UMR IRISA

Activity Report 2024

Team Shadoc

Systems for Hybrid Analysis of DOCuments

D6 – Signal, Image, Language

Université de Rennes | ENS 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

Bertrand Coüasnon, Associate Professor HDR, Insa, head of the team
Eric Anquetil, Professor, Insa
Pierre Beust, Professor, Univ. de Rennes
Denis Coquenet, Associate Professor, Univ. de Rennes
Nathalie Girard, Associate Professor, Univ. de Rennes
Aurélie Lemaitre, Professor, Univ. Rennes 2
Ivan Leplumey, Associate Professor, Insa
Yann Ricquebourg, Associate Professor, Insa
Yann Soullard, Associate Professor, Univ. Rennes 2

Associate members

Jean Camillerapp, Retired Professor, Volunteer Researcher, Univ. de Rennes

Research engineers, technical staff

Hugo Hazard, Insa Research Engineer, until February 29th
Bruno Hortollary, Insa Research Engineer
Ali Yesilkanat, Insa Research Engineer
Sarah Bucquet, Insa Research Engineer
Erell Choulette, Insa Research Enginee

PhD students

Islam Barchouch, Insa PhD student
Florent Imbert, Insa PhD student, until 31st December 2024
Florent Meyer, CIFRE/Insa PhD Student

PostDoc

Killian Barrere, Assistant Professor (ATER), ENS, until 31st August 2024
Omar Krichen, Insa Post Doc, until 29th March 2024

Administrative assistant

Stéphanie Gosselin-Lemaile, until 31st August 2024
Gwenaelle Lannec, from September 1st 2024

2 Overall objectives

2.1 Overview

The Shadoc team focuses on *modelling man-made data for written communication*: handwriting, gesture (2D and 3D), and documents, under various aspects: analysis, recognition, composition, interpretation.

The objective is to achieve a continuum between paper and digital documents with a certain readability. We mainly focus on the following topics:

- Intelligent recognition of handwritten content: documents, writings, gestures;
- Analysis of the semantic/structural content: document structure, stages of production of diagrams, drawings, musical scores, sketches, architectural plans;
- Design of new AI, combining recognition and analysis: offer enriched experiences for digital humanities or e-education.

The roadmap of the Shadoc team is on the frontier of several research axes: Pattern Recognition, Machine Learning, Artificial Intelligence, Human-Machine Interaction, Uses and Digital Learning.

Our research is characterized by the hybridization of several AI approaches: two-dimensional grammars, deep learning, fuzzy inference systems... This hybridization aims at guaranteeing, beyond performance, important aspects such as: explicability, genericity, adaptability, data frugality.

Beyond hybridization, the originality of this research is to focus on user interaction. This strategy aims at answering the limits of the current approaches which are based on non-interactive treatments. The concept is to reinforce the decision processes by relying on the implicit validations or explicit corrections of a user to avoid the propagation of errors throughout the analysis. The notions of interpretation, adaptation and incremental learning are at the heart of this research, the objective being to design efficient, robust and self-evolving systems.

The studied data take two main aspects: image recognition and analysis of sequences (time series) in different forms, from sensor signals to document collections.

Image recognition The first field of interest is image recognition of documents. Nowadays, some commercial OCR (Optical Character Recognition) systems are available for automatic document recognition. However, those systems present their limits for the recognition of ancient, handwritten or heterogeneous documents. We work on scanned images of historical and recent documents with complex structures. We also consider digital native documents, such as PDFs, the structure of which is not always directly interpretable.

Analysis of Sequences / Time series The team works on time series and information sequences in the field of analysis, interpretation and recognition according to several granularities and modalities.

We consider first of all low-level time series associated with trajectories formed by handwritten traces or 2D/3D gestures. They come from different types of sensors: inertial, Pen-based and (Multi-)Touch Capture on touch screen, Motion capture, Kinect or Leap Motion sensor. The objective here is the reconstruction, analysis, synthesis or interpretation of these time series, like for on-line handwritten scripts recognition [R9, R13]. Handwritten text recognition in document images are also processed sequentially and considered as time series [R3][2].

At a higher level, time series are studied to provide context (temporal, spatial and semantic) and to develop evolutionary or incremental analysis and learning approaches. The objective is for instance to detect concept changes in a data stream (a sequence of documents, a sequence of gestures, or more generally a sequence of actions) in order to adapt recognition models to concept drift [R11, R6]. Another concept is for example to design an on-the-fly analysis of a document composition (stroke by stroke) [R7].

We can also consider many sequences in collections of documents. Thus, with historical degraded documents, it is sometimes possible to improve the recognition using other pages of the document, when some information is repeated on different pages of the collection. We proposed to work in an original way by automatically transforming the different unit data (like text fields, titles, column widths...) found on the pages of a collection of documents, into different sequences of these unit data. These sequences are then analyzed for stabilities and breaks, in order to use the context of a collection of documents to improve the recognition quality [R1].

2.2 Scientific foundations

2.2.1 Combination / Hybridization

In the field of document recognition, recent approaches based on deep learning have shown results that outperforms the state of the art. However, those approaches present two main limitations: first, they require a large amount of labeled data for training; second, the trained systems can be seen as black boxes, and the results are often difficult to interpret and correct.

On another hand, the previous Intuidoc team has been working for a long time on the development of two-dimensional grammars that enables a physical, syntactic, and semantic description of the contents. The interest of these syntactical approaches is that they do not require labeled data for training.

The originality of Shadoc team is to propose a combination between deep learning based systems and syntactical ones. We study different implementations of combination:

- The syntactical part brings contextual information to generative neural networks to make them able to converge [R4];
- Some low level elements can be extracted using deep learning systems: text-lines, simple gestures, symbols... They are then combined using two dimensional grammars. This type of combination builds hybrid systems with greater generalization capabilities than neural-only systems, while requiring a smaller amount of annotated data [R10];

- Combination of document structure recognition and handwriting recognition;
- Combination of syntactical language models with transformers neural networks [R15];
- Combination of handwriting recognition with explicit segmentation with Seq2Seq recognition [R8].
- Strong combination of two dimensional grammars and transformers, where syntactical rules drives the transformer architecture.

This exploration of different mechanisms of combination between syntactic and neural models allows to reduce as much as possible the expression of a priori knowledge in syntactic form on elements that are difficult to learn for deep neural networks (or at the cost of very large amounts of annotated data), while taking advantage of the modeling capabilities of deep learning on elements that require less annotated data. This is a way to simplify the adaptation of a system to a new corpus, while increasing its generalization capabilities. Another interest of using combined approaches is to keep the systems interpretable. We can also formalize how the user interaction and the recognition system combine to keep the human in the loop.

2.2.2 Learning with few data

Deep learning methods become state-of-the-art approaches for many tasks. This is the case in the field of the Shadoc team for online and offline handwriting recognition and document image analysis. As discussed before, such methods are widely explored in many of ours works. However, such models require a lot of training examples to perform well.

Learning with few data is a regular limitation in our applications. On the one hand, works of the team are done with humans. Thus, data have to be acquired with users, which limits the amount of data that can be acquired. In particular, recently, several projects have been done in the team for students and doing data acquisition in schools is not easy for various reasons. On the other hand, other works of the team are focused on historical documents such as register, journals, books ... Having labeled examples related to the document is difficult as it may be hard to annotate examples, even if the user is an expert of the domain. This may be due to the old language, the handwriting style, or degraded documents. Thus, one has to deal with only a limited amount of labeled examples.

Various approaches can be investigated to overcome this limitation. One way is to design network architectures which build a relevant latent representation of data, even if it is trained on a small training set [R5]. Another way is to design a semi-supervised approach. These approaches allow to benefit of large set of unsupervised data when only a small amount of labeled examples is available. The users can be involved in the labeling process through a semi-automatic approach, called active learning, for which a model selects data examples of interest which will be manually labeled by the user [R11].

Those approaches can be combined with syntactical methods that do not require label data. The syntactical methods can be used to model the need of interaction when content recognition requires the intervention of an expert. They can also give the contextual information needed by generative neural networks (like IsolatingGAN, see section 2.2.3) to automatically generate labeled data of symbols [R4].

2.2.3 Self-adaptive systems

Building self-adaptive systems which can automatically adapt themselves to a new corpus of document without any or with only few labeled data is a challenging objective. It can be reached by combining syntactic and unsupervised deep learning methods. We propose to first work on a self-adaptive system for Optical Music Recognition (OMR) capable of improving its performance on degraded old scores. This method will be built on the IsolatingGAN [R4] proposed in a previous PhD work of the team where the GAN generator is able to generate labeled data of musical symbols on real images using only unlabeled musical scores and examples of isolated symbols. With this data, the system will be able to adapt to the unlabeled corpus by successive unsupervised learning, producing annotated data with the IsolatingGAN. These automatically annotated data can then be used to adapt the musical symbol detectors.

