



Internship proposal - Computer Vision & Machine Learning Intern

<https://www.vialink.fr>

Join Vialink as a Computer Vision & Machine Learning R&D Intern !

Vialink is a REGTECH software company specializing in security and compliance. We develop SaaS software offering to fully digitize the customer relationship process: authentication, first contact, contracting.

In this context, the Know Your Customer (KYC) process helps to identify people and protect against identity theft and risk of fraud. This process is based on the collection, verification and automated management of the elements constituting a customer file (ID documents, RIB, supporting documents), and, moreover on the collection of information from external data sources (Taxes, Sirene, Infogreffe, etc).

There are 2 potential topics for this internship:

1. **Forgery Detection and Spotting in Document Images based on Deep learning approaches.**
2. **Complex document Layout Analysis based on Deep Learning approaches.**

You will join a team of security experts, data scientists, machine learning engineers whose values are passion, quality, rigor and benevolence.

We are looking for people who are passionate, enthusiastic.

M2 student / last year engineering school specializing in Machine Learning and / or Computer Vision

- Having experience of at least one academic project or external, with focus in Machine Learning and / or Computer Vision
- Solid maths background : linear algebra, statistics
- Knowledge of at least one deep learning architecture : CNN, UNET, RNN, BERT
- Knowledge of programming language : Python. Java / C++ are considered a plus
- Knowledge of common libraries and at least one framework : scikit-learn, TensorFlow, Keras, PyTorch

5 good reasons to join Vialink and the KYC team

- Security is a high-growth field
- KYC needs are exponential in companies
- Vialink is a profitable start-up
- R&D teams are autonomous and proactive
- Our processes are agile and constantly evolution

For more information: <https://www.welcometothejungle.co/fr/companies/vialink>

Contact

Mail : ludovic.toublanc@vialink.fr

Vialink

33 Place Ronde, 92800 Puteaux
Station La Défense (M1, RER A, T2)



Internships description

1. Forgery Detection and Spotting in Document Images based on Deep learning approaches.

The rapid growth of technology, especially in image processing softwares, has given rise to fraudulent image documents, with no obvious traces, generating a high demand for automatic fraud detection. Modification of sensitive information documents like first, last names of the owner of the document to impostor ones have strong consequences in bank and financial areas, for instance.

The aim of this internship is to propose one detection and localization method of falsified information (sometimes sensitive) done on document image. The problem could be approached as (but not limited) anomaly detection based on deep learning approaches.

The main goals are:

- *State of art to detect fraudulent documents.*
- *Find and implement the suitable model.*
- *Evaluate accuracy and performance of the model of detection of fraudulent documents and localization of forged parts.*

The research will be conducted on a public dataset as well as on a private one.

2. Complex document Layout Analysis based on Deep Learning approaches

Layout analysis from a document image plays an important role in document content understanding and information extraction systems, such as knowledge / information extraction and handwriting recognition. Then, key information needed for our client will be well extracted.

The aim of this internship is to improve our information extraction method using the new methods to analyze and represent complex document structures based on deep learning approaches. The problem could be addressed by using Visual features and Multi-modal features approaches.

The main goals are:

- *State of art on Document Layout Analysis.*
- *Find and implement the suitable architecture to segment and represent a document in coherent and independent blocks of data.*
- *Evaluate accuracy and performance of the model.*

The research will be conducted on a public dataset as well as on a private one.

For more information: <https://www.welcometothejungle.co/fr/companies/vialink>

Contact

Mail : ludovic.toublanc@vialink.fr

Vialink

33 Place Ronde, 92800 Puteaux
Station La Défense (M1, RER A, T2)