UMR IRISA

Activity Report 2019

Team INTUIDOC

Intuitive User Interaction for Documents

D6 – Media and Interactions

CentraleSupélec cnrs rennes IMT Atlantique Bretagne-Pays de la Loire École Mines-Télécom Inria INSA RENNES Université Bretagne Sud ubs: UNIVERSITÉ DE RENNES 1

1 Team composition

Researchers and faculty

Eric Anquetil, Professor, Insa, head of the team
Jean Camillerapp, Retired Professor
Bertrand Coüasnon, Associate Professor, Insa, HDR
Nathalie Girard, Associate Professor, Uni. Rennes 1
Aurélie Lemaitre, Associate Professor, Uni. Rennes 2
Ivan Leplumey, Associate Professor, Insa
Yann Ricquebourg, Associate Professor, Insa
Yann Soullard, External Colloborator, Uni. Rennes 2, from December 1st

Research engineers, technical staff

Simon Bouvier, Insa Research Engineer
Morgane Carry, Insa Research Engineer, from February 11th
Simon Corbillé, Insa Research Engineer, until September 30th
Pauline Nerdeux, Insa Research Engineer, from June 24th
Solène Tarride, Rennes 1 Research Engineer, until January 31st
Sébastien Thomas, Insa Research Engineer, until September 27th

PhD students

Said-Yacine Boulahia, Insa PhD student, Government of Algeria grant, until August 31st
Simon Corbillé, Uni. Rennes 1, PhD student, from October 1st
Kwon-Young Thomas Choi, Insa PhD student,
Camille Guerry, Insa PhD student
Omar Krichen, Insa PhD student
Clément Leroy, Insa PhD student
Arnaud Lods, CIFRE/Insa PhD Student
Solène Tarride, CIFRE PhD Student, from February 1st

PostDoc

Mohammed Hindawi, Insa Post-Doc, until February 1st
Damien Simonnet, Insa Post-Doc

Administrative assistant

Nadia Derouault

2 Overall objectives

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:

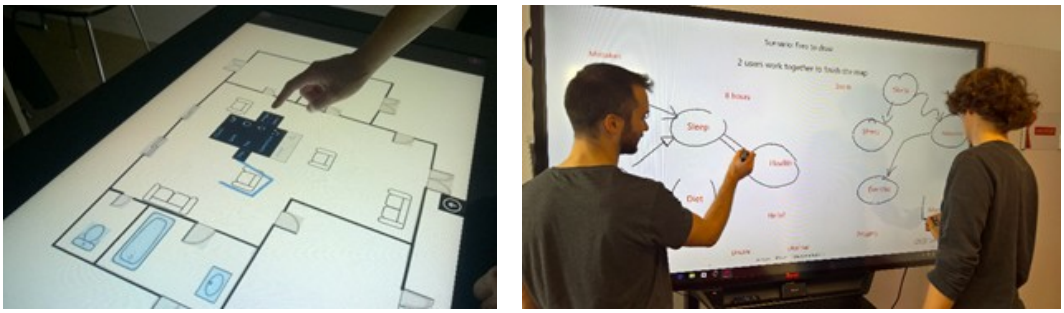
- "On-line" evolving cross-learning of 2D (touch and pen –based) and 3D gestures (Kinect and Leap Motion);
- "On-line" analysis of drawing, sketching and handwriting with pen-based tablet for digital learning (e-education);
- Interactive learning of document structure without ground-truth;
- Document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

3 Scientific Foundations

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

3.1.1 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.



(a) (b)

Figure 1: Multi-Touch and Multi-User Interaction

3.1.2 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 are going to 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.

3.1.3 Multiple users 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 participations of multiple users. This is a really new research topic to automatically recognize 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 [6] to recognize 3D action gesture. This new research topic will be conducted in collaboration with MIMETIC project team of Inria. We base this proposition on the observation that 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 modeling such human motions. We hypothesize that both recognition problems could be addressed in similar ways.

3.2 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 5 major research projects over the past 5 years with a funding budget of 1.456 Million euros.

3.2.1 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 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 customize "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.

3.2.2 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 is 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-centered 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.

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

IntuiDoc team from IRISA laboratory has been a leader for more than 20 years in online analysis and recognition of writing and handwritten documents. Thanks to its expertise in Artificial Intelligence and more specifically in Machine Learning, it designs recognition and analysis engines allowing automated interpretation of online traces using pen-based tablets : handwriting, gestures, symbols, sketches, architecture plans... Learn&Go company, related to Script&Go, is specialized in developing digital solutions on tablets for mobility and e-education,

Figure 2: Handwriting analysis for digital learning

aiming to improve learning through digital tools, from pre-schools to graduate higher education as well as in professional training.

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 [11].

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 three 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 - Intelligence Project Artificial (P2IA). The 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.

3.2.3 "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 [9] (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).

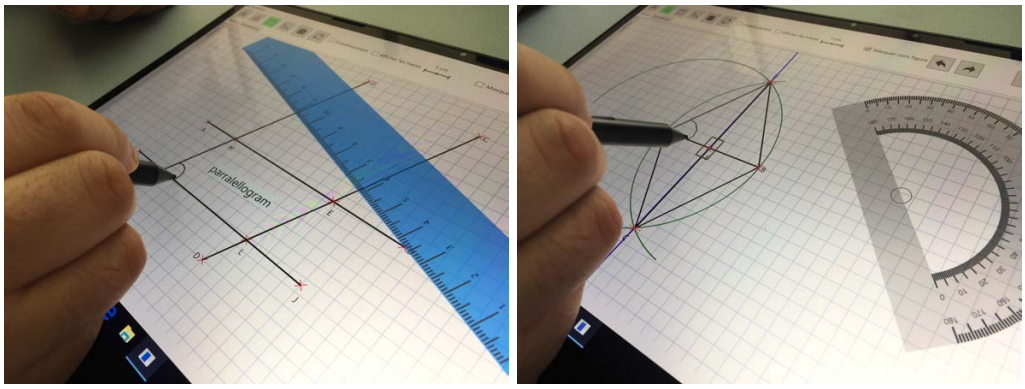


Figure 3: Pen-based and Gesture-based software for geometry learning.

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 developed the new innovative four year national project "ACTIF" (BPI – e-FRAN) with the support of the Brittany Region. The partners for this axe are LP3C and LOUSTIC laboratories, Learn&Go company, educational experts (ESPE) and Brittany region.

3.3 Interactive learning of document structure without ground-truth

3.3.1 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 organizations 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.

3.3.2 Combination Deep Learning / Syntactical

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 modeling 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.

3.3.3 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.

3.4 Document collection analysis for historical big-data

3.4.1 Handwritten Historical Registers

We start collaboration with Irccyn - University of Nantes on the layout recognition of registers of the Théâtre-Italien from the 18th century. The University of Nantes is a partner of the ANR CIREFI and will work with the DMOS-PI method, proposed by IntuiDoc, to build a document structure recognition system for these handwritten registers, which will drive handwritten text recognizers, to make a complete information retrieval system.

3.4.2 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 modeling 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 modeling 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.

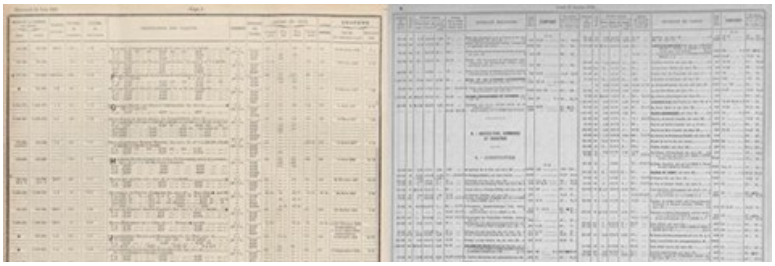


Figure 4: Examples of daily listing on the Paris Stock Exchange

3.4.3 Adaptive Document Layout Analysis

We propose to integrate the interactive document structure learning without ground truth and the collection modeling 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.

4 Scientific achievements

4.1 Fine Handwriting Quality Analysis: going further with handwriting

Participants: Éric Anquetil, Damien Simonnet, Sébastien Thomas, Pauline Nerdeux, Nathalie Girard.

The LabCom "Script&Labs" research was prioritized this year on the design of new handwriting analysis engines for the specifics of learning to write in England. The development of these engines required several campaigns in English schools (in London) to collect data. These new engines were transferred to Learn&Go company, which was able to release an English version of its product "KALIGO UK" in 2019 on the English market. The LabCom Script&Labs worked this year on four main axes: the improvement of the analyse of English (UK) handwriting characters, the management of UK words, a better management of line spacing feedback and the improvement of french handwriting recognition by increasing training dataset.

4.1.1 Analysis of UK characters and words

Following what has been done for French handwritten digits and letters [13], further work has been done to implement and improve analysis engines for UK digits, capitals, manuscript, precursive and cursive letters.

The existing word analyser has been extended to manage UK handwriting. First, specific recogniser and analyser have been trained to manage letters with high and low ligatures (see figure 5) and with specific UK cursive handwriting such as the x and z (see figure 5). Then, the building of the segmentation graph has been adapted to managed letters specific to UK handwriting due the introduction of the ligatures and the specific letters x and z. Some qualitative results are presented in figure 6.

