

영상 분석에서의 AI

머신 러닝 및 딥 러닝 기반의 분석을 위한 고려사항
3월 2021

목차

1	요약	3
2	서론	4
3	AI, 머신 러닝 및 딥 러닝	4
3.1	머신 러닝	4
3.2	딥 러닝	6
3.3	기존 머신 러닝과 딥 러닝의 비교	6
4	머신 러닝의 단계	7
4.1	데이터 수집 및 데이터 주석	7
4.2	훈련	7
4.3	테스트	8
4.4	배치	9
5	에지 기반 분석	9
6	하드웨어 가속	9
7	AI는 여전히 개발 초기 단계입니다.	10
8	최적의 분석 성능을 위한 고려 사항	10
8.1	이미지 가용성	11
8.2	감지 거리	11
8.3	알람 및 녹화 설정	12
8.4	유지관리	12
9	개인 정보 보호 및 개인 무결성	13
10	부록	15
10.1	신경망	15
10.2	합성곱 신경망(CNN)	16

1 요약

AI 기반 영상 분석은 영상 감시 업계에서 가장 많이 논의되는 주제 중 하나입니다. 일부 애플리케이션은 데이터 분석 속도를 크게 높이고 반복 작업을 자동화할 수 있습니다. 그러나 오늘날의 AI 솔루션은 인간 운영자의 경험과 의사 결정 기술을 대체할 수 없습니다. 그 대신 AI 솔루션을 활용하여 인간의 효율성을 개선하고 증가시키는 조합에 강점이 있습니다.

AI 개념은 머신 러닝 알고리즘과 딥 러닝 알고리즘을 통합합니다. 두 유형 모두 특별히 프로그래밍되지 않고도 결과를 계산하는 능력을 얻을 수 있도록 상당한 양의 샘플 데이터(*훈련 데이터*)를 사용하여 수학적 모델을 자동으로 구축합니다. AI 알고리즘은 반복적인 과정을 통해서 개발됩니다. 즉 훈련 데이터를 수집하고, 훈련 데이터를 라벨링하고, 라벨링된 데이터를 사용하여 알고리즘을 훈련하고 훈련된 알고리즘을 테스트하는 일련의 주기가 원하는 품질 수준에 도달할 때까지 반복됩니다. 이후, 보안 감시 사이트에서 구입하여 배치할 수 있는 분석 애플리케이션에서 알고리즘을 사용할 수 있습니다. 이때, 모든 훈련이 완료되고 애플리케이션은 새로운 것을 배우지 않습니다.

AI 기반 영상 분석의 일반적인 작업은 비디오 스트림에서 사람과 차량을 시각적으로 감지하고 구별하는 것입니다. *머신 러닝* 알고리즘은 이러한 객체를 정의하는 시각적 특징의 조합을 학습했습니다. *딥 러닝* 알고리즘은 더 정교하고(훈련된 경우) 훨씬 더 복잡한 객체를 감지할 수 있습니다. 그러나 개발 및 학습에 훨씬 더 많은 노력이 필요하고 완성된 애플리케이션이 사용될 때 훨씬 더 많은 계산 리소스가 필요합니다. 따라서 잘 지정된 감시 요구사항의 경우, 전용의 최적화된 머신 러닝 애플리케이션으로도 충분히 감시할 수 있는지 고려해야 합니다.

알고리즘 개발과 카메라의 처리 능력 향상으로 인해 서버에서(서버 기반) 계산을 수행할 필요 없이 카메라에서(에지 기반) 직접 고급 AI 기반 영상 분석을 실행할 수 있게 되었습니다. 이를 통해 더 나은 실시간 기능이 가능해졌습니다. 애플리케이션이 압축되지 않은 영상 자료에 즉시 액세스할 수 있기 때문입니다. MLPU(머신 러닝 처리 장치) 및 DLPU(딥 러닝 처리 장치)와 같은 전용 하드웨어 가속기를 카메라에 탑재하면, 에지 기반 분석을 CPU 또는 GPU(그래픽 처리 장치)에서 수행할 때보다 더 높은 전력 효율로 수행할 수 있습니다.

AI 기반 영상 분석 애플리케이션을 설치하기 전에, 알려진 전제 조건 및 제한 사항에 기반하여 제조사의 권장 사항을 주의하여 검토하고 따라야 합니다. 모든 감시 시스템 설치에 고유하고, 각 사이트에서 애플리케이션의 성능을 평가해야 합니다. 분석 품질이 예상보다 낮은 경우, 분석 애플리케이션 자체에만 집중하지 말고 전체적인 수준에서 조사를 수행해야 합니다. 영상 분석의 성능은 카메라 하드웨어, 카메라 구성, 비디오 품질, 장면 역학 및 조명과 관련된 여러 요소에 따라 달라집니다. 대부분의 경우, 이러한 요소의 영향을 알고 그에 따라 최적화하면 설치된 감시 시스템의 영상 분석 성능을 높일 수 있습니다.

AI가 감시에서 점점 더 많이 적용됨에 따라, 운영 효율성과 새로운 사용 사례의 장점을 기술 적용 시점과 장소에 대한 신중한 논의를 통해 검토해야 합니다.

2 서론

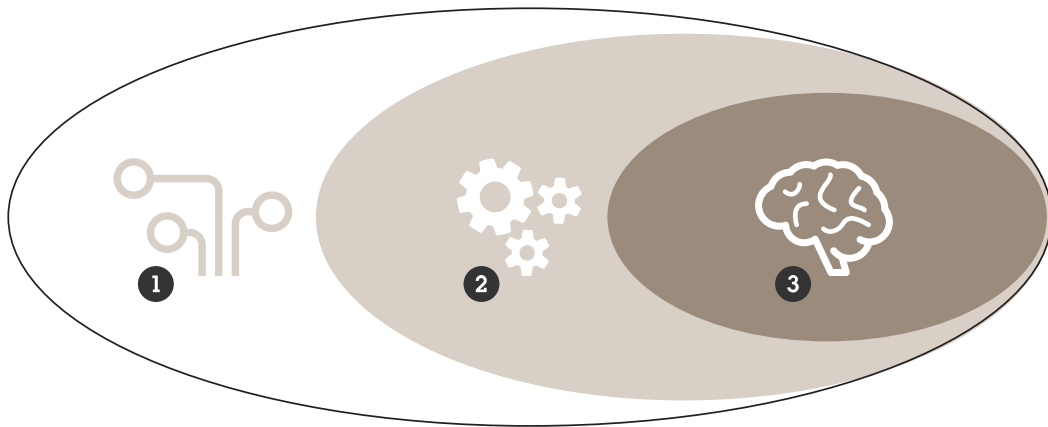
AI 즉 인공지능은 최초의 컴퓨터가 발명된 이래로 발전되어 왔고 논쟁을 거쳐 왔습니다. 가장 혁신적인 구현체는 아직 등장하지 않았지만, 오늘날 AI 기반 기술은 음성 인식, 검색 엔진 및 가상 비서와 같은 애플리케이션에서 명확하게 정의된 작업을 수행하는 데 널리 사용됩니다. AI는 예를 들어 X-레이 진단 및 망막 스캔 분석과 같이 중요한 리소스를 제공하는 의료 분야에서도 점점 더 많이 사용되고 있습니다.

