

세상이 어떻게 이미지를 보고, 만들고, 고치는지를 알려줄게

- 기계학습, 딥러닝, 이미지 인식, 영상 보정, 그리고 그림 그리는 AI 까지!

### 1. 기계학습(Machine Learning)과 딥러닝(Deep Learning)은 뭐가 달라?

이 두 개는 모두 '컴퓨터가 스스로 뭔가를 배우는 기술'이야. 근데 배우는 방식과 구조가 달라.

◆ 기계학습(Machine Learning)은 "사람이 알려준 규칙을 보고 배우는 컴퓨터"야.

예: "고양이는 귀가 뾰족하고 눈이 크다"  
→ 컴퓨터가 이 규칙을 기억해서, 다음에 고양이 사진을 보면 알아보는 방식

◆ 딥러닝(Deep Learning)은 "컴퓨터가 스스로 규칙을 찾는 똑똑한 방식"이야.

예: 수천 장의 고양이 사진을 보여주면, 컴퓨터가 스스로 "이런 게 고양이구나" 하고 규칙을 찾아냄  
→ 딥러닝은 훨씬 복잡하고, 사람보다 더 정밀하게 패턴을 찾을 수도 있어!

### 2. 딥러닝은 기계학습 안에 포함된 개념이야

기계학습은 더 큰 개념이고, 딥러닝은 그 안에 들어가는 한 갈래야.

기계학습 (Machine Learning)  
├── 전통적인 머신러닝 (SVM, Decision Tree, K-NN 등)  
└── 딥러닝 (CNN, RNN, Transformer 등)

→ 그래서 딥러닝도 기계학습의 한 종류라고 보면 돼!  
차이는, 딥러닝은 뇌처럼 생긴 신경망 구조를 써서, 더 깊고 복잡한 학습을 할 수 있다는 점이야.

### 3. 그럼 이미지 인식에서는 어떻게 다를까?

🧠 전통적인 머신러닝 기반 인식

사람이 먼저 "이런 특징을 보면 고양이야" 하고 알려줘야 해. (예: 귀 모양, 눈 위치, 수염의 방향 등)

컴퓨터는 그걸 숫자로 바꿔서 구분해  
→ 장점: 빠르고 단순함  
→ 단점: 사람이 알려준 정보만 사용할 수 있음

🧠 딥러닝 기반 인식 (특히 CNN)

CNN(합성곱 신경망)은 이미지 전체를 직접 보고, 선 → 모서리 → 형태 → 얼굴 전체처럼 점점 복잡한 특징을 자동으로 찾아내

→ 장점: 사람이 안 알려줘도 스스로 학습함  
→ 그래서 더 정확하고, 더 유연하게 인식 가능

### 4. 감독학습(지도학습)과 비지도학습(비지도학습)은 어떻게 연결될까?

딥러닝과 전통적인 머신러닝은 둘 다 지도학습(Supervised) / 비지도학습(Unsupervised) 방식으로 학습할 수 있어.

◆ 지도학습은?

- 문제 + 정답(라벨)이 같이 있는 경우
- 예: "이건 고양이 사진이야!"라고 알려주는 학습

→ 전통 머신러닝에서는 이 방식이 거의 대부분을 차지해. 왜냐하면 사람이 미리 특징을 뽑아서, 정답을 알려주는 데에 적합하거든.

→ 딥러닝에서도 CNN, RNN 같은 건 지도학습 기반이 많아. 하지만 딥러닝은 비지도학습도 훨씬 잘해!

◆ 비지도학습은?

- 정답 없이, 데이터만 주고 스스로 구조를 파악하는 방식

→ 전통 머신러닝에서는 K-means, PCA 같은 군집화나 차원 축소 정도만 가능했어.

→ 그런데 딥러닝은 오토인코더, GAN, BERT 사전학습처럼 복잡한 구조도 스스로 학습할 수 있어!

## 5. 흔들린 영상 보정(Digital Image Stabilization, DIS)도 발전 중이야

영상을 찍을 때 손이 흔들리면 화면이 출렁이지? 그걸 보정해주는 기술이 DIS 야.

### 🔍 옛날 방식은?

전통 머신러닝처럼, 특정한 '특징점'을 찾아서 "이 점이 여기서 저기로 움직였구나!" 하고 따라가며 보정했어.

→ 사람이 어떤 부분이 중요한지 미리 알려줘야 하니까, 지도학습처럼 작동한 거야.

### 🔍 요즘 방식은?

딥러닝을 써서 전체 영상을 보면서 "흔들림의 패턴"을 스스로 배워서 보정해.