Overflow/underflow feedback have been improved when guidelines are used, providing more precise feedback especially relative to English ligature models.

Finally, the execution time and memory consumption of the words analysis engines have been greatly improved, both for French and English.

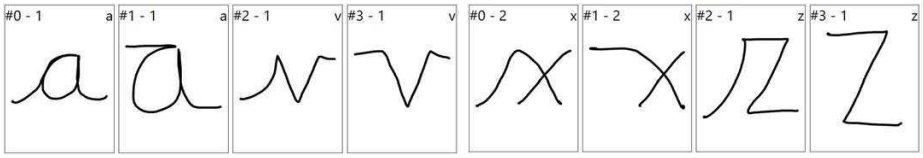


Figure 5: Letters a and v with low/high left ligatures and Letters x and z specific to the UK cursive handwriting with low/high left ligature.

4.1.2 French handwriting recognition

In order to improve the performances of the FR recognition engine, tests on the data synthesis have been carried out. Deformations are performed randomly on all the seeds of the original base in order to increase the size of the learning base and also to bring as much diversity as

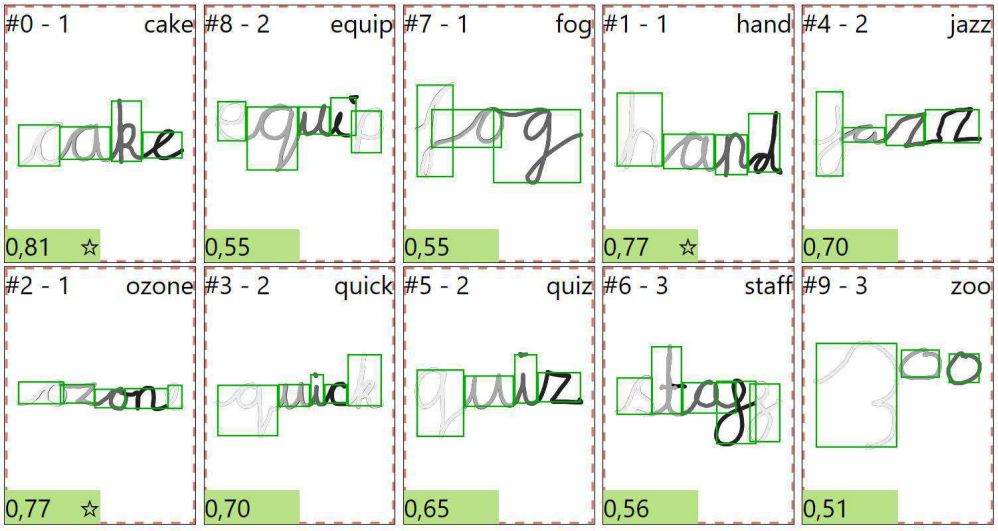


Figure 6: Qualitative results about the UK word analysis engine.

possible to the training dataset. New data letters extracted from words have also been added to the training database.

4.2 Real-time interpretation of geometric shapes for digital learning

Participants: Omar Krichen, Eric Anquetil, Nathalie Girard, Simon Corbillé.

Keywords: Online recognition of Hand-drawn sketches, Intelligent Tutoring Systems.

This work is in the context of e-FRAN ACTIF projects (see 6.2.3, which aims to use pen-based tablets in an educational context to foster active and collaborative learning in French middle-schools. The partners are LP3C and LOUSTIC laboratories, Learn&Go company, educational experts (ESPE) and Brittany region.

"IntuiGeo" is a tutoring system for geometry learning in middle-schools on pen-based tablets. Our approach is based on three main principles: the online recognition of the user's hand-drawn sketches, the supervision of the pupil's problem resolution strategy, and the generation of real-time visual, corrective and guidance feedback. Our system is composed of two main engines. The 2D recognition engine is responsible for the interpretation of the hand-drawn strokes (and has already been validated in [12]). The supervision engine is responsible for the automated generation of a geometry construction problem from a solution example drawn by the teacher, as well as the supervision of the resolution strategies chosen by the pupil. The tutor architecture is inspired by the traditional ITS (Intelligent Tutoring System) architecture and is illustrated in Figure 7.

- **The problem generation module**: takes as input the teacher figure interpreted by the the 2D engine and creates a model of the problem, in the form of a knowledge graph, storing all the geometric elements and their links, mathematical constraints such as length, angle, or perpendicularity;

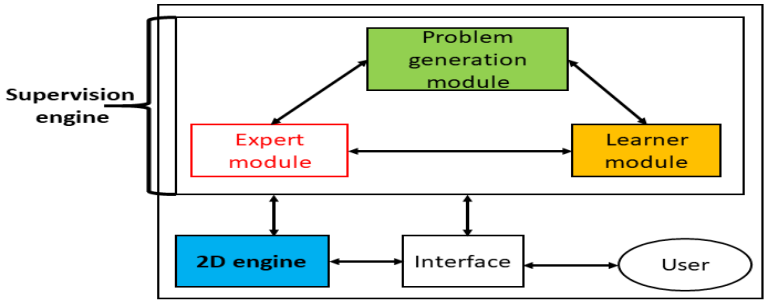


Figure 7: IntuiGeo architecture

- **The learner module**: is responsible for the semantic interpretation of the student production, in real-time, by matching the recognized hand-drawn strokes with the generated model of the problem (spatial relations, mathematical constraints, etc.). More broadly, from an intelligent tutoring point of view, the learner module refers the dynamic representation of the student knowledge and skill;
- **The expert module**: To be able to provide corrective and guidance feedback, the expert has to be able to synthesize strategies to solve a given construction problem. This ability is based on the definition of a dynamic planning environment. The knowledge here is defined by the set of ruler-compass drawing actions that will be the basis of the strategy chosen to solve the problem.

The pen and touch interaction allows for the simulation of the pen and paper traditional setting. The latest version of the interface is illustrated in Figure 8. Thanks to the users feedback (teachers and pupils), the developed virtual tools have greatly evolved (*c.f* Figure 9). This represents one of the main advantages of our user-centered design. A replay mode has also been integrated to the tool, as well as a feedback mode in textual format. The tutor has been tested in five pilot middle schools in the region of Brittany. 54 pupils were given four construction exercises with varying difficulty. This deployment in the schools had as a result the total of 203 "digital sheets" containing all the saved pupils actions (*.e.g* handwritten strokes, virtual tools usage) when solving one of the given exercises. Over this set, the supervision engine is able to interpret the child resolution strategy (state of knowledge, errors, realized steps) with a 99.07% accuracy and 97.72% recall. The tutor is also able to synthesize resolution strategies given any resolution state, which demonstrate the ability to provide personalized feedback, dependent on the pupil skills and advancement. The later work on the tutoring aspect of our system has been accepted [KAG20] to be presented in ECAI (European Conference on Artificial Intelligence).

4.3 An anticipation strategy to manage non stationary data stream

Participants: Clement Leroy, Éric Anquetil, Nathalie Girard.

Keywords: Online learning, Non stationary datastream, Evolving fuzzy system.

[KAG20] O. Krichen, E. Anquetil, N. Girard, "IntuiGeo: Interactive tutor for online geometry problems resolution on pen-based tablets", *in : European Conference on Artificial Intelligence (ECAI) 2020, In Press*, Santiago de compostela, Spain, August 2020, `https://hal.archives-ouvertes.fr/hal-02544384`.

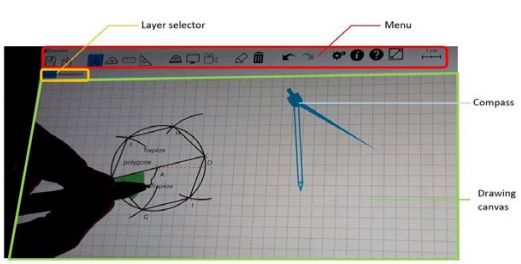


Figure 8: IntuiGeo's interface

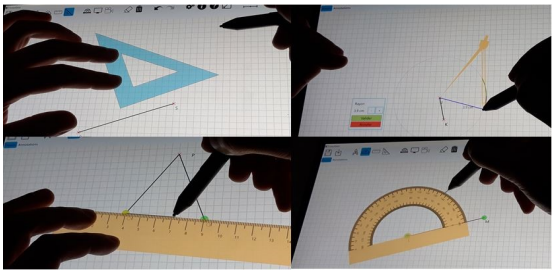


Figure 9: Virtual tools

The following results come from the first PhD years of Clement Leroy entitled: "Incremental learning in non stationary stream: Application to the on-the-fly gesture recognition".