AI 기반 영상 분석은 영상 감시 업계에서 가장 많이 논의되는 주제 중 하나이며 기대가 높습니다. AI 알고리즘을 사용하여 데이터 분석 속도를 높이고 반복 작업을 자동화하는 애플리케이션이 시장에 나와 있습니다. 그러나 더 광범위한 감시 맥락에서, 현재와 가까운 미래의 AI는 정확한 솔루션을 구축하는 과정에서 여러 요소 중 하나의 요소로 간주되어야 합니다.

이 백서에서는 머신 러닝 알고리즘과 딥 러닝 알고리즘에 대한 기술적 배경과 이를 영상 분석을 위해 발전시키고 적용할 수 있는 방법을 제공합니다. 여기에는 AI 가속 하드웨어에 대한 간략한 설명과 서버와 비교하여 에지에서 AI 기반 분석을 실행하는 것의 장단점이 포함됩니다. 이 백서에서는 광범위한 요인을 고려하여 AI 기반 영상 분석 성능의 전제 조건을 최적화할 수 있는 방법도 살펴봅니다.

3 AI, 머신 러닝 및 딥 러닝

인공지능(AI)은 겉보기에는 지능적인 특성을 보여주면서 복잡한 작업을 해결할 수 있는 기계와 관련된 광범위한 개념입니다. 딥 러닝과 머신 러닝은 AI의 하위 세트입니다.



- 1 인공지능
- 2 머신 러닝
- 3 딥 러닝

3.1 머신 러닝

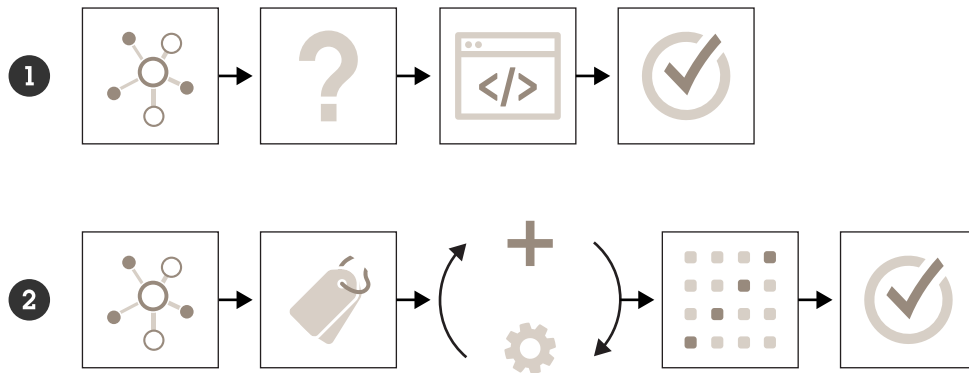
머신 러닝은 명시적으로 프로그래밍되지 않고 훈련 중 자동으로 학습하고 개선할 수 있는 시스템을 구축하기 위해 통계 학습 알고리즘을 사용하는 AI 내의 하위 집합입니다.

이 단원에서는, 컴퓨터가 이미지 또는 비디오를 분석하여 장면 내의 상황을 이해하도록 하는 원리인 *컴퓨터 비전*의 맥락에서 기존 프로그래밍과 머신 러닝을 구분합니다.

기존 방식으로 프로그래밍된 컴퓨터 비전은 이미지의 특징을 계산하는 방법, 예를 들어 뚜렷한 가장자리와 꼭짓점을 찾는 컴퓨터 프로그램에 기반해 있습니다. 이러한 특징은 이미지 데이터에서 중요한 것이 무엇인지 알고 있는 알고리즘 개발자가 수동으로 정의해야 합니다. 그런 다음 개발자는 알고리즘을 위해 이러한 특징을 조합하여 장면에서 무엇을 찾았는지 추정합니다.

머신 러닝 알고리즘은 특별히 프로그래밍되지 않고도 결과를 계산하여 결정을 내릴 수 있는 능력을 얻을 수 있도록 상당한 양의 샘플 데이터 – *훈련 데이터* – 를 사용하여 자동으로 수학적 모델을 구축합니다. 특징은 여전히 사람이 정의하지만 이러한 특징을 조합하는 방법은 알고리즘 자체가 많은 양의 라벨링된 또는 *주석*이 있는 훈련 데이터에 노출되어 학습합니다. 이 백서에서는 사람이 정의한 특징을 학습된 조합에서 사용하는 이러한 방법을 *기존 머신 러닝*이라고 언급합니다.

즉, 머신 러닝 애플리케이션의 경우, 원하는 프로그램을 얻으려면 컴퓨터를 훈련시켜야 합니다. 데이터는 사람이 수집하여 주석을 달고, 때로는 서버 컴퓨터에 의한 사전 주석으로 지원을 받습니다. 결과는 시스템에 입력되고 이 과정은 애플리케이션이 특정 유형의 차량과 같이 원하는 것을 감지할 수 있을 만큼 충분히 학습할 때까지 계속됩니다. 훈련된 모델이 프로그램이 됩니다. 프로그램이 완료되면 시스템은 새로운 것을 학습하지 않습니다.



1 기존 프로그래밍:

데이터가 수집됩니다. 프로그램 기준이 정의됩니다. 프로그램이 코딩됩니다(사람에 의해). 완료됩니다.

2 머신 러닝:

데이터가 수집됩니다. 데이터가 라벨링됩니다. 모델이 반복적인 훈련 과정을 거칩니다. 완성된 훈련된 모델이 프로그램이 됩니다. 완료됩니다.

컴퓨터 비전 프로그램을 구할 때 기존 프로그래밍 대비 AI의 장점은 광범위한 데이터를 처리할 수 있다는 것입니다. 컴퓨터는 집중력을 잃지 않고 수천 개의 이미지를 살펴볼 수 있는 반면 인간 프로그래머는 잠시 후 피곤해져서 집중력을 잃습니다. 이러한 방식으로, AI가 애플리케이션을 훨씬 더 정확해지게 할 수 있습니다. 그러나, 애플리케이션이 복잡할수록, 원하는 결과를 시스템이 생성하는 것은 더 어려워집니다.

