

1. 다음 빈칸에 들어갈 알맞은 단어는?

GAN(Generative Adversarial Network)은 서로 다른 두 개의 신경망, 즉 ( ① )와 ( ② )를 적대적으로(adversarial) 학습시키며 실제 데이터와 비슷한 데이터를 생성해내는 모델이다. 이 구조에서 ( ① )는 ( ② )를 속이기 위해, 즉 가짜 데이터를 진짜처럼 보이도록 만들어내는 역할을 수행하며, ( ② )는 입력으로 들어온 데이터가 실제(real)인지, 아니면 ( ① )가 만든 가짜(fake)인지를 구별하려고 노력한다.

① Generator / Discriminator

② Decoder / Encoder

③ Encoder / Decoder

④ Discriminator / Generator

⑤ Synthesizer / Verifier

2. 다음은 디지털 영상 처리의 주요 분류와 그 정의를 설명한 것이다. 이 중 옳지 않은 설명을 고르시오.

① 화소점 처리 (Point Processing) - 화소점의 원래 값이나 위치를 기반으로 화소값을 변경하는 처리 방식이다.

②. 영역 처리 (Area Processing) - 화소의 원래 값과 이웃하는 화소의 값들을 함께 고려하여 새로운 화소값을 계산하는 처리 방식이다.

③ 기하학 처리 (Geometric Processing) - 여러 프레임 간 연산을 통해 움직임이나 차이를 계산하는 처리 방식이다.

④ 프레임 처리 (Frame Processing) - 두 개 이상의 서로 다른 디지털 영상을 조합하여 새로운 화소값을 생성하는 처리 방식이다.

3. 다음 중 샘플링(Sampling) 또는 양자화(Quantization)에 대한 설명과 예시가 잘못 짝지어진 것은?

A. 샘플링 - 연속적인 영상의 위치 정보를 일정 간격으로 나누어 디지털 이미지의 픽셀 좌표를 얻는다.

B. 양자화 - 연속적인 밝기 값을 256단계로 나누어 각각의 정수 값에 대응시킨다.

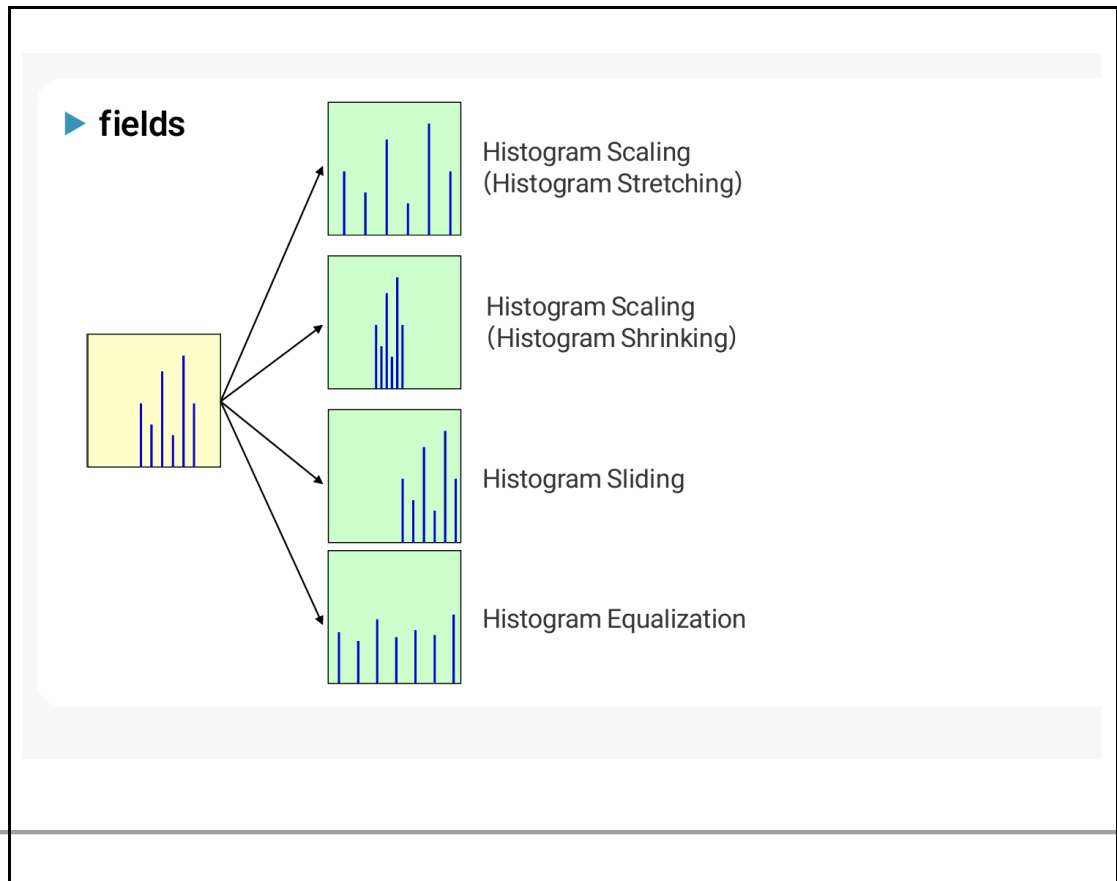
C. 양자화 - 영상의 밝기값을 전체 픽셀 중 최소값과 최대값을 기준으로 0~1 범위로 비율화한다.

