

Deep Learning Analytical Thinking – Fluendo AI Intern Assessment

En 2024 varios conductores británicos recibieron multas de hasta £105 por no pagar el peaje del Dartford Crossing... aunque jamás pasaron por allí. La causa fue que las cámaras ANPR confundieron matrículas muy parecidas (un solo carácter de diferencia o coches del mismo color y modelo).

1. What do you think it occurred during this model development (training & evaluation)?

El error que ocurrió en el entrenamiento y la evaluación del sistema ANPR pueden haber ocurrido estos fallos:

1. Sesgo y falta de dataset

El peso del entrenamiento no fuera correcto debido a la falta de ejemplos para el entrenamiento que contengan ejemplos “difíciles” de distinguir (por ejemplo: 0/ O, C/G) o también añadir imágenes que tengan poca iluminación, reflejos...

2. Etiquetado incorrecto

Algunas imágenes puede que no tuvieran la matrícula completa o estaban mal rotuladas, eso conlleva a que no aprenda de manera correcta como es la matrícula.

3. Sobreajuste y validación poco “real”

Se midió la calidad del sistema con situaciones que nos son reales, puede haber sido por utilizar cámaras diferentes, clima nocturno...

4. Métrica de optimización inadecuada

Por ejemplo se maximizo la exactitud de caracter a caracter y no se penalizo suficiente a la hora de que se equivoque en un carácter (que para este caso es realmente importante).

2. How would you fix this behavior? Please provide at least 2 options explaining their pros and drawbacks

Opción	Descripción	Ventajas	Desventajas
Ampliar y limpiar el dataset	Recopilar más imágenes reales y realizar también variaciones extremas para un acierto exhaustivo	+Mejor generalización +Reduce el sesgo a largo plazo	-Coste de captura y etiquetado -Necesidad de volver a entrar periodico
Pipeline híbrido con reglas de verificación	Tras el OCR, comparar la salida con • Regex de formato UK • Distancia de edición ≤ 1 frente a la base de matrículas reales	+ Rápido de implementar + Filtra errores groseros sin tocar la red	– Puede bloquear matrículas válidas poco comunes – Añade dependencia de bases externas
Arquitectura multi-etapa (CRAFT → CRNN + confianza)	1) CRAFT para detectar texto 2) CRNN/TrOCR para reconocer 3) Descartar predicciones con $confidence < \tau$	+ Estado del arte + Fácil de extender a otros dominios	– Más latencia y memoria – Mayor complejidad de mantenimiento

3. What do you think it will occur when running this AI in a different country with different plates formats? How would you ensure system accuracy?

Hay que tener en cuenta muchos factores, el principal son las matrículas ya que muchos países cambia la tipografía, el número de letras y números y todo esto conlleva que exista una menor precisión, para todo esto nuevo habría que realizar un entrenamiento exhaustivo. Plan para garantizar exactitud:

- Fine-tuning con ejemplos locales bien anotados.
- Generar datos sintéticos para cubrir casos raros.
- Lanzar en modo sombra (A/B) y medir F_1 antes del cambio.
- Monitorizar deriva y lanzar alertas si suben los falsos positivos.

4. Do you know any OCR (Optical Character Recognition) algorithms (Deep learning based) that could be used here?

Modelo	Idea clave	Uso típico
CRNN (Guo et al., 2016)	CNN + RNN + CTC	Secuencias cortas (matrículas, captchas)
ASTER / MORAN	Añaden rectificación geométrica previa	Texto curvado o distorsionado
CRAFT + CRNN (Baek et al., 2019)	Detección de texto + reconocimiento	Pipelines industriales
EAST (Zhou et al., 2017)	Detector FCN extremadamente rápido	<i>Edge</i> devices, tiempo real
Tesseract 5 (LSTM	OCR open-source re-entrenable	Soluciones con bajo presupuesto

5. Explain a Computer Vision / Artificial Intelligence project in which you have participated (goals, your role, difficulties you found, how they were solved, ...)

A lo largo del curso he desarrollado varias prácticas relacionadas con AI o Computer Vision la que yo he encontrado más parecida a este ejemplo es sobre Aprendizaje por Refuerzo (RL). Esta es el **estudio comparativo de algoritmos en CliffWalking-v0 (modo *slippery*)**.

Un breve resumen es:

Aspecto	Detalle
Título	Comparativa de Value Iteration, Estimación Directa, Q-Learning y REINFORCE en CliffWalking-v0
Objetivo	El objetivo era analizar ventajas e inconvenientes de cada método en un MDP mayor y estocástico. Esto se realizó evaluando cómo influyen los hiper parámetros.
Rol personal	Este fue un trabajo en equipo en mi caso lo que implementé fue es el diseño experimental, además de la implementación en Gymnasium + PyTorch, y por último la redacción del informe final
Entorno	El entorno se basaba en Grid 4× más grande que FrozenLake, las recompensas eran densas negativas al caer por el acantilado

Dificultades encontradas y soluciones

- **Espacio de estados/acciones 4× mayor.** Esto conllevaba tiempos de cómputo altos la solución encontrada fue una vectorización de la dinámica, la paralelización de episodios y **numba** para acelerar Value Iteration.
- **Alta varianza debida al ruido “slippery”.** Nuestra solución fue implementar una penalización exponencial al caer (logarítmico y una media de 100 episodios).
- **Sensibilidad a hiperparámetros.** Nuestra solución se basó en la búsqueda automatizada con Optuna (Q-Learning) y programación de *learning-rate* (REINFORCE)
- **Comparabilidad entre métodos.** Para poder comparar los métodos utilizamos una métrica común: recompensa media por episodio, distancia de Levenshtein respecto a la política óptima y tiempo total de entrenamiento.