The driving of these auto-adaptation mechanisms to a corpus is possible by using the ISICA method (Interactive Strategy for Iterative Collection Analysis), validated on the European project EurHisFirm and the HBDEX ANR project, on cross-validation mechanisms on a collection of documents, applied to stock exchange quotation lists of the 19^{th} and 20^{th} century [R1]. Thus, at each iteration, a new set of automatically produced annotated data will be built for a subset of musical symbols. This dataset will be used to learn a new detector for this subset of musical symbols. Then this detector will be integrated into the parsing in the next iteration, thus producing the necessary data for a new subset of musical symbols that would not have been accessible in the previous iteration. This progressive iterative process will stop when all classes of musical symbols have been covered. This approach allows self-adaptation on symbols, detected by deep learning.

More and more data are being produced continuously. In order to analyze this data, it is necessary to integrate it continuously, which is often referred to as learning on data streams. The problem is that often the environment can be non-stationary, resulting in concept drifts, or the data stream is potentially infinite, which requires the system not to save the data. We explore different approaches of incremental learning based on evolutionary fuzzy inference systems that have the ability to develop both generative and discriminative modeling. This work will be applied to continuous gesture recognition allowing the user to evolve his gesture set on the fly [R11].

In our work, including those oriented to help learning writing and geometry, the production of feedback is an essential element [R12]. To be relevant, these feedbacks must be personalized, in fact the system must adapt to the current user. In the context of work on learning aid tools, the modeling of the process of solving a problem by knowledge graphs seems to us to be an avenue to explore in order to define new self-adaptive models.

2.2.4 Rejection capabilities

The construction of recognition systems with rejection capabilities is important both for the integration of these systems in interactive processes, with humans or other systems, but also to be able to exploit automatically generated annotations, and integrate them in semi-supervision processes. Indeed it is important to select through rejection, when a human expert will be solicited to answer questions in an interactive system. We will also study rejection capacities of deep neural networks to be able to select unsupervised annotated data to be used as new training data.

For example we will work on rejection in the CollabScore project for building a self-adaptive OMR system and in the ANTAI project on license plate recognition. Rejection capabilities of deep neural networks are also important to build hybrid systems to make decisions at the interface between the syntactic and deep part. Rejection is also necessary for hybrid systems to explain their decisions.

2.3 Application domains

The application contexts are very numerous, which is important to access real and large datasets, with real applications which lead to strong scientific challenges, while studying generic solutions.

Among the different types of documents studied by the team, there are ancient documents. Thus, there are many possible applications with archive services and digital humanities. Depending on the field of application, the analysis can be made in cooperation with various experts: economists, historians, musicologists, geologists, geographers... to work on documents such as administrative archives, musical scores, old newspapers, stock exchange price lists, financial yearbooks, geological section plans... One difficulty for ancient document recognition is that the documents are often degraded, which complicates their recognition. The second limit is that the recognition is often difficult, even for humans, while few labeled data are generally available. The objective is then to design recognition systems for structure and writings, able to be trained with few annotated data and able to improve themselves by self-adaptive recognition mechanisms. Some applications to historical documents require a very high quality of information extraction which implies the development of methods able to reach the best recognition rates in very difficult contexts. One way to improve the recognition quality is to use the sequence of documents and their redundancies which can be found in collections of documents, like in serial documents such as statistical data registers (weather, population), financial documents, administrative records... The collection of documents can also help in reducing the number of user interactions while improving the recognition.

Another type of documents studied by the team are those produced online. For these documents, the objective is to design interpretation systems and intelligent tutors based on artificial intelligence. Some of the online documents studied by the team come from productions written in a school learning context, thus making it possible to innovate in the field of e-education. The work carried out in this area is part of educational innovation projects supported by the academy and the Ministry of Education. We rely on the scientific foundations acquired in the fields of artificial intelligence ("pattern recog-

nition", "machine learning") and human-machine interaction. One of the challenges for education is to guarantee the transferability of the acquired knowledge via the digital solution (tablet) to the traditional use (paper/pencil) and vice versa. For this, we have focused our digital tablet interactions on "pen" interactions to allow the student to write and draw on the tablet as on paper. The objective is to design automated interpretation systems (intelligent tutor) of the students' productions: writing, arithmetic operations, geometric diagrams. This scientific know-how is the basis for the design of new "e-Learning" solutions that will allow more autonomy and personalization in the learning of each student.

The analysis of these online handwritten productions (2D gestures) makes it possible to imagine new gesture-interaction systems, and in particular 3D gesture-interactions. This other field of research interests the team, with the objective to obtain reactive and natural interactions with tactile devices as well as 3D sensor like Kinect. Gesture interaction allows the user to manipulate naturally the device, but they are today often limited to very basic functionalities like zooming, rotating and scrolling. The difficulty of adding new gestures is two folds: recognition accuracy and system reactivity. Increasing the number of gestures increases the probability of having gestures with common beginning. As a consequence, the system cannot predict the gesture from first traces without potentially executing undesirable commands. In this application domain, we design methods for the challenging task of early recognition of untrimmed gestures.

3 Scientific achievements

3.1 Lightweight Transformer-based Architectures for Historical Handwritten Text Recognition

Participants: Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Coüasnon, Ali Yesilkanat.

Transformer architectures could alleviate many concerns related to historical documents (degradations, specific handwritings for which few examples are available and ancient languages that vary over time), thanks to their ability to have a global view of textual images and their language modeling capabilities. However, they require a significant amount of annotated data to achieve competitive results.

Killian Barrere proposed in his PhD (defended in December 2023) a lightweight Transformer architectures (5–8M parameters) for handwriting recognition in historical documents [2]. To train these architectures, we introduced realistic looking synthetic data reproducing the style of historical handwritings (Figure 1). We proposed a specific strategy, both for training and predicting, to deal with historical documents, where only a limited amount of training data is commonly available. We evaluated our approaches on the ICFHR 2018 READ dataset which is dedicated to handwriting recognition in specific historical documents.

The results show that our Transformer-based approaches is able to outperform existing methods.

The code associated with this work, called Modular Light Transformer (MLT), and

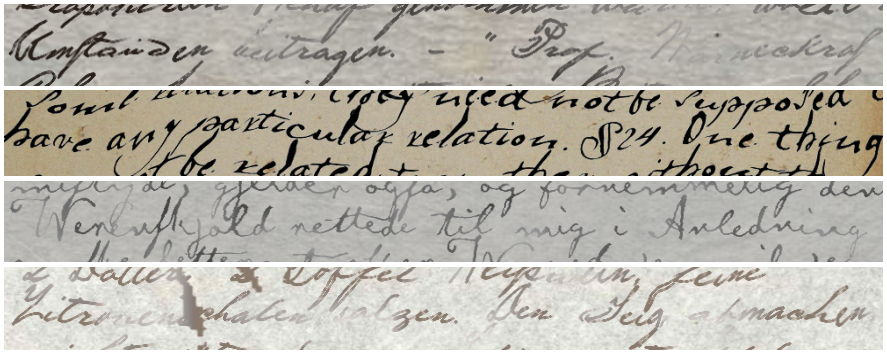

Figure 1: Examples of generated synthetic text-lines, trying to match the style of the ICFHR 2018 READ dataset.

the Handwriting generator software are available for research purposes under the CLIC license (see section 6.2.4). The MLT has also been optimized and developed for use on Windows as part of a collaboration agreement with the company IMDS. Different models have been also produced during this collaboration. An exploitation licence of the MLT has been granted to IMDS, the MLT is now part of the IMDS Solution for Document Analysis.

3.2 Optical Music Recognition of full-page orchestra scores recognition

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

In the context of the ANR project Collabscore 5.2.1, we proposed several contributions. We develop an OMR (Optical Music Recognition) system, based on the combination of deep learning approaches, traditional image processing approaches, and g grammatical syntactic rules.

