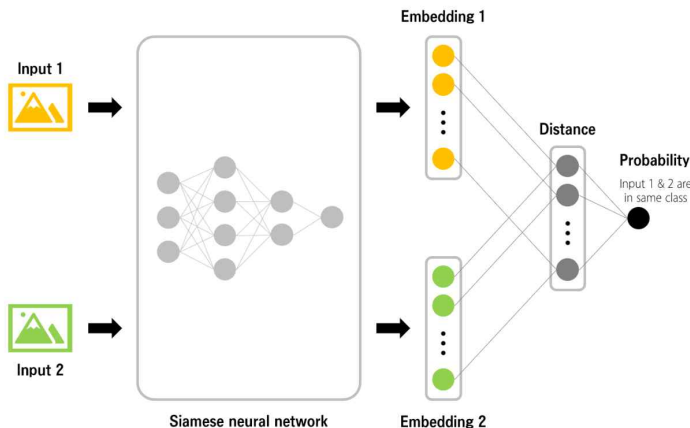


## 2024-1학기 창의학기제 주간학습보고서 (3주차)

창의과제	세종대학교 집현캠퍼스를 개선시킨 웹서비스 개발				
이름	신찬영	학습기간	4월 1일 ~ 4월 12일		
학번	23012094	학습주차	3	학습시간	3
학과(전공)	인공지능	과목명	자기주도창의전공1	수강학점	3
※ 수강학점에 따른 회차별 학습시간 및 10주차 이상 학습 준수					
금주 학습목표	얼굴의 비슷한 정도를 나타내주는 SiamNetwork모델 학습시키고 성능 테스트 하기				
학습내용	<p>삼 신경망(siamese neural network)는 가중치와 구조가 같은 인공 신경망에 두 입력을 넣은 뒤 출력을 비교하는 인공 신경망이다. 이 네트워크 구조 자체는 2005년 에 Yann LeCun 이라는 교수 연구팀에 의해 발표되었다. 그리고 2015년에 이 네트워크에 신경망을 접목시킨 삼 신경망이 발표되었다.</p> <p>삼 네트워크를 이해하기 위해서는 One-shot learning 뭔지 알아야한다. 딥러닝은 기존 모델들에 비해 강력한 모델이다. 그러나 이 딥러닝 모델을 학습시키기 위해 정말 많은 양의 파라미터를 생성하고 학습시켜야 한다. 이를 위해서는 많은 데이터도 필요하다. 반면 사람은 적은 데이터로도 학습이 가능하다. 딥러닝 모델이 인간처럼 소량의 데이터만으로 학습할 수 있게하는 것이 Few-shot learning이고 극단적으로 한 장의 데이터 만으로 학습 할 수 있게 하는 것을 One-shot learning이라고 한다.</p> <div data-bbox="335 1220 1029 1646">  </div> <p>위의 사진은 삼 네트워크의 구조를 그림으로 표현한 것이다. 위 사진의 과정을 설명하자면 다른 두 사진을 모델에 입력하고 모델의 출력값을 이용해 이 두 사진이 얼마나 유사한지 출력한다.</p> <p>학습과정</p> <p>1.먼저 두 개의 입력데이터의 모음을 준비한다.(여기서 첫 번째 데이터 모음에는 하나의 물체나 생물등의 사진을 준비한다. 두번째 데이터 모음에는 첫 번째 데이터의 사진 일부와 첫번째 데이터 사진과는 다른 별개의 사물이나 생물의 사진을 준비한다.)</p>				



2. 각 입력에 대한 출력값을 얻는다. (첫 번째 데이터가 통과한 값, 두 번째 데이터가 통과한 값)
3. 이 두 값 사이의 거리는 L1 norm이나 L2 norm 등 의 방법을 사용하여 계산한다.
4. 두 입력이 같은 사물이나 생물의 사진이라면 거리를 가깝게, 아니면 거리를 멀게하는 방식으로 삼 네트워크를 학습시킨다.

이 삼 네트워크는 손실함수로는 Contrastive loss라는 손실 함수를 이용한다.

$$\frac{1}{2} \left[ \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \lambda \right]$$

(어떤사람)      (같은 사람)      (다른사람)

이 손실함수는 유클리드의 거리를 이용한다.

### 학습방법



필자의 얼굴데이터를 첫 번째 데이터로 선정하고 필자의 데이터와 필자 조원의 사진을 섞은 데이터를 두 번째 데이터로 선정했다.

```
class SiameseNetwork(nn.Module):
    def __init__(self):
        super(SiameseNetwork, self).__init__()
        self.cnn1 = nn.Sequential(
            nn.ReflectionPad2d(1),
            nn.Conv2d(3, 64, kernel_size=3),
            nn.ReLU(inplace=True),
            nn.BatchNorm2d(64),

            nn.ReflectionPad2d(1),
            nn.Conv2d(64, 64, kernel_size=3),
            nn.ReLU(inplace=True),
            nn.BatchNorm2d(64),

            nn.ReflectionPad2d(1),
            nn.Conv2d(64, 64, kernel_size=3),
            nn.ReLU(inplace=True),
            nn.BatchNorm2d(64),

            )

        self.fc1 = nn.Sequential(
            nn.Linear(64*100*100, 500),
            nn.ReLU(inplace=True),

            nn.Linear(500, 500),
            nn.ReLU(inplace=True),

            nn.Linear(500, 5))

    def forward_once(self, x):
        output = self.cnn1(x)
        output = output.view(output.size()[0], -1)
        output = self.fc1(output)
        return output

    def forward(self, input1, input2):
        output1 = self.forward_once(input1)
        output2 = self.forward_once(input2)
        return output1, output2
```

삼 네트워크의 신경망 구조를 파이토치로 나타낸 코드이다. 삼네트워크 안의 CNN은 임의로 만든 것이다.



	<p>Dissimilarity: 8.23</p>  <p>Dissimilarity: 4.68</p>  <p>위의 사진은 출력값이다. 위의 Dissimilarity값이 낮을수록 유사한 것이다. 몇몇 사진은 잘 분류 했지만 대다수의 사진을 잘 분류하지 못했다.</p>
학습성과 및 목표달성도	필자의 데이터 외의 데이터가 한종류 밖에 없어서 값이 이상하게 나왔다. 또한 얼굴 이외의 배경까지 고려하여 모델이 잘 분류를 못한다.
참고자료 및 문헌	<a href="https://tyami.github.io/deep%20learning/Siamese-neural-networks/">https://tyami.github.io/deep%20learning/Siamese-neural-networks/</a> <a href="https://hipolarbear.tistory.com/25">https://hipolarbear.tistory.com/25</a>
내주 계획	두 번째의 데이터셋을 바꾸기 위해 모델을 찾아보고 학습되는 데이터의 양을 늘릴 것이다. 또한 모델에 입력되기전에 harrcascade모델에을 거친후 입력되게 코드를 구성할 것이다. 또한 핸드폰이나 사진으로 모델을 속여 보안인식을 통과할 수도 있으므로 이것을 방지하는 모델도 찾아볼 것이다.

2024 년 4 월 2 일

지도교수

(인)