital music scores

Participant: Bertrand Coüasnon, Aurélie Lemaitre, Yann Soullard, Ali Yesilkanat, Hugo Hazard, Jean Camillerapp, Nathalie Girard.

- Partners: Cnam, INSA, BnF, Antescofo, IReMus, Fondation Royaumont
- 48 months (2020-2024)
- Contract: INSA, ANR

The project is dedicated to the collaborative digitization of music scores currently available only as images in museum collections. It will combine OMR (Optical Music Recognition) and a crowdsourcing correction phase of remaining recognition errors. A reconciliation step of the different versions will be automatised with specifically developed software, to obtain a reference score. This fulcrum notation will then be used in conjunction with several sources, to enhance the user experience. For instance, listening could be assisted by the synchronized scrolling of the score, and augmented with musicological annotations. The project aims at solving some scientific challenge, first by guiding and controlling an OMR process with musical knowledge, then by elaborating an automated crowdsourcing process.

5.2.3 Directory of Musical Writings of the Music Department of the BnF

Participant: Bertrand Coüasnon, Aurélie Lemaitre.

- Partners: Université de La Rochelle, INSA, BnF
- 48 months (2020-2024)
- Contract: INSA

Collaboration project on the constitution of a directory of musical writings of the collections of the Music Department of the BnF. Taking into account both autograph manuscripts and manuscripts by identified or anonymous copyists, the project will study the indexing of scripts on graphic characteristics.

5.2.4 ANR SKETCH : Sketches analysis and interpretation for the design of an intelligent tutorial system for medical studies

Participant: Eric Anquetil, Nathalie Girard, Islam Barchouch, Omar Krichen.

- Role : LP3C is the project leader, we are scientific leader of the AI part.
- Partner: *LP3C*, *IFPS*, *IFPEK*, *INSA*
- 42 months (2022-2026).
- Contract: INSA

Several recent studies have demonstrated that educational activities based on drawing can have positive effects on the learning of scientific concepts. The advent of innovative devices such as pen-based tablets means that new types of scaffolding involving artificial intelligence can now be designed and assessed. This opens up interesting avenues of research, as these devices make it possible to provide learners not only with support that can be parameterized by the instructor, but also with automatic and personalized realtime feedback during the drawing task.

The twofold aim of the SKETCH project is to 1) collaboratively design an intelligent tutoring system (ITS) that can analyze learners' actions in real time during the freehand production of a complex scientific drawing on a tablet, and 2) assess and optimize the effects of this system and the feedback it provides on learning. This project will be carried out jointly by two research teams in Rennes (France): the Psychology, Cognition, Behavior & Communication Laboratory (LP3C), and the IntuiDoc team at the Computer Science Laboratory (IRISA). It will focus on drawing activities intended to enhance learning about anatomy. Two of the partners in the project are paramedical colleges: IFPEK and IFPS. This will allow instructors and students to be involved in the project.

5.2.5 ANR(e-Fran) TRIANGLE : Working with Intelligent Feedback from a Digital Geometry Application for Student Engagement

Participant: Eric Anquetil, Nathalie Girard.

- Role: LP3C is the project leader, we are scientific leader of the AI part.
- Partner: *LP3C*, *Rennes and Poitiers academies*, *Rennes and Niort INSPE*, *INSA*
- 24 months (2022-2024).
- Contract: INSA

The objective of the TRIANGLE project is to consolidate the IntuiGéo application (intelligent tutorial system for geometry learning assistance on a tablet with a pen) by improving the impact of correction and guidance feedbacks, notably by adding a virtual pedagogical agent. We will also study the effects of this type of assistance on students' performance and engagement. Finally, in terms of dissemination, evaluation studies will be conducted in two academies. The deployment of a free multi-platform version in schools is planned at the end of the project. The consortium is made up of a research team in computer science (Intuidoc/IRISA, Rennes), a research team in psychology and ergonomics of learning (LP3C, Rennes), two INSPEs in close collaboration with the Academic Delegations for Digital Education (DANe) in two academies (Rennes and Poitiers).

5.2.6 AMI "Digital demonstrators in higher education" / AIR project- Increasing Interaction in Rennes

Participant: Eric Anquetil, Nathalie Girard, Bruno Hortollary.