비율화는 정규화에 해당함

D. 양자화 - 16비트의 밝기 데이터를 8비트로 줄여 표현 가능한 밝기 단계를 감소시킨다.

E. 샘플링 - 아날로그 영상에서 일정 시간 간격으로 프레임을 추출해 디지털 영상으로 저장한다.

4. 오른쪽은 왼쪽을 처리 후 그린 히스토그램이다. 다음 그림처럼 처리한 이후에 대한 설명으로 알맞은 것을 고르시오.



A. **Stretching** – 중앙에 몰린 히스토그램을 좌우로 퍼뜨려 명암 대비를 증가시킨다.

B. **Shrinking** – 원래 전체 밝기 범위를 사용하던 히스토그램을 좁은 구간에 몰리게 하여 대비를 감소시킨다.

C. **Sliding** – 히스토그램 전체를 좌우로 이동시켜 영상의 명암 대비를 조절한다.

D. **Equalization** – 특정 밝기에 집중된 히스토그램을 고르게 분산시켜 픽셀 분포를 균등하게 만든다.

5. 다음은 PNG 이미지에서 알파 채널과 BGR 색상 채널을 분리한 코드이다. 이 코드와 관련된 설명 중 옳지 않은 것을 고르시오.

```
src = cv2.imread("glasses.png", cv2.IMREAD_UNCHANGED)

mask = src[:, :, -1]

glass = src[:, :, 0:3]
```

- A. cv2.IMREAD\_UNCHANGED는 알파 채널까지 포함한 이미지를 읽는다.
- B. mask는 투명도 정보를 나타내는 알파 채널이다.
- C. glass는 R, G, B 3채널을 추출한 결과이다.
- D. mask는 BGR 이미지의 R 채널을 의미한다.
- E. 이 코드는 합성을 위한 전처리(채널 분리) 과정으로 해석할 수 있다.

6. 다음은 영상 필터링 기법에 대한 설명이다. 이 중 틀린 것을 고르시오.

- A. 평균값 필터는 커널 내 모든 픽셀에 동일한 가중치를 주어 계산하며, 주변 잡음을 줄이지만 경계 또한 흐려진다.
  - B. 가우시안 필터는 픽셀 간 거리에 따라 가중치를 주고, 가우시안 분포의  $\sigma$  값을 조절하여 블러의 정도를 조정할 수 있다.
  - C. 가우시안 필터는  $\pm 3\sigma$  범위 내 픽셀을 모두 고려하므로, 연산 성능은 평균값 필터보다 항상 뛰어나다.
- 평균값 필터보다 무거움
- D. 샤프닝 필터는 원본 영상에서 라플라시안(2차 미분)의 강조값을 더하거나 빼는 방식으로 경계를 강화한다.
  - E. 양방향 필터는 공간적 거리와 색상(강도) 차이 두 가지 조건을 함께 고려하여 경계는 보존하면서 노이즈를 줄이는 비선형 필터이다.

