

<기능화 캠퍼스 프로젝트 SW2 공간조사 2023 254019 이준혁>

1. 불균형 데이터셋의 문제점과 함께 그것을 해결하는 방법을 설명하라.

• 데이터가 적은 클래스에 대해 학습정확도가 낮아질, 적은 데이터로 인해 overfitting 과적합 발생

• 방법: 데이터 전처리, 원본 데이터를 무작위로 증강시킴으로써 데이터셋의 다양성과 크기를 모두 향상시킨다.

• 기하학적: 수평/수직대칭, 수평, 수직 이동, 회전/확대, 축소 → 불균형 크기, 위치.

• 노이즈, 변형에 대한 내성을 갖도록 함.

• 화소적: 밝기, 대조, 채도, 노이즈 등.

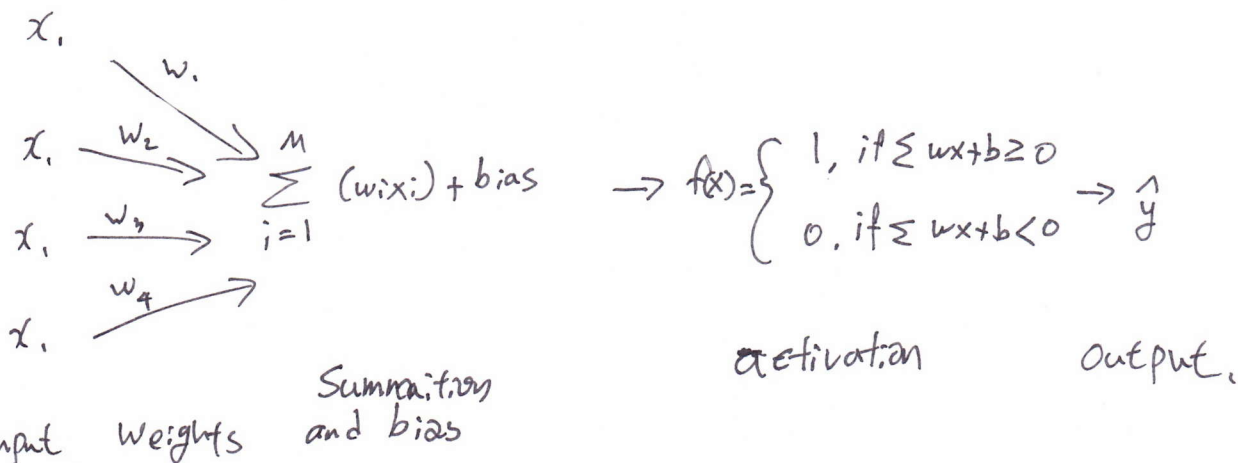
• off-line data augmentation.

Translate / Rotate / Shear.

해결방법

- 데이터 추가 수집, 데이터 증강, 손실함수 계산시 클래스가중치 부여, 클래스 오버샘플링, 클래스 언더샘플링

2. 퍼셉트론의 구조를 그림으로 설명하시오



3. 전이학습에 대하여 설명하시오

- 모든 가중치를 초기화하고 처음부터 학습하는 것. (training from scratch)
- Transfer learning (전이학습) : 사전학습된 모델을 이용하여 학습하는 것. \rightarrow Data, time 이 적음.
- 사례 1) pretrained network as a classifier
- 사례 2) 전이학습 : pretrained network as a feature extractor.
- 사례 3) 미세조정 (fine-tuning) 학습률을 작게 설정.

4. CNN 아키텍처의 특징과 주요 구성요소 및 구조를 서술하시오.

- 다층 퍼셉트론의 발전된 형태이며 이미지 처리에 적합한 신경망.
- 인간과 같은 시각 처리와 유사하고 2D 특징을 추출하고 학습.
- 같은 크기의 DNN이 비해 파라미터 수가 매우 적다.

<구조>

- 특징 추출 네트워크 : convolution layer (+ Activation function) + pooling layer \rightarrow feature map 생성
- 분류 네트워크 : Fully connected layers + Sigmoid / softmax \rightarrow 다중/단일 class (완전연결 레이어)
- 히든 네트워크 : fully connected layers \rightarrow 수리함.