- Role : Univ Rennes is the project leader, we are member of the consortium
- Partners: Univ. Rennes, INSA, Univ. Rennes 2
- 24 months (2022-2024).
- Contract: INSA

In this project we will consolidate and experiment in collaboration with the LP3C (UR2) and the University of Rennes, the KASSIS software suite which is a digital device we designed for pen based tablets to support synchronous active learning in class and remotely.

5.3 Bilateral industry grants

5.3.1 Research contract Interdigital company

Participant: Christian Raymond, Bertrand Coüasnon, Yann Ricquebourg.

- Partners: *Interdigital company*
- Since 2020
- Contract: INSA

Intuidoc team started to work with Interdigital company on latent representations in deep learning. This collaboration is based on the CIFRE grant for the PhD of Martin Dornier.

Current activities are described in section 3.9.

5.3.2 Research contract AriadNext company

Participant: Aurélie Lemaitre, Bertrand Coüasnon, Yann Soullard.

- Partners: *AriadNext company*
- Since 2021
- Contract: INSA

Intuidoc team started to work with AriadNext company in the field of document recognition applied to identity documents. This collaboration is based on a CIFRE grant for the PhD of Timothée Neitthoffer.

Current activities are described in section 3.8.

5.3.3 Research contract Learn&Go company : KALUX (Kaligo Luxembourg and Swiss)

Participant: Eric Anquetil, Nathalie Girard, Omar Krichen, Pauline Nerdeux.

- Partner: *Company Learn&Go*
- 12 months (2021-2022).
- Contract: INSA

After having developed the handwriting analysis engines for the French and English languages for the company Learn&Go within the scope of the Labcom Script&Labs, we are now developing the analysis engines for numbers, upper and lower case letters for the German language in order to allow the company Learn&Go to release their Kaligo application in Luxembourg and Swiss.

5.3.4 Industrial software licensing and technology transfer with Learn&GO company

Participant: Eric Anquetil, Nathalie Girard, Omar Krichen, Pauline Nerdeux.

- Partners: *Learn&GO company*
- 2017-2022
- Contract: INSA

The IntuiDoc team has close links with the Learn&GO company for transferring its research results for e-education. This partnership is now based on several licensing agreements. They cover various technologies of the Intuidoc team such as handwriting recognition and analysis.

This partnership is also supported by several collaborative projects: in particular the ANR LabCom "Script&Labs", the *IntuiScript* BPI Project (`http://intuiscript.com/`) and the P2IA project. This collaboration is also based on a CIFRE grant (A. Lods PhD Student) in relation with the ANR LabCom "Script&Labs".

The developments resulting from Labcom and P2IA have enabled the KALIGO product to be marketed in France, England and soon in Luxembourg and Germany, with high added value functionalities. The SATT OUEST VALORISATION is also supporting this technology transfer by participating in the equity of the company Learn&Go in 2021. The effective work of the INTUIDOC team has been fundamental to the technical evolution of the KALIGO product.

Current activities are described in sections 3.3, 3.2.

6 Dissemination

6.1 Promoting scientific activities

6.1.1 Scientific Events Organisation

- N. Girard, A. Lemaitre and Y. Soullard organized the SIFED national symposium in Rennes, in October 2022. `https://project.inria.fr/sifed2022/`

6.1.2 Scientific Events Selection

Member of Conference Program Committees

- A. Lemaitre and E. Anquetil are members of the program committee of the International Conference on Frontiers in Handwriting Recognition (ICHFR 2022).
- A. Lemaitre, B. Coüasnon and E. Anquetil are members of the program committee of the Document Analysis System workshop (DAS 2022).

Reviewing Activities

- A. Lemaitre is reviewer for International Conference on Pattern Recognition (ICPR) 2022.

6.1.3 Journal

Reviewing Activities

- E. Anquetil, B. Coüasnon, A. Lemaitre and Y. Soullard are reviewers for the International Journal on Document Analysis and Recognition (IJDAR) in 2022.

6.1.4 Scientific Expertise

- A. Lemaitre was a reviewer in 2022 of one project for the French National Research Agency (ANR).
- E. Anquetil was a scientific expert in 2022 to one project for the French National Research and Technologie Agency (ANRT).
- B. Coüasnon was a scientific expert in 2022 on a project for Innoviris, Brussels.
- A. Lemaitre was a member of two recruitment committees of assistant professor: in May 2022 (IUT Nantes), and in November 2022 (IUT Nantes).
- N. Girard was a member of one recruitment committee of assistant professor: in November 2022 (La Rochelle University, Computer Science departement, L3i).