→ 마치 CNN 이 이미지를 인식하듯,  
→ DIS 에서도 딥러닝이 등장해서 훨씬 더 부드럽고 자연스러운 보정이 가능해졌어!

## 6. 이제는 컴퓨터가 이미지를 '그리기'까지 해! (이미지 생성 기술)

사진을 분석하는 걸 넘어서, 이젠 그림을 만들어내기까지 해!

예: "할머니가 웃고 있는 사진을 그려줘"  
→ 컴퓨터가 진짜 같은 사진을 만들어줌

### 🎨 GAN (생성적 적대 신경망)

화가(Generator)와 감정사(Discriminator)가 싸우는 구조야. 화가가 그림을 그리고, 감정사가 "이건 가짜야"라고 잡아냄 → 둘이 경쟁하면서 점점 더 진짜 같은 그림이 나와

- ✅ 장점: 빠르고 선명한 이미지 생성
- ❌ 단점: 훈련이 어렵고, 가끔 이상한 그림이 나올 수도 있어

### 🎨 확산 모델(Diffusion) (25 수능 기출)

아무것도 없는 상태(노이즈)에서 점점 그림을 복원해나가는 방식이야.

예: 노이즈 → 점 → 선 → 형태 → 진짜 그림

→ 그림 퀄리티가 훨씬 자연스럽고 정교해

→ 요즘 인기 있는 이미지 생성 AI 들 (DALL·E 2, Stable Diffusion, Midjourney)은 다 이 방식이야

## 7. 그럼 GAN 은 이제 쓸모없어졌어?

❌ 아니야! GAN 은 여전히 아주 강력한 도구야.

GAN 이 잘 쓰이는 분야는 이래:

- 실시간 생성: 모바일, 영상 실시간 처리에 적합 (빠름)
- 데이터 증강: 훈련 데이터가 부족할 때 GAN 으로 가짜 데이터를 만들어서 보완
- 이미지 복원/변환: 흑백 사진을 컬러로, 얼굴 합성, 스타일 전환 등

→ GAN 은 상황에 따라 훨씬 더 좋은 선택이 될 수 있어!

## 8. 딱 요약해줄게

항목	GAN	확산 모델
학습 구조	경쟁 (화가 vs 감정사)	점점 그림을 복원
생성 속도	빠름	느림
그림 품질	선명하지만 불안정할 수 있음	디테일 풍부하고 자연스러움
훈련 난이도	민감하고 까다로움	안정적이고 일관됨
텍스트 → 이미지	상대적으로 약함	훨씬 강함 (대세 기술)

## 9. 마무리!

- 기계학습은 컴퓨터가 사람이 알려준 정보로 배우는 방식,
- 딥러닝은 컴퓨터가 스스로 정보 속에서 규칙을 찾아내는 방식이야.
- 딥러닝은 기계학습 안에 들어가는 더 강력한 방법이고,
- 전통적인 머신러닝은 지도학습 중심, 딥러닝은 지도/비지도 모두 가능
- 이미지 인식, 영상 보정, 이미지 생성 기술도 딥러닝 덕분에 훨씬 정교하고 강력하게 발전하고 있어.
- GAN 과 Diffusion 은 이미지를 만들어내는 두 가지 방식인데,  
요즘은 Diffusion 이 더 인기 많지만,  
GAN 도 상황에 따라 여전히 최고일 수 있어!

## 🤖 인공 신경망 완전 이해하기 - 1 편

“사람 뇌를 따라하는 컴퓨터?”

### 🧠 뇌에는 뉴런이 있다

사람의 뇌에는 \*\*뉴런(신경세포)\*\*이라는 게 있어.

그 수는 **약 860억 개!**

(진짜 많지? 사람 한 명당 10개씩 나눠줘도 부족해)

이 뉴런들은 서로 **전기 신호**를 주고받으면서  
우리가 생각하고 기억하고 움직이는 모든 걸 처리해.  
말하자면, **세상에서 제일 복잡한 컴퓨터가 바로 뇌야.**

### 🤖 컴퓨터도 그걸 흉내 냈다고?

맞아. 과학자들은 이렇게 생각했어:

“이 뉴런이라는 걸 흉내 내면  
컴퓨터도 사람처럼 뭔가 판단하고 배울 수 있지 않을까?”

그래서 만든 게 바로 **퍼셉트론(perceptron)**이야.

