

A low-angle, black and white photograph of a modern building's exterior. The facade is composed of numerous vertical slats or louvers, creating a rhythmic pattern of light and shadow. The building appears to be made of concrete or a similar solid material. The perspective is looking up, emphasizing the height and architectural details of the structure.

Detection of Wildlife Animals using Deep Learning Approaches: A Systematic Review

저자

Vigneshwaran Palanisamy :
Department of Computing and Information Systems
Sabaragamuwa University of Sri Lanka

Nagulan Ratnarajah :
Department of Physical Science University of Vavuniya Sri Lank

Abstract

카메라는 동물의 이미지를 수집한다.

컨볼루션 신경망은 이미지를 감지하고 인식하는 방법.

딥러닝의 응용을 연구하고 분석.

동물이 특정 영역에 진입하면 카메라가 동물의 이미지를 수집한다.

수집된 이미지는 컨볼루션 신경망(Convolutional Neural Network, CNN)을 통해 감지하고 식별한다.

이 연구는 딥러닝 기술의 응용 사례를 탐구하고 이를 분석 및 비교하는 것을 목표로 한다.

Introduction

합성곱 신경망(Convolutional Neural Network, CNN)은 야생동물의 종을 식별하고 수량을 통계적으로 분석하는 데 활용된다.

전통적인 야생동물 검출 방법인 인공 관찰이나 적외선 트리거 카메라는 시간 소모가 크고 정확도가 낮다.

야생동물 탐지 및 모니터링에 대한 딥러닝의 응용

(a)강력한 하드웨어 구조

(b)잘 다룰 수 있는 소프트웨어 알고리즘

(c)대량의 훈련 데이터

Overview Of Deep Learning Techniques

A.CNN in Object Detection:

You Only Look Once (YOLO) 및 Single shot detector (SSD)와 같은 분류 기반 방법.

가장 정확한 것은 Faster R-CNN이다.

YOLO는 관찰 대상의 다양한 구성 요소를 하나의 신경망으로 결합하는 방법이다

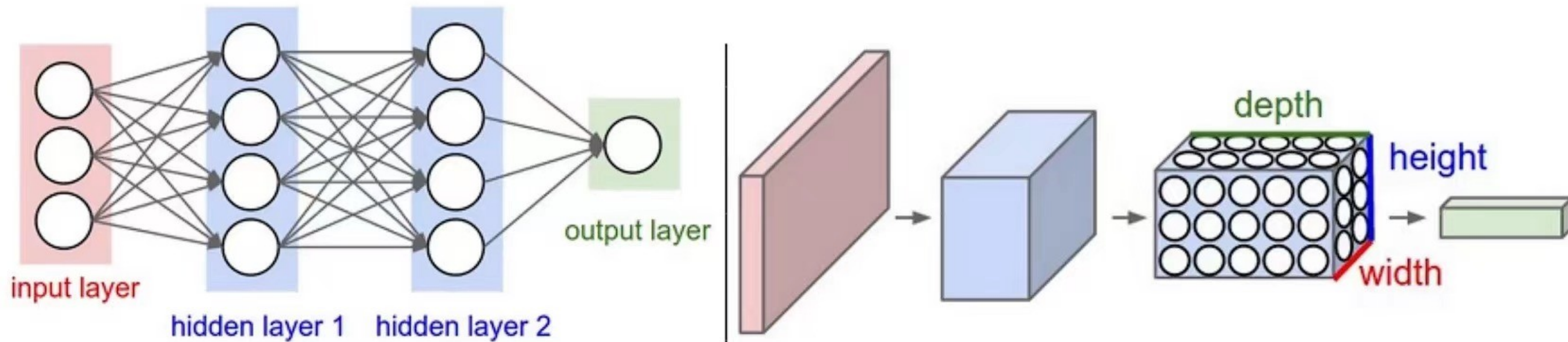
B.CNN Architecture and Models:

왼쪽 그림은 전통적인 신경망(NN)이고 오른쪽 그림은 컨볼루션 신경망(CNN)

예를 들어, 전통적인 신경망에서는 입력되는 그래프에 784개의 픽셀이 있기 때문에 입력층에는 784개의 뉴런이 존재한다.