6.1.5 Research Administration

- N. Girard, A. Lemaitre and Y. Soullard are members of the executive committee of the society GRCE : " Groupe de Recherche en Communication Écrite ".
- Intuidoc members are members of the AFRIF (Association Française pour la Reconnaissance et l'Interprétation des Formes) and IAPR (International Association for Pattern Recognition) associations.
- E. Anquetil is a member of the educational committee of the "DIGISPORT" University Research School (EUR).
- E. Anquetil is project manager for "Innovation and Entrepreneurship" at INSA Rennes.
- E. Anquetil is the manager of the incubator for innovative projects at INSA Rennes.
- E. Anquetil is the co-manager of the inter-institutional student incubator for innovative projects from 12 higher education institutions in Rennes: Station Rennes Innovation.

- E. Anquetil is an elected member of the administration council of INSA Rennes.
- E. Anquetil is a member of the administration council of INSA Group Foundation.
- B. Coüasnon is an elected member of the scientific council of INSA Rennes.
- B. Coüasnon is an elected member of the laboratory council of the INSA component of IRISA.
- B. Coüasnon is member of the Gender Equality Commission of IRISA.
- B. Coüasnon is member of the board of Valconum (Centre Européen de Valorisation Numérique).
- B. Coüasnon is member of the laboratory council of IRISA.
- N. Girard is an elected member of the administration council of UFR ISTIC, Univ. Rennes 1.
- Y. Ricquebourg is an elected member of the scientific council of INSA Rennes.

6.2 Teaching, supervision

6.2.1 Teaching

The team is mainly made up of teachers who are very implied in activities of teaching. But a majority of lectures are not attached to this research topic, so they are not mentioned here.

- E. Anquetil is program manager of the Master of Science *"Innovation and Entrepreunership"* of INSA and Rennes School of Business (RSB).
- E. Anquetil and N. Girard give lectures at *Research in Computer Science (SIF)* Master of University of Rennes 1, University of Southern Brittany, ENS Rennes, INSA Rennes and CentraleSupélec.
- E. Anquetil is in charge of the module "Analysis, Interpretation and Recognition of 2D (touch) and 3D Gestures for New Man-Machine Interactions" (AIR) of the *Research in Computer Science (SIF)* Master of University of Rennes 1, University of Southern Brittany, ENS Rennes, INSA Rennes and CentraleSupélec.
- E. Anquetil is in charge of the module "Motion Analysis and Gesture Recognition (2D / 3D)" (AMRG) of the Computer Science Dept. of INSA Rennes.
- B. Coüasnon is co-Head with A. Termier of the *Research in Computer Science (SIF)* Master of University of Rennes 1, University of Southern Brittany, ENS Rennes, INSA Rennes and CentraleSupélec (https://master.irisa.fr).
- B. Coüasnon was invited for two courses at Master-Research “New technologies applied to History” of the Ecole nationale des Chartes on : “Digital Documents: Textual Documents” and “Automatic Access to Old Documents”, Paris, France.

- B. Coüasnon is in charge (with M. Babel) of the module "Image & Video Analysis" (TIV) of the Computer Science Dept. *(Medias & Interactions section)* of INSA Rennes.

- Y. Ricquebourg and E. Anquetil are in charge of the module "Recognition and Interpretation of Images & Videos" (RIV) of the Computer Science Dept. *(Medias & Interactions section)* of INSA Rennes.

- Y. Soullard is in charge of the part "Text Mining and Deep Learning" of the module "Introduction to the Text Mining" at Master MAS *(Mathématiques Appliquées, Statistique (Science des Données))* of Rennes 2 University.

- Y. Soullard is in charge of the part "Recurrent Networks and Time Series Analysis" of the module "Deep Learning" at Master TELENVI *(Télédétection-Environnement)* of Rennes 2 University and Agrocampus Ouest.

- N. Girard was invited for two courses at Master and Master-Research "Découverte de la recherche" at La Rochelle University on : "IA for e-education", La Rochelle, France.

