

077 합성곱 신경망/CNN

Convolutional Neural Network

이미지·영상 등 시각 데이터를 인식·분석하는 딥러닝 신경망 구조

- 입력 이미지의 특징을 자동으로 추출해 패턴을 인식하는 딥러닝 모델
- 시각 정보 처리에 뛰어나 컴퓨터 비전의 핵심 기술로 사용

합성곱 신경망의 개념

CNN은 시각적 데이터를 효율적으로 분석하기 위해 고안된 인공신경망 구조입니다. 기존 완전연결 신경망이 모든 입력 특징을 동일하게 처리하는 것과 달리, CNN은 이미지의 국소 영역에서 특징을 계층적으로 추출하여 점차 더 복잡한 전체 패턴을 학습합니다. 즉, 전체 이미지를 한 번에 보는 대신 특징이 집중된 영역(국소 패턴)을 중심으로 학습합니다. 예를 들어 얼굴 이미지를 학습할 때 눈·코·입 등 개별 요소를 인식한 뒤, 이를 종합해 '얼굴'이라는 개념을 추론하는 식입니다. 이러한 구조는 인간의 시각 피질이 사물을 인식하는 방식과 유사하며, 이미지 분류·객체 탐지·자율주행 등 다양한 응용 분야에서 사용됩니다.



합성곱 신경망의 작동 원리

CNN은 작은 부분에서 큰 의미로 나아가며 이미지를 이해하는 계층적 구조로 이미지를 이해합니다. 먼저 입력층에서 사진이 픽셀 단위로 들어오면, 합성곱층이 작은 창(필터)을 움직이며 이미지 곳곳을 살펴보고 윤곽선, 색 변화, 질감 같은 기본 특징을 찾아냅니다. 다음으로 풀링층이 비슷한 정보를 묶어 크기를 줄이면서 핵심만 남겨 계산을 단순하게 만듭니다. 이러한 과정을 계층적으로 반복하며 CNN은 저차원 특징(모서리·윤곽)에서 고차원 특징(객체·얼굴 등)으로 발전된 표현을 학습합니다. 마지막 단계인 완전연결층에서는 앞서 모은 특징들을 종합해, 전체가 '사람 얼굴'인지 '자동차'인지 같은 최종 판단을 내립니다.

합성곱 신경망의 활용

CNN은 컴퓨터 비전(Computer Vision)의 중심 기술로, 이미지 분류·얼굴 인식·의료 영상 분석 등 시각적 정보 처리 전반에 활용됩니다. 또한 음성 스펙트로그램 등 비시각적 데이터에도 적용되며, 데이터를 직접 가공하지 않아도 스스로 특징을 학습하는 것이 장점입니다. CNN은 딥러닝 기반 컴퓨터 비전 기술의 혁신을 이끈 핵심 구조로, AI가 시각 정보를 이해하는 기반 구조로 자리 잡았습니다. 최근에는 트랜스포머 기반의 비전 모델이 등장했지만, CNN은 여전히 효율성과 안정성이 검증된 시각 인식의 표준 기술로 평가됩니다.