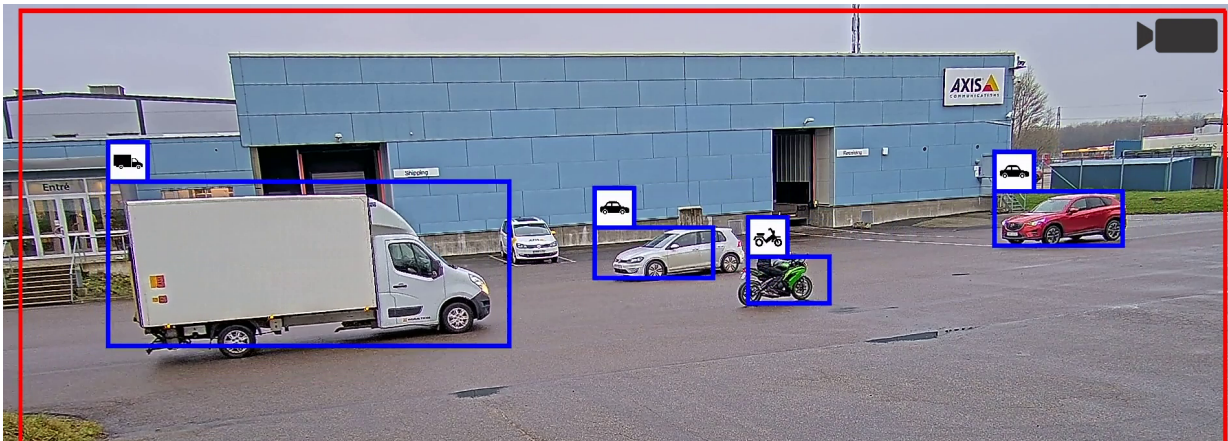
3.2 딥 러닝

딥 러닝은 특징 추출과, 출력을 생성하는 규칙의 심층 구조에서 이러한 특징을 조합하는 방식을 데이터 중심적 방식으로 학습하는 개량된 버전의 머신 러닝입니다. 이 알고리즘은 학습 데이터에서 찾을 특징을 자동으로 정의할 수 있습니다. 또한 특징의 연쇄 조합의 매우 심층적인 구조를 학습할 수 있습니다.

딥 러닝에서 사용되는 알고리즘의 핵심은 뉴런이 작동하는 방식과 뇌가 뉴런 출력을 연쇄 규칙의 심층적 계층구조 또는 *네트워크*에서 조합하여 더 높은 수준의 지식을 형성하는 방법에서 영감을 얻었습니다. 뇌는 조합 자체도 뉴런에 의해 형성되는 시스템으로, 특징 추출과 특징 조합 사이의 구분을 없애고 어떤 의미에서는 동일하게 만듭니다. 이러한 구조는 딥 러닝에서 가장 널리 사용되는 알고리즘 유형인 *인공 신경망*이라고 하는 것으로, 연구자들에 의해 시뮬레이션되었습니다. 신경망에 대한 간략한 개요는 이 문서의 부록을 참조하십시오.

딥 러닝 알고리즘을 사용하면, 복잡한 시각적 감지기를 구축하고 이를 자동으로 훈련하여 매우 복잡한 객체를 크기, 회전 및 기타 변화에 탄력적으로 대응하여 감지하도록 할 수 있습니다.

이러한 유연성은 딥 러닝 시스템이 기존 머신 러닝 시스템보다 훨씬 더 많은 양의 데이터와 훨씬 더 다양한 데이터에서 학습할 수 있다는 것에서 비롯합니다. 대부분의 경우 사람이 정의한 컴퓨터 비전 알고리즘보다 성능이 훨씬 뛰어납니다. 따라서 딥 러닝은 이미지 분류, 언어 처리 및 객체 감지와 같이 인간 전문가가 특징 조합을 쉽게 구성할 수 없는 복잡한 문제에 특히 적합합니다.



딥 러닝에 기반한 객체 감지는 복잡한 객체를 분류할 수 있습니다. 이 예에서 분석 애플리케이션은 차량을 감지할 수 있을 뿐만 아니라 차량 유형도 분류할 수 있습니다.

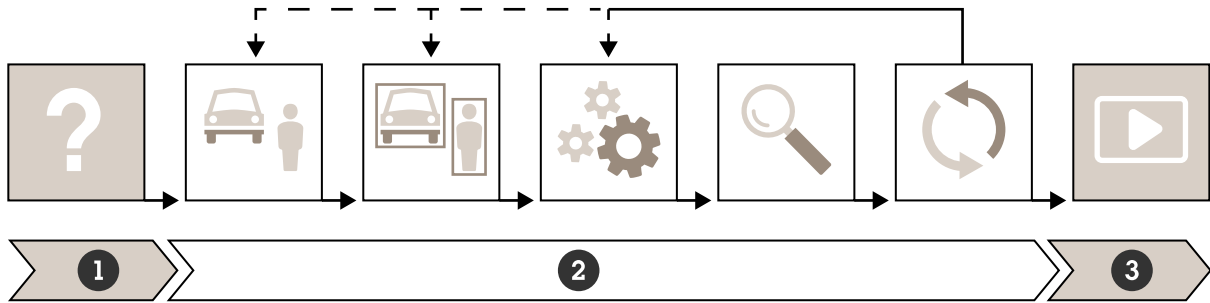
3.3 기존 머신 러닝과 딥 러닝의 비교

기존 머신 러닝과 딥 러닝은 알고리즘 유형이 비슷하지만, 딥 러닝 알고리즘은 일반적으로 기존 머신 러닝 알고리즘이 사용하는 것보다 훨씬 더 커다란 세트의 학습된 특징 조합을 사용합니다. 이는 딥 러닝 기반 분석이 더 유연할 수 있고, 훈련될 경우 훨씬 더 많은 복잡한 작업을 수행하는 것을 학습할 수 있다는 것을 뜻합니다.

그러나 특정 감시 분석은 전용의 최적화된 기존 머신 러닝 알고리즘으로 충분히 가능할 수 있습니다. 잘 지정된 범위 내에서는, 수학적 연산이 더 적게 필요하면서도 딥 러닝 알고리즘과 비슷한 결과를 제공할 수 있고 따라서 비용 효율이 더 높고 소모 전력이 더 적을 수 있습니다. 추가적으로 훨씬 더 적은 훈련 데이터를 요구하고 이는 개발 노력을 대폭 감소시킵니다.

