

010 광학 문자 인식 / OCR

Optical Character Recognition

이미지나 문서 속 문자를 인식해 디지털 텍스트로 변환하는 기술

- 사진, 손글씨 등 비정형 문서의 문자를 식별해 컴퓨터가 처리 가능한 데이터로 전환하는 인식 기술
- 문서 자동화, 스캔 데이터 분석, 자율주행 등에서 시각 정보를 언어 정보로 바꾸는 핵심 기반 기술

OCR의 배경

광학 문자 인식(OCR)은 이미지나 스캔된 문서 속 문자를 분석해 디지털 텍스트로 바꾸는 기술로, 1950년대 규칙 기반 문자 판독 장치에서 출발했습니다. 초기 OCR은 패턴 데이터베이스와 문자 형태를 비교해 일치 여부를 판단하는 방식이었으며, 인쇄된 활자나 단순 문서에 한정된 인식만 가능했습니다. 이후 통계적 분류와 전처리 기법이 도입되며 다양한 글꼴과 인쇄 상태를 처리할 수 있게 되었지만, 시스템이 스스로 학습하지는 못했습니다. 딥러닝과 시각 인식 기술이 발전하면서 OCR은 AI를 활용한 학습형 인식 시스템으로 전환되었고, 이제는 손글씨·표·사진 속 텍스트 등 비정형 데이터까지 정밀하게 인식하는 단계로 진화했습니다.

OCR의 작동 방식

현대의 OCR은 시각 인식과 언어 이해 기술을 결합해 작동합니다. AI 모델은 이미지에서 문자 영역을 탐지하고, 픽셀 단위로 형태를 분석해 문자를 분리합니다. 이후 딥러닝 알고리즘이 문자의 특징을 학습하고, 자연어 처리(NLP)가 문맥을 분석해 잘못 인식된 부분을 교정합니다. 이 통합 구조를 통해 AI OCR은 문자의 형태와 의미를 동시에 인식하며, 복잡한 배경이나 비표준 글꼴에서도 높은 정확도를 유지합니다.

OCR의 활용과 의의

OCR은 시각적 정보를 언어 데이터로 전환하는 핵심 기술로, 산업 전반의 자동화와 디지털 전환을 촉진합니다. 금융·행정 분야에서는 계약서·청구서·신분증 정보를 자동 인식하고, 물류·제조에서는 송장과 라벨을 읽어 분류 효율을 높입니다. 자율주행 차량은 표지판을 인식하고, 시각장애인 보조기기는 글자를 음성으로 안내합니다. 이러한 응용은 비정형 정보를 구조화해 AI 서비스의 확장을 가능하게 하는 기반으로 평가됩니다.

AI OCR로의 발전

초기의 OCR은 규칙과 패턴에 의존한 비학습형 인식 기술로, 제한된 글꼴과 명확한 인쇄물만 처리할 수 있었습니다. 반면 AI 활용 이후의 OCR은 딥러닝을 통해 문자의 형태뿐 아니라 맥락과 의미를 학습하며, 문서 구조와 내용까지 이해합니다. AI 모델이 데이터를 통해 스스로 인식 정확도를 개선하기 때문에, OCR은 '글자를 읽는 기술'을 넘어 '문서를 이해하는 지능형 인식 기술'로 발전했습니다.