

정 석 우 한국과학기술원 항공우주공학과 석사과정

| e-mail : swjung92@kaist.ac.kr

심 현 철 한국과학기술원 항공우주공학과 부교수

| e-mail : hcshim@kaist.ac.kr

이 글에서는 사람이 직접 운전하지 않아도 주행이 가능하도록 하는 자율주행기술에 적용된 인공지능기술들에 대해 소개하고자 한다. 최근에는 사람 수준 또는 그 이상의 인공지능기술이 발달함으로써 자동차업계뿐만 아니라 많은 IT업계 또한 활발한 연구를 진행하고 있다.

자율주행 자동차는 기본적으로 ‘인지-판단-제어’의 3단계를 거치면서 작동하게 된다. 인지단계에서는 카메라·라이더(LiDAR)·레이더(RADAR) 등의 센서로 주변 환경을 인지한다. 판단단계에서는 인지된 환경 정보를 해석하여 안전하고 원활한 주행이 가능한 주행 경로를 생성한다. 마지막 제어단계에서는 판단 단계에서 나온 경로를 추종하도록 하는 가·감속 조향 제어를 수행한다. 각 단계에서 풀어야 할 문제들은 상당히 복잡한데, 이를 해결하기 위해서 적용된 인공지능기술을 소개한다.

### 심층 컨볼루션 신경망을 이용한 주행 환경 정보 인지

카메라는 사람의 눈과 가장 유사한 환경정보를 취득할 수 있기 때문에 테슬라, 메르세데스-벤츠 등 많은 자동차 회사들이 주력 센서로 사용하고 있다. 카메라는 기본적으로 색상정보뿐만 아니라 여러 카메라를 사용하여 다중 시점(Multi-view)을 구현하면 주변 환경의 3D 정보 또한 추출할 수 있는데, 정보가 풍부한 만큼 이를 분석하기 위해서는 고도의 알고리즘과

고성능 컴퓨터가 필요하다. 최근에는 딥러닝의 일종으로 인간의 시신경 작동 원리를 모방한 컨볼루션 레이어(Convolutional neural network)를 여러 층으로 중첩하여 사용하는 인공신경망기술이 영상인식분야에서 높은 정확도를 보이고 있다. 기존의 고전적인 영상처리에서는 영상 내에서 ‘특징점(Feature point)’을 추출한 뒤 분류하는 연산을 수행하였기 때문에, 환경을 인지하기 위해 주의 깊게 봐야 할 특징 점들을 선별하는 것에 큰 어려움이 있었다. 컨볼루션 신경망은 학습을 거듭함으로써 영상 내에서 주의 깊게 봐야 하는 점들을 추출한 특징 맵(feature map)을 적절하게 생성한다. 그림 1은 차량이 포함된 영상이 컨볼루

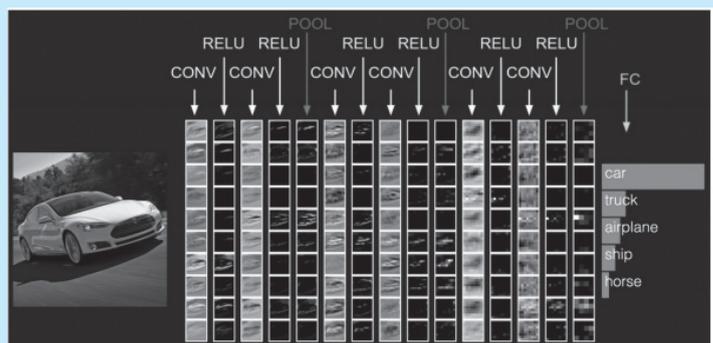


그림 1 컨볼루션 신경망에 의한 영상 내 물체 분류(Stanford University, CS231n)



그림 2 심층 컨볼루션 신경망을 이용한 표지판 탐지



그림 3 SegNet을 이용한 주행 환경 영상 분할(Badrinarayanan, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation")

선 레이어(그림 내 CONV로 표시)를 통과함에 따라 생성되는 특징 맵을 보여준다. 여러 층의 특징 맵 변환 과정을 통해 결과적으로 영상 내 물체가 어떤 물체인지 성공적으로 분류할 수 있다.

컨볼루션 레이어 기반의 네트워크를 설계하는 방법에 따라 다양한 인지 결과를 얻을 수 있는데, 일반적으로 영상 내에 많은 부분을 차지하는 물체의 종류를 분류(Classification)하는 데 많이 사용되고, 이 분류 알고리즘을 확장하여 영상 내에 특정 물체의 위치까지 탐지(Detection)하는 것도 가능하다. 그림 2는 심층 컨볼루션 신경망을 이용하여 전방 주행 영상에서 표지판을 탐지하는 알고리즘의 결과를 나타낸다. 또 다른

**딥러닝 기반의 인지 기술은 다양한 환경에서의 많은 데이터를 확보하는 것과 고성능 병렬 컴퓨팅이 가능한 GPU를 활용하는 것을 필요로 한다.**

예로 그림 3과 같이 입력 영상 내 모든 픽셀을 의미 있는 클래스로 분류하는 영상분할기법(Segmentation)이 있다. 영상 내의 모든 영역을 분류함으로써 동시에 여러 물체를 인지하고, 차선 및 주행 가능 영역을 판단할 수 있다. 2차원 영상 내에서 분류된 물체들은 라이다나 레이더 등 다른 센서와 융합하여 사용됨으로써 장애물들과의 거리 및 주변 환경의 3D 정보 또한 얻을 수 있다.

