الجمهور ية الجزائر ية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للاعلام الالي 8 ماي 1945

École Supérieure en Informatique 8 Mai 1945 Sidi Bel Abbès

MÉMOIRE

En vue de l'obtention du diplôme de I**ngénieur**
Filière: **Informatique**
Spécialité: **Intelligence Artificielle et Sciences de données (IASD)**

Thème

AI INVOICE CAPTURE SOLUTION

Présenté par: Saidi Mouslim

- Soutenu le: 30/09/2025
- devant le jury composé de :

Dr. ELOUALI Nadia	MCA	Présidente
Dr. DAOUDI Sara	MCA	Examinatrice
Dr. MALKI Abdelhamid	MCA	Encadrant

Année Universitaire : 2024-2025

ABSTRACT

Invoice capture remains a significant bottleneck in business operations, characterized by labor-intensive, error-prone manual data entry from diverse and unstructured document formats. This engineering thesis presents the AI Invoice Capture Solution, a robust system designed to automate this critical task. The approach is multidisciplinary, drawing heavily on Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), and advanced Computer Vision techniques.

Specifically, the solution integrates traditional Optical Character Recognition (OCR) with the cognitive power of Large Language Models (LLMs) to enhance extraction accuracy and adaptiveness. The implementation leverages modern transformer-based models, including LayoutLMv3 and the OCR-free Donut model, alongside Gemini AI Interaction, after essential image preprocessing steps (e.g., noise removal, table enhancement).

The primary objective is to develop a scalable, high-performance system capable of handling diverse invoice layouts, ensuring high data extraction accuracy, and offering seamless integration with existing software management systems via a Flask API. This work demonstrates the practical potential of combining state-of-the-art AI architectures to solve long-standing, real-world document processing challenges, ultimately contributing to greater efficiency and reduced operational costs.

Keywords : Large Language Models, Deep Learning, NLP, Computer Vision, Optical Character Recogintion, Pre-trained Models, Fine-tuning,

RÉSUMÉ

La capture de factures demeure un goulot d'étranglement important dans les opérations commerciales, caractérisée par une saisie manuelle de données laborieuse et sujette aux erreurs, provenant de formats de documents divers et non structurés. Ce mémoire de fin d'études d'ingénieur présente la Solution de Capture de Factures par IA (AI Invoice Capture Solution), un système robuste conçu pour automatiser cette tâche critique. L'approche est multidisciplinaire, s'appuyant fortement sur l'Apprentissage Automatique (AA), l'Apprentissage Profond (DL), le Traitement du Langage Naturel (TLN) et des techniques avancées de Vision par Ordinateur.

Plus précisément, la solution intègre la Reconnaissance Optique de Caractères (OCR) traditionnelle avec la puissance cognitive des Grands Modèles de Langage (LLMs) pour améliorer la précision de l'extraction et l'adaptabilité. L'implémentation utilise des modèles modernes basés sur les transformateurs, notamment LayoutLMv3 et le modèle Donut (sans OCR), ainsi que l'Interaction Gemini AI, après des étapes essentielles de prétraitement d'image (par exemple, suppression du bruit, amélioration des tableaux).

L'objectif principal est de développer un système évolutif et performant capable de gérer diverses mises en page de factures, d'assurer une grande précision dans l'extraction des données et d'offrir une intégration transparente avec les systèmes de gestion de logiciels existants via une API Flask. Ce travail démontre le potentiel pratique de la combinaison d'architectures d'IA de pointe pour résoudre des défis de traitement de documents réels, anciens et complexes, contribuant ainsi à une plus grande efficacité et à une réduction des coûts opérationnels.

Mots clés : Grands modèles de langage, Apprentissage profond, Traitement du langage naturel (NLP), Vision par ordinateur, Reconnaissance optique de caractères (OCR), Modèles pré-entraînés, Fine-tuning

مُلخَّص

لايزال التقاط الفواتير يمثل عنق الزجاجة الرئيسي في العمليات التجارية، ويتسم بإدخال البيانات يدويًا الذي يتطلب جهدًا كبيرًا تقدم أطروحة الهندسة هذه حل التقاط الفواتير بالذكاء .وعرضة للأخطاء، وذلك من تنسيقات المستندات المتنوعة وغير المهيكلة يعتمد هذا النهج على أساس .، وهو نظام قوي مصمم لأتمتة هذه المهمة الحاسمة(AI Invoice Capture Solution) الاصطناعي ، وتقنيات (NLP)، ومعالجة اللغة الطبيعية (DL)، والتعلم العميق (ML)متعدد التخصصات، مستندًا بشكل كبير إلى التعلم الآلي .متقدمة في رؤية الكمبيوتر

(LLMs)التقليدي مع القوة المعرفية للنماذج اللغوية الكبيرة (OCR)على وجه التحديد، يدمج الحل التعرف البصري على الحروف Donutونموذج LayoutLMv3يستخدم التنفيذ نماذج المحولات الحديثة، بما في ذلك .لتعزيز دقة الاستخراج والقدرة على التكيف مثل إزالة الضوضاء وتحسين)، بعد خطوات معالجة الصور الأولية الضرورية Gemini AI، إلى جانب تفاعل (OCRالخالي من) .(الجداول

قادر على التعامل مع تخطيطات الفواتير المتنوعة، وضمان دقة عالية في الهدف الأساسي هو تطوير نظام قابل للتطوير وذي أداء عالٍ يوضح هذا العمل .(API) Flaskاستخراج البيانات، وتوفير تكامل سلس مع أنظمة إدارة البرامج الحالية عبر واجهة برمجة تطبيقات الإمكانات العملية للجمع بين أحدث بنى الذكاء الاصطناعي لحل تحديات معالجة المستندات الواقعية والقديمة والمعقدة، مما يساهم في نهاية المطاف في زيادة الكفاءة وخفض التكاليف التشغيلية

الكلمات المفتاحية :

النماذج اللغوية الكبيرة، التعلم العميق، معالجة اللغات الطبيعية ،النماذج المدربة مسبقاً، الرؤية الحاسوبية، التعرف البصري على الحروف، الضبط الدقيق.