Working with a non-stationary stream of data requires for the analysis system to evolve its model over time. In particular, it should incorporate drift detection and adaptation methods to continuously adapt to the data stream in a long-life learning context. Based on inference rules set, Evolving Fuzzy System - EFS - has proved effective in addressing this issue. However, tackle jointly the concept drift detection, the novelty detection as well as the forgetting is still an open question.
Throughout the year, two main axis have been explored. An anticipation module have been proposed to replace old not suitable rules by new anticipated rules when brutal drifts are detected. The module ensures reactivity of the system to the drifts while maintaining stability with time when no change is detected. This works consolidates previous works and have been reviewed in the FUZZ-IEEE 2019 conference at New Orleans [16]. The benefits in term of stability and reactivity of the anticipation module are shown Figure 10 and Table 1. Later, improvements have been bring by integrating forgetting capacity in the consequent part of the rules by the use of multiple sliding windows. This work has been submitted in the next 25th International Conference on Pattern Recognition (ICPR2020).

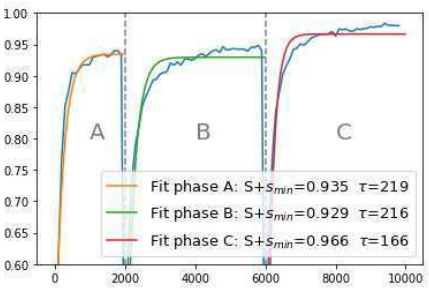


Figure 10: Artificial brutal drifts are generated over real datasets (e.g Letters) over three phase A,B,C. A handcraft fit is carried out on the prequential accuracy score over the three phases using the following equation : $y(t) = S(1 - e^{-\frac{t}{\tau}}) + s_{min}$ with τ, S, s_{min} three fitted parameters characterizing the drift.

The second axis concern the proposition of a new anticipation strategy to manage the inference rule deletion in EFS. The strategy is built upon two key concepts, the spatial and temporal scales analysis and the measure of utility of rules that mix local and global information,

		$< S + s_{min} >$	$< \tau >$	$< Acc >$
Letters	With Anticipation	**94.3**	200	**90.3**
	Without Anticipation	92.2	214	88.1
PenDigits	With Anticipation	98.8	**56**	97.9
	Without Anticipation	98.8	103	97.2
Laviola	With Anticipation	**98.2**	**74**	**96.2**
	Without Anticipation	97.7	76	95.7

Table 1: Mean of the fitted parameters and mean accuracy score are given for different data in an EFS with and without the anticipation module

both made effective thanks to the anticipation module. The proposed strategy is still under development but preliminary results suggested that the deletion mechanism improve robustness of the system, while reducing its complexity.

4.4 Mathematical expression recognition and analysis

Participants: Arnaud Lods, Éric Anquetil, Sébastien Macé from Learn&Go company.

This part sum up the work done for the second thesis' year of A. Lods that began on January 2018 on the subject of On-line analysis of handwritten arithmetic operation on digital tablet: Design of an innovative educational solution to improve learning arithmetic calculations in elementary school. This thesis is carried out in collaboration with the LabCom "Script&Labs" (see section 6.2.2) and is financed as a CIFRE by the ANR n°ANR- 16-LVC2-0008-01.

The objective of the thesis is to produce a pen-based tablet application to provide children in elementary school with an application to practice solving arithmetic operation (addition, subtraction...). Given a mathematical problem proposed by a teacher, the child is expected to solve this problem the same way he would using a pen on a paper. The system then analyzes the student handwritten input and analyse it knowing the expected solution. If errors are detected because of miscalculation, algorithm misunderstanding, omission (if he forgot to carry over in addition or subtraction) or misplacement, the system guides the student to fix his mistakes.

The first step of this system is the construction of a stroke graph with symbolic links. The idea is to create link between strokes through extracted characteristics that can approximate as much as possible the human vision. We use geometrical features to combine pair of strokes into symbols. We use a CNN-based neural network to classify each pair of strokes into a mathematical symbol. Then we use fuzzy logic to link related symbols. We obtain a graph of mathematical symbols linked by their mathematical relationships that can be parsed for either the recognition the analysis of mathematical expression. The recognition process has been applied on a set of already segmented symbols and presented in [17] at the ICDAR 2019. The table 4.4 presents the results obtained on the CROHME dataset by the system presented in the paper.

We can produce, given a mathematical problem given by a teacher, an expected graph of symbols which corresponds to the answer. Given our previous work we are now able through our representation to compare both graphs and highlight the differences and thus the mistakes made by the children. We realize this by computing the Graph-Edit Distance (GED) between the two graphs which gives us the optimal way to transform one graph into another (adding/removing a symbols, adding a relationship...). This way we can deduce what divert from the expected

Table 2: Structure recognition with provided symbols on CROHME 2014 and CROHME 2016 datasets.

	Structure	**Structure + Labels**		
	Rec. Rate	Rec. Rate	<= 1 err	<= 2 err
CROHME 2014				
MyScript	**90.67**	**84.38**	**85.90**	**87.62**
Wiris	86.61	78.80	80.42	82.75
Our solution	87.76	78.43	81.54	83.28
MST	76.66	67.44	-	-
BLSTM	69.27	64.81	67.34	70.69
CYK	70.99	61.46	63.89	66.84
CROHME 2016				
Our solution	91.11	85.79	88.84	90.23

solution. We were first using an A* algorithm, which compute a perfect but time-consuming GED. This preliminary work has been presented in [18] at the GREC workshop 2019. The current objective is to apply an optimized search algorithm and find an alternative matching representation to speed-up this process to carry out the matching in a matter of seconds even on dense graphs.

4.5 Active learning with pen-based tablet

Participants: Morgane Carry, Sébastien Thomas, Mickaël Renault, Nathalie Girard, Éric Anquetil.

The eFil Cominlabs project (see 6.2.4) is based on CAMIA/KASSIS which is a new digital learning environment (workbook) for active learning with pen-based tablet. The goal of the eFIL Project is to optimize and evaluate this digital learning environment in classroom settings during lectures (Optimizing with user-centered design). Traces will be collected and analyzed both to explore learning dynamics in classroom, and to provide a dashboard allowing instructors to monitor their activities during lectures as well as in-between sessions. The goal is to explore and analyze traces to monitor learning activities in classroom with Educational Data Mining and Learning Analytics.

The KASSIS software has been extended to propose different dashboards about users(students) activity to teachers. To implement this extension we provide a client-server solution and collect information about the way people use it. It is important to mention that, we don't have to collect what people are writing or drawing but only their overall activity (which page they are looking at, the fact they wrote a note on some page...). The design of these dashboards was carried out by our partner, the LP3C laboratory, following a user-centered design (UCD), and the LS2N laboratory proposed different analyzes of traces of user activity, we enriched the information collected according to their need. The purpose of these dashboards is to follow the students 'work in real time during the activity (to help the teacher during his course by following the students' activities) or after (for a reflexive analysis of the educational strategy, measure student progress, to develop the course, see fig. 11 the retrospective dashboard).

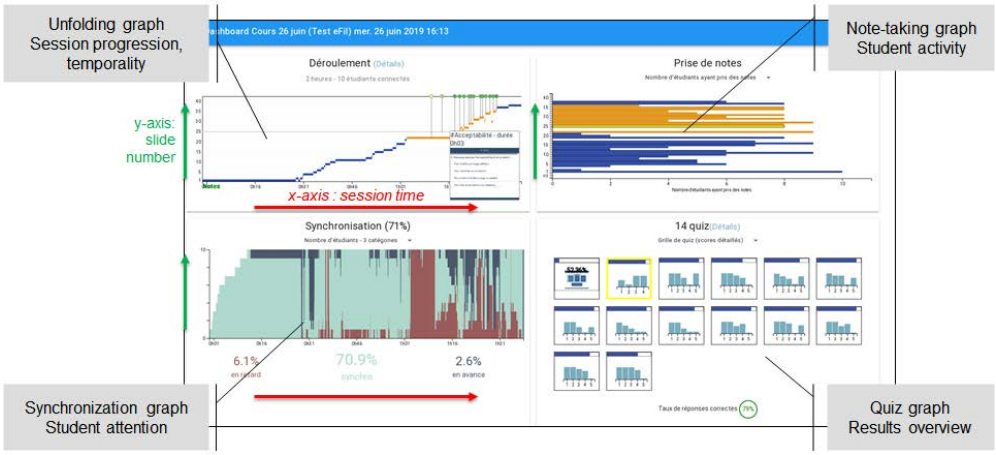


Figure 11: Kassis retrospective dashboard

The KASSIS solution offers the possibility to compose handwritten advanced graphical quizzes for teachers: questions can be expressed on-the fly, by sketching or drawing directly on the tablet display and responses are also handwritten by students. All the graphical responses are collected in real time to produce immediate feedback for students and teacher. Along with these trace data, we added the ability for the teacher to indicate the correct answer to quizzes in the software, which enables to display visual corrective feedback to the students and to compute quiz statistics in the dashboard (see fig. 12 and fig. 13).

The KASSIS solution has been also extended with the ability to make a real time aggregation, analysis and restitution of the graphical responses of the student by two ways (First experiments have been presented in [21]):

- for hand-drawn responses as emphasize, surround...: we aggregate the graphical responses in the form of a heat-map in which areas containing a majority of answers are highlighted (see the workflow in fig. 12).

