|  |  |
| --- | --- |
| 교육 제목 | **딥러닝 GAN, autoencoder, transfer learning** |
| 교육 일시 | 2021년 12월 8일 수요일 |
| 교육 장소 | 영우글로벌러닝 2층 |
| **교육 내용** | |
| 오전 | 1. GAN    1. 생성적 적대 신경망(Generative Adversarial Networks), 줄여서 ‘GAN(간)’이라고 부르는 알고리즘    2. GAN은 딥러닝의 원리를 활용해 가상의 이미지를 생성하는 알고리즘    3. GAN이라는 이름에는 적대적(Adversarial, 서로 대립 관계에 있는)이란 단어가 들어 있는데, 이것은 GAN 알고리즘의 성격을 잘 말해 줍니다. 진짜 같은 가짜를 만들기 위해 GAN 알고리즘 내부에서는 ‘적대적’인 경합을 진행하기 때문    4. 생성자는 판별자가 구별할 수 없을 만큼 정교한 가짜를 만드는게 목표    5. 페이스북의 AI 연구팀이 만들어 발표한 DCGAN(Deep Convolutional GAN)    6. padding = 'same'이라는 설정을 통해 입력과 출력의 크기가 다를 경우 자동으로 크기를 확장해 주고, 확장된 공간에 0을 채워 넣을 수 있다    7. 배치 정규화(Batch Normalization)       * 배치 정규화란 입력 데이터의 평균이 0, 분산이 1이 되도록 재배치하는 것       * 다음 층으로 입력될 값을 일정하게 재배치하는 역할을 합니다.       * 이 과정을 통해 층의 개수가 늘어나도 안정적인 학습을 진행할 수 있습니다.       * 케라스는 이를 쉽게 적용할 수 있게끔 BatchNormalization() 함수를 제공합니다. |
| 오후 | 1. UpSampling2D() 함수는 이미지의 가로, 세로 크기를 2배씩 늘려줍니다. 7x7이 ➍레이어를 지나며 그 크기가 14x14가 되고, ➑레이어를 지나며 28x28이 되는 것입니다. 2. 이렇게 작은 크기의 이미지를 점점 늘려 가면서 컨볼루션 레이어(➎, ➒)를 지나치게 하는 것이 DCGAN의 특징입니다 3. Autoencoder    1. 인코더(Auto-Encoder, AE)    2. 오토인코더는 GAN과 비슷한 결과를 만들지만, 다른 성질을 지니고 있다. GAN이 세상에 존재하지 않는 완전한 가상의 것을 만들어 내는 반면에, 오토인코더는 입력 데이터의 특징을 효율적으로 담아낸 이미지를 만들어 냅니다.    3. 영상 의학 분야 등 아직 데이터 수가 충분하지 않은 분야에서 사용될 수 있습니다. 학습 데이터는 현실 세계의 정보를 담고 있어야 하므로, 세상에 존재하지 않는 가상의 것을 집어넣으면 예상치 못한 결과를 가져올 수 있습니다. 하지만 데이터의 특징을 잘 담아내는 오토인코더라면 다릅니다. 부족한 학습 데이터 수를 효과적으로 늘려주는 효과를 기대할 수 있지요. 4. transfer learning    1. 하지만 데이터가 충분하지 않은 상황도 발생합니다.    2. 딥러닝의 데이터 양이 충분하지 않을 때 활용할 수 있는 방법    3. 여러 방법 중에서 수만 장에 달하는 기존의 이미지에서 학습한 정보를 가져와 내 프로젝트에 활용하는 것을 전이 학습(transfer learning)이라고 합니다.    4. 방대한 자료를 통해 미리 학습한 가중치(weight)값을 가져와 내 프로젝트에 사용하는 방법으로 컴퓨터 비전, 자연어 처리 등 다양한 분야에서 전이 학습을 적용해 예측율을 높이고 있습니다. 5. 데이터의 수를 늘리는 ImageDataGenerator() 함수 6. 폴더에 저장된 데이터를 불러오는 flow\_from\_directory() 함수 7. 이미지넷    1. 1,000가지 종류로 나뉜 1백 20만 개가 넘는 이미지를 놓고 어떤 물체인지를 맞히는 ‘이미지넷 이미지 인식 대회(ILSVRC)’에 사용되는 데이터셋.    2. MNIST와 더불어 가장 유명한 데이터셋 중 하나이지요. 전체 크기가 200GB에 이를 만큼 커다란 이 데이터를 놓고 그동안 수많은 그룹이 경쟁하며 최고의 분류기를 만들기 위해 노력해 왔습니다.    3. 치매/일반인 뇌 사진 분류 프로젝트를 하고 있는 우리에게도 이 자료가 중요한 이유는, 지금부터 이 방대한 양의 데이터셋에서 추출한 정보를 가져와서 우리의 예측율을 극대화하는 ‘전이 학습’을 할 것이기 때문입니다.    4. ImageDataGenerator() 함수는 주어진 데이터를 이용해 변형된 이미지를 만들어 학습셋에 포함시키는 편리한 기능을 제공합니다. 이미지 데이터의 수를 확장할 때 효과적으로 사용할 수 있습니다. |