**퍼셉트론은 계산을 하는 뉴런 역할을 해.**

그리고 퍼셉트론을 수천 개, 수만 개, 많게는 수억 개  
연결하면

**인공 신경망(Artificial Neural Network)**이 되는 거야!

### 🧠 퍼셉트론 하나는 어떻게 생겼을까?

퍼셉트론은 이렇게 생겼어:

1. **입력값**들을 받는다 (예: 색깔 점수, 모양 점수)
2. 각각의 입력값에 **\*\*가중치(weight)\*\***를 곱한다  
→ 이걸 **\*\*이 정보가 얼마나 중요한지\*\***를 나타내는 숫자야
3. 다 더해서 **\*\*가중합(weighted sum)\*\***을 만든다
4. 그 합이 어떤 기준값(**임계치**)을 넘으면  
→ **1** 출력  
넘지 못하면 → **0** 출력

이런 방식으로 작동하는 게 퍼셉트론이야.

이 퍼셉트론이 진짜 뉴런처럼 모이면,

**이미지도 보고, 말도 듣고, 얼굴도 알아보고, 숫자도 분류할 수 있게 되는 거지!**

## 🧠 신경망은 어떤 구조일까?

신경망은 보통 이렇게 생겼어:

층 이름	하는 일
입력층	정보를 받아들이는 첫 번째 층
은닉층	받은 정보를 가공하고 정리하는 중간 층
출력층	최종 판단 결과를 내보내는 마지막 층

입력층은 정보만 받아.

출력층은 정리된 정보로 **결과만** 뱉어.

근데 그 사이에 있는 **은닉층(hidden layer)!**

애가 **진짜 핵심, 사고하는 부분**이야.

### 🔥 비유하자면?

- **입력층:** 감각기관 (눈, 귀처럼 정보를 받아오는 역할)
- **은닉층:** 뇌 (받은 정보를 조합하고 의미를 파악)
- **출력층:** 입 (결과를 말해주는 역할)

## 🤖 인공 신경망 완전 이해하기 - 2 편

“컴퓨터도 피드백 받고 성장한다고?”

### 🎮 처음엔 컴퓨터도 바보야

처음에 신경망은 **완전 무지 상태**야.  
“이게 사과인지 배인지?” 전혀 몰라.  
그냥 입력받은 숫자들을 **대충 계산해서** 결과를 내는 정도.

그래서 **정답을 알려주면서** 훈련시켜야 해.

### 🧠 정답을 주고 훈련하는 방식: 감독학습

이런 방식은 **감독학습(Supervised Learning)**이라고 불러.

예를 들어:

- “이건 사과야!” → 정답: 1
- “이건 바나나야!” → 정답: 0

컴퓨터는  
**입력값(색깔, 모양)과 정답값(사과: 1, 바나나: 0)을 보고**  
**출력값이 얼마나 틀렸는지** 계산해.

그럼 틀렸으면 어떻게 해야 해?

### 📦 정답과 다르면? → 역전파!

출력값이 정답과 다르면  
**\*\*그 차이(= 오차, error)\*\*를 계산해서**  
거꾸로! 다시 퍼트려!

→ 이 과정을 **\*\*역전파(Backpropagation)\*\***라고 해.

**출력층 → 은닉층 → 입력층** 방향으로  
“이 연결선(가중치)이 잘못돼 있었네” 하고  
**조금씩 가중치를 수정하는** 거야.

🔴 이게 바로 컴퓨터가 피드백 받는 방식이야!  
역전파 = 수학적으로 설계된 ‘피드백 시스템’이야.

사람이 “여기 틀렸네~” 하고 말해주는 대신,  
컴퓨터는 **수학 공식으로 자기가 스스로 오차를 계산해서**  
**스스로 더 잘하도록 고치는** 구조인 거지.

## 🔧 가중치는 어디서 조정할까?

질문! 사고는 은닉층에서 한다고 했는데,  
그럼 출력층, 입력층 가중치는 바꿀 필요 없지 않나?

→ **아니야! 꼭 필요해!**

- 출력층 연결이 이상하면 → 결과가 엉망됨
- 입력층 연결이 엉성하면 → 애초에 잘못된 정보가 들어감

즉, 사고는 은닉층에서 하지만  
**사고가 잘 되려면 입력 정보도 제대로 들어와야 하고,**  
**결과도 정확하게 나가야 해.**

🔴 그래서 **입력층-은닉층, 은닉층-은닉층, 은닉층-출력층**  
**모든 연결선의 가중치가 다 조정 대상이야!**

## 🤖 인공 신경망 완전 이해하기 - 3 편

“사고는 어디서 일어나고, 딥러닝은 뭐가 다른 걸까?”

### ✿ 사고는 은닉층에서 일어난다!

퍼셉트론이 모여 만들어진 은닉층(hidden layer),  
여기가 바로 생각을 정리하는 핵심 장소야.