- for hand-drawn responses as sketches, handwritten formula...: in conjunction with the eFran Project ACTIF (see section 4.2) the previous prototype that analyzes the student's answers which are freehand graphical drawing has been reinforced. Indeed, the clustering of the students answers has been improved by reduction of the feature space. We apply a Correlation-based Feature Subset Selection approach [Hal98] to reduce the feature space from 70 (see enriched HBF49 features [11]) to 50 (see the workflow in fig. 13). This selection allows the improvement of clustering in different class configurations (less than 30 students to more then 100). Students answers are automatically clustered by similarity and are projected in a 3D display, allowing the teacher to quickly see the different kind of answers in a large group (see fig. 14). This 3D graphical feedback has been integrated in KASSIS software to be evaluated in classroom by LP3C laboratory.

[Hal98] M. A. Hall, *Correlation-based Feature Subset Selection for Machine Learning*, PdD Thesis, Hamilton, New Zealand, 1998.

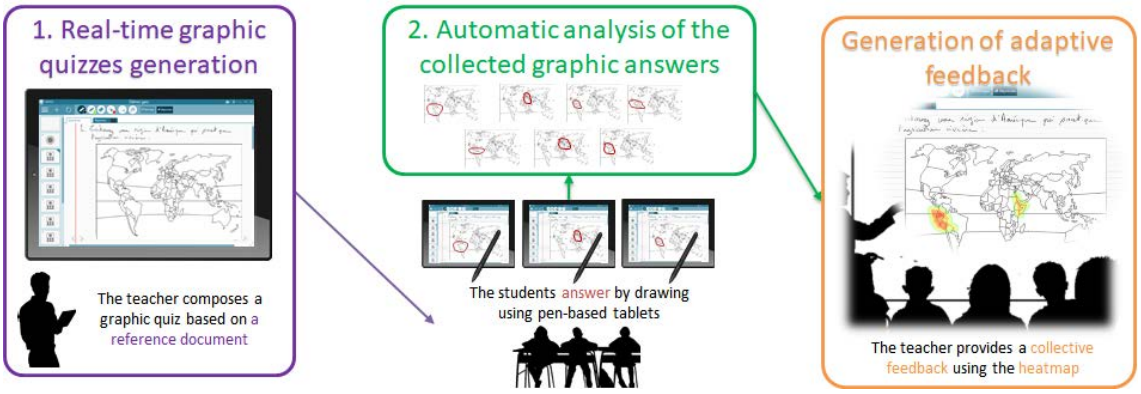


Figure 12: Kassis workflow for heat-map generation

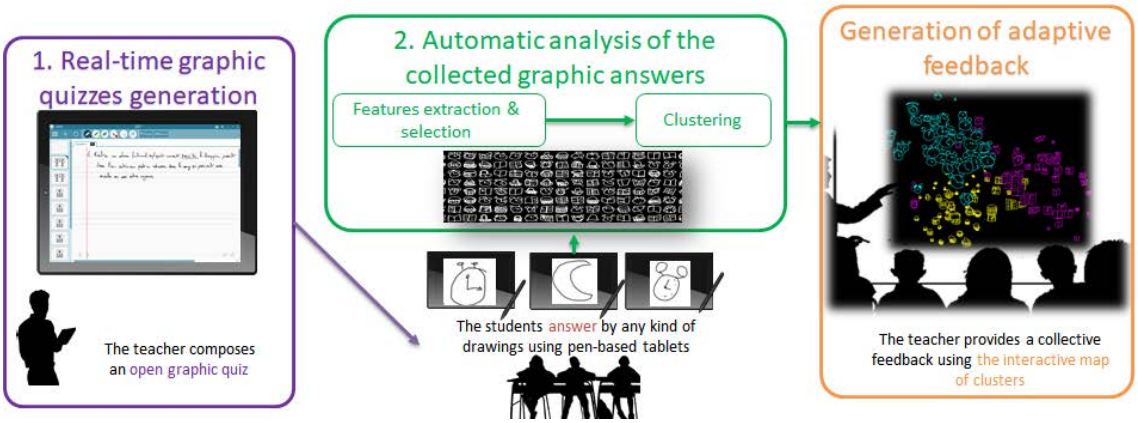


Figure 13: Kassis workflow for heat-map generation

4.6 Interactive combination of deep learning and syntactical methods for contextual segmentation and structure learning, without annotated data, in musical scores recognition

Participants: Kwon-Young Choi, Bertrand Coüasnon, Yann Ricquebourg, Jean Camillerapp.

This work is done in the context of a collaboration with Richard Zanibbi from the Rochester Institute of Technology (see 6.4.1). We propose a new small and fast deep learning detector adapted to the detection of music symbols in dense and noisy historical music scores. The use of a neural network detector allows us to merge the segmentation and classification task which performs better with broken and touching symbols 4.6.

The particularity of this detector is the seamless integration of a Spatial Transformer layer that allows to also use the localization information as an attention mechanism. Using this architecture, we obtained a mAP of 94.8% with an IoU threshold of 0.75. We compare with

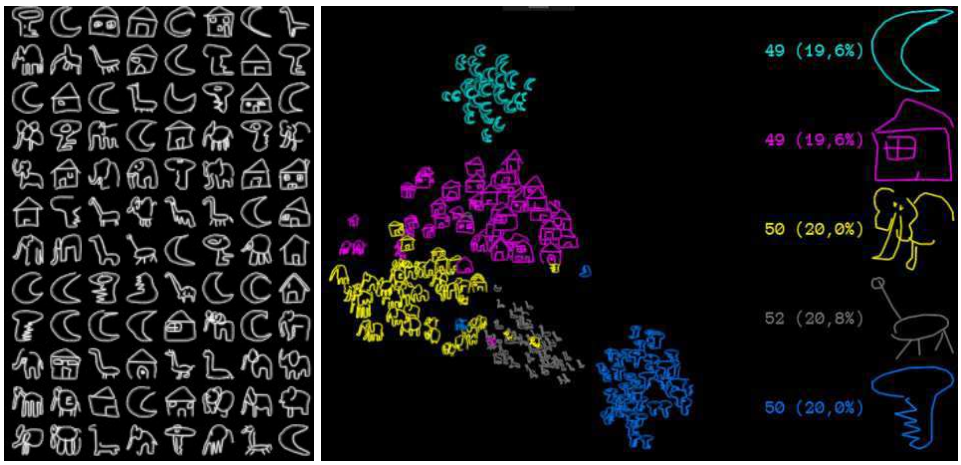


Figure 14: Raw data and Answer clustering

(a) Broken symbol

(b) Touching symbols

Figure 15: Examples of segmentation problems

state-of-the-art deep learning object detectors like Faster R-CNN, R-FCN and SSD producing respectively 98.3%, 98.7% and 97.8% mAP with an IoU threshold of 0.75. These results were published at ICDAR 2019 [14].

A new unsupervised training method for the detection of music symbols was developed using a combination of three components. First the DMOS-P syntactical method was used to focus the search music symbols into areas of high presence probability of symbols. Then, a Generative Adversarial Network (GAN) was used to transfer small patches of images coming from real images of music scores into a simpler representation, removing background noise while retaining symbol shapes and position. Finally, a simple SSD detector was used to detect remaining music symbols in the generated images of the GAN generator. The GAN training algorithm has been modified to introduce shapes information on symbols, in the training of the generator and the discriminator networks. This introduction is done with an already existing dataset of isolated music symbols.

Using this method on a small accidental detection dataset, we obtain a mAP of 94.2% with an IoU threshold of 0.75 while using no manually annotated detection groundtruth like bounding boxes annotations. Current work is done to replicate these results on a bigger dataset and on more classes of music symbols.

Finally, a lot of work has been done in the low-level graphical processing of the DMOS-P method in order to improve the recognition of staff-lines as well as simple music graphical primitives like stems, beams and noteheads.

4.7 Field localisation and recognition in administrative forms

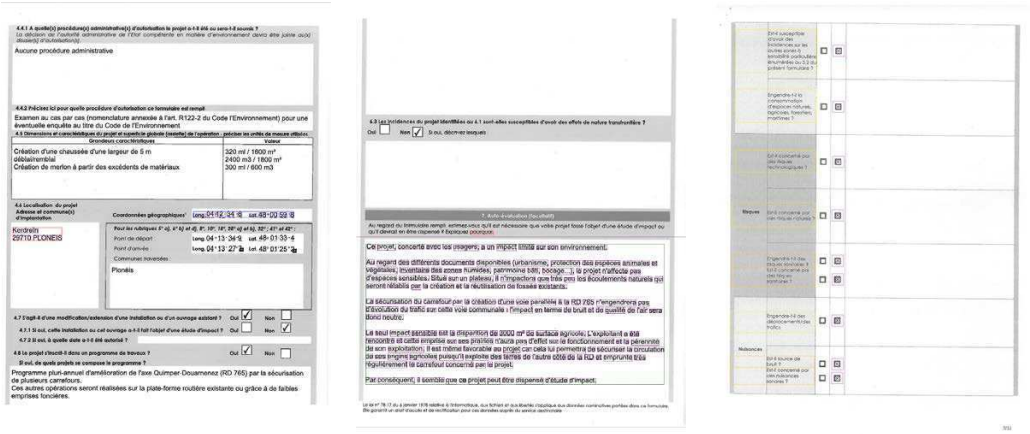
Participants: Solène Tarride, Bertrand Coüasnon, Aurélie Lemaitre.

