

ANDREA GEMELLI

Ph.D. Student at AI Lab, University of Florence

📍 Florence, Italy
✉ gemelli.andrea@hotmail.com
☎ +39 331 995 2716
🌐 andrea-gemelli
🐙 andreagemelli
🐦 Andr3aGemelli
🆔 0000-0002-6149-8282



EXPERIENCE

Visiting Researcher

Computer Vision Center

📅 March 2022 - ongoing 📍 Barcelona, Spain

I am deepening my knowledge in GNNs and Document Analysis during my visiting research period in one of the best research centers in Europe. We are applying deep learning techniques (including Computer Vision models and Transformers) to the task of key information extraction from challenging documents, such as receipts, forms and bills.

Ph.D. Student

University of Florence

📅 November 2020 - Ongoing 📍 Florence, Italy

So far, I have mainly worked on the **Document Analysis** field, exploiting **Graph Neural Networks** and **Natural Language Processing** techniques. I have attended courses and conferences about the latest discoveries on several Deep Learning fields and state-of-the-art models.

A list of the courses I have attended during my Ph.D. period:

- *Probabilistic Graphical Models* by Prof. Manfred Jaeger;
- *Sequence Learning and Explainable AI* by Prof. Paolo Frasconi;
- *Memory Networks* (AIDA course) by Prof. Alberto Del Bimbo;
- *Graph neural networks and neural-symbolic computation* (UCA Deep Learning School 2021) by Prof. Marco Gori.
- *Towards developmental learning*, by Prof. Marco Gori

Summer School of Machine Learning

Cambridge University - Ellis Unit

📅 11rd – 15th Jul 2022 📍 Cambridge, UK

I attended the **Cambridge ELLIS Machine Learning Summer School** and poster presentation, selected with 130 other students and professionals (260 total applications). I was able to attend presentations by some of the most prominent figures in the Machine Learning community (DeepMind, Google, and Cambridge University), as well as meet other remarkably talented colleges. I also presented my research work at the poster session provided by the summer school.

Summer School of Document Analysis

Luleå University of Technology

📅 23rd – 27th Aug 2021 📍 Luleå, Sweden

I have taken part to the **4th IAPR Summer School on Document Analysis**, a premier annual summer school focused on recent developments in document analysis, following expert's talks and taking part to hands-on laboratory sessions and an hackathon competition.

PUBLICATIONS

👤 Conference Proceedings

- Bimbo, D. D., Gemelli, A., & Marinai, S. (2022). **Data augmentation on graphs for table type classification**. In *Iapri joint international workshops on statistical techniques in pattern recognition (spr) and structural and syntactic pattern recognition (sspr)*, august 26-27, Montréal Québec: **S+SSPR22**.
- Gemelli, A., Biswas, S., Civitelli, E., Lladós, J., & Marinai, S. (2022). **Doc2Graph: a Task Agnostic Document Understanding Framework based on Graph Neural Networks**. In *Text in everything workshop (tie) @ eccv 2022*, august 21-25, Tel Aviv, Israel: **ECCV22**.
- Gemelli, A., Vivoli, E., & Marinai, S. (2022). **Graph Neural Networks and Representation Embedding for Table Extraction in PDF Documents**. In *26th international conference on pattern recognition, (icpr 2022)*, august 21-25, Montréal, Québec: **ICPR22**.

📖 Journals

- Pisaneschi, L., Gemelli, A., & Marinai, S. (2022). **Automatic generation of scientific papers for data augmentation in document layout analysis**. *Under Review*.

EDUCATION

Master in Computer Science

University of Florence

📅 Sep 2017 – Sep 2020

Thesis: "Structural layout analysis of document images using Graph Convolutional Neural Networks" | 110/110

Bachelor in Computer Science

University of Florence

📅 Sep 2014 – Sep 2017

Thesis: "Exploiting Convolutional Neural Networks for the detection of walls in floorplan images"

ADDITIONAL EXPERIENCE

Thesis supervision

University of Florence

📅 November 2020 - ongoing

📍 Florence, Italy

I am supervising some students on their bachelor and master graduation thesis, about these topics:

- *Graph Neural Networks per il riconoscimento del layout di documenti scientifici*: we exploited GNNs to recognise layout objects inside scientific papers;
- *Classificazione di tabelle in articoli scientifici utilizzando Graph Neural Networks*: we attempted to classify different kind of tables in scientific documents given their structure and content;
- *Entity Grouping in scientific papers by Link Prediction on Graphs*;
- *Automatic document generation* (master thesis): we exploited GANs to generate new scientific papers, for augmentation purposes.)

Bachelor Course Teaching

Polo Universitario Aretino

📅 March 2022 - June 2022

📍 Arezzo, Italy

I am teaching the bachelor course of *Introduzione di Sistemi Informativi* (Introduction to Informative Systems) as a tutor, held by the University of Arezzo and Politecnico di Milano. Main contents:

- Components of an Informative System and the Enterprise Architecture;
- the BOAT and BPMN frameworks;
- Data Warehouse and Data Mining;
- Distributed architectures and Cloud Computing;
- Security in Informative Systems;
- ArchiMate model.

Erasmus+ Traineeship

La Rochelle Université

📅 Jan 2020 - Mar 2020

📍 La Rochelle, France

During my stay at L3i lab of La Rochelle University I have studied the history of graphs in structural pattern recognition and developed a **GNN model to automatically detect structural layouts inside invoice documents** (part of my master thesis work).

Erasmus+ Study

Ghent University

📅 Sep 2018 - Feb 2019

📍 Ghent, Belgium

At the Ghent University I was able to deepen my knowledge and acquire new skills in one of the best universities in Europe, attending courses such as **Machine Learning, Software Hacking and Protection, Parallel and Distributed Software Systems** and **IoT**.

Data Scientist

Alias2k

📅 Apr 2018 - Sep 2018

📍 Arezzo, Italy

I have applied **machine learning techniques** to the company's already developed tools. I have extended an online tool for evaluating the prices of construction machinery rental. The application has been presented personally to the CEO of the Avesco Rent company in Lausanne.

CERTIFICATES



Esame di Stato

I obtained the Italian qualification / license for the profession of Information Engineer with maximum score (60/60)
· Released by University of Florence in September 2021.



English C1

Released by C.L.A. UNIFI (University of Florence), in November 2017



French B1

Released by C.L.A. UNIFI (University of Florence), in November 2017

SKILLS

Programming Languages

Python

Typescript

C++



Frameworks

Pytorch

DGL

Spacy

Angular



SOFT SKILLS

Communication & Public Speech

Flexible

Team worker

Enthusiastic learner

LANGUAGES

Italian

English

Spanish

French