The localization and classification of musical symbols on scanned or digital music scores pose significant challenges in Optical Music Recognition, such as similar musical symbol categories and a large number of overlapping tiny musical symbols within high-resolution music scores. Recently, deep learning-based techniques show promising results in addressing these challenges by leveraging object detection models. However, unclear directions in training and evaluation approaches, such as inconsistency between usage of full-page or cropped images, handling image scores at full-page level in high-resolution, reporting results on only specific object categories, missing comprehensive analysis with recent state-of-the-art object detection methods, cause a lack of benchmarking and analyzing the impact of proposed methods in music object recognition. To address these issues, we perform intensive analysis with recent object detection models, exploring effective ways of handling high-resolution images on existing benchmarks. Our goal is to narrow the gap between object detection models designed for common objects and relatively small images compared to music scores, and the unique challenges

of music score recognition in terms of object size and resolution. We achieve state-of-the-art results across mAP and Weighted mAP on two challenging datasets, namely DeepScoresV2 and the MUSCIMA++ datasets, by demonstrating the effectiveness of this approach in both printed and handwritten music scores [15]. This paper received the Nakano Best Paper Award at DAS 2024 (section 6.1.8).

Then, we develop an OMR system (Optical Music Recognition) using a combination of data guided by syntactic rules (Figure2). This work relies on DMOS method. It enables to validate the presence of elements localised by the musical object detector, to produce a global recognition of the score, and to point out the elements that do not respect a global consistency (red boxes in Figure2).

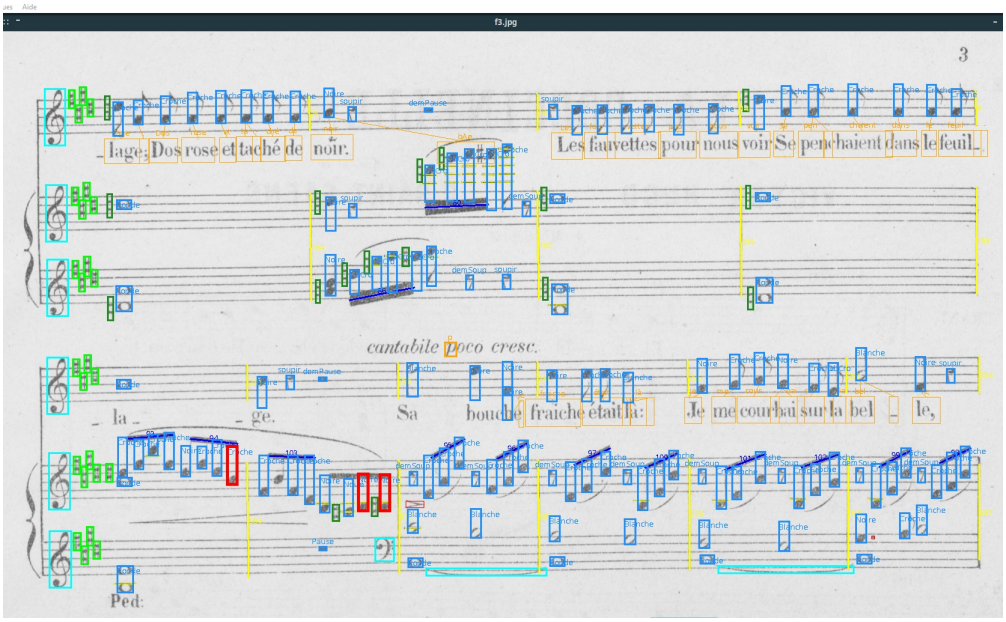

Figure 2: Using syntactic rules to combine object detected by deep learning

The new OMR approach developed in COLLABSCORE is part of a larger goal, namely that of interlinking multimodal documents related to music works. In this perspective, the music notation obtained from the OMR process is seen as a pivot that associates related fragments of images, audio, video, XML, or text sources. As an application of this principle, COLLABSCORE supports the synchronization of sources, leveraging the raw content of digital libraries with listening and visualization experiences [10].

3.3 IntuiSketch: Pen-Based Tutoring System for Anatomy Learning

Participants: Islam Barchouch, Eric Anquetil, Nathalie Girard.

IntuiSketch is a pen-based Intelligent Tutoring System (ITS) conceived to help students to learn anatomy by evaluating their drawings. This work is part of the ANR "Sketch" project (section 5.2.3) led by LP3C at the University of Rennes 2, where we contribute to its development. The system combines online handwriting recognition

approaches with tutoring techniques to guide students through the drawing process.

The system operates in two modes: the author mode, used by teachers to create exercises, and the student mode, where students complete the exercises and receive feedback to improve their understanding and skills. IntuiSketch's architecture is based on two main engines: the recognition engine and the supervision engine. The recognition engine is based on the CD-CMG (Context-Driven Constraint Multiset Grammar) formalism, used to recognize and interpret the students' drawings. It is supported by fuzzy logic and an incremental classifier, Evolve, which can learn from a few examples. This engine interprets semi-structured sketches on the fly, using shape, spatial, and geometric constraints to assess their accuracy. On the other hand, the supervision engine receives the interpreted sketches and evaluates them according to a Knowledge Graph (KG) generated from the teacher's exercise. The KG represents the solution: its nodes correspond to the elements of the exercise, while its edges define the constraints between these elements. The supervision engine compares the student's work to the KG, identifying any violations of the constraints to generate corrective feedback (Figure 3).

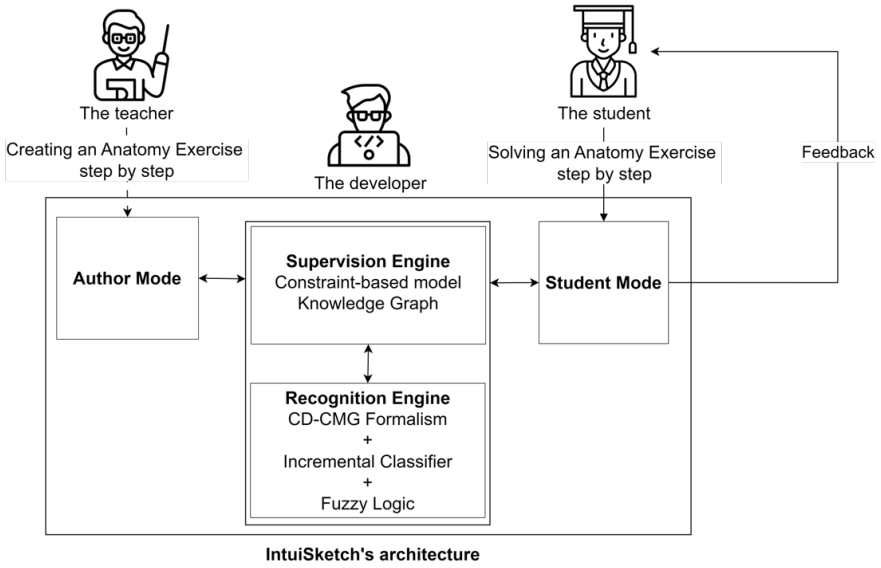

Figure 3: IntuiSketch Architecture

3.3.1 Extension of IntuiSketch Architecture

The architecture of IntuiSketch has been significantly extended to improve its ability to support learning anatomy by drawing. The system can now provide real-time and delayed feedback during drawing process, based on the extended CD-CMG formalism that introduces primary and secondary constraints to analyze semi-structured sketches. Primary constraints are used to make vital corrections, while secondary constraints take into account minor inaccuracies, thus promoting a balance between precision and flexibility. This approach ensures adaptability to a variety of drawing styles, due to the flexibility provided by the extension of the CD-CMG formalism, and enhances the system's ability to generate appropriate feedback according to the student's progress

and the error made (Figure 4).