7. 다음 중 보간법(Interpolation)의 종류 및 원리에 대한 설명으로 틀린 것은?

- ① 최근접이웃보간법(Nearest neighbor)은 대상 위치에서 가장 가까운 정수 좌표의 픽셀 값을 그대로 복사하여 사용하며, 가장 빠르지만 영상 품질은 낮다.
- ② 이웃평균보간법(Neighbor averaging)은 주위 4개의 픽셀 값을 단순 평균하여 새로운 값을 계산하며, 최근접이웃보간법보다 품질이 향상된다.
- ③ 양선형보간법(Bilinear interpolation)은 가로·세로 방향으로 각각 1차 보간을 수행하여 총 4개 픽셀에 가중치를 곱해 결과를 구한다.
- ④ 고차보간법(Higher-order interpolation)은 보간 정확도를 높이기 위해 8개 또는 16개 이상의 픽셀을 사용하고, 속도가 빨라서 실시간 시스템에 적합하다.

#### 속도가 느림

- ⑤ 보간법은 일반적으로 정확히 대응되는 입력 픽셀이 존재하지 않을 때, 주변 픽셀 정보를 바탕으로 값을 추정하여 확대나 회전 등 기하학적 변환 시 활용된다.

8. 다음 중 OpenCV의 cv2.warpAffine() 함수에 대한 설명으로 틀린 것은 무엇인가?

- ① cv2.warpAffine() 함수는 2x3 크기의 아핀 변환 행렬을 입력받는다.
- ② 아핀 변환은 직선성, 평행성, 비율을 보존한다.
- ③ cv2.getRotationMatrix2D() 함수는 회전 중심과 각도를 바탕으로 아핀 행렬을 생성한다.
- ④ dsize 인자는 결과 이미지의 채널 수를 설정하는 값이다.
- dsize는 결과 이미지의 출력 크기 (너비, 높이)를 지정하는 인자입니다.
- ⑤ borderMode 인자는 이미지 외곽 부분의 보간 방식을 결정한다.

9. 다음 중 Hough Transform을 이용한 직선 및 원 검출에 대한 설명으로 틀린 것은?

- ① 허프 변환은 이미지 공간의 한 점을 파라미터 공간의 곡선으로 변환하여, 여러 점에서 발생한 곡선이 교차하는 지점에서 공통 직선을 찾는 방식이다.
- ② cv2.HoughLinesP()는 Canny Edge와 같은 에지 영상에 대해 직선을 검출하며, minLineLength는 선분의 최소 길이, maxLineGap은 선을 연결할 수 있는 최대 간격을 나타낸다.
- ③ Hough 변환에서 일반적인 직선 방정식  $y = ax + b$ 는 수직선에서 기울기 무한대가 되어 표현이 어렵기 때문에 OpenCV에서는 극좌표 형태인  $\rho = x\cos\theta + y\sin\theta$ 를 사용한다.
- ④ cv2.HoughCircles() 함수는 Grayscale 영상에서 원을 검출할 수 있으며, 내부적으로는 Canny Edge와 유사한 경계 정보를 이용하여 원의 중심과 반지름을 추정한다.

⑤ 허프 변환은 머신러닝 기반 기법으로 직선이나 원을 학습하여 검출하는 방식이며, 사전 학습된 모델이 필요하다.

10. 다음 중 Optical Flow에 대한 설명으로 틀린 것은?