6.2.2 Supervision

- PhD: S. Tarride, Combination of logical and textual knowledge for recognition of ancient register images, B. Coüasnon, A. Lemaitre, S. Tardivel (Doptim), INSA de Rennes, defended April 2022 [3].
- PhD: C. Leroy, Incremental learning and evolving recognition system: application to on-the-fly recognition of handwritten gestures, E. Anquetil, N. Girard, INSA Rennes, defended June 2022 [2].
- PhD: C. Guerry, Historical big data: modelization of strategies to analyse collections of documents, B. Coüasnon, A. Lemaitre, S. Adam (Univ Rouen), INSA de Rennes, defended December 2022 [1].
- PhD in progress: I. Barchouch, Intelligent tutorial system for sketch-based learning (SKETCH), E. Anquetil, N. Girard, started October 2022.
- PhD in progress: F. Imbert, Design of a deep neural network architecture dedicated to the synthesis of handwriting from kinematic sensors of a sensors of a digital pen, E. Anquetil, Y. Soullard, R. Tavenard, INSA Rennes, started October 2021.
- PhD in progress: T. Neitthoffer, Joint analysis of handwriting localization and recognition in structured documents, B. Coüasnon, A. Lemaitre, Y. Soullard, A. M. Awal (AriadNext), INSA Rennes, started May 2021.
- PhD in progress: K. Barrere, Deep neural networks and attention mechanisms for handwritten text recognition, B. Coüasnon, A. Lemaitre, Y. Soullard, INSA Rennes, started October 2020.
- PhD in progress: M. Dornier, Deep Learning in the Wild, B. Coüasnon, P.H. Gosselin, C. Raymond, Y. Ricquebourg, INSA Rennes, started October 2020.
- PhD in progress: W. Mocaër, Spatio-Temporal Convolutional Neural Network for early action detection and analysis, E. Anquetil, R. Kulpa, INSA Rennes, started October 2020.

- PhD in progress: S. Corbillé, Hybridization of "Transparent" and "Deep Learning" AI approaches for automated handwriting analysis of children in the context of education, E. Anquetil, E. Fromont, Univ. Rennes 1, started October 2019.
- PhD in progress: A. Lods, On-line analysis of handwritten arithmetic operation on digital tablet: Design of an innovative educational solution to improve learning arithmetic calculations in elementary school, E. Anquetil, S. Macé, INSA Rennes, started Febuary 2018.

6.2.3 Juries

- E. Anquetil was reviewer in the thesis committee of Jordan Gonzalez PhD, Apprentissage incrémental en données : Application à la reconnaissance d'émotions personnalisée, HESAM University (ED SMI-432), October 2022.
- N. Girard was member of the thesis committee of Joris Voerman's PhD, Classification automatique à partir d'un flux de documents, La Rochelle Université, June 2022.
- N. Girard was member of the thesis committee of Salah Eddine Boukhetta's PhD, Analyse de séquences avec GALACTIC – Approche générique combinant analyse formelle des concepts et fouille de motifs, La Rochelle Université, August 2022.
- E. Anquetil is member of mid-term evaluation committee of the PhD candidate: Y. XIE (Université de Nantes, LS2N/IPI).
- B. Coüasnon is member of mid-term evaluation committee of the PhD candidate: Thomas Constum (Université de Rouen-Normandie, LITIS).
- Y. Ricquebourg is member of mid-term evaluation committee of the PhD candidate: Tsiry Mayet (Insa Rouen, LITIS).
- Y. Soullard is member of mid-term evaluation committee of the PhD candidate: Simon Corbillé (IRISA, université Rennes 1).

6.3 Popularization

- A. Lemaitre published a paper in a journal for human sciences [5].
- E. Anquetil was invited to participate in the panel discussion: "Rennes and Brittany, digital territories for education and innovation" in January 2022.

7 Bibliography

Doctoral dissertations and "Habilitation" theses

[1] C. Guerry, *Combinaison de connaissances logiques et textuelles pour la reconnaissance d'images de registres anciens*, Theses, INSA RENNES, December 2022.

[2] C. Leroy, *Apprentissage en environnement non stationnaire de systèmes d'inférence floue évolutifs: Application pour un système gestes/commandes*, Theses, INSA RENNES, June 2022.

[3] S. Tarride, *Big-data historique : modélisation de stratégies d'analyse de collections de documents*, Theses, INSA RENNES, April 2022.

Articles in referred journals and book chapters

[4] O. Krichen, S. Corbille, E. Anquetil, N. Girard, E. Fromont, P. Nerdeux, "Combination of explicit segmentation with Seq2Seq recognition for fine analysis of children handwriting", *International Journal on Document Analysis and Recognition*, November 2022, https://hal.science/hal-03845144.