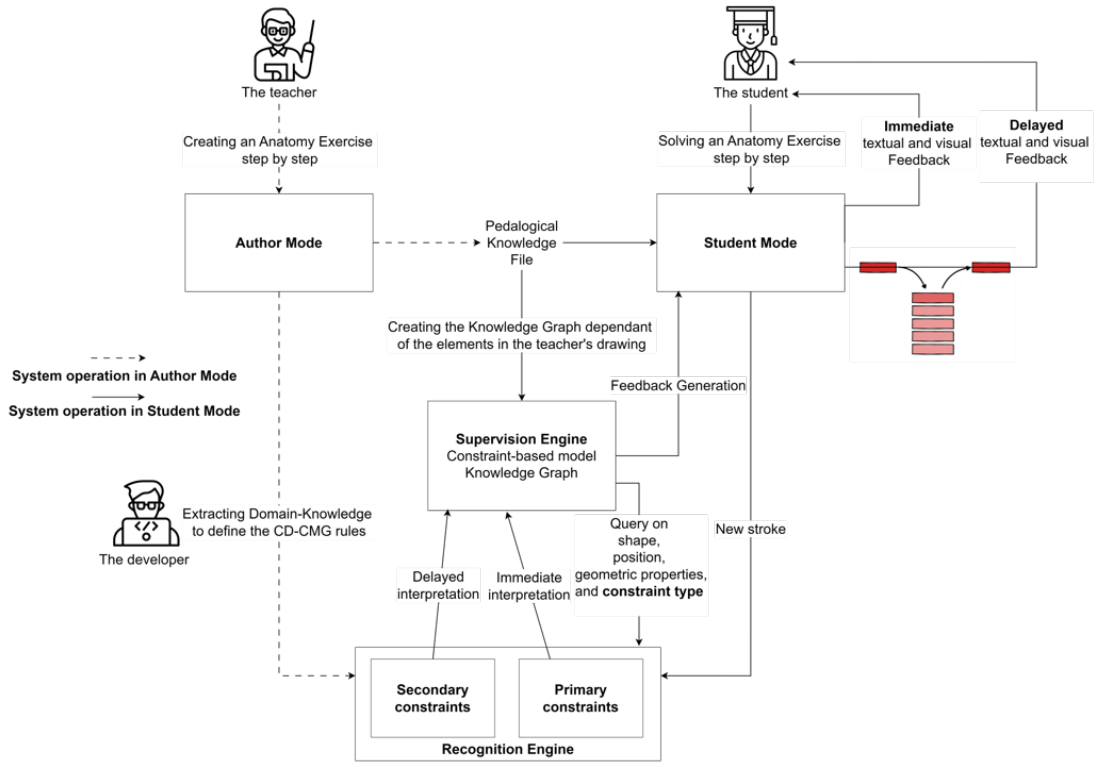

Figure 4: The extended architecture of IntuiSketch with its Author and Student modes relying on the new version of the recognition and supervision engines

3.3.2 Multi-Stroke Analysis

To address challenges associated to multi-stroke analysis, CD-CMG formalism has been further extended and the Evolve incremental classifier has been adapted. Frequently, students use multiple strokes to represent a single element that is supposed to be drawn using a single stroke and defined in the CD-CMG rules as a mono-stroke element, creating variability in drawing styles. The new hybrid approach reintegrates rejected strokes not recognized into the analysis process, and aggregates them to recognize the desired shapes. The CD-CMG formalism has been extended, and Evolve has been adapted to recognize both shapes drawn using a single stroke and complex shapes drawn using multiple strokes, whatever their number, order or drawing sense. By addressing such complexities, IntuiSketch maintains robust recognition capabilities and can recognize strokes that are usually rejected and often result in failure to recognize the following dependent elements.

3.4 IntuiGeo : an intelligent pen-based tutoring system for geometry learning in middle school - Triangle ANR Project

Participants: Eric Anquetil, Nathalie Girard, Omar Krichen, Sarah Bucquet, Erell

Choulette.

This work is part of the ANR project "TRIANGLE" (e-Fran transfert, section 5.2.4), for which we are responsible for the AI/intelligent tutor workpackages.
The aim is to support students learning geometry in secondary schools. It includes experimentation and deployment in the Rennes and Poitiers academies (see figure 5). The final product (IntuiGeo) will be used in French schools.

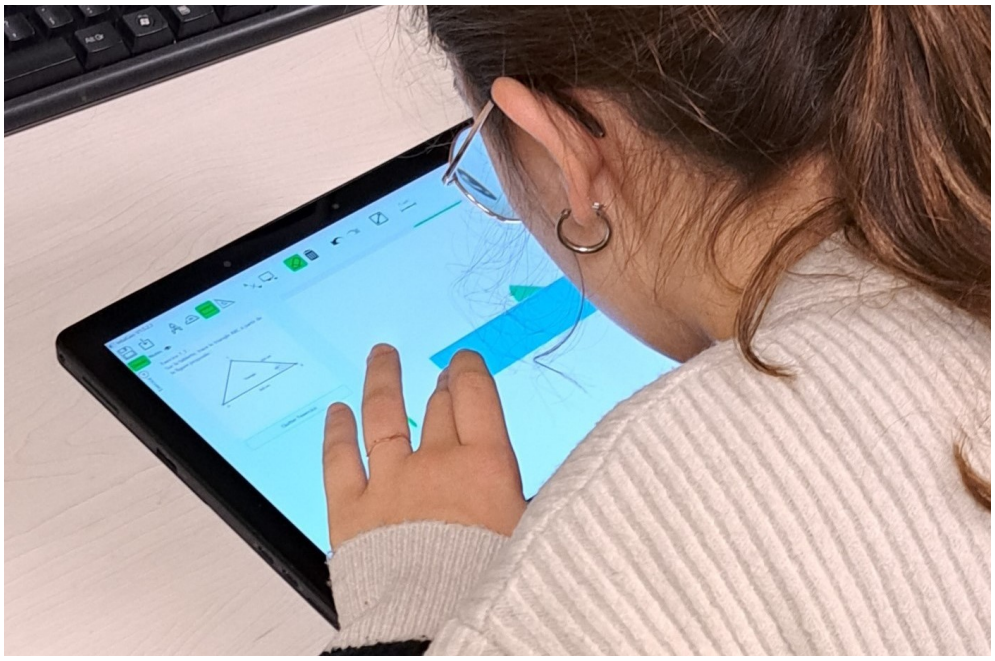

Figure 5: Student using IntuiGeo during a study.

IntuiGeo is a pen-based intelligent tutoring system for geometry learning in middle school. The objective is to combine pattern recognition and tutoring techniques to provide real-time feedback for the construction of a geometry problem.

This year, with the research team in psychology and ergonomics of learning (LP3C, Rennes), training sessions were held with teachers to introduce them to IntuiGeo. Their advice and comments have been noted for future versions of the application.

We have implemented three kinds of feedbacks: color, descriptive and guidance feedbacks. The color feedbacks, displayed on the recognized geometry shaped, and the descriptive feedbacks, written in a speech bubble, informs the student of the accuracy of his drawings. The guidance feedbacks, also displayed in a speech bubble linked to the avatar, is aimed to assist the student resolving the next action he should realise. Both kinds of written feedbacks are adaptive to the way the student decides to construct the figure. The results of the impact of descriptive feedback have been published in [3]. The ultimate results obtained from new field testing show that the guidance feedback has also significant impact on the kids capacity of solving more complex problems. The projet have been prensented in conferences [7, 8, 9].

Subsequent to the experimentations and the LP3C advises, the feedbacks' texts given by the intelligent tutor have been revised to a better understandably child-oriented

explanation.

An alternative way of drawing is also being developed. It has been observed that IntuiGeo's automatic point naming can be distracting to some students when solving an exercise. Also, on paper, students have to name their points themselves. It was therefore decided to encourage the manual naming of each point.

Simultaneously, a Flutter source-code is being developed with the Wyatt company to allow the app to be compatible with both Windows, iOS and Android platforms. The free-edition mode of this version of IntuiGeo should be tested soon (see an illustration figure 6).

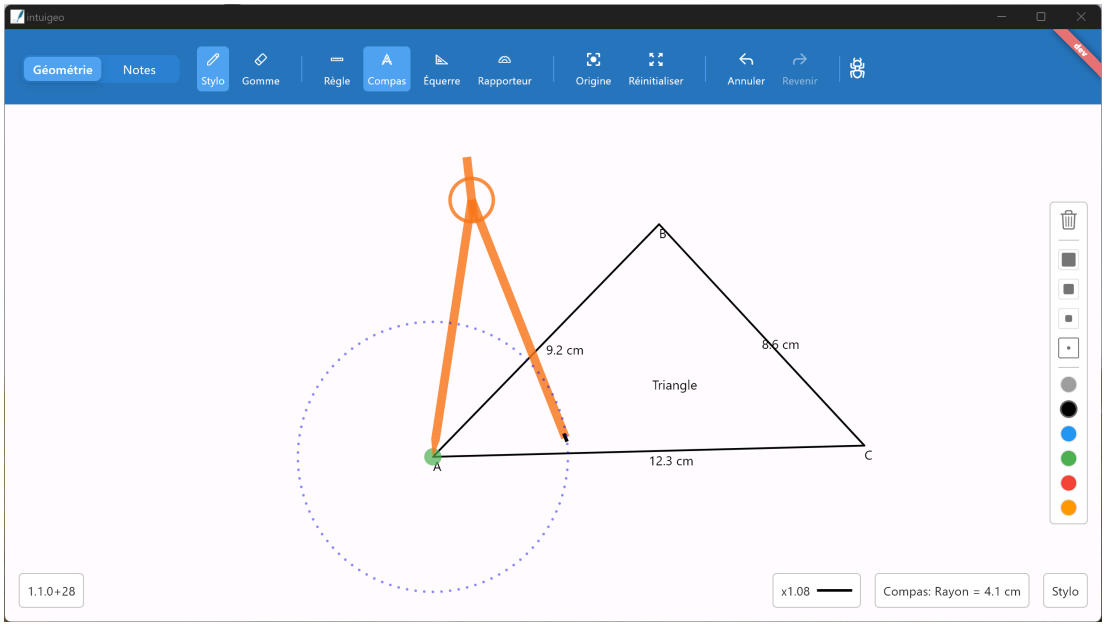

Figure 6: Screenshot of the in progress new IntuiGeo version free-edition in Flutter.