**Optical Flow**는 컴퓨터 비전에서 연속된 두 영상(Frame) 사이에서 각 픽셀이 어떻게 이동했는지를 추정하는 기법

① Optical Flow는 연속된 두 이미지 간의 픽셀 단위 이동을 벡터 형태로 추정하는 기법이다.

② Optical Flow는 시간에 따른 픽셀의 밝기 변화를 이용하며, 같은 물체는 시간에 따라 밝기가 변한다고 가정한다.

**Optical Flow의 핵심 가정 중 하나는 밝기 불변성(Brightness Constancy)**

③ Lucas-Kanade 방법은 Optical Flow를 추정할 때 윈도우 기반으로 작은 영역의 움직임을 계산한다.

④ Optical Flow에서 각 픽셀의 움직임은  $(\Delta x, \Delta y)$  벡터로 표현되며, 전체 장면의 이동 정보로 시각화할 수 있다.

⑤ Optical Flow는 객체 추적, 모션 분석, 비디오 안정화 등 다양한 영상 처리 분야에 활용된다.

11. 다음 중 Depthwise Convolution 및 Pointwise Convolution에 대한 설명으로 틀린 것은?

① Depthwise Convolution은 각 채널마다 독립적으로 커널을 적용하며, 커널의

수는 입력 채널 수와 동일하다.

② Pointwise Convolution은 1x1 커널을 사용하여 채널 간 정보를 혼합하는 연산이며, 출력 채널 수만큼 커널을 가진다.

③ 일반적인 Convolution의 연산량은  $K^2 \cdot H \cdot W \cdot C_{in} \cdot C_{out}$ 이며, 이는 Depthwise + Pointwise를 합친 것보다 크다.

④ Depthwise Convolution은 채널 간 연산이 포함되어 있어 연산량은 Pointwise보다 많다. 빈도(frequency)만을 고려하여 각 문장을 **희소 벡터**로 변환합니다. **단어 간의 의미적 유사성은 전혀 반영하지 않습니다.**

⑤ Depthwise Separable Convolution은 MobileNet 등 경량화 모델에서 사용되며, 모델 크기와 계산량을 크게 줄이는 데 효과적이다.



12. 다음 중 자연어 처리(NLP)에서 틀린 설명을 고르시오.

- ① 토큰화(Tokenizing)는 문장을 단어 또는 의미 있는 최소 단위로 나누는 과정으로, 문장의 구조를 이해하기 위한 전처리 단계이다.
- ② **Word cloud**는 문서에 자주 등장하는 단어를 큰 글씨로 시각화하며, 데이터의 전반적인 특징을 파악하는 데 유용하다.
- ③ **CountVectorizer**는 단어의 빈도를 기반으로 벡터를 생성하며, 단어 간의 의미적 유사성도 반영한다.
- ④ **TfidfVectorizer**는 단어 빈도(TF)와 역문서 빈도(IDF)를 곱하여 단어의 중요도를 수치화하는 방법이다.
- ⑤ **Cosine similarity**는 문장을 벡터 공간에 투영한 후, 두 벡터의 방향 간의 유사도를 기반으로 문장 간 유사도를 계산한다.

13. 다음 코드와 설명을 보고, 이에 대한 설명 중 틀린 것을 고르시오.

```
texts = ["I love AI", "Deep learning is fun"]

tokenizer = Tokenizer()

tokenizer.fit_on_texts(texts)

sequences = tokenizer.texts_to_sequences(texts)
```

```
padded = pad_sequences(sequences)
```

```
model = Sequential()
```

```
model.add(Embedding(input_dim=1000, output_dim=64))
```

```
model.add(SimpleRNN(32))
```

```
model.add(LSTM(64))
```

```
model.add(GRU(32))
```

```
vector1, vector2 = [[0.1, 0.2, 0.3]], [[0.1, 0.2, 0.4]]
```

```
sim = cosine_similarity(vector1, vector2)
```

①Tokenizer와 pad\_sequences는 자연어 문장을 정수 시퀀스로 변환하고 길이를 맞춰주는 RNN 전처리 과정이다.

