



# Valida-pdf

## AI-based document forgery detection for fraud prevention

**Valida-pdf by Gradiant assesses the trustworthiness of digital documents to prevent fraud in documents such as invoices, payroll slips, bank statements, KYC documents, etc.**

Our technology allows companies to offer an extra level of security and prevent fraud by analysing documents and detecting forgeries through AI-based forensic techniques.

- ✓ **Automatically detects forgeries in pdf documents: invoices, payslips, etc.**
- ✓ **Does not require connection to external databases (e.g. fraud databases) to detect modifications**
- ✓ **Does not require the original document to detect modifications**
- ✓ **Easy and fast integration via API**

## Functionalities

✓ <b>File forensics</b>	Examine every detail of a document including the file structure and pixels to uncover evidence of fraud.
✓ <b>Metadata inconsistencies</b>	Uncover traces of suspicious behaviour and attributes in the metadata of documents.
✓ <b>Intra-word inconsistencies</b>	Examine characteristics of the words such as spacing or size and find anomalies from fraudulent modifications.
✓ <b>Overlapped content</b>	Find overlapping contents in order to detect malicious editions.
✓ <b>Pdf reference based traces</b>	Examine anomalies in the pdf code language.

*All the modules described, except for the metadata inconsistencies, allow the detection of the exact region where the document has been modified*

## Performance

### Accuracy on real-world data

- ✓ Forgery detection:
  - Accuracy = **92.27%**
  - TPR: 87,05%
  - FPR: 2.5%

### Time consumption

- ✓ For a reference server:
  - i7 4790S
  - RAM: 16 Gbytes.
- ✓ **PDF: ~0.6 s** for a single page PDF

## Integration

REST API with sample code for multiple platforms and languages

Code samples provided for Shell, HTTP, JavaScript, Node.js, Ruby, Python, Java, and Go  
SaaS / Dedicated Cloud / On-Premises dockerized Deployment

## Recommended deployment requirements

**OS:** Linux 64 bits (recommended Ubuntu 18.04) with Docker installed.

**Hardware:** Google Cloud n1-standard-4<sup>1</sup> (4 CPUs, 15 GB RAM) or equivalent

---

1 <https://cloud.google.com/compute/docs/machine-types>