반면, CNN에서는 원본 이미지가 $28 \times 28 \times 1$ 형태로 입력되며, 이는 3차원 매트릭스 형태이다.

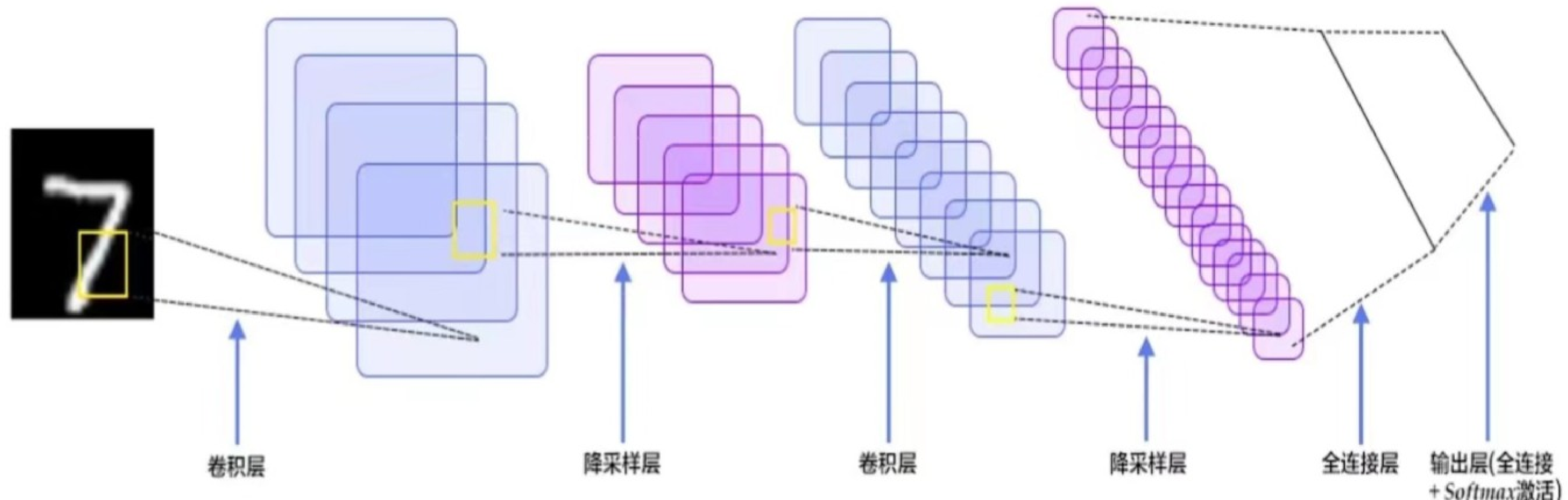
오른쪽 그림에서 3차원 이름 정의 'height*width*depth' 약칭 'h*w*d'.



CNN 프레임: 입력층, 컨볼루션층, 풀링층, 전체 연결층

(a) 입력층 :

맨 왼쪽 그림의 가정은 바로 우리가 입력한 것입니다. 가정은 " $28*8*3$ "이며, 각각 " $h*w*d$ "에 해당
일반적으로 RGB 3채널 형식으로 입력되므로 d 는 3이고 그림에서 두 번째 이미지는 뒷면의 3채널 이미지를 통해 겹쳐집.



(b)컨볼루션층:

입력은 5*5*1 크기의 이미지이고, 중간의 3*3*1 영역은 정의된 컨볼루션 코어이다.

원본 입력 이미지와 컨볼루션 코어를 이용해 연산을 수행하면 녹색 부분의 결과를 얻을 수 있다.

어두운 부분에서 중앙에 위치한 숫자는 이미지의 픽셀이며, 오른쪽 아래에 있는 숫자는 컨볼루션 코어의 값이다.이 숫자들을 곱하고 더한 결과가 최종적으로 도출된다.

