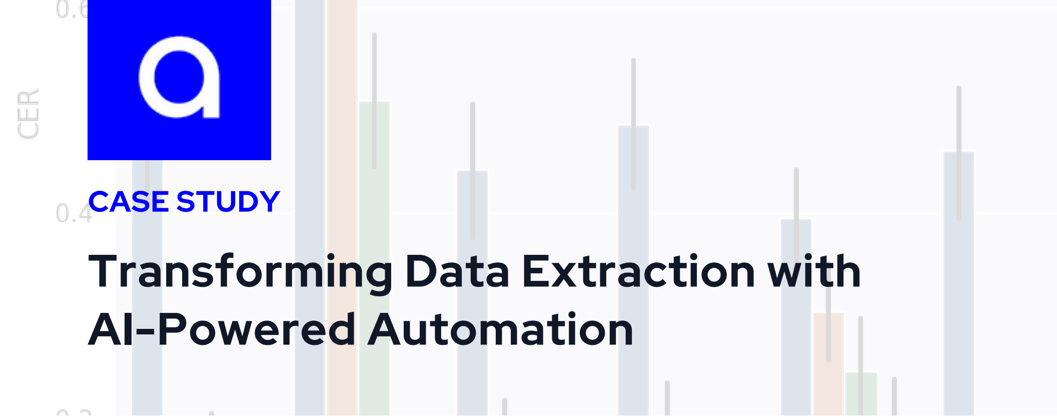


CASE STUDY

Transforming Data Extraction with AI-Powered Automation

CENTEGIX

Nearshore Solutions

AI Development
AI Consulting

Technologies

YOLO8
YOLO11
GOT_OCR
Tesseract
EasyOCR
PaddleOCR
Python
Jupyter Notebooks
GPU servers for training and inference

BACKGROUND

CENTEGIX is a leading provider of rapid incident response safety solutions designed to minimize identification, notification, and response times during emergencies. Their Safety Platform™ integrates dynamic digital mapping, real-time location tracking, user-friendly wearable panic buttons, visitor management, and reunification features to enhance preparedness and response across various sectors.

THE CHALLENGE

Centegix faced two primary obstacles in developing an automated solution for extracting data from driver’s licenses: technical complexities and budget feasibility.

Technical Challenges

Varied Data Structures: Driver’s licenses vary significantly in format, font type, and layout across regions, making universal data extraction complex.
Image Quality: Poor resolution, inconsistent lighting, and low text-background contrast posed challenges for OCR accuracy.
Real-Time Processing: The system needed to operate efficiently under real-time constraints, requiring optimized inference times without sacrificing accuracy.
Limited Training Data: The project relied on a relatively small synthetic dataset of 600 images, requiring creative augmentation techniques to improve model performance.
Error-Prone Fields: Variability in fields like dates and addresses made them particularly difficult to extract accurately.

Budget Feasibility Challenges

Centegix encountered difficulty finding a partner who could address these challenges within a manageable budget. Many providers proposed hourly billing models, leading to unpredictable and potentially unpayable costs.
They needed a solution that was not only effective but also aligned with their financial constraints. Azumo stood out as the only provider offering a practical, cost-effective solution tailored to Centegix’s specific needs. This approach ensured both technical success and financial feasibility.

SOLUTION

To address these challenges, Azumo implemented a Proof of Concept (POC) with the following approach:

Object Detection

Model Selection: Evaluated YOLO8 (nano and small) and YOLO11 (nano and medium) for their balance of speed and accuracy.

Training Process:

- Used a synthetic dataset of 600 driver’s licenses generated from the Roboflow dataset.
- Applied dynamic augmentation techniques during training to simulate diverse scenarios.
- Split the dataset into training (70%), validation (20%), and testing (10%) subsets.
- Conducted 80 epochs of training using minimal hyperparameter tuning for rapid iteration.

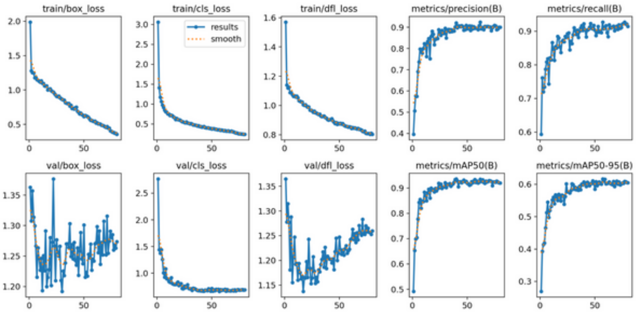


Figure 1. Training metrics like box_loss and mAP.