3.5 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 [11]. This project is composed of two academic partners, IRISA (France) and the Karlsruhe Institut of Technology (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, with which we work.
Our goal is to reconstruct the handwriting trajectory from the Digipen. The Stabilo Digipen captures time series from IMU sensors (accelerometers, gyroscope, magnetometer, force sensors) during writing, and we aim at reconstructing the online handwriting trace from the sensors (Fig. 7).

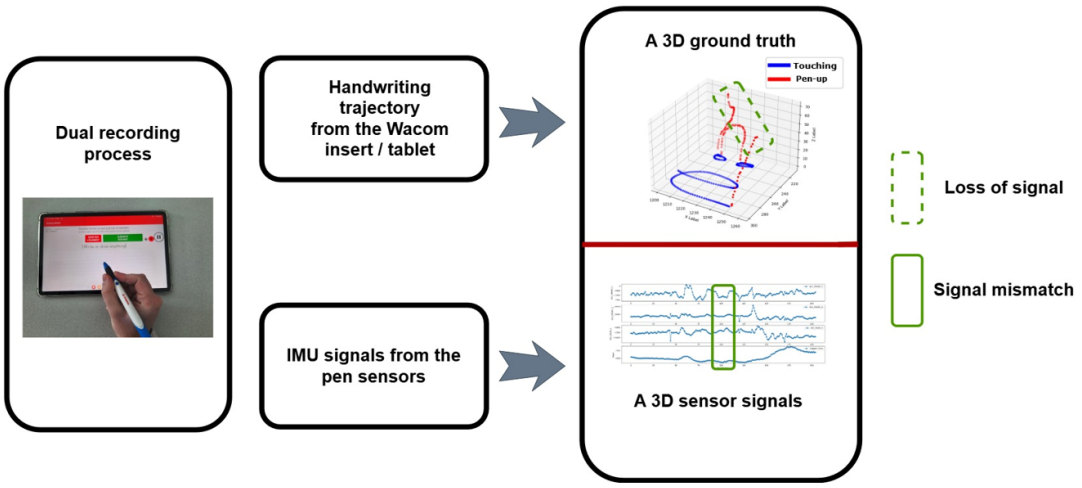

Figure 7: Dual recording process to generate learning database

We propose a new approach learning to finely reconstruct the touching trajectories while precisely analyzing the pen-up part in order to position the next touching trace correctly. This relies on a Mixture-Of-Experts (MOE) approach (Fig. 8). The first expert is dedicated to pencil touch, and is named Touching Expert Model (TEM). The second one is dedicated to deal with pen-up trajectory, and is named Pen-up Expert Model (PEM). We propose two learning strategies, one for each expert. The first one is to add additional context in input of the Touching expert model (TEM-C) and the second one consists in training the Pen-up expert model on specific examples acquired on inclined plane to learn 3D input signals (PEM-I). In addition we introduce a novel public benchmark dataset, to enable future research and comparisons in the field of handwriting reconstruction. The results demonstrates a significant enhancement compared to its primary competitors. This work has been accepted in the Pattern Recognition journal [4], and presented at SIFED 2024 [12]

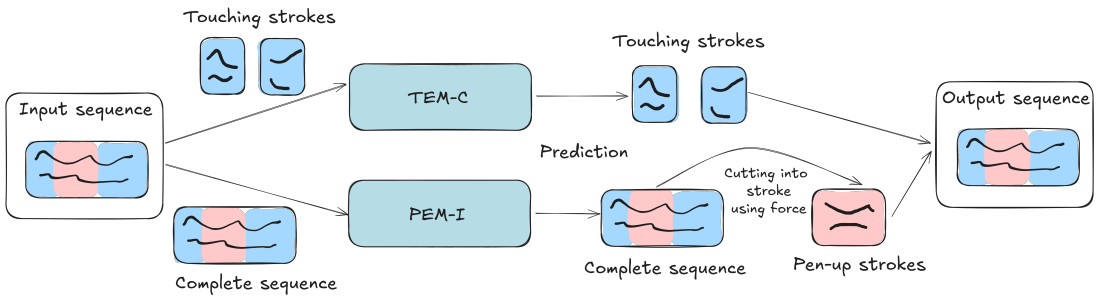

Figure 8: Our MOE-CI approach combines two models and their enhancements: TEM-C and PEM-I. TEM-C is dedicated to the touching part. PEM-I is dedicated to the pen-up with a training on complete sequences.

We also explore domain adaptation methods for processing child data. A major issue is based on the difference in captured signals between adults and children. For similar handwriting trace, we have large differences in sensor signals due to differences in speed and confidence in the handwriting gesture of children. To address this, we investigate a domain adversarial training to build a unified intermediate feature representation for both adult and children data. We compare our domain adaptation approach with two other methods: training the model from scratch and fine-tuning the model from one domain to the other. We show that our proposed domain adversarial training allows to get a system that is as robust on children data as on adults data, while other methods fail. This work has been accepted in the ADAPDA workshop of the ICDAR 2024 conference [13].

This work is part of Florent Imbert's PhD [1], which was successfully defended on 22 November 2024.

3.6 External language models and rejection capabilities for text recognition in difficult conditions

Participants: Florent Meyer, Yann Soullard, Bertrand Coüasnon, Guillaume Gravier, Denis Coquenet, Laurent Guichard (ANTAI).

Vehicles license plate recognition is an OCR application for which the tremendous volume of contraventions recorded accounts for the need of a very high performance in the reading. Unfortunately, additionally to common image quality-related difficulties emerging from weather conditions or camera settings, the distribution of the characters on the plates is constantly evolving due to vehicles being put in and out of service which makes it harder for systems to preserve peak performance over time. A lever towards obtaining a predictor capable of adapting nearly daily is to first ignore the bias in the plates in order to better reinject a useful, up-to-date bias later on without fully retraining the model. We have conducted numerous experiments which enabled us to both observe how a model comes to absorb various biases and to empirically demonstrate the inability for a Transformer encoder-decoder like TrOCR to predict known characters in new positions on a plate. Work is now being conducted on a CNN-encoder and Transformer-decoder architecture so as to train an OCR model while not learning any character ordering or positional biases.

3.7 Projet AIR : Intuinote

Participants: Éric Anquetil, Bruno Hortollary, Pierre Beust, Nathalie Girard.

IntuiNote is a new digital active learning environment for synchronous teaching (face-to-face and distance learning). Its aim is to optimize the possibilities for interaction between teacher and learners. In particular, it is based on handwritten note-taking and graphical MCQs developed on the fly on a on a touch tablet/stylus.

This work is part of the AIR project (ANR-21-DMES-0001, section 5.2.5) set up by

Rennes University.

This year we have been working on developing a multi-platform 'student' client so that students can interact during the course from different devices: smartphone, pen-based tablet and computer. This development has been prototyped in an initial web client version and is currently being completely rethought with a multi-platform Flutter design. Experiments are currently underway to determine the impact of Intuinote on student learning. These experiments are being carried out in collaboration with the LP3C laboratory at the University of Rennes 2.

3.8 Projet IntuiNaxe

Participants: Éric Anquetil, Bruno Hortollary, Omar Krichen.

The collaboration between INAXE compagny and the IntuiDoc/Shadoc team goes back several years. This resulted in the transfer of the IntuiNaxe technology (cf. section 5.4.1).

IntuiNaxe, which has different modes (author and technician), is used on a Windows pen tablet and is perfectly suited to the daily work of INAXE's technicians, who are reinforcing the digital shift in their working methods to capture information on the scale of a building, whatever its architectural configuration. IntuiNaxe software can be used to import building plans to create projects that incorporate all the information gathered on site. Thanks to IntuiNaxe, these plans become the digital medium for all the reports that the technicians create intuitively, freehand, by writing and drawing on the tablet. This information is linked, for example, to the presence of asbestos or the state of disrepair of the building (cf. fig. 9).