4 머신 러닝의 단계

머신 러닝 알고리즘의 개발은 완성된 분석 애플리케이션을 배치할 수 있기 전에 아래에 대략적으로 시각화된 일련의 단계와 반복을 따릅니다. 분석 애플리케이션의 핵심에는 객체 감지와 같은 하나 이상의 알고리즘이 있습니다. 딥 러닝 기반 애플리케이션의 경우 알고리즘의 핵심은 딥 러닝 모델입니다.



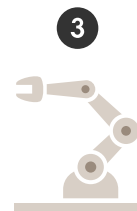
- 1 준비: 애플리케이션의 목적 정의.
- 2 훈련: 훈련 데이터 수집. 데이터 주석 달기. 모델 훈련. 모델 테스트. 품질이 예상과 다르면 이전 단계가 반복적인 개선 주기로 다시 수행됩니다.
- 3 배치: 완성된 애플리케이션의 설치 및 사용.

4.1 데이터 수집 및 데이터 주석

AI 기반 분석 애플리케이션을 개발하려면 많은 양의 데이터를 수집해야 합니다. 영상 감시에서 이것은 일반적으로 사람과 차량 또는 기타 관심 객체의 이미지와 비디오 클립으로 구성됩니다. 시스템 또는 컴퓨터에서 데이터를 인식할 수 있도록 하려면 관련 객체를 분류하고 라벨링하는 데이터 주석 과정이 필요합니다. 데이터 주석은 주로 수동적이고 노동 집약적인 작업입니다. 준비된 데이터는 분석 애플리케이션이 사용되는 맥락에 적합한 충분히 다양한 샘플을 포함해야 합니다.

4.2 훈련

훈련 또는 학습은 모델에 주석이 달린 데이터를 제공하고, 학습 프레임워크를 사용하여 원하는 품질에 도달할 때까지 모델을 반복적으로 수정하고 개선하는 것입니다. 즉, 정의된 작업을 수행하도록 모델이 최적화됩니다. 훈련은 세 가지 주요 방법 중 하나에 따라 수행할 수 있습니다.



- 1 지도 학습: 모델이 정확한 예측을 하는 방법을 학습합니다.
- 2 비지도 학습: 모델이 클러스터를 식별하는 방법을 학습합니다.

3 강화 학습: 모델이 실수로부터 학습합니다.

4.2.1 지도 학습

지도 학습은 오늘날 머신 러닝에서 가장 많이 사용되는 방법입니다. 지도 학습이란 예제를 통해 학습하는 것이라고 설명할 수 있습니다. 훈련 데이터에는 명확하게 주석이 달려 있습니다. 즉, 입력 데이터가 이미 원하는 출력 결과와 쌍을 이루고 있습니다.

지도 학습은 일반적으로 매우 많은 양의 주석 데이터가 필요하며 훈련된 알고리즘의 성능은 훈련 데이터의 품질에 직접적으로 의존합니다. 가장 중요한 품질 측면은 실제 배치 상황에서 제공하는 모든 잠재적 입력 데이터를 나타내는 데이터 세트를 사용하는 것입니다. 객체 감지기의 경우, 개발자는 다양한 객체 사례, 방향, 배율, 조명 상황, 배경 및 주의 혼란을 사용하여 다양한 이미지로 알고리즘을 훈련시켜야 합니다. 훈련 데이터가 계획된 사용 사례를 대표하는 경우에만, 최종 분석 애플리케이션이 훈련 단계에서 볼 수 없는 새로운 데이터를 처리할 때도 정확한 예측을 할 수 있습니다.

4.2.2 비지도 학습

비지도 학습은 알고리즘을 사용하여 라벨링되지 않은 데이터 세트를 분석하고 그룹화합니다. 여전히 품질을 예측할 수 없는 상태이지만 이 모델은 많은 보정과 테스트가 필요하기 때문에 감시 업계에서 일반적인 훈련 방법이 아닙니다.

데이터 세트는 분석 애플리케이션에 적합해야 하지만 명확하게 라벨링하거나 표시할 필요는 없습니다. 수동 주석 작업은 제거되지만 훈련에 필요한 이미지 또는 비디오의 수를 몇 배로 크게 늘려야 합니다. 훈련 단계에서, 훈련해야 하는 모델은 훈련 프레임워크의 지원을 받아 데이터 세트의 공통 기능을 식별합니다. 이를 통해 배치 단계에서 패턴에 따라 데이터를 그룹화할 수 있는 동시에 학습된 그룹에 맞지 않는 이례적인 것을 감지할 수 있습니다.

4.2.3 강화 학습

강화 학습은 로봇 공학, 산업 자동화 및 비즈니스 전략 계획 등에 사용되지만 대량의 피드백이 필요하기 때문에 오늘날 이 방법은 감시 부문에서 제한적으로 사용됩니다. 강화 학습은 특정 상황에서 잠재적 보상 즉 모델이 올바른 선택을 할 때 더 커지는 보상을 극대화하기 위해 적절한 조치를 취하는 것에 관한 것입니다. 이 알고리즘은 훈련에 데이터/라벨 쌍을 사용하지 않고 대신 보상을 측정하는 동안 환경과의 상호 작용을 통해 결정을 테스트하여 최적화됩니다. 이 알고리즘의 목표는 보상을 극대화하는 데 도움이 되는 조치를 위한 정책을 학습하는 것입니다.

4.3 테스트

모델이 훈련되면 철저한 테스트를 거쳐야 합니다. 이 단계에는 일반적으로 실제 배치 상황에서 광범위한 테스트로 보완되는 자동화된 부분이 포함됩니다.

자동화된 부분에서, 애플리케이션은 훈련 중에 모델이 본 적 없는 새로운 데이터 세트로 벤치마킹됩니다. 이러한 벤치 마크가 예상되는 위치에 있지 않으면, 이 과정이 다시 시작됩니다. 새로운 훈련 데이터가 수집되고 주석이 작성되거나 개선되고 모델이 재훈련됩니다.

원하는 품질 레벨에 도달하면, 필드 테스트가 시작됩니다. 이 테스트에서 애플리케이션은 실제 시나리오에 노출됩니다. 양과 변경은 애플리케이션의 범위에 따라 다릅니다. 범위가 좁을수록 테스트 해야 할 변경 사항이 적습니다. 범위가 넓을수록 더 많은 테스트가 필요합니다.

결과는 다시 비교되고 평가됩니다. 그런 다음 이 단계로 인해 이 과정이 다시 시작될 수 있습니다. 또 다른 잠재적 결과는 애플리케이션을 사용하지 않거나 부분적으로만 사용하도록 권장되지 않는 알려진 시나리오를 설명하면서 전제조건을 정의하는 것일 수 있습니다.