This work is part of the IAT Project "Towards AI for local administration" (see 6.2.6) that aims to provide the territorial administration with a proof of concept for a decision-making tool.

The goal of this project is to lead a first experiment on a specific corpus: the tool is expected to learn from a corpus of impact studies in order to predict whether a new construction project would likely have a significant negative impact on its territory. The corpus is composed of highly compressed scans of printed documents (Figure 16).

The administration identified the fields that are the most relevant in their decision-making process. Our role is to localize and recognize these fields in order to build a training database. First, we focus on the "Localization" section of the form where we extract the "Address" and "GPS coordinates" fields. Then, we extract the "Self-evaluation" field where the project manager argues on why its project should be exempted of an environmental assessment. Finally, we analyze the "Environmental sensitivity" and "Potential impact of the project on the environment and human health" fields which are a series of questions that can be answered with "Yes" or "No" by ticking a box.

We used DMOS-method to develop a logical description of these documents in order to localize the fields of interest. Depending on the field, we used an OCR to recognize the text or we performed a pixel-level analysis to identify the boxes that are ticked. We have processed over 1000 documents using this approach.



(a) Analysis of "localisation" fields (b) Analysis of "self-evaluation" field (c) Analysis of check-boxes

Figure 16: Examples of processed pages

4.8 Analysis and recognition of historical parish registers

Participants: Solène Tarride, Bertrand Coüasnon, Aurélie Lemaitre, Jean Camillerapp.

This work is done in the context of a CIFRE PhD (n°2018/0896) in collaboration with

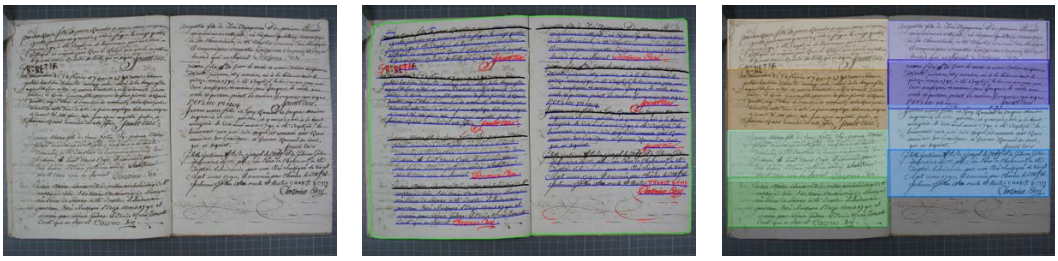
Doptim 6.3.1. We focus on the analysis of pages from French parish registers from 16th to 18th century. Parish registers are documents in which were recorded the acts of baptisms, marriages and burials. Each act contains valuable demographic information that can be useful to genealogists willing to find information about their ancestors. The first step toward parish register analysis is delimiting each act. However these documents are so poorly-structured that the visual separation between the acts is not always clearly visible. One of the main visual indication of separation is the signature of the priest at the end of each act (Figure 17(a)).

We have designed a strategy to delimit each act. First, we train fully convolutional neural networks to recognize four main objects : page borders, first text-lines of each act, other text-lines and signatures. The post-processing step includes the use of text lines detection in blurred images.

Then, we use DMOS-PI to define logical rules in order to design a description of the acts based on these objects. We have designed two approaches so far:

- a description based on borders, text-lines and signatures, where an act is defined as a list of text-lines followed by a signature. This work has been published in [20].
- a description based on border, the first text-lines of each act and the other text-lines, where an act is defined as a first text-line followed by other text-lines.

We have processed 200 images using these methods. We plan to improve the logical description by taking into account all detected objects (borders, first text-lines, other text-lines and signatures) as well as recurrent keywords.



(a) Original parish register

(b) Predicted borders (green), first text-lines (black), other text-lines (blue) and signatures (red)

(c) Predicted acts based on gramatical description

Figure 17: Illustration of the prediction process

4.9 Analysis and recognition of yearbooks

Participants: Simon Bouvier, Bertrand Coüasnon, Aurélie Lemaitre, Yann Ricquebourg, Nathalie Girard.

This work is part of the EURHISFIRM project (see 6.1.1). It aims to extract data from historical stock documents, and in particular company yearbooks of the 20th century. In 2019 we focused on the 1960 Desfossés Yearbook including two types of documents:

- Administrators lists: second half of the tome 1, each page is a table containing 3 columns with the name, address and roles of each administrator (figure 18).

- Issuing company pages: entirety of the tome 2, contains information on each company including rubrics such as counselors, assets etc. and tables such as a complete balance sheet of the past years (figure 19).

The DMOS-PI method is used to describe and analyze the structure of these documents. Text lines are extracted by a neural network (ARU-Net) and textual transcriptions are generated with the commercial OCR Abbyy FineReader 11.

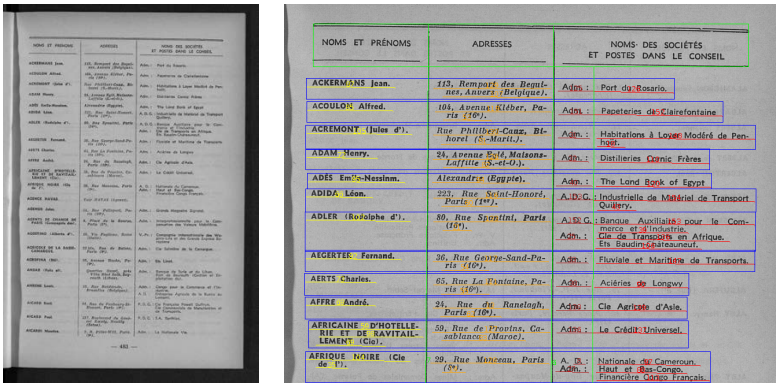


Figure 18: Example of administrators list extraction

The administrators lists (fig. 18) detection is guided by the table rulings to differentiate the 3 columns, and rows are delimited by blank spaces. The current output is formatted in JSON, where each entry corresponds to an administrator with the location and transcription of their information.

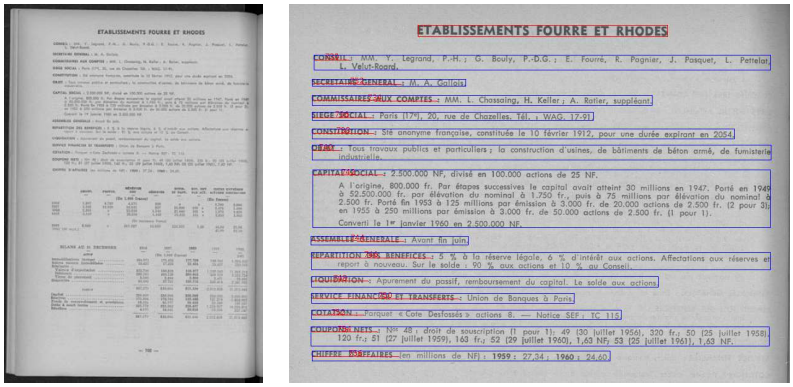


Figure 19: Example of issuer rubrics extraction

On the issuers pages (fig. 19), the information is split in rubrics. The structure detection is based on the page layout elements such as the alignment between rubrics and indentations.

Below the rubrics are various tables that either sum up the information contained in the rubrics, or add financial data, for example about previous years (fig. 20). These tables are not delimited with rulings so the columns detection relies on alignments. Other structural elements include key words and delimiting lines. In a first step, we focused on the balance sheet recognition. Next work will rely on a more generic description of the table in the grammar.

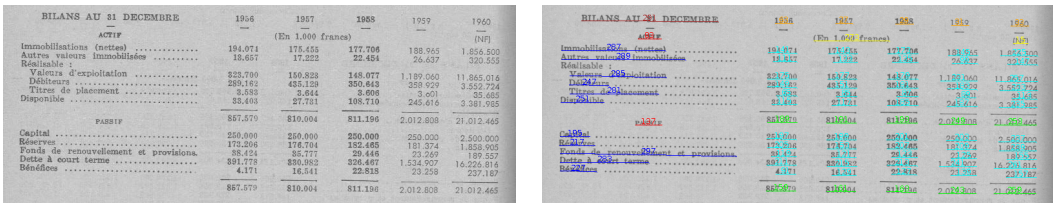


Figure 20: Example of balance sheet extraction

4.10 Analysis and recognition of price lists

Participants: Camille Guerry, Bertrand Coüasnon, Aurélie Lemaitre.

This work is done in the context of the ANR project HBDEX (see part 6.2.5). The aim is to model and test different strategies that automatically exploit the context of the collection to improve documents recognition. We work on a first strategy composed of different iterations. The aim of each iteration is to recognize and validate a structural element of the documents. An iteration consists of:

1. a first structural analysis of each document. This analysis relies on a combination of deep-learning and syntactical approach (see method described in [15]). The syntactical approach consists of a generic description of price lists documents. With this analysis we extract different information from the documents.

2. a validation phase based on knowledge coming from the context of the collection. Our system applies rules on information extracted at the previous step and validates or corrects them.