하지만, 심층 컨볼루션 신경망을 학습하기 위해서는 많은 데이터를 필요로 하기 때문에 자율주행 자동차를 개발하는 회사들은 많은 시간 주행 데이터를 취득하면서 학습을 거듭하여 인식 정확도를 높이는 작업에 매진한다. 또한, 고화질의 카메라 영상을 이용한 실시간 인지를 수행하기 위해서는 고성능 병렬 컴퓨팅이 가능한 장치를 필요로 한다. 그래픽 카드로 유명한 엔비디아 사는 올해 Drive PX2라는 자율주행차량용 슈퍼컴퓨터를 발표하였는데, 도시락 크기의 이 장치는 자그마치 맥북 프로 150개의 연산 능력을 가지고 있어 여러 센서로부터 데이터를 동시에 받고 복잡한 딥러닝 계산도 실시간으로 수행할 수 있다.

## SLAM을 이용한 차량 정밀 측위 기법

자율주행차가 정해진 목적지까지 안정적으로 가기 위해서는 현재 차량의 위치가 정확히 어디에 위치하고 있는지 알아야 한다. 미국의 포드 사는 3D 라이다 여러 개를 차량에 장착하여 주행함으로써 GPS 값, 차량 주행거리 정보, 레이저 점군 데이터를 융합하여 현재 차량 주변 환경의 3D 지도를 만들면서 동시에 현재 차량이 어디에 있는지를 계산한다. 이 기술은 SLAM(Simultaneous localization and mapping)이라고 부르는데, 부정확한 센서 값 및 환경의 변화들에 대응하여 정확한 위치를 추정하기 위해서는 인공지능의 일종인 기계학습 알고리즘이 요구된다. 이때 사용되는 알고리즘은 가우시안 혼합 모델(Gaussian mixture model)인데, 이 알고리즘을 사용함으로써 차량 주변 환경의 레이저 점군 데이터의 분포 특성을 파악하여 현재 차량의 위치를 비교적 정확하게 계산할 수 있다. 포드는 눈 오는 날이나 비 오는 날 등 날씨에 상관없이 자신들이 개발한 차량이 정밀하게 위치를 추정할 수 있고, 이를 기반으로 자율주행을 수행하였다는 것을 보여주었다.

### 인자-판단-제어를 한 번에 수행하는 End-to-End 딥러닝 기반 자율 주행

앞서 설명한 기술들의 경우에는 인식시스템과 경



그림 4 가우시안 혼합 모델을 이용한 차량 정밀 측위(포드, GTC 2016)

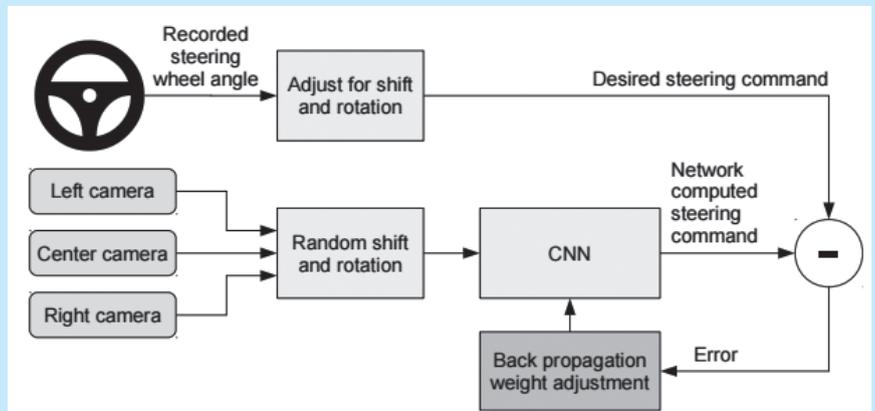


그림 5 심층 컨볼루션 신경망을 이용한 영상 입력에 대한 핸들 조향각 학습(엔비디아)

로 설계 및 제어 알고리즘을 각각 개별적으로 설계해야 한다. 또한, 주변에 대한 정보 및 경로 계획을 정확하게 계산해야 자율주행을 안정적으로 수행할 수 있다. 하지만, 사람의 경우에는 눈으로 들어오는 정보만을 이용하여 주변 정보를 정확하게 알지 못해도 안정적으로 운전을 한다. 자율주행 자동차 또한 이와 같은 방식으로 자율적인 운전을 해야 진정한 의미로 '인간과 유사한' 인공지능이라 할 수 있을 것이다. 엔비디아 사는 주행 상황에서 전방 카메라 영상과 그 순간의 핸들 조향각도를 짝지어 데이터 셋으로 제작하고 딥러닝을 이용한 학습을 진행하였다. 즉, 자율주행차가 '이 장면에서는 핸들을 그만큼 돌려야 한다'라는 것을 알 수 있도록 사람이 가르쳐주는 것이다. 학습된 딥러