②Embedding 레이어는 단어 인덱스를 고정된 차원의 밀집 벡터로 바꾸며, 학습 중 자동으로 조정된다.

③SimpleRNN, LSTM, GRU는 모두 시퀀스 데이터를 처리하며, LSTM과 GRU는 장기 의존성 문제를 보완한 구조이다.

④cosine\_similarity는 단어의 순서를 고려하여 문장 유사도를 계산하는 방식이다. cosine\_similarity는 두 벡터 간의 방향 유사도를 계산할 뿐, 단어의 순서나 문장의 구조는 고려하지 않음

⑤위 모델은 자연어를 벡터로 변환하고, 순환 신경망으로 처리한 후, 문장 간 유사도 계산까지 수행할 수 있는 흐름을 나타낸다.

14. 스타일이 내용 이미지에 전이(Transfer)되는 방식에 대한 설명으로 틀린 것은?

①출력 이미지  $y^{\text{What}\{y\}}y^{\wedge}$ 는 학습 중 스타일 손실과 내용 손실이 동시에 최소화되도록 Image Transform Network에 의해 생성된다.

②스타일 손실은 VGG-16의 여러 중간 레이어에서 추출한 Gram Matrix를 비교하여 계산되며, 이미지의 질감과 색감 등을 반영한다.

③내용 손실은 VGG-16의 깊은 레이어에서 추출한 특징 맵(Feature Map) 간의 차이를 기준으로 계산된다.

④학습 과정에서는 출력 이미지  $y^{\text{What}\{y\}}y^{\wedge}$ 의 픽셀 값을 조정하여, VGG-16이 예측한 스타일 분류 확률이 Style Target과 같아지도록 만든다.

분류를 하지 않음

⑤최종적으로 스타일 손실과 내용 손실을 결합한 total loss를 기반으로 Image Transform Net의 파라미터가 업데이트된다.

## 15. 다음은 Neural Radiance Fields (NeRF)의 구조 및 작동 원리에 대한 설명이다. 이 중 틀린 것은 무엇인가?

①NeRF는 3D 장면을 표현하기 위해 볼륨 렌더링 기법을 사용하며, 해당 위치에서의 색과 밀도를 예측하는 신경망을 학습한다.

②NeRF는 위치와 시점을 입력받아 RGB 색상과 밀도( $\sigma$ )를 출력하며, 이를 통합하여 픽셀 색을 계산한다.

③위치 정보를 효과적으로 인코딩하기 위해 Fourier 기반의 Positional Encoding을 적용한다.

④NeRF는 각각의 픽셀마다 레이(ray)를 추적하며, 그 레이와 교차하는 3D 오브젝트 표면을 삼각형(mesh) 기반으로 직접 모델링한다.

⑤ 학습된 NeRF는 novel view synthesis에 활용되어, 학습에 사용되지 않은 시점에서도 장면을 생성할 수 있다.

16. 다음은 설명가능한 AI (XAI)에 대한 설명이다. 이 중 틀린 것은 무엇인가?

① Model-Agnostic 기법인 LIME과 SHAP은 다양한 모델에 적용 가능하며, 모델 내부 구조를 알 필요가 없다.

② Grad-CAM은 CNN의 마지막 convolution layer의 gradient를 기반으로 class-specific heatmap을 생성한다.

③ CAM은 FC layer 뒤에 global max pooling layer를 사용해 feature map을 시각화하는 방식이다.

CAM은 Global Average Pooling (GAP)을 사용합니다.

④ Integrated Gradient는 입력과 기준값(Base Line) 사이의 경로상의 변화량을 적분하여 feature 기여도를 계산한다.

⑤ DeepLIFT는 기울기 기반 기법보다 gradient 소실 문제를 완화할 수 있다는 장점을 갖는다.

17. 강화학습에서 사용하는 용어에 대한 설명으로 옳지 않은 것은 무엇인가?

① 정책(Policy)은 에이전트가 어떤 상태에서 어떤 행동을 선택할지 결정하는 전략이다.