[5] A. Lemaitre, "L'analyse de cartes anciennes par l'intelligence artificielle", *Palimpseste. Sciences, humanités, sociétés Printemps-été 2022*, 7, June 2022, p. 45–48, https://hal.inria.fr/hal-03697549.

[6] T. Vayer, R. Tavenard, L. Chapel, N. Courty, R. Flamary, Y. Soullard, "Time Series Alignment with Global Invariances", *Transactions on Machine Learning Research*, October 2022, https://arxiv.org/abs/2002.03848, https://hal.science/hal-02473959.

Publications in Conferences and Workshops

[7] K. Barrere, Y. Soullard, A. Lemaitre, B. B. Coüasnon, "A Light Transformer-Based Architecture for Handwritten Text Recognition", *in : 15th IAPR International Workshop on Document Analysis Systems (DAS 2022)*, *Lecture Notes in Computer Science*, *13237*, Springer International Publishing, p. 275–290, La Rochelle, France, May 2022, https://hal.science/hal-03685976.

[8] K. Barrere, Y. Soullard, A. Lemaitre, B. Coüasnon, "A Light Transformer-Based Architecture for Handwritten Text Recognition", *in : Reconnaissance des Formes, Image, Apprentissage et Perception (RFIAP) 2022*, Vannes, France, July 2022, https://hal.science/hal-03857723.

[9] A. Chapin, Y. Soullard, A. Lemaitre, B. B. Coüasnon, "Handwritting Recognition using Multi-Hypothesis based Semi-Supervised Learning for adaptation to specific documents", *in : Symposium International Francophone sur l'Ecrit et le Document (SIFED'2022)*, Rennes, France, October 2022, https://hal.science/hal-03847153.

[10] M. Dornier, P.-H. Gosselin, C. Raymond, Y. Ricquebourg, B. Coüasnon, "SCAF: Skip-Connections in Auto-encoder for Face Alignment with Few Annotated Data", *in : ICIAP 2022 - International Conference on Image Analysis and Processing*, *Lecture Notes in Computer Science*, *13231*, Springer International Publishing, p. 425–437, Lecce, Italy, May 2022, https://hal.science/hal-03687091.

[11] F. Imbert, E. Anquetil, R. Tavenard, Y. Soullard, W. Swaileh, "Toward Deep neural network for pen trajectory reconstruction from kinematic sensors", *in : Symposium International Francophone sur l'Ecrit et le Document (SIFED'2022)*, Rennes, France, October 2022, https://hal.science/hal-03895960.

[12] W. Mocaër, E. Anquetil, R. Kulpa, "Early Recognition of Untrimmed Handwritten Gestures with Spatio-Temporal 3D CNN", *in : ICPR 2022 - International Conference*

on Pattern Recognition, p. 1636–1642, Montréal, Canada, August 2022, https://hal.science/hal-03683441.

[13] W. Mocaër, E. Anquetil, R. Kulpa, "Réseau Convolutif Spatio-Temporel 3D pour la Reconnaissance Précoce de Gestes Manuscrits Non-Segmentés", *in : RFIAP 2022 - Congrès Reconnaissance des Formes, Image, Apprentissage et Perception*, p. 1–9, Vannes, France, July 2022, https://hal.science/hal-03682604.

[14] W. Swaileh, M. Jordan, D. Kotzinos, "3D Modelling Approach for Ancient Floor Plans' Quick Browsing", *in : DAS 2022 - 15th IAPR International Workshop on Document Analysis Systems*, La Rochelle, France, May 2022, https://hal.science/hal-03657982.

[15] S. Tarride, A. Lemaitre, B. B. Coüasnon, S. Tardivel, "A comparative study of information extraction strategies using an attention-based neural network", *in : 15th IAPR International Workshop on Document Analysis Systems*, La Rochelle, France, May 2022, https://hal.science/hal-03677908.

Miscellaneous

[16] K. Barrere, B. Coüasnon, A. Lemaitre, Y. Soullard, "Architecture Transformer Légère pour la Reconnaissance de Textes Anciens", SIFED 20222 - Symposium International Francophone sur l'Ecrit et le Document, October 2022, Poster, https://hal.science/hal-03857509.

[17] S. Corbillé, E. Fromont, E. Anquetil, "Recognition and segmentation of children handwriting", SIFED 20222 - Symposium International Francophone sur l'Ecrit et le Document, October 2022, Poster, https://hal.science/hal-03879307.