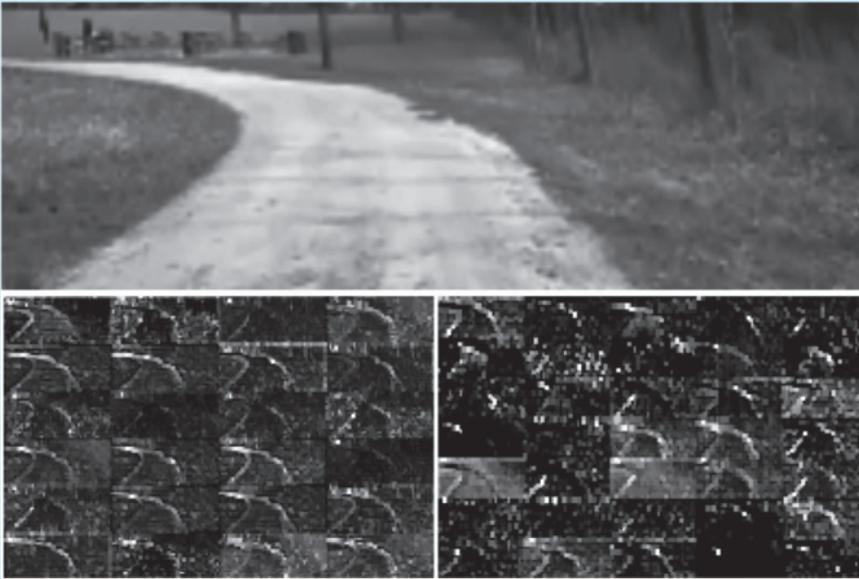


그림 6 컨볼루션 레이어에 의해 생성된 특징 맵(엔비디아) 도로 경계가 강조된 모습을 볼 수 있다.

닝 네트워크는 전방 영상 데이터가 입력으로 들어오면 올바르게 주행하기 위한 핸들 조향 각도를 출력 값으로 내보내기 때문에, 차량은 MDPS나 별도의 모터를 이용하여 핸들을 이 각도에 도달할 때까지 돌리기만 하면 된다. 이와 같이 인지부터 제어 명령까지 전체 시스템의 학습 가능한 변수를 동시에 학습할 수 있는 구조를 End-to-End 러닝시스템이라 부른다. 엔비디아 연구진에 따르면, 인식시스템과 판단 및 제어 시스템이 따로 나뉘어 있는 것보다 End-to-End 러닝시스템으로 동시에 학습을 진행하는 것이 효율적이고 성능도 더 좋다고 한다.

엔비디아 연구진은 이 연구를 진행하면서, 또 다른 흥미로운 결과를 얻을 수 있었다. 학습을 진행함으로써 자율주행 자동차가 전방 영상에서 주의 깊게 봐야 할 부분을 스스로 인지하게 되는 것이다. 전방 영상이 컨볼루션 레이어를 통과하여 나오는 특징 맵을 본 결과, 도로 경계면이 강조되게 생성됨을 알 수 있었다(그림 6). 즉, 자율주행 자동차는 핸들 조향과 관련해서 전방 영상에서 도로 경계를 주의 깊게 봐야 된다는

것을 알게 된 것이다. 영상에 대한 라벨링은 진행하지 않고 단지 사람의 운전 데이터와 전방 영상을 짝지어 학습시켰을 뿐인데, 영상 내에서 중요하게 봐야 할 요소들에 대한 학습도 내부적으로 진행된 것이다.

엔비디아는 72시간에 해당하는 수동 운전 데이터를 End-to-End 딥러닝을 이용하여 학습시켜 Garden State Parkway에서 10마일에 해당하는 고속도로를 운전자의 개입 없이 성공적으로 자율주행하였다.

## 결론 및 향후 전망

미국자동차공학회(SAE)가 제시하는 자율주행 단계는 0단계부터 5단계까지 있는데, 현재 판매 중인 차량에 탑재된 자율주행기술은 3단계가 최고 기술이다. 이는 일정 구간이 자율주행이 가능하고, 운전자가 주변상황을 주시해 돌방상황에 대비해야 하는 첨단 운전자보조시스템(ADAS)에 불과하다. 하지만, 구글은 이미 일부 구간에 대해서 레벨 5의 자율주행이 가능하고, 테슬라는 지난 해 10월 24일 레벨 5에 해당하는 자율주행기술을 모든 차종에 탑재할 것이라고 언급했다. 현재 각 업체들이 자율주행에 이용하는 기술들은 모두 크게 다르지 않다. 많은 데이터와 딥러닝을 이용한 학습이 기반이 되어 있기 때문에, 자율주행에 필요한 하드웨어를 탑재한 차량이 많이 보급될수록 자율주행 자동차의 지능은 눈부시게 발전할 것이라 기대된다.