3. a second structural analysis. For this step, the recognition system exploits the information validated by step 2 to specify the grammatical description used in step 1.

In a first instance, we focus on daily price list coming from Paris stock exchange markets of the end of the 19th and begin of the 20th century. We will forward applied the same strategy on other kind of documents like companies' yearbooks (see part 6.1.1). Figure 21 presents some qualitative results of the improvement obtained thanks to the cross-validation for the first iteration: recognition of the columns. Table 3 shows some quantitative results for this iteration. For the cross-validation step we use the fact that the columns width and name are stable on slot of times. This same rule will be used for each "stable" elements of prices lists: columns, sections and stables fields.

Number of pages	Number of error without the contextual analysis	Number of error with the contextual analysis
311	14	0

Table 3: Results of the column extraction phase, before and after the use of contextual analysis

resume

Figure 21: Example of improvements obtain thanks to the global strategy exploiting contextual analysis

5 Software development

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).

5.1 RESIF: Handwriting recognition by hierarchical fuzzy inference systems

Contact: Eric Anquetil

Keywords: Handwriting Recognition, smartphone, fuzzy logic.

RESIF technology is composed of three main software to analyze, model and recognize handwritten characters and words:

- RESIFCar is specialized to recognize isolated handwritten characters: Latin alphabet, digits and special symbols.
- RESIFMot is the software for unconstrained cursive handwritten word recognition.
- RESIFApp is the automatic learning process that generates from a handwritten character database the hierarchical fuzzy models used by the recognition systems: RESIFCar and RESIFMot.

RESIFCar and RESIFApp are already in their fifth version. Through industrial collaborations, RESIFCar has been successfully integrated into mobile devices (smartphones) which are characterized by their limited computing and memory resources.

ResifCar has been integrated in the educational software *Toutaki* of Evodia/Script&Go Company. This Tablet PC software helps the young children to learn how to write. Toutaki has been licensed to HITACHI Company to be embedded to their Electronic Whiteboards.

Principles of Resif technology have inspired the conception of the new handwriting analyse software (DAP, ISA, ISF, IWA presented in section 5.2).

5.2 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 6.2.2 and 4.1), 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.

5.3 EVOLVE++ / EVOLVE TOUCH: Evolving recognition engine

Contact: Eric Anquetil

Keywords: Incremental recognition, Evolving system, Gestures and Symbols Recognition.

Evolve++ is an evolving recognition engine, that can be trained incrementally, starting from few data samples. Evolve++ is based on a fuzzy inference systems that learn incrementally and cope with class adding.

Evolve-Touch is a derived software based on Evolve++ for the application domain of graphical gesture recognition for multi-touch devices. *Evolve-Touch* offer a complete framework to allow user to manage and customize his gesture sets for different application contexts in simple and user-friendly manner. An intuitive mechanism is adopted to get user feedback on recognizer answers, which allows the latter to continuously enhance its performance. In 2014 we focused on bringing a qualitative evaluation of gestures. To demonstrate the main features of Evolve-Touch system, a showcase application is presented in this video : *http://youtu.be/qOx4IY6uYf8*.

This work is supported by a *European Regional Development Fund* (FEDER), and protected by a *European and US patents* (N° 2995704 / 14/429,649).

In 2014, EvolveTouch was made available for *Android*, *WinRT* and *iOS* tablet systems.

Since 2018, Evolve is the IA engine integrated in the educational software *Kaligo* of Learn&Go company. Evolve engine has also been integrate in 2018 in the "IntuiDiag" software of Innax company.

The research on Evolving system based on Evolve technology is going on today with the new Phd student Clément Leroy (see 4.3).

5.4 Varchitect: Windows Store application based on Evolve++/EvolveTouch

Contact: Eric Anquetil

Keywords: Incremental recognition, Evolving system, Gestures and Symbols Recognition, Interior design, Tablet, Windows Store.

Varchitect is a Windows Store application that was developped as part of the effort to port the *Evolve++/EvolveTouch* system to current tablet operating systems. It is available for free at http://apps.microsoft.com/windows/en-us/app/aa0889d0-2097-4a91-aa28-2a74df7e206c

With Varchitect, users can define their own set of gesture commands to insert furniture or architectural elements in a plan, and then design their interior with a stylus or fingers. The users can use a picture (taken from the tablet's builtin camera) as a base and set the scale of their plan to be sure everything fits. Plans made using Varchitect can be shared or printed.

This work is supported by a development fund from *SATT Ouest Valorisation*.

In 2015, Varchitect was downloaded more than 7000 times.

5.5 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 [9, 10]. *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 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 integrate in 2018 in the "IntuiDiag" software transferred to the Innax company. This transfer has been supported by a development

fund from *SATT Ouest Valorisation*.

The research on DALI framework is going on today with the new Phd student Omar Krichen who is working on online recognition of the user's hand-drawn sketches (see 4.2).

5.6 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 [5]. 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:

- ScoRead: a prototype for musical scores recognition;
- MathRead: a prototype for mathematical formulae recognition;
- TennisRead: a prototype for tennis court detection in videos;
- TabRead: a prototype for table structures recognition;
- FormuRead: a software for reading military forms of the 19th century recognition despite their deterioration. This software has been successfully tested on more than 480,000 pages of the *Archives de la Mayenne* and *Archives des Yvelines*;
- NatuRead: a software for recognition of naturalization decree registers from 1883 to 1930. This software has been applied on 85,088 pages of the *Centre Historique des Archives Nationales*;
- LettRead: a software for extracting structure of mail documents. It has been applied on 1150 images provided by the French project RIMES.
- BanglaRead: a software for extracting headline in Bangla script. This extraction represent a pre-processing tool for handwritting recognition. This work has been realized in collaboration with University of Kolkata and applied on 1922 words from 26 writers.
- FlowRead: a sofwtare for on-line flow-chart segmentation and structure recognition.
- JournRead: a software for the recognition of old newspaper content structured in headlines, articles with title and author, graphics and tables. This software has been developed with a *SATT Ouest Valorisation* development fund;

- MaurdorRead: a software for the structure recognition of heterogeneous and multi-language documents, with handwritten, printed or mixed content. This software has been developed in the context of a PEA (Programme d'Etude Amont - upstream study program) from the DGA (French Ministry of Defense).

5.7 Precoce: Library to extract visual cues

Jean Camillerapp

Keywords: Early vision, adaptative binarization, vectorization, Kalman filtering, squeletonization, multi-resolution.

It is the library on which the DocRead software described in the section 5.6 is based.

This library is composed of various methods of adaptive binarization. Those use either local determinations of threshold or edge detection and region growing cooperation. Some of these methods can provide results with a sub-pixel resolution.

Using Kalman filters, this library carries out the detection of rectilinear segments in complex images.

It also provides a skeletonization method from gray level images.

The selected resolution for the digitization of the documents is not always adapted at best to the recognition of all the structures. This is why the library also provides possibilities of multi-resolution processing.

This library has versions depending on the information available as input:

- gray-scale images,
- on-line stroke combined with the associated generated off-line image,
- native PDF file.

It also has more specific parts:

- detection of visual cues used in musical score recognition,
- detection of handwritten text lines using blurred images.

6 Contracts and collaborations

6.1 International Initiatives

6.1.1 EURHISFIRM: Historical high-quality company-level data for Europe

Participant: Bertrand Coüasnon, Aurélie Lemaitre, Nathalie Girard, Simon Bouvier

.

- Partners: Paris School of Economics (PSE)(Coordinator), Universiteit Antwerpen, Johann Wolfgang Goethe Universitat Frankfurt am Main, Erasmus Universiteit Rotterdam,

Uniwersytet Ekonomiczny we Wroclawiu, The Queen's University of Belfast, Koninklijke Nederlandse Akademie van Wetenschappen – Knaw, Universidad Carlos III de Madrid, Université de Rouen Normandie, Institut National des Sciences Appliquées de Rennes, Gesis Leibniz-Institut Fur Sozialwissenschaften

- 36 months (2018-2020)
- Contract: INSA, H2020 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.

EURHISFIRM enables researchers, policymakers, and other stakeholders to develop and evaluate effective strategies to promote investment, economic growth and job creation. The RI provides the tools for long-term analysis highlighting the dynamics of the past and the way those dynamics structure our present and future.

EURHISFIRM develops innovative models and technologies to spark a "Big data" revolution in historical social sciences and valorize Europe's cultural heritage.

In this project we work on building 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. We first started to work on the structure recognition of companies yearbooks.

6.2 National Initiatives

6.2.1 P2IA: Project of Ministry of Education (DNE): Project of Innovation - Intelligence Artificial (P2IA) for e-eduction

Participant: Eric Anquetil, Nathalie Girard, Pauline Nerdeux, Damien Simonnet, Simon Corbillé.

- Partner: *Company Learn&Go*, *LP3C*, *Loustic*, *INSPE*, *IRISA-Expression*, *Academies of Rennes and Caen*
- 18 + 24 months (2019-2021).
- Contract: INSA

This project was launched by the Ministry of Education (DNE) as part of the project of Innovation - Intelligence Artificial (P2IA). It is a project funded on several instalments; the first one (12+6 months) is founded with an amount of 240K€ for the IntuiDoc team. The project is led by the Learn&Go company. The objective is to design a software environment for helping to learn French (writing / spelling) based on artificial intelligence for teachers and pupils 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.