The software was enhanced in 2024 with new interpretation features and is now used daily in the field by 25 INAXE technicians as part of 8,000 building inspections per year. IntuiNaxe's innovative functionalities extend its scope to residential, commercial, retail, educational, healthcare, ERP and industrial buildings.

4 Software development

4.1 Automatic edition of Journals on 18th-century soldiers

Participants: Ivan Leplumey, Yann Ricquebourg.

The "Revues sur les soldats du XVIII^{e} siècle" (Journals on 18th-century soldiers) project involves a direct collaboration with the "Mémoire des Hommes" website[1], the cultural portal of the French Ministry of the Armed Forces, and up to 60 departmental archives. We use manual annotations of old documents to create several types of digital journals able to evolve: the first of a geographical type for associations and departmental archives, the second of a regimental type for the "Mémoire des Hommes" website or for the ten annotators taking part in the project (see Figure 10). These reviews, pro-

[1] `https://www.memoiredeshommes.sga.defense.gouv.fr`

Figure 9: Pen-based interaction with IntuiNaxe software

duced in LaTeX, mainly contain ordered lists of soldiers linked to their original image on the web by hypertext links. Regimental histories, statistics and indexes complete the content of these journals. To date, the project has produced 130 journals, 42 of them geographical, totaling 40,000 pages.

4.2 Modular Light Transformer (MLT)

Participant: Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Coüasnon.

The code associated with our publication on training transformer architectures on few annotated data with an application to historical handwritten text recognition [2] is available for research purposes. Called Modular Light Transformer (MLT), it is distributed under the CLIC license and deposited with the Agency for the Protection of Programs (APP)(section 6.2.4). It has also been optimized and developed for use on Windows as part of a collaboration agreement with the company IMDS. An exploitation licence has been granted to IMDS, the MLT is now part of the IMDS Solution for Document Analysis.

4.3 Handwriting generation

Participant: Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Coüasnon.

The code associated with our publication on training transformer architectures on few annotated data with an application to historical handwritten text recognition [2] for

(a)

(b)

(c)

(d)

Figure 10: (a) Cover of the Ligerian soldiers under the Ancien Régime journal produced with the Loire departmental archives, (b) Cover of the Crown regiment journal, (c) Cover of the Bretillian soldiers under the Ancien Régime journal produced with the Loire departmental archives, (d) Cover of the Troop control and matriculation registers

generating realistic looking synthetic handwritten data for training the Modular Light Transformer (MLT). It is distributed under the CLIC license and deposited with the Agency for the Protection of Programs (APP)(section 6.2.4).

4.4 IntuInaxe App: pen-based report on building plan

Participants: Eric Anquetil, Omar Krichen, Bruno Hortollary.

We have launched a new research and transfer contract with Inaxe company in 2023. The aim is to design a new pen-based application on a Windows tablet based on AI Technologies of our research Team (DALI, Evolve...) for freehand transfer (gesture recognition) of asbestos and lead surveys on building plans.

This transfer work led in December 2023 to an APP registration and a licence with the company. The software was enhanced in 2024 with new interpretation features and is now used daily in the field by 25 INAXE technicians as part of 8,000 building inspections per year.

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, IRISA).

- Project type: French-German bilateral ANR project KIHT - Kaligo-based Intelligent Handwriting Teacher
- Dates: 2021–2025
- Role : We are the project leader of the french part.
- PI institution: IRISA (French part - ANR) and Stabilo (German Part - BMBF)
- Other partners: STABILO International GmbH, Karlsruher Institut für Technologie Institut für Technik der Informations verarbeitung, LearnAndGo company

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 CollabScore: Shared spaces for digital music scores

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

- Project type: ANR CollabScore
- PI institution: CNAM
- Other partners: Cnam, INSA, BnF, Antescofo, IReMus, Fondation Royaumont
- 48 months (2020–2025)
- Contract: INSA

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 automated 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.

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

Participant: Bertrand Coüasnon, Aurélie Lemaitre.

- Partners: Université de La Rochelle, INSA, BnF
- PI institution: Université de La Rochelle
- 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.

5.2.3 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 ShaDoc (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.

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

Participant: Eric Anquetil, Nathalie Girard, Sarah Bucquet, Erell Choulette, Omar Krichen.

- 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*
- 2022-2025.
- 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 (ShaDoc (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).

5.2.5 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
- 2022-2025
- 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 Regional Initiatives

5.3.1 AIS funding

Denis Coquenet received the "Aide à l'Installation Scientifique" from Rennes Metropole for a budget of 40,000€. It led to the purchase of two H100 GPUs, supplemented by a team budget for the server part, to enrich the local hardware computational infrastructure in the context of deep neural network training for handwritten document analysis and recognition.

5.4 Bilateral industry grants

5.4.1 Research contract Inaxe company

Participant: Eric Anquetil, Omar Krichen, Bruno Hortollary.

- Partners: *Inaxe company*

- 2023-2025
- Contract: INSA Rennes

Following on from a 1st transfer with the company Inaxe in 2018, we have launched a new research and transfer contract with the same company in 2023. The aim is to design a new pen-based application on a Windows tablet based on AI Technologies of our research Team (DALI, Evolve...) for freehand transfer (gesture recognition) of asbestos and lead surveys on building plans. This transfer work led in December 2023 to an APP registration and a licence with the company. The software was enhanced in 2024 with new interpretation features and is now used daily in the field by 25 INAXE technicians as part of 8,000 building inspections per year

5.4.2 Research contract ANTAI (CIFRE)

Participant: Yann Soullard, Bertrand Coüasnon, Guillaume Gravier, Denis Coquenet.

- Partners: *ANTAI*
- Since 2023
- Contract: INSA

Shadoc team started to work with ANTAI on recognition with rejection capabilities and external language models with application on licence plate recognition. This collaboration is based on a CIFRE grant for the PhD of Florent Meyer.

Current activities are described in section 3.6.

5.4.3 Research contract IMDS company

Participant: Killian Barrere, Ali Yesilkanat, Yann Soullard, Bertrand Coüasnon.

- Partners: *IMDS company*
- 2024
- Contract: INSA

A research collaboration has been done with IMDS company to build new models for the MLT (Modular Light Transformer), and to integrate it in the Windows IMDS Solution for Document Analysis. At the end of this contract an exploitation licence has been granted to IMDS, and the MLT is now part of the IMDS Solution (see 3.1).

6 Dissemination

6.1 Promoting scientific activities

6.1.1 Scientific Events Organization

Organization of International Workshop

- ADAPDA 2024 (Greece): Eric Anquetil Co-organized the international workshop for ICDAR2024, "Automatically Domain-Adapted and Personalized Document Analysis (ADAPDA)" with Prof. Rita Cucchiara, prof. Eric Anquetil, and Christopher Kermorvant.

General Chair, Scientific Chair

- A. Lemaitre organized the Arts, Culture and Heritage cross-disciplinary seminar, at IRISA, on April 11th 2024.

6.1.2 Scientific Events Selection

Member of Conference Program Committees

- B. Coüasnon is senior member of the program committee of the International Conference on Document Analysis and Recognition (ICDAR 2024)
- E. Anquetil, N. Girard and A. Lemaitre are members of the program committee of the International Conference on Document Analysis and Recognition (ICDAR 2024)
- D. Coquenet and N. Girard are members of the program committee of the 27th European Conference on Artificial Intelligence (ECAI 2024).
- D. Coquenet is member of the program committee of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2024, Research Track).

Reviewer

- D. Coquenet, N. Girard and Y. Soullard are reviewers for the International Conference on Document Analysis and Recognition (ICDAR) in 2024.
- D. Coquenet and N. Girard are reviewers for the International Conference on Pattern Recognition (ICPR) in 2024.
- D. Coquenet and N. Girard are reviewers for the European Conference on Artificial Intelligence (ECAI) in 2024
- D. Coquenet is reviewer for the International Joint Conference on Artificial Intelligence (IJCAI) in 2024.

6.1.3 Journal

Reviewer - Reviewing Activities

- A. Lemaitre is reviewer for the International Journal on Document Analysis and Recognition (IJDAR) in 2024.
- A. Lemaitre is reviewer for a book publication in PUR (Presses Universitaires de Rennes) in 2024.
- D. Coquenet is reviewer for the International Journal on Document Analysis and Recognition (IJDAR) in 2024.
- D. Coquenet is reviewer for IEEE Transactions on Artificial Intelligence (TAI) in 2024.
- Y. Soullard is reviewer for Neural Processing Letters (NPL) in 2024.
- N. Girard is reviewer for the International Journal on Document Analysis and Recognition (IJDAR) in 2024.