5. Convolution 연산의 가중치

parameter의 총 수를 계산하는 과정은 그림으로 설명하시오

- 입력 영상 : $h \times w + 1C$
- 필터 크기 : $k_h \times k_w$
- 출력 영상 : $o_h \times o_w \times o_c$

가중치 매개변수의 총 수.

$$(k_h \times k_w \times \text{입력채널} + 1) \times \text{출력채널}$$

$$= (k_h \times k_w) \times C_i + 1 \times C_o$$

$$\sum_{i=1}^{C_i} \text{cube}_{k_h, k_w} \times \text{cube}_{k_h, k_w} = \text{cube}_{o_h, o_w}$$

$$\text{cube} * \begin{pmatrix} \text{cube} \\ \vdots \\ \text{cube} \end{pmatrix}_{C_i \text{개의 채널}} = \text{cube}$$

6. 2차 역전파와 가중치 소멸문제를

설명하고 이에 해결방법을 설명하시오.

- 2차 역전파법은 딥러닝 모델의 가중치와 편향을 순차적으로 조정하는 방법으로 주어진 입력 데이터에 대한 예측값과 실제값의 차이인 오차를 역전파하여 각층의 가중치와 편향을 조정하는 알고리즘이다.

단점! 가중치 소멸문제 : 심층 신경망에서 입력층으로 갈수록 2차 gradient가 줄어들어 학습이 안되는 현상.

해인책으로 그라디언트를 계산하는 과정과 학습률의 피보라이 각자 균등히 이어가게 되실.

\rightarrow 시그모이드 \rightarrow Relu로 대체한다.

7. Confusion matrix를 그림으로 그려
정밀도, 재현율, F1-Score에 대해
숙달하기 설명하시오.

	positive	negative	
positive	true positive (TP)	False negative (FN)	Sensitivity $\frac{TP}{TP + FN}$
negative	False positive (FP)	true negative (TN)	Specificity $\frac{TN}{TN + FP}$
	precision $\frac{TP}{TP + FP}$	negative predictive value $\frac{TN}{TN + FN}$	Accuracy $\frac{TP + TN}{TP + TN + FP + FN}$

정밀도와 재현율의 조화평균

$$F_1\text{-Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

8. Gradient 기반 최적화 알고리즘의 종류와 그 특징에 대해 설명하시오

- SGD: 각 파라미터를 점사 하강법으로써 Gradient Decent를 동일하게 동작한다. 현재 가중치
매개 변수의 값에서의 손실 함수 미분값을 이용하여 가중치를 업데이트 한다.
- Momentum: SGD는 local minimum에 매우 취약한 알고리즘이다. 속도 개선은 위하여
모멘텀 힘을 사용하여 지역 최소값을 더 적게, SGD보다 빠르게 수렴한다.
- NAG: Momentum을 반영한 기법으로써 기하학적 속도를 동시에 고려하여
다음 이동 속도와 방향을 결정한다면 NAG는 일단 현재 속도를 이동
기하학적 고려하여 다음 이동 방향을 정한다.
- Adagrad: 각각의 매개 변수에 서로 다른 학습률 적용하여 학습한 그라디언트에
대해 큰 업데이트를 수행한다. 학습률이 강한 R치에 유용하다.
- RMSprop: Adagrad의 학습률이 지나치게 감소하는 문제를 해결하기 위해 최근
최근 그라디언트에 대해서만 지수 평균을 사용하여 학습률을 조정,
Adagrad의 효율적인 업데이트와 유사하지만 빠른 수렴과 함께 학습률 감소
조절.

8. 학습이 overfitting을 방지하기 위한 규제 기법에 대하여 설명하시오.

- L1, L2 규제는 기계학습 모델이 훈련 데이터에 너무 맞추어져서 일반화 성능이 저하되는 현상-드레퍼

• L1 규제: 모델의 손실함수에 가중치(w)의 절대값의 비례하는 항을 추가, 희소성 증가

• L2 규제: 모델의 손실함수에 가중치(w)의 제곱에 비례하는 항을 추가 복잡도를 줄이고 일반화 성능을 향상.

• 드롭아웃: 학습과정 중 네트워크의 일부 뉴런을 랜덤하게 제거하여 모델을 강제로 단순화. 과적합을 줄이고 일반화를 성능강화한다.

• 조기종료: 모델의 과적합될때 일반화 성능을 향상한다.

• 앙상블 학습: 여러 다른 모델의 예측을 결합하여 더 강력한 모델을 만드는 기법.