6.2.2 Script&Labs/ANR joint Laboratory (LabCom): On-line handwriting and drawing recognition and interpretation for active Learning in e-education

Participant: Eric Anquetil, Nathalie Girard, Damien Simonnet, Pauline Nerdeux, Sébastien Thomas, Morgane Carry.

- Partner: *Société Learn&Go*
- 36+12 months (2017-2021).
- Contract: INSA
- Website: `https://scriptandlabs.irisa.fr/`

With the great success of the IntuiScript Project (BPI/PIA-2) (`https://www-intuidoc.irisa.fr/projet-intuiscript/`), we want to consolidate our implication in e-education to achieve innovative contribution based on pen-based tablet devices. In that way, we build a ANR joint laboratory (LabCom) between IntuiDoc and Learn&Go Company on this promising topic for a total amount of 1 562K€(6.15 ETP) with 300K€of grant.

This new structured partnership would have the goal to design new educational learning experience at school by using tablet and tactile digital devices: handwriting learning, generative drawing/sketching, on-line handwritten arithmetic expression, collaborative interaction using 2D gestures, document annotation...

The ScriptAndLabs Laboratory has been selected (from the 122 national LabCom) for an oral presentation of its activities for the "National LabCom days" in Paris in September 2018.

Following this presentation, we have been invited by the ANR President and CEO to present the ScriptAndLabs LabCom at the administration council of ANR the November 22, 2018.

The ScriptAndLabs Laboratory has presented its first research results by an inauguration on October 18, 2018 at the "Jacobin Convent" Palais des congrès in Rennes, as part of the "Learning Show" event.

Current activities are described in section 4.1.

6.2.3 Actif: Active Learning and Collaboration with Tablet Computer, Interactions and Feedbacks (e-Fran - National Innovative BPI Project)

Participant: Eric Anquetil, Nathalie Girard, Simon Corbillé, Omar Krichen, Morgane Carry.

- Partners: *LP3C*, *Loustic*, *Société Learn&Go*, *Saooti*, *Région Bretagne*, *Espe*, *espace des sciences*
- 48 months (2017-2020).
- Contract: INSA

The project "ACTIF" is one of the 22 selected national project from the "e-fran / innovative national project " call. This is a 4 years project (2017-2020) for a total amount of 1 143 856€, led

by the LP3C of the University of Rennes 2. It carries on: "Active Learning and Collaboration with Tablet Computer, Interactions and Feedbacks".

One of the key topics in this project is to investigate how to encourage new learning strategy based on "active learning" and especially on "generative drawing" using pen-based numerical devices. The concept is to ask students to create drawings while reading text for causing generative processing that leads to better learning outcomes. The goal is to improve the learning of student by considering "learning" as a generative activity especially with collaborative/cooperative interaction. In the scope of digital learning, the potential induces by pen-based devices (tablet, TNI and new interactive screen) is really interesting. The goal is to investigate how we can automatically generate intelligent "corrective" or "predictive" individual and collective feedbacks to users by interpreting their drawing process: summarizing, mapping, sketching...

The IntuiDoc team will ensure all the design related to the analysis of graphic productions with one engineer, one postdoctoral researcher and one PhD student. They will work on the issues of "Generative Drawing" in the field of Geometry at secondary school.

We have presented the first research results of the e-Fran projects on January 30, 2019 in Rennes in the presence of Mrs Le Pellec Muller, Rector of the Brittany academic region.

Current activities are described in section 4.2.

6.2.4 eFil: e-Feedback for Interactive Lecture - Labex CominLabs project

Participant: Eric Anquetil, Nathalie Girard, Sébastien Thomas, Mickael Renault.

- Partner: *LP3C (University of Rennes 2)*, *LS2N (University of Nantes)*
- 24 months (2017-2019).
- Contract: INSA

The e-Fil Cominlabs project is based on CAMIA/KASSIS which is a new digital learning environment (workbook) for active learning with pen-based tablet. The goal of the eFIL Project is to optimize and to evaluate this digital learning environment in classroom settings during lectures (Optimizing with user-centered design). Traces will be collected and analyzed both to explore learning dynamics in classroom, and to provide a dashboard allowing instructors to monitor their activities during lectures as well as in-between sessions. The goal is to explore and analyze traces to monitor learning activities in classroom with Educational Data Mining and Learning Analytics.

Current activities are described in section 4.5.

6.2.5 HBDEX : Exploitation of Historical Big Data for Digital Humanities

Participant: Bertrand Coüasnon, Aurélie Lemaitre, Camille Guerry.

- Partners: Paris School of Economics (PSE), Université de Rouen Normandie (LITIS), CAMS-EHESS
- 48 months (2017-2021)

- Contract: INSA, ANR

This project focuses on massive historical data extraction for digital humanities, applied on 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.

Current activities are described in section 4.10.

6.2.6 IAT: Artificial Intelligence for Administration

Participant: Bertrand Coüasnon, Aurélie Lemaitre, Solène Tarride.

- Partners: Ministère de l'intérieur
- 2 months (12/2018-01/2019)
- Contract: Université de Rennes 1

This contract, founded by the Ministère de l'intérieur and leaded by the Druid research team of Irisa, is focused on the application of Artificial Intelligence to administration. We work on the automatic extraction of information in digitized administrative forms.

Results are described in section 4.7.

6.3 Bilateral industry grants

6.3.1 Research contract Doptim company

Participant: Aurélie LemaitreBertrand Couasnon.

- Partners: *Doptim company*
- Since 2019
- Contract: INSA

Intuidoc team started to work with Doptim company in the field of document recognition applied to parish registers. This collaboration is based on a CIFRE grant.

Current activities are described in section 4.8.

6.3.2 Industrial software licensing with Learn&GO company

Participant: Eric Anquetil.

- Partners: *Learn&GO company*

- Since 2017
- 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".

Current activities are described in section 4.1.

6.4 Collaborations

6.4.1 Rochester Institute of Technology, USA

Participants: Kwon-Young Choi, Bertrand Coüasnon, Yann Ricquebourg, Richard Zanibbi (RIT).

- Partner: Rochester Institute of Technology (Rochester, NY, USA).
- Since 2016

In collaboration with Richard Zanibbi from the Rochester Institute of Technology (RIT), Rochester, NY, USA, we work on interactive learning by combining deep learning technics, syntactical analysis and user interaction to introduce learning of segmentation. 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. This collaboration is done through the co-supervising (Bertrand Coüasnon, Yann Ricquebourg and Richard Zanibbi) of the PhD of Kwon-Young Choi (see section 4.6).

7 Dissemination

7.1 Promoting scientific activities

7.1.1 Scientific Events Selection

- A. Lemaitre is member of the program committee of the International Conference of Document Analysis and Recognition (ICDAR 2019)
- A. Lemaitre is member of the program committee of the 2nd International Workshop on Open Services and Tools for Document Analysis (ICDAR-OST 2019)
- E. Anquetil is member of the program committee of the International Conference of Document Analysis and Recognition (ICDAR 2019).

- E. Anquetil is member of the program committee of LNCS post-proceedings of the paper of the International Workshop on Graphics Recognition and Graphical Document Analysis (GREC 2019).
- B. Coüasnon is member of the program committee of the International Conference of Document Analysis and Recognition (ICDAR 2019)
- B. Coüasnon is member of the program committee of the 5th International Workshop on Historical Document Imaging and Processing (HIP 2019).
- B. Coüasnon is member of the program committee of the 2nd International Workshop on Open Services and Tools for Document Analysis (ICDAR-OST 2019).
- B. Coüasnon is member of the program committee of the Workshop on Industrial Applications of Document Analysis and Recognition (2019).

7.1.2 Journal

- E. Anquetil is a reviewer in 2019 of: Pattern Recognition Letters (PRL) international journal.

7.1.3 Invited Talks

- Nov. 2019: E. Anquetil presented the "e-Fil" Project at the "CominLabs Days" (Labex) day, Rennes.
- Oct. 2019: E. Anquetil presented the eFran "ACTIVE" project: Learning and Collaboration on Tablets, Interactions and Feedback ", at the national conference of ministerial projects in Paris.
- Sept. 2019: E. Anquetil gave an invited talk at the conf'Lunch conference at the IRISA laboratory: "How AI and can help improve learning in school, college and higher " (https://videos-rennes.inria.fr/video/Hk8K8qJDS/), Rennes
- June 2019: E. Anquetil gave an invited talk at the Conference of the "Grandes Ecoles" (CGE) "human and social sciences in "Grandes Ecoles" in Paris: the title of the talk was: "Digital technology and AI at the service of education".

7.1.4 Scientific Expertise

- E. Anquetil was a reviewer in 2019 of one project for the French National Research Agency (ANR).