6.1.4 Invited Talks

- A. Lemaitre presented "Combinaison d'intelligences artificielles pour la reconnaissance d'images de documents" at Seminar "L'Art des données, les données de l'art", Université Rennes 2, 2024/12/02.
- D. Coquenet and K. Barrere presented "Reconnaissance d'écriture manuscrite appliquée aux documents historiques" at Séminaire de l'axe art, culture et patrimoine, IRISA, 2024/04/11.
- B. Coüasnon presented "Introduction à l'IA" in the context of "Le jeudi de l'éloquence - Concours d'éloquence national inter-INSA", INSA Rennes, 2024/10/21.

6.1.5 Leadership within the Scientific Community

6.1.6 Scientific Expertise

- B. Coüasnon was a scientific expert in 2024 on a project for Innoviris, Brussels.
- Y. Soullard took part in the evaluation process of AAPG 2024 as a Scientific Expert for the scientific panel CE38 - Interfaces: mathematics, digital sciences – Humanities and social sciences for the National Research Agency (ANR)
- Y. Soullard was a scientific expert to evaluate a project submitted to the ANR, seeking certification from the "Images & Réseaux" competitiveness centre - 2024
- N. Girard was a member of four recruitment committees of assistant professor (COS MCF27 & ECC): Université de Lorraine Nancy (Télécom Nancy - LORIA ; MCF 0951 - May 2024), University of Rennes (UFR ISTIC - IRISA ; MCF 0452 & MCF 0457 - May 2024), La Rochelle University (IUT Computer Science - L3i ; June 2024 - ECC-Info).

6.1.7 Research Administration

- Shadoc 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 10 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.
- E. Anquetil is a member of the « Science-Society committee » of TISSAGE Project.
- B. Coüasnon is the head of the computer science lab of INSA Rennes (INSA component of IRISA) since April 2024 (about 65 scientists, including 21 faculty members).
- B. Coüasnon is an elected member of the scientific council of INSA Rennes.
- B. Coüasnon is an elected member of the computer science lab council of INSA Rennes (INSA component of IRISA).
- B. Coüasnon is member of the laboratory council of IRISA.
- B. Coüasnon and N. Girard are members of the Gender Equality Commission of IRISA.
- B. Coüasnon is member of the board of Valconum (Centre Européen de Valorisation Numérique).
- N. Girard, A. Lemaitre and Y. Soullard are elected members of the executive committee of the society GRCE : " Groupe de Recherche en Communication Écrite ".
- N. Girard is an elected member of the administration council of UFR ISTIC, Univ. of Rennes.
- A. Lemaitre is responsible for the Arts, Culture and Heritage transversal theme at IRISA.
- Y. Ricquebourg is an elected member of the scientific council of INSA Rennes.

6.1.8 Awards

- Nakano Best Paper Award, 16th IAPR International Workshop on Document Analysis Systems 2024, Ali Yesilkanat, Yann Soullard, Bertrand Coüasnon, Nathalie Girard, "Full-page music symbols recognition: state-of-the-art deep models comparison for handwritten and printed music scores" [15], September 2024.

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, N. Girard and D. Coquenet give lectures at *Research in Computer Science (SIF)* MASTER of University of Rennes, 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, 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), until July 2024.
- 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.
- N. Girard is in charge (with L. Guého - PEMI, MSHB) of the module "Graphical User Interaction" (GUI) of the *Software Engineering* MASTER of ISTIC, University of Rennes.
- N. Girard is in charge of the module "Analyse Descriptive et Exploratoire de Données" (ASED) of the *Artificial Intelligence* MASTER of ISTIC, University of 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 *(Applied Mathematics, Statistics (Data Science))* of Rennes 2 University.
- Y. Soullard is in charge of the part "Deep Learning for Sequential Analysis" of the module "Deep Learning" of the DIGISPORT MASTER *(Digital and Sport Sciences)* of the University Research School (EUR) DIGISPORT.
- Y. Soullard is in charge of the module "Data Mining & Clustering" of the DIGISPORT MASTER *(Digital and Sport Sciences)* of the University Research School (EUR) DIGISPORT.
- P. Beust is Director of the SUPTICE : pedagogical support service at the University of Rennes (https://suptice.univ-rennes.fr/).

6.2.2 Supervision

- PhD in progress: S. Serre, Semi-supervised learning of an intelligent tutorial system for e-education through the production of drawings/sketches, E. Anquetil, N. Girard, started October 2024.
- PhD in progress: I. Barchouch, Intelligent tutorial system for sketch-based learning (SKETCH), E. Anquetil, N. Girard, started October 2022.
- PhD: 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 and successfully defended on 22 November 2024.
- PhD in progress: F. Meyer, External language models and rejection capabilities for text recognition in difficult conditions, B. Coüasnon, G. Gravier, Y. Soullard, L. Guichard (ANTAI), INSA Rennes, started June 2023.
- PhD in progress: A. N. E. Sahbi, Peuplement automatique d'ontologies à partir d'approches hybrides (LLM et sémantiques), P. Beust, C. Alec (University of Caen), University of Caen Normandy, started december 2022.

6.2.3 Juries

- E. Anquetil was reviewer in the thesis committee of Elmokhtar MOHAMED MOUSSA's PhD, Offline to online handwriting conversion with deep neural networks, Nantes University, January 2024.
- E. Anquetil was reviewer in the thesis committee of Florian Marchal-Bornert's PhD, Towards narrative explanations for recommendations within an augmented territory, Lorraine University, September 2024.
- E. Anquetil was president in the thesis committee of Arthur Hoarau's PhD, Active learning from uncertain and imprecise data, Rennes University, June 2024.
- A. Lemaitre was president in the thesis committee of Tiange Zhu's PhD, Content-based symbolic music retrieval : a study on pattern and similarity, CNAM, Paris, October 2024.
- A. Lemaitre was president in the thesis committee of Ibrahim Souleiman's PhD, Apprentissage automatique, et sans modèle a priori, des liens sémantico-structurels entre les champs dans un document, La Rochelle University, November 2024.
- A. Lemaitre was reviewer in the thesis committee of Thomas Constum's PhD, Extraction d'information dans des documents historiques à l'aide de grands modèles multimodaux, Rouen University, November 2024.
- 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).
- N. Girard is member of mid-term evaluation committee of the PhD candidate: T. Legay (Insa Rennes, IETR).
- Y. Ricquebourg is member of mid-term evaluation committee of the PhD candidate: Tsiry Mayet (Insa Rouen, LITIS).

6.2.4 Patent and Deposit of digital creations (APP)

- E. Anquetil, O. Krichen, B. Hortollary deposited (APP) the IntuInaxe App in 2023/2024 :IntuiNaxe - IntuiNaxe IDDN.FR.001.030020.000.S.P.2024.000.10000.
- "Modular Light Transformer" Software: Lightweight Transformer Architectures for the Recognition of Ancient Handwritten Texts. Deposited with the Agency for the Protection of Programs (APP) the 7 March 2024 n° IDDN1.FR2.0013.1000244.0005.S6.P7.20248.00 (réf. N° DV 5346) and the 30 September 2024 n° IDDN.FR.001.400005.000.S.P.2024.000.30000. Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Coüasnon
- "Handwriting generation" Software : Synthetic Handwritten text and printed text generation used in the training of the MLT. Deposited with the Agency for the Protection of Programs (APP) the 30 September 2024 n° IDDN.FR.001.400006.000.S.P.2024.000.30000 (réf. N° DV 5464). Licensed to IMDS company. Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Coüasnon

6.3 Popularization

- A. Lemaitre was interviewed for a popularization podcast AZERTY : The AI that could read music [2]
- A. Lemaitre and B. Coüasnon were interviewed for the popularization science magazine "Science Ouest", June 2024 : An AI for deciphering old music scores [3]

7 Bibliography

Major publications by the team in recent years

[R1] S. Adam, J. Annaert, F. Buelens, B. Coüasnon, B. Cule, A. de Vicq, C. Guerry, P.-C. Hautcoeur, T. Paquet, A. R. Camacho, I. Le Floch, A. Lemaitre, P. Karapanagiotis, J. Poukens, A. Riva, "Data extraction and matching The EurHisFirm experience", *in : Methodological Advances in the Extraction and Analysis of Historical Data, Methodological Advances in the Extraction and Analysis of Historical Data*, Kellogg School of Management - Northwestern University, Chicago/Virtual, United States, December 2021, https://hal.archives-ouvertes.fr/hal-03828381.