4.4 배치

배치 단계는 추론 또는 예측 단계라고도 합니다. 추론 또는 예측은 훈련된 머신 러닝 모델을 실행하는 과정입니다. 이 알고리즘은 훈련 단계에서 학습한 것을 사용하여 원하는 출력을 생성합니다. 감시 분석 맥락에서 추론 단계는 실제 장면을 모니터링하는 감시 시스템에서 실행되는 애플리케이션입니다.

오디오 또는 비디오 입력 데이터에 대한 머신 러닝 기반 알고리즘을 실행할 때 실시간 성능을 얻으려면, 일반적으로 특정 하드웨어 가속이 필요합니다.

5 에지 기반 분석

고성능 영상 분석은 카메라가 제공할 수 있는 것보다 더 많은 전력과 냉각이 필요했기 때문에 서버 기반으로 사용되었습니다. 그러나 최근 몇 년 동안 알고리즘 개발과 에지 장치의 처리 능력 향상으로 에지에서 고급 AI 기반 영상 분석을 실행할 수 있게 되었습니다.

에지 기반 분석 애플리케이션에는 분명한 이점이 있습니다. 매우 짧은 지연 시간으로 압축되지 않은 영상 자료에 액세스 할 수 있어서 실시간 애플리케이션을 지원하는 동시에 계산을 위해 데이터를 클라우드로 이동하는 데 따른 추가 비용과 복잡성을 피할 수 있습니다. 에지 기반 분석은 감시 시스템에 필요한 서버 리소스가 더 적기 때문에 하드웨어 및 배치 비용이 더 낮습니다.

일부 애플리케이션은 에지 기반 및 서버 기반 처리의 조합을 사용하여 카메라에서 사전 처리하고 서버에서 추가 처리하는 것을 통해 이점을 얻을 수 있습니다. 이러한 하이브리드 시스템은 여러 카메라 스트림에서 작업하여 분석 애플리케이션의 비용 효율적인 확장을 촉진할 수 있습니다.

6 하드웨어 가속

여러 유형의 플랫폼에서 특정 분석 애플리케이션을 종종 실행할 수 있지만, 전력이 제한적일 때 전용 하드웨어 가속을 사용하여 훨씬 더 높은 성능을 얻을 수 있습니다. 하드웨어 가속기는 분석 애플리케이션을 전력 효율적으로 실행할 수 있도록 합니다. 하드웨어 가속기는 적합할 경우 서버 및 클라우드 컴퓨팅 리소스로 보완할 수 있습니다.

- **GPU(그래픽 처리 장치).** GPU는 주로 그래픽 처리 애플리케이션을 위해 개발되었지만 서버 및 클라우드 플랫폼에서 AI를 가속하는 데도 사용됩니다. 때때로 임베디드 시스템(에지)에서도 사용되지만, 전력 효율성 관점에서 GPU는 머신 러닝 추론 작업에 최적이지 않습니다.
- **MLPU(머신 러닝 처리 장치).** MLPU는 매우 높은 전력 효율로 컴퓨터 비전 작업을 수행하기 위해 특정한 기존 머신 러닝 알고리즘의 추론을 가속할 수 있습니다. MLPU는 사람과 차량과 같은 제한된 수의 동시적 객체 유형에 대한 실시간 객체 감지를 위해 고안되었습니다.

- **DLPU(딥 러닝 처리 장치).** DLPU가 내장된 카메라는 일반적 딥 러닝 알고리즘 추론을 높은 전력 효율로 가속하여 더 세분화된 객체 분류를 할 수 있도록 합니다.

7 AI는 여전히 개발 초기 단계입니다.

AI 솔루션의 잠재력과 인간이 달성할 수 있는 것을 비교하는 것은 매력적입니다. 인간 영상 감시 운영자는 짧은 시간 동안만 완전하게 경계를 할 수 있지만 컴퓨터는 지치지 않고 대량의 데이터를 매우 빠르게 처리할 수 있습니다. 그러나 AI 솔루션이 인간 운영자를 대체할 것이라고 가정하는 것은 근본적인 오해입니다. 진정한 강점은 현실적인 조합에 있습니다. AI 솔루션을 활용하여 인간 운영자의 효율성을 높일 수 있습니다.

머신 러닝 또는 딥 러닝 솔루션은 종종 자동으로 경험을 통해 학습하거나 향상될 수 있는 능력을 갖고 있는 것으로 묘사됩니다. 그러나 오늘날 사용 가능한 AI 시스템은 배치 후 새로운 기술을 자동으로 학습 *하지 않고* 발생한 특정 이벤트를 기억 *하지 않습니다*. 시스템의 성능을 개선하기 위해서, 지도 학습 세션 중에 시스템을 더 향상되고 더 정확한 데이터로 재훈련해야 합니다. 비지도 학습은 일반적으로 클러스터를 생성하는 데 많은 데이터가 필요하므로 비디오 감시 애플리케이션에서 사용되지 않습니다. 대신 오늘날 주로 대량의 데이터 세트를 분석하여 금융 거래 등에서 이례적인 것을 찾는 데 사용됩니다. 영상 감시에서의 "자기 학습"이라고 알려진 대부분의 접근 방식은 통계적 데이터 분석에 기반해 있고 실제로 딥 러닝 모델을 재교육하는 데 기반해 있지 않습니다.

인간의 경험은 여전히 많은 감시 목적용 AI 기반 분석 애플리케이션을 능가합니다. 특히 매우 일반적인 작업을 수행해야 하는 사람들이 그러하고 상황에 맞는 이해가 중요한 경우에도 그러합니다. 머신 러닝 기반 애플리케이션은 특별히 훈련된 경우 "달리는 사람"을 성공적으로 감지할 수도 있습니다. 하지만 데이터를 상황에 맞게 적용할 수 있는 사람과 달리, 애플리케이션은 사람이 달리는 이유를 이해하지 못합니다(버스를 잡기 위해서 달리는 걸까요 아니면 주변에서 달려오는 경찰관한테서 도망가기 위해서 달리는 걸까요?). 감시용 분석 애플리케이션에 AI를 적용하는 기업의 약속에도 불구하고, 애플리케이션은 아직 영상에서 원격으로 보는 것을 인간과 동일한 통찰력으로 이해할 수 없습니다.

동일한 이유로, AI 기반 분석 애플리케이션은 거짓 알람을 트리거하거나 알람을 놓칠 수도 있습니다. 이는 일반적으로 움직임이 많은 복잡한 환경에서 발생할 수 있습니다. 그러나 예를 들어, 커다란 객체를 운반하는 사람으로 인한 알람일 수 있습니다 — 애플리케이션이 사람의 특성을 인식하는 것이 효과적으로 차단되어 올바른 분류 가능성이 감소합니다.