② 상태 가치 함수  $V(s)$ 는 현재 상태에서 출발하여 최적으로 행동했을 때 받을 수 있는 누적 보상의 기대값이다. >> 최적 가치 함수

③ 감가율(Discount Factor)은 미래 보상의 중요도를 현재 기준으로 얼마나 낮출지를 결정하는 값이다.

④ 상태 전이 확률은 어떤 상태에서 특정 행동을 했을 때 다음 상태로 갈 확률을 의미한다.

⑤ 행동 가치 함수  $Q(s, a)$ 는 특정 상태에서 어떤 행동을 했을 때 기대되는 누적 보상을 의미한다.

18. 강화학습을 위한 환경을 설정하고 동작시키는 다음 설명 중 틀린 것은 무엇인가?

- ① `env = gym.make("FrozenLake-v1", map_name="4x4", is_slippery=False)`는 4x4 크기의 FrozenLake 환경을 생성하며, `is_slippery=False`로 설정하면 에이전트는 미끄러지지 않고 정확한 방향으로 이동한다.
- ② `state, _ = env.reset()` 코드를 통해 환경을 초기화하며, 이때 반환되는 `state`는 에이전트의 시작 위치를 의미하는 정수 값이다.
- ③ `env.action_space.n`을 통해 현재 환경에서 가능한 행동의 개수를 확인할 수 있으며, FrozenLake 환경에서는 4개의 방향(좌, 하, 우, 상)이므로 4가 출력된다.
- ④ `obs, reward, terminated, truncated, info = env.step(2)`를 호출하면 에이전트는 왼쪽 방향으로 이동하며, 다음 상태, 보상, 종료 여부 등을 받을 수 있다.
- ⑤ `env.unwrapped.P`는 환경의 내부 전이 확률 정보를 담고 있으며, 상태와 행동에 따른 결과 확률을 확인하는 데 사용된다.

19. 강화학습에서 정책을 정의하는 함수 `pi(state)`가 다음과 같이 주어졌다고 할 때 이 정책의 목적과 작동 방식으로 틀린 설명은 무엇인가?

```
def pi(state): return 0 if state[0] >= stick_threshold else 1
```

- ① `state[0]`이 `stick_threshold` 이상일 경우 행동 0(Stick)을 선택한다.
- ② `stick_threshold` 값이 클수록 플레이어는 더 적극적인 전략을 취한다.
- ③ 행동 0과 1은 각각 Stick과 Hit을 의미하며, 상태값에 따라 결정된다.
- ④ 정책 `pi`는 deterministic(결정적)하며, 동일한 `state`에 대해 항상 같은 행동을 선택한다.
- ⑤ `stick_threshold`를 0으로 설정하면 무조건 Hit만 하게 된다.

20. 다음 중 의료 영상 분석에서 폐렴(Pneumonia) 진단을 위한 딥러닝 모델 구성 방식으로 적절하지 않은 것은?

(1) xray\_dataset = ImageFolder(data\_path, transform=data\_transforms)와 같이 ImageFolder를 사용하여 X-ray 이미지 데이터를 폴더 구조에 따라 분류 데이터셋으로 불러올 수 있다.

(2) models.resnet18(pretrained=True)로 불러온 모델의 fc layer를 nn.Linear(512, 2)로 교체하여 정상/폐렴 2개 클래스를 분류할 수 있도록 수정해야 한다.

(3) DICOM 포맷은 .dcm 확장자를 가지며, 일반적으로 torchvision.datasets.ImageFolder를 통해 직접 로드할 수 있다.

(4) Chest X-ray는 폐렴 진단에 주로 사용되는 의료 영상이며, CNN 기반의 모델은 픽셀 강도의 차이를 활용해 병변 유무를 학습한다.

(5) train\_transforms = transforms.Compose([...]) 구문을 활용해 이미지 크기 변환, 정규화 등의 전처리를 수행하여 학습 성능을 향상시킬 수 있다.