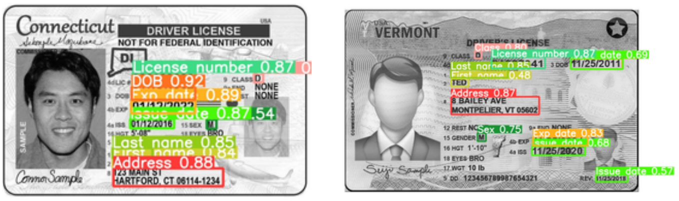
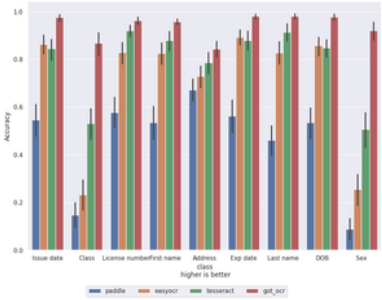
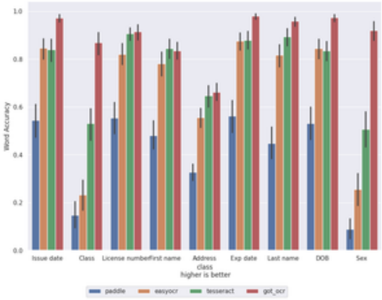


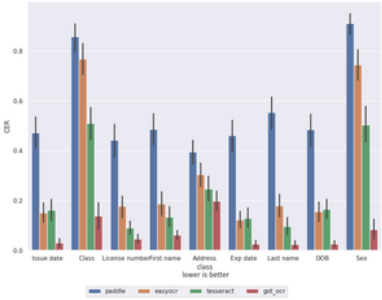
Figure 2. Example of detected boxes on an image.

CASE STUDY

SOLUTION







Optical Character Recognition (OCR):

Library Evaluation: Tested four OCR libraries (Tesseract, EasyOCR, GOT_OCR, PaddleOCR) for text extraction accuracy and speed.

Preprocessing Techniques:

- Ran tests with raw images and preprocessed versions (OTSU binarization) to assess the impact on accuracy.
- Observed that preprocessing improved performance in some cases, but fine-tuning preprocessing techniques would be necessary for production readiness.

Performance Metrics:

- **Character Error Rate (CER):** GOT_OCR had the lowest error rate, making it the most accurate.
- **Levenshtein Ratio:** Showed high similarity between extracted text and ground truth for most fields.
- **Sequence Ratio:** Highlighted challenges with multi-line fields like addresses.
- **Inference Time:**
  - EasyOCR (GPU version) was the fastest.
  - GOT_OCR, despite being slower, achieved the highest accuracy.

Evaluation and Insights:

- Detailed confusion matrices highlighted strong detection performance, with minimal misclassification.
- For fields like names and addresses, GOT_OCR outperformed others, but errors persisted in low-quality images or those with non-standard fonts.

Recommendations for Future Work:

- Enhance training datasets with greater diversity in license types, illumination conditions, and text-background contrasts.
- Optimize preprocessing and model configurations for specific data extraction tasks.
- Experiment with hybrid OCR setups (e.g., combining a fast library for initial results with a slower, more accurate library for refinement).

RESULTS

The POC delivered key outcomes:

Detection Accuracy: YOLO11m achieved a mean average precision (mAP) exceeding 80%, demonstrating the viability of object detection for extracting license fields.

OCR Performance: GOT_OCR emerged as the most accurate OCR library, achieving over 80% accuracy for critical fields.

Efficiency Trade-offs: EasyOCR provided the fastest inference times, while GOT_OCR offered the best balance of accuracy and robustness.

Actionable Recommendations: The analysis identified specific improvements for future iterations, such as fine-tuning hyperparameters, enhancing datasets, and testing under more diverse conditions.

Foundation for MVP: The findings validated the feasibility of automating license data extraction and provided a roadmap for transitioning to a Minimum Viable Product (MVP).

>80%

Detection accuracy with the YOLO11m model.

80%

Sccuracy for critical fields on GOT_OCR.