오늘날 AI 기반 분석은, 예를 들어 대응에 대한 결정을 내릴 것을 인간 운영자에게 경고하기 전에 사고가 얼마나 관련성이 있는지 대략적으로 결정하기 위해 보조적으로 사용되어야 합니다. 이런 식으로 AI는 확장성에 도달하는 데 사용되고 인간 운영자는 잠재적 사고를 평가하기 위해 존재합니다.

8 최적의 분석 성능을 위한 고려 사항

AI 기반 분석 애플리케이션의 품질 기대치를 파악하려면, 일반적으로 애플리케이션 문서에 나열된 알려진 전제 조건 및 제한 사항을 주의하여 검토하고 이해하는 것이 좋습니다.

모든 감시 시스템 설치에 고유하고 각 사이트에서 애플리케이션의 성능을 평가해야 합니다. 품질이 기대 수준 또는 예상 수준에 미치지 못하는 경우, 애플리케이션 자체에만 조사를 집중하지 않을 것을 적극 권장합니다. 분석 애플리케이션의 성능은 매우 많은 요인에 따라 달라지고 대부분의 요인은 그 영향을 알고 있을 경우 최적화할 수 있기 때문에 모든 조사는 전체적인 수준에서 이루어져야 합니다.

어제야 합니다. 이러한 요인에는 카메라 하드웨어, 비디오 품질, 장면 역학, 조명 수준뿐만 아니라, 카메라 구성, 위치 및 방향 등이 포함됩니다.

8.1 이미지 가용성

이미지 품질은 카메라의 고해상도와 높은 감광도에 따라 달라지는 경우가 많습니다. 이러한 요인의 중요성은 의심할 수 없지만, 이미지 또는 비디오의 실제 *가용성*에 영향을 주는 다른 요인이 분명히 있습니다. 예를 들면, 야간에 장면의 조명이 충분하지 않거나, 카메라의 방향이 변경되거나, 시스템 연결이 끊어지면, 가장 비싼 감시 카메라가 제공하는 최고 품질의 비디오 스트림도 소용이 없게 될 수 있습니다.

카메라 설치에 배치 전에 신중하게 고려해야 합니다. 영상 분석이 예상대로 수행되려면 보려는 장면을 장애물 없이 선명하게 볼 수 있도록 카메라를 배치해야 합니다.

이미지 가용성은 사용 사례에 따라 달라질 수도 있습니다. 사람의 눈으로 보기 좋은 영상은 영상 분석 응용 프로그램의 성능을 위한 최적의 품질을 갖고 있지 않을 수 있습니다. 실제로 사람이 볼 수 있도록 비디오 화질을 향상시키는데 일반적으로 사용되는 많은 이미지 처리 방법은 영상 분석을 사용할 때 권장되지 않습니다. 예를 들어 적용된 노이즈 감소 방법, 광역 역광 보정(WDR) 방법 또는 자동 노출 알고리즘이 여기에 포함될 수 있습니다.

오늘날 비디오 카메라는 종종 완전한 어둠 속에서 작동할 수 있는 통합 IR 조명과 함께 제공됩니다. 이것은 조명이 어두운 곳에 카메라를 배치할 수 있고 추가 조명 설치 필요성을 줄일 수 있으므로 긍정적입니다. 그러나 사이트에 폭우나 강설이 예상되는 경우, 카메라 또는 카메라와 매우 가까운 위치에서 나오는 빛에 의존하지 않을 것을 강력히 권장합니다. 너무 많은 빛은 빔방울과 눈송이에 부딪쳐 카메라로 직접 반사되어 분석을 수행하지 못할 수 있습니다. 반면에 주변 조명은 낯선에도 분석 결과를 제공할 가능성이 더 높습니다.

8.2 감지 거리

AI 기반 분석 애플리케이션의 최대 감지 거리를 결정하는 것은 어렵습니다. 정확한 데이터 시트 값(미터 또는 피트 단위의)이 절대적인 값일 수는 없기 때문입니다. 이미지 품질, 장면 특성, 날씨 조건 및 색깔 및 밝기와 같은 객체 속성은 감지 거리에 크게 영향을 줍니다. 예를 들어 맑은 날에는 어두운 배경의 밝은 객체가 비오는 날의 어두운 객체보다 훨씬 먼 거리에서 시각적으로 감지될 수 있다는 것은 분명합니다.

감지 거리는 감지할 객체의 속도에 따라서도 달라집니다. 정확한 결과를 얻으려면, 영상 분석 애플리케이션이 충분히 오랜 시간 동안 객체를 "볼" 필요가 있습니다. 이 기간은 플랫폼의 처리 성능(프레임 레이트)에 따라 달라집니다. 처리 성능이 낮을수록 객체는 더 오래 표시되어야만 감지될 수 있습니다. 카메라의 셔터 시간이 객체 속도와 제대로 일치하지 않으면, 이미지에 나타나는 모션 블러도 감지 정확도를 떨어뜨릴 수 있습니다.

빠르게 움직이는 객체는 카메라와 더 가까운 거리에서 지나가면 더 쉽게 놓칠 수 있습니다. 예를 들어, 카메라에서 멀리 떨어져서 달리는 사람은 잘 감지될 수 있지만, 같은 속도로 카메라와 매우 가까운 거리에서 달리는 사람은 너무 빨리 화각에 들어왔다가 벗어날 수 있어서 알람이 트리거되지 않을 수 있습니다.

모션 감지에 기반한 분석에서 카메라를 향해 직접 이동하거나 카메라에서 멀어지는 객체는 또 다른 문제를 제기합니다. 느리게 움직이는 객체의 경우 감지가 특히 어려우며, 장면 전체의 움직임에 비해 이미지에 아주 작은 변화만 발생합니다.

카메라 해상도가 높아져도 감지 거리가 길어지지는 않습니다. 머신 러닝 알고리즘을 실행하는 데 필요한 처리 능력은 입력 데이터의 크기에 비례합니다. 이는 4K 카메라의 전체 해상도를 분석하는 데 필요한 처리 능력이 1080p 카메라보다 4배 이상 높다는 것을 의미합니다. 카메라 처리 능력의 한계로 인해 카메라 또는 스트림이 제공할 수 있는 것보다 더 낮은 해상도에서 AI 기반 애플리케이션을 실행하는 것이 매우 일반적입니다.