입력층은 그냥 정보만 받아.  
출력층은 정리된 정보를 꺼내서 최종 판단을 내려.  
근데 진짜 '판단을 위한 계산'은 은닉층에서 벌어져!

### 🎯 은닉층이 하는 일, 예를 들어볼게

예를 들어 어떤 이미지를 보고 판단하려 해.

- 입력층:  
→ 빨간색 점수 0.9, 둥근 모양 점수 0.8
- 은닉층:  
→ “이거 과일일 가능성 높아 보인다”  
→ “사과 쪽에 가까운 패턴이야”
- 출력층:  
→ “이건 사과야!” (결정)

즉, 입력은 원재료,  
은닉층은 요리,  
출력층은 완성된 음식

### 🧠 은닉층이 많아지면 생기는 일

은닉층이 한 개만 있으면  
→ 간단한 문제만 해결 가능

은닉층이 여러 개 겹겹이 있으면?  
→ 훨씬 복잡한 문제도 해결 가능!

예:

- 손글씨 숫자 구별하기
- 고양이 vs 강아지 구별하기
- 사람 얼굴 인식하기
- 번역, 자율주행, 챗봇 등...

이처럼 은닉층이 많은 구조를  
딥러닝(Deep Learning)이라고 불러!  
→ 말 그대로 “깊은 학습”

### 🔥 비유로 말하자면?

층	역할	비유
입력층	재료를 받음	눈, 귀처럼 정보만 가져오는 감각기관
은닉층	가공, 분석	뇌처럼 생각하고 정리하는 중간 단계
출력층	판단 결과 전달	입처럼 말하는 곳 (결과 발표!)

### 👉 그럼 퍼셉트론과 뉴런은 1:1 대응?

뉴런 하나가 하는 일은 진짜 대단해.  
한 뉴런이 수천~수만 개의 다른 뉴런과 연결돼 있고,  
그 복잡한 신호를 조합해서 뇌 전체를 작동시켜.

그에 비해 퍼셉트론은  
입력 받고 계산 한 번 하고 출력 내보내는 구조.

→ 그래서 뉴런 하나 ≈ 퍼셉트론 수천~수만 개 분량이라고 보면 돼!

🔥 한 줄 정리: 사고는 은닉층에서 벌어지고,  
은닉층이 많아지면 컴퓨터도 복잡한 걸 이해할 수 있게 된다!  
그게 바로 딥러닝의 본질이야.

## 🤖 인공 신경망 완전 이해하기 - 4 편

“컴퓨터도 피드백 받고 성장한다고?”

### 📄 컴퓨터도 틀릴 수 있다

신경망이 처음 학습을 시작할 땐 **정답을 잘 맞히지 못해**.  
왜냐면 가중치도 엉망이고, 판단 기준도 대충이거든.

예를 들어:

- 사과 사진 넣었는데 → “이건 고양이야!” 라고 하면...?  
→ 당연히 틀렸지!

이때 **정답과 비교해서**

\*\*얼마나 틀렸는지(오차)\*\*를 계산해.

이걸 바탕으로 **가중치를 조정**하는데,

그 조정 신호가 바로 **역전파(Backpropagation)**야.

### 📄 역전파 = 수학으로 만든 피드백 시스템

역전파는 **피드백(feedback)**이야.

그런데 말 대신 **숫자로 피드백**을 주는 거야.

“방금 결과 틀렸어! 이 가중치 줄이고, 저쪽은 늘려!”

→ 이런 식으로 신경망은 자기 실수를 **조용히, 수학적으로 고쳐 나감**

그래서 역전파는 **학습의 핵심 엔진**이야.

### 🤖 어디까지 피드백이 전달될까?

질문!

“사고는 은닉층에서 일어나는데,  
입력층이나 출력층까지 피드백할 필요 있어?”

**정답: 당연히 있어!**

- 출력층 연결이 잘못돼 있으면 → 결과 자체가 뒤바뀜
- 입력층 연결이 이상하면 → 애초에 엉뚱한 정보가 들어옴

**은닉층만 똑똑하면 소용 없어.**

→ 결국 **입력은닉출력 전체 흐름이 정리**되어야

정확한 판단이 가능해지는 거야.

### 📄 정리 표: 누가 어떤 역할?

위치	사고함?	피드백 필요?	이유
입력층	✗	✓	들어오는 정보 품질 결정
은닉층	✓	✓	실제로 생각하고 판단하는 부분
출력층	✗	✓	결과 해석 정확도 결정

### ✓ 진짜 한 줄 정리!

👉 인공 신경망은

**사고는 은닉층에서,**

**성장은 역전파로,**

**정확도는 모든 층의 조정으로 완성된다!**