[R2] K. Barrere, Y. Soullard, A. Lemaitre, B. Coüasnon, "Transformers for Historical Handwritten Text Recognition", *in : 16th International Conference on Document Analysis and Recognition (ICDAR 2021) Doctoral Consortium*, 2021.

[R3] K. Barrere, Y. Soullard, A. Lemaitre, B. Coüasnon, "A Light Transformer-Based Architecture for Handwritten Text Recognition", *in : International Workshop on Document Analysis Systems*, Springer, p. 275–290, 2022.

[2] https://www.c-lab.fr/emission/azert/lia-qui-savait-lire-la-musique-avec-aurelie-le-maitre.html

[3] https://www.espace-sciences.org/sciences-ouest/420/actualite/une-ia-pour-dechiffrer-d-anciennes-partitions

[R4] K.-Y. Choi, 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.

[R5] 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.archives-ouvertes.fr/hal-03687091.

[R6] C. Guerry, *Big-data historique : modélisation de stratégies d'analyse de collections de documents*, Theses, INSA de Rennes, December 2022.

[R7] 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*, p. 1842 – 1849, Santiago de compostela, Spain, August 2020, https://hal.archives-ouvertes.fr/hal-02544384.

[R8] 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.archives-ouvertes.fr/hal-03845144.

[R9] O. Krichen, S. Corbillé, E. Anquetil, N. Girard, P. Nerdeux, "Online analysis of children handwritten words in dictation context", *in : 14th International Workshop on Graphics Recognition*, Lausanne, Switzerland, September 2021, https://hal.science/hal-03448357.

[R10] A. Lemaitre, J. Camillerapp, C. Carton, 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.

[R11] C. Leroy, E. Anquetil, N. Girard, "Drift anticipation with forgetting to improve evolving fuzzy system", *in : 25th International Conference on Pattern Recognition (ICPR2020)*, Milan, Italy, January 2021, https://hal.archives-ouvertes.fr/hal-02974253.

[R12] N. Michinov, E. Anquetil, E. Michinov, "Guiding the use of collective feedback displayed on heatmaps to reduce group conformity and improve learning in Peer Instruction", *Journal of Computer Assisted Learning*, June 2020, https://hal.univ-rennes2.fr/hal-02875166.

[R13] 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*, 2022.

[R14] R. Plamondon, G. Pirlo, E. Anquetil, C. Rémi, H.-L. Teulings, M. Nakagawa, "Personal Digital Bodyguards for e-Security, e-Learning and e-Health:A Prospective Survey", *Pattern Recognition 81*, September 2018, p. 633–659, https://hal.science/hal-01767055.

[R15] S. Tarride, A. Lemaitre, B. Coüasnon, S. Tardivel, "Combination of deep neural networks and logical rules for record segmentation in historical handwritten registers using few examples", *International Journal on Document Analysis and Recognition*, January 2021, https://hal.archives-ouvertes.fr/hal-03160212.

[R16] C. Truong, L. Oudre, N. Vayatis, "Selective review of offline change point detection methods.", *in : Signal Processing, 167, 107299*, IEEE, p. 770–778, 2020.

Doctoral dissertations and "Habilitation" theses

[1] F. Imbert, *Design of a deep neural network architecture dedicated to handwriting synthesis using kinematic sensors from a digital pen.*, Theses, INSA RENNES, November 2024, https://hal.science/tel-04887058.

Articles in referred journals and book chapters

[2] K. Barrere, Y. Soullard, A. Lemaitre, B. Coüasnon, "Training transformer architectures on few annotated data: an application to historical handwritten text recognition", *International Journal on Document Analysis and Recognition 27*, 4, 2024, p. 553 – 566, https://hal.science/hal-04447488.

[3] T. Colliot, O. Krichen, N. Girard, É. Anquetil, É. Jamet, "What makes tablet-based learning effective? A study of the role of real-time adaptive feedback", *British Journal of Educational Technology*, February 2024, p. 18, https://inria.hal.science/hal-04448708.

[4] F. Imbert, E. Anquetil, Y. Soullard, R. Tavenard, "Mixture-of-experts for handwriting trajectory reconstruction from IMU sensors", *Pattern Recognition 161*, 2025, p. 111231, https://hal.science/hal-04811975.

[5] W. Mocaër, E. Anquetil, R. Kulpa, "Early gesture detection in untrimmed streams: A controlled CTC approach for reliable decision-making", *Pattern Recognition*, June 2024, p. 110733, https://hal.science/hal-04634678.

[6] A. Sahbi, C. Alec, P. Beust, "Semantic vs. LLM-based approach: A case study of KOnPoTe vs. Claude for ontology population from French advertisements", *Data and Knowledge Engineering 156*, March 2025, p. 102392, https://hal.science/hal-04823061.

Publications in Conferences and Workshops

[7] T. Colliot, E. Jamet, N. Michinov, N. Bonneton-Botté, A. Cherbonnier, O. Krichen, N. Girard, E. Anquetil, "Guider les élèves dans la résolution de problèmes : étude à partir d'un logiciel de géométrie sur tablette", *in : 16ème colloque international de l'association Ripsydeve*, Reims (France), France, June 2024, https://hal.science/hal-04606573.

[8] G. de Pereyra, T. Colliot, A. Cherbonnier, N. Michinov, N. Bonneton-Botté, O. Krichen, N. Girard, E. Anquetil, E. Jamet, "Bénéfices du guidage d'un logiciel de géométrie (IntuiGéo) sur les performances d'élèves de cycle 4", *in : Journées Scientifiques Jeunes Chercheurs JSJC*, Poitiers, France, July 2024, https://hal.science/hal-04636156.

[9] G. de Pereyra, T. Colliot, M. Robert, A. Cherbonnier, N. Michinov, N. Bonneton-Botté, O. Krichen, N. Girard, E. Anquetil, E. Jamet, "Effet de feedbacks adaptatifs basés sur l'Intelligence Artificielle pour l'apprentissage de la géométrie sur tablette", *in : L'Intelligence Artificielle dans la recherche en psychologie*, Université de Tours, Tours, France, December 2024, https://hal.science/hal-04868348.

[10] C. Guillotel-Nothmann, P. Rigaux, B. B. Coüasnon, M. Giraud, A. Lemaitre, "The CollabScore project – From Optical Recognition to Multimodal Music Sources", *in : WoRMS 2024: 6th International Workshop on Reading Music Systems*, Jorge Calvo-Zaragoza and Alexander Pacha and Elona Shatri, p. 33–37, Online, France, November 2024, https://hal.science/hal-04925968.

[11] T. Harbaum, A. Serdyuk, F. Kress, T. Hamann, J. Barth, P. Kämpf, F. Imbert, Y. Soullard, R. Tavenard, E. Anquetil, J. Delahaie, "KIHT: Kaligo-based Intelligent Handwriting Teacher", *in : DATE 2024*, Valencia, Spain, March 2024, https://inria.hal.science/hal-04359877.

[12] F. Imbert, E. Anquetil, Y. Soullard, R. Tavenard, "Mixture-of-experts for handwriting trajectory reconstruction from IMU sensors", *in : Symposium International Francophone sur l'Ecrit et le Document (SIFED'2024)*, Nantes (France), France, June 2024, https://hal.science/hal-04811994.

[13] F. Imbert, R. Tavenard, Y. Soullard, E. Anquetil, "Domain adaptation for handwriting trajectory reconstruction from IMU sensors", *in : ICDAR 2024 Workshops, ADAPDA*, *Lecture Notes in Computer Science*, *14935*, Springer Nature Switzerland, p. 3–11, Athènes, Greece, August 2024, https://hal.science/hal-04605593.

[14] A. Sahbi, C. Alec, P. Beust, "Automatic Ontology Population from Textual Advertisements: LLM vs. Semantic Approach", *in : 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES)*, *246*, p. 3083–3092, Séville, Espagne, France, September 2024, https://hal.science/hal-04814520.

[15] A. Yesilkanat, Y. Soullard, B. Coüasnon, N. Girard, "Full-page music symbols recognition: state-of-the-art deep models comparison for handwritten and printed music scores", *in : Document Analysis Systems. DAS 2024*, p. 327–343, Athens, Greece, August 2024, https://hal.science/hal-04268139.

Miscellaneous

[16] I. Barchouch, E. Anquetil, N. Girard, "Interpretation of semi-structured hand-drawn sketches for the conception of a Pen-Based Intelligent Tutorial System", SIFED 2024 : Symposium International Francophone sur l'Écrit et le Document, June 2024, Poster, https://inria.hal.science/hal-04876641.