8.3 알람 및 녹화 설정

적용되는 다양한 수준의 필터로 인해 객체 분석은 거짓 알람을 거의 생성하지 않습니다. 그러나 객체 분석은 나열된 전제 조건이 모두 충족될 때만 수행해야 합니다. 대신에 다른 경우에는 중요한 이벤트를 놓칠 수 있습니다.

모든 조건이 항상 충족될지 절대적으로 확실하지 않은 경우에는, 보수적인 접근 방식을 취하고 특정 객체 분류가 유일한 알람 트리거가 되지 않도록 시스템을 설정하는 것이 좋습니다. 이것은 더 많은 거짓 알람을 유발하겠지만 중요한 것을 놓칠 위험을 줄여주기도 합니다. 알람 또는 트리거가 알람 모니터링 센터로 직접 전달되면, 각 거짓 알람은 매우 많은 비용을 유발합니다. 원치 않는 알람을 걸러 내려면 신뢰할 수 있는 객체 분류가 분명히 필요합니다. 그러나 녹화 솔루션을 객체 분류에만 의존하지 않도록 설정할 수 있으며 그렇게 설정해야 합니다. 실제 알람이 누락된 경우, 이 설정을 사용하면 알람 누락 이유를 평가한 다음 전체 설치 및 구성을 개선할 수 있습니다.

사고 검색 중에 서버에서 객체 분류가 수행되는 경우에는, 초기 녹화를 전혀 필터링하지 않고 시스템을 연속 녹화로 구성하는 것이 좋습니다. 연속 녹화는 많은 저장 공간을 사용하지만, 이는 Zipstream과 같은 최신 압축 알고리즘으로 어느 정도 보완됩니다.

8.4 유지관리

감시 시스템은 정기적으로 유지관리 해야 합니다. 시야를 방해하거나 차단할 수 있는 모든 것을 발견하고 제거하려면 VMS 인터페이스를 통해 영상을 보는 것뿐만 아니라 물리적 검사를 수행하는 것이 좋습니다. 이는 표준 녹화 전용 시스템에서도 중요하지만 분석을 사용할 때는 더욱 중요합니다.

기본적인 비디오 모션 디텍션의 맥락에서, 바람에 흔들리는 거미줄과 같은 일반적인 장애물은 알람 수를 증가시켜 필요 이상의 저장 공간을 소비할 수 있습니다. 객체 분석을 통해, 거미줄은 기본적으로 감지 영역에 제외 영역을 생성합니다. 거미줄은 객체를 모호하게 하고 감지 및 분류 가능성을 크게 줄입니다.



거미줄은 감시 카메라의 화각을 방해할 수 있습니다.

카메라 전면 유리의 먼지나 작은 물방울은 주간에는 문제를 일으키지 않습니다. 그러나 저조도 조건에서는 측면에서 더러운 작은 물방울에 부딪히는 빛(예: 자동차 전조등)은 감지 정확도를 떨어뜨릴 수 있는 예기치 않은 반사를 일으킬 수 있습니다.

장면 관련 유지관리는 카메라 유지관리 만큼이나 중요합니다. 카메라의 수명 기간 동안, 카메라가 모니터링하는 장면에서 많은 일이 발생할 수 있습니다. 간단한 전후 이미지 비교를 통해 잠재적인 문제를 확인할 수 있습니다. 카메라가 배치되었을 때의 장면은 어떠했으며 지금은 어떠합니까? 감지 영역을 조정할 필요가 있습니까? 카메라의 화각을 조정해야 합니까? 아니면 카메라를 다른 위치로 이동해야 합니까?

9 개인 정보 보호 및 개인 무결성

보안 및 감시 업무를 수행하려면 개인 정보 보호 및 개인 무결성에 대한 개인의 권리와 범위를 예방하거나 법의학 조사를 가능하게 하여 안전을 높인다는 목표가 균형을 이루도록 해야 합니다. 특정 설치 및 사용 사례에서는, 현지 법률의 이해 및 적용뿐만 아니라 신중한 윤리적 고려도 필요합니다. 예를 들어 사이버 보안을 보장하고 영상 자료에 대한 우발적 액세스를 방지하기 위한 요구사항도 솔루션에 적용됩니다. 동시에 에지 기반 분석과 통계적 목적을 위한 메타데이터 생성은 익명화된 데이터만 나중에 처리하기 위해 전송되는 경우 개인 정보 보호를 강화할 수 있습니다.

감시 시스템에서 자동화된 분석의 적용이 증가함에 따라 몇 가지 새로운 측면을 고려해야 합니다. 분석 애플리케이션에는 거짓 감지 위험이 있기 때문에 결정 과정에 숙련된 운영자 또는 사용자가 포함되는 것이 중요합니다. 종종 이것을 가리켜 "인간의 개입"을 유지하는 것이라고 합니다. 또한 알람이 생성되고 표시되는 방식이 사람의 결정에 영향을 미칠 수 있음을 인식하는 것이 중요합니다. 분석 솔루션의 기능에 대한 적절한 훈련과 인식 없이는 잘못된 결과를 도출할 수 있습니다.

딥 러닝 알고리즘이 개발되는 방식으로 인해 추가적인 우려가 발생할 수 있으며, 일부 사용 사례에서는 기술을 적용할 때 신중한 접근 방식이 필요합니다. 이러한 알고리즘의 품질은 기본적으로 알고리즘 훈련에 사용되는 데이터 세트, 즉 비디오 및 이미지와 연결되어 있습니다. 테스트에 따르면 해당 자료를 신중하게 선택하지 않으면 일부 AI 시스템이 감지 시 인증 및 성별 편향을 모두 나타낼 수 있습니다. 이것은 공개적인 논의를 촉발시켰고, 그러한 측면이 시스템 개발 중에 다루어지도록 보장하기 위한 입법적 한계와 활동 모두를 야기했습니다.

AI가 감시에 점점 더 많이 적용됨에 따라, 기술을 적용할 장소와 시기에 대한 신중한 논의와 더불어 운영 효율성의 이점과 새로운 잠재적 사용 사례의 균형을 유지하는 것이 중요합니다.