7.1.5 Research Administration

- E. Anquetil is a member 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, B. Coüasnon, J. Camillerapp and A. Lemaitre, take part in the activities of the society GRCE : " Groupe de Recherche en Communication Écrite ".
- E. Anquetil is a member of the steering committee of LOUSTIC laboratory of Rennes (laboratoire d'observation des usages des technologies de l'information et de la communication).
- 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. He is in charge of the construction of the student incubator project in Rennes "Station B" which brings together 11 higher education establishments in Rennes.
- E. Anquetil is an elected member of the 27e section of the CNU council of INSA.
- E. Anquetil is an elected member of the administration council of INSA.
- B. Coüasnon is member of the board of Valconum (Centre Européen de Valorisation Numérique).
- B. Coüasnon is an elected member of the laboratory council of the INSA component of IRISA.
- B. Coüasnon is scientific head of the Media and Interactions Department of IRISA.
- B. Coüasnon is member of the scientific board of Irisa.

7.2 Teaching, supervision

7.2.1 Teaching

The team is mainly made up of teachers who are very implied in activities of teaching. But a majority of lectures are not rattached 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 Master-Research *d'informatique* of University of Rennes 1.
- 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 Master-Research *d'informatique* of University of Rennes 1.
- 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 in charge of the module on professionalization adapted to research (PROF) of the Master-Research *d'informatique* of University of Rennes 1.
- 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.

7.2.2 Supervision

- 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, Uni. Rennes 1, started October 2019.
- PhD in progress: K.Y. Choi, Interactive combination of deep learning and syntactical methods for contextual segmentation and structure learning in document recognition, B. Coüasnon, Y. Ricquebourg, R. Zanibbi (RIT, Rochester, USA), INSA de Rennes, started October 2016.
- PhD in progress: O. Krichen, Real-time interpretation of geometrical sketches production for e-feedback generation in digital learning, E. Anquetil, N. Girard, INSA Rennes, started September 2017.
- PhD in progress: Camille Guerry, Historical big data: modelization of strategies to analyse collections of documents, B. Coüasnon, A. Lemaitre, S. Adam (Univ Rouen), INSA de Rennes, started October 2018.
- PhD in progress: Solene Tarride, Combination of logical and textual knowledge for recognition of ancient register images, B. Coüasnon, A. Lemaitre, S. Tardivel, INSA de Rennes, started February 2019.
- PhD in progress: C. Leroy, Incremental learning and evolving recognition system: application to on-the-fly recognition of handwritten gestures, E. Anquetil, N. Girard, INSA Rennes, started October 2018.
- 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.

7.2.3 Juries

- B. Coüasnon was a reviewer of the PhD of Riyadh Benammar, Détection non-supervisée de motifs dans des partitions musicales manuscrites, Université de Lyon (operated at INSA Lyon), November 2019.

7.3 Popularization

The ScriptAndLabs LabCom activities (see 6.2.2) have been featured in several news bulletin:

- Le Monde: (02/09/2019) : À Rennes, tablette et stylet à l'école maternelle.

  `https://www.lemonde.fr/mperso/article/2019/09/02/arennestabletteetstyletalecolematernelle_5505334_4497916.html`

- Radio France Inter: (23/01/19) 7h15: Comment l'intelligence artificielle aide l'apprentissage de l'écriture et de la lecture.

  `https://www.franceinter.fr/emissions/le-zoom-de-la-redaction/le-zoom-de-la-redaction-23-janvier-2019`

7.4 Patent and Deposit of digital creations (APP)

7.4.1 Patents

- E. Anquetil, A. Almaksour and G. Richard obtain in february 2019 the grant of the US Patent on "Evolve Touch" technology - N° 2995704 / 14/429,649 : "Méthode de sélection de mode d'interactivité" (see section 5.3)

7.4.2 Deposit of Digital creations (APP)

- E. Anquetil, D. Simmonet and M. Renault deposited a V2 of the digital creation, ISA - Isolated Symbol Analysis.(ISA V2 IDDN.FR.001.060006.000.S.A.2020.000.31235)
- E. Anquetil, D. Simmonet and M. Renault deposited a V2 of the digital creation, IWA – Isolated Word Analysis.(IWA V2 IDDN.FR.001.060005.S.A.2020.000.31235)

8 Bibliography

Major publications by the team in recent years

[1] A. Almaksour, E. Anquetil, "Improving premise structure in evolving Takagi-Sugeno neuro-fuzzy classifiers", *Evolving Systems 2*, 2011, p. 25–33, 10.1007/s12530-011-9027-0, `http://dx.doi.org/10.1007/s12530-011-9027-0`.

[2] A. Almaksour, E. Anquetil, "ILClass: Error-Driven Antecedent Learning For Evolving Takagi-Sugeno Classification Systems", *Applied Soft Computing*, 0, 2013, `http://www.sciencedirect.com/science/article/pii/S1568494613003414`.

[3] M. Bouillon, E. Anquetil, "Online active supervision of an evolving classifier for customized-gesture-command learning", *Neurocomputing 262*, November 2017, p. 77 – 89, `https://hal.inria.fr/hal-01575805`.

[4] B. Coüasnon, J. Camillerapp, I. Leplumey, "Access by Content to Handwritten Archive Documents: Generic Document Recognition Method and Platform for Annotations", *International Journal on Document Analysis and Recognition, IJDAR 9*, 2-4, 2007, p. 223–242, `http://springerlink.com/content/5843461264501u81/`.

[5] 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.

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

[7] A. Ghorbel, A. Lemaitre, E. Anquetil, S. Fleury, E. Jamet, "Interactive interpretation of structured documents: Application to the recognition of handwritten architectural plans", *Pattern Recognition 48*, 8, 2015, p. 2446 – 2458, `http://www.sciencedirect.com/science/article/pii/S0031320315000473`.

[8] A. Lemaitre, J. Camillerapp, B. Coüasnon, "Multiresolution Cooperation Improves Document Structure Recognition", *International Journal on Document Analysis and Recognition (IJDAR) 11*, 2, November 2008, p. 97–109.

[9] 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.

[10] S. Macé, E. Anquetil, "Exploiting on-the-fly interpretation to design technical documents in a mobile context", *Journal on Multimodal User Interfaces 4*, 2011, p. 129–145.

[11] 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.

Articles in referred journals and book chapters

[12] O. Krichen, E. Anquetil, N. Girard, M. Renault, "Online analysis of hand-drawn strokes for Geometry learning", *in : Frontiers in Pattern Recognition and Artificial Intelligence*, M. Blom, N. Nobile, and C. Y. Suen (editors), *Language Processing, Pattern Recognition, and Intelligent Systems*, *5*, World Scientific, 2019, p. 129–149, https://hal.archives-ouvertes.fr/hal-02162405.

[13] D. Simonnet, N. Girard, E. Anquetil, M. Renault, S. Thomas, "Evaluation of Children Cursive Handwritten Words for e-Education", *Pattern Recognition Letters 121*, 2019, p. 133–139, https://hal.archives-ouvertes.fr/hal-01863235.

Publications in Conferences and Workshops

[14] K.-Y. Choi, B. B. Couasnon, Y. Ricquebourg, R. Zanibbi, "CNN-Based Accidental Detection in Dense Printed Piano Scores", *in : 15th International Conference on Document Analysis and Recognition*, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02430041.

[15] C. Guerry, B. B. Coüasnon, A. Lemaitre, "Combination of deep learning and syntactical approaches for the interpretation of interactions between text-lines and tabular structures in handwritten documents", *in : 15th International Conference on Document Analysis and Recognition (ICDAR)*, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02303293.

[16] C. Leroy, E. Anquetil, N. Girard, "ParaFIS:A new online fuzzy inference system based on parallel drift anticipation", *in : FUZZ-IEEE*, New Orleans, United States, June 2019. https://arxiv.org/abs/1907.09285, https://hal.archives-ouvertes.fr/hal-02183142.

[17] A. Lods, E. Anquetil, S. Macé, "Fuzzy Visibility Graph for Structural Analysis of Online Handwritten Mathematical Expressions", *in : 15th IAPR International Conference on Document Analysis and Recognition (ICDAR2019)*, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02281462.

[18] A. Lods, E. Anquetil, S. Macé, "Graph matching over Hypothesis Graphs for the Analysis of Handwritten Arithmetic Operations", *in : 13th IAPR International Workshop on Graphics Recognition (GREC 2019)*, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02281511.

[19] M. Mehri, P. Héroux, R. Mullot, J.-P. Moreux, B. B. Couasnon, B. Barrett, "ICDAR2019 Competition on Historical Book Analysis -HBA2019", *in : 15th International Conference on Document Analysis and Recognition*, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02490897.

[20] S. Tarride, A. Lemaitre, B. B. Couasnon, S. Tardivel, "Signature detection as a way to recognise historical parish register structure", *in : HIP 2019*, ACM Press, p. 54–59, Sydney, Australia, September 2019, https://hal.archives-ouvertes.fr/hal-02433964.

Miscellaneous

[21] N. Michinov, E. Anquetil, E. Michinov, "Réduire l'effet de conformité au groupe pour améliorer l'apprentissage dans l'enseignement par les pairs", Journées thématiques de l'ADRIPS : éducation et psychologie sociale, September 2019, Poster, https://halshs.archives-ouvertes.fr/halshs-02293720.