10 부록

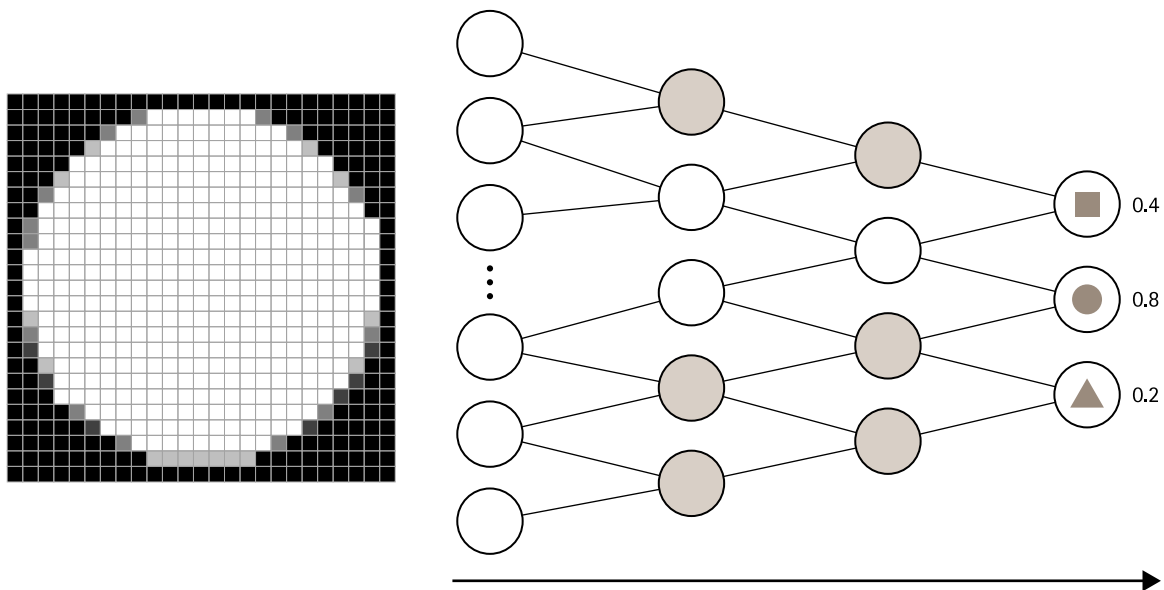
이 부록에서는 딥 러닝의 기반을 형성하는 인공 신경망에 대한 배경 정보를 제공합니다.

10.1 신경망

신경망은 인간의 뇌가 작동하는 방식과 다소 유사한 과정을 통해 데이터 세트의 관계를 인식하는 데 사용되는 알고리즘의 집합체입니다. 신경망은 상호 연결된 이른바 노드나 뉴런의 여러 계층으로 구성된 계층으로 구성되며, 정보는 입력 계층에서 네트워크를 거쳐 출력 계층으로 연결부를 따라 전달됩니다.

신경망이 작동하기 위한 가정은 입력 데이터 샘플을 유한한 특징 세트로 축소하여 입력 데이터를 잘 표현할 수 있다는 것입니다. 그런 다음 이러한 특징을 조합하여 입력 데이터를 분류하는 데 도움이 됩니다(예: 이미지 내용 설명).

아래 그림은 입력 이미지가 속한 클래스를 식별하기 위해 신경망을 사용하는 예를 보여줍니다. 이미지의 각 픽셀은 하나의 입력 노드로 표시됩니다. 모든 입력 노드는 첫 번째 계층의 노드에 결합됩니다. 이들은 두 번째 계층에 입력값으로 전달되는 출력값을 생성하고, 그 다음 계층들에 대해서도 이런 식으로 계속됩니다. 각 계층에서 가중 함수, 바이어스 값 및 활성화 함수도 과정에 포함됩니다.



입력 이미지(왼쪽) 및 신경망(오른쪽)의 예. 출력 계층에 도달하면 네트워크는 가능한 각 범주(정사각형, 원형 또는 삼각형)에 대한 확률을 추정합니다. 확률 값이 가장 높은 범주는 입력 이미지의 가장 가능성이 높은 모양입니다.

이 과정을 **순전파**라고 합니다. 순전파의 결과 불일치가 발생할 경우 네트워크 매개 변수가 **역방향 전파**를 통해 약간 수정됩니다. 이 반복적인 훈련 과정에서 네트워크의 성능이 점차 향상됩니다.

배치 후 일반적으로 신경망에는 이전의 순전파의 메모리가 없습니다. 이것은 신경망이 시간이 지남에 따라 개선되지 않으며 훈련된 객체 유형만 감지하거나 작업 유형을 해결할 수 있음을 의미합니다.

10.2 합성곱 신경망(CNN)

합성곱 신경망(CNN)은 컴퓨터 비전 작업에 특히 적합한 것으로 증명된 인공 신경망의 하위 유형이며, 딥 러닝의 급속한 발전의 핵심에 있습니다. 컴퓨터 비전의 경우, 네트워크는 이미지 전체에 걸쳐 객체 모양을 식별하는 과정에서 가장자리, 모서리 및 색상 차이와 유사한 독특한 이미지 특징을 자동으로 찾도록 훈련됩니다.

이를 수행하기 위한 주요 연산은 *합성곱*이라고 하는 수학적 연산입니다. 이것은 각 개별 노드의 출력이 전체 입력 데이터 볼륨을 사용하기보다는 이전 계층에서 생성한 입력 데이터의 제한된 주변 환경에만 의존하기 때문에 매우 효율적인 연산입니다. 즉 CNN에서 각 노드는 이전 계층의 모든 노드에 연결되지 않고 작은 하위 세트에만 연결됩니다. 합성곱은 가장 유용한 정보를 유지하면서 데이터 크기를 줄이는 다른 연산으로 보완됩니다. 표준 인공 신경망에서와 마찬가지로, 데이터는 네트워크로 더 깊숙이 들어갈수록 점점 더 추상화됩니다.

훈련 단계에서 CNN은 계층을 적용하는 가장 좋은 방법을 학습합니다. 즉, 네트워크의 출력이 훈련 데이터의 주석과 최대한 일치하도록 합성곱이 이전 계층의 특징을 조합하는 방법을 학습합니다. 추론하는 동안 훈련된 합성곱 신경망은 훈련의 결과인 합성곱의 계층을 순차적으로 적용합니다.

Axis Communications에 대하여

Axis는 보안 개선과 새로운 비즈니스 수행 방식에 대한 통찰력을 제공하는 네트워크 솔루션을 개발하여 보다 스마트하고 안전한 세상을 만들 수 있도록 지원합니다. 네트워크 비디오 업계의 선도 기업인 Axis는 비디오 감시 및 분석, 접근 제어, 인터콤, 오디오 시스템 분야의 제품과 서비스를 제공합니다. 50개 이상의 국가에서 3,800명이 넘는 Axis 임직원이 파트너와 협력하여 전세계 고객에게 최적의 솔루션을 제공하고 있습니다. 1984년에 설립된 Axis는 스웨덴에 본사를 두고 있습니다.

Axis에 대한 자세한 정보는 axis.com에서 확인하실 수 있습니다.