' $3*0+1*1+2*2+2*2+0*2+0*0+2*0+0*1+0*2=9$ '

3	3	2	1	0
0	0	1	3	1
3_0	1_1	2_2	2	3
2_2	0_2	0_0	2	2
2_0	0_1	0_2	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

(c)풀링층 :

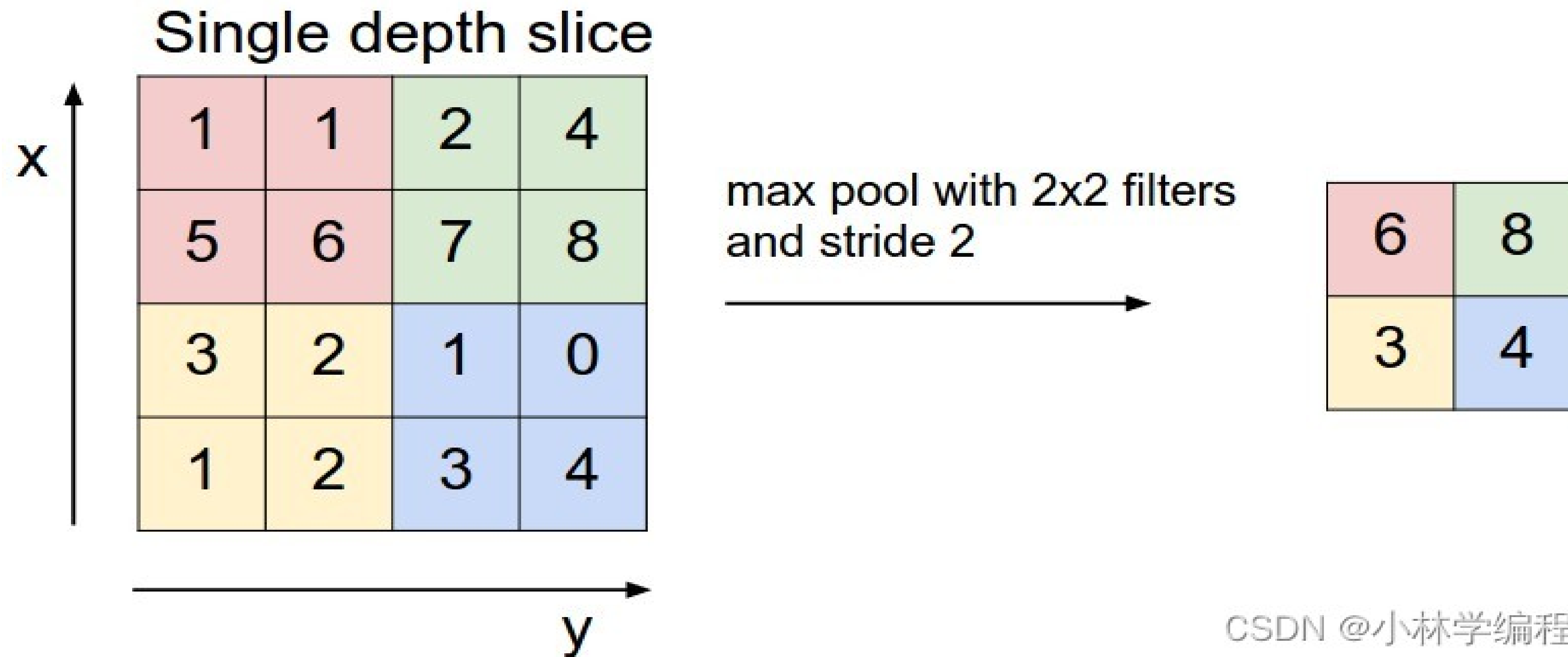
그림은 max-pooling (최대 풀링)

분홍색 영역에 있는 '1, 1, 5, 6' 네 개의 숫자 중 최대값은 6이다.

또 하나는 mean-pooling(평균 풀링)

해당 지역 내의 숫자들을 더한 후 평균을 구하면 된다.

블루존은 바로 $(1+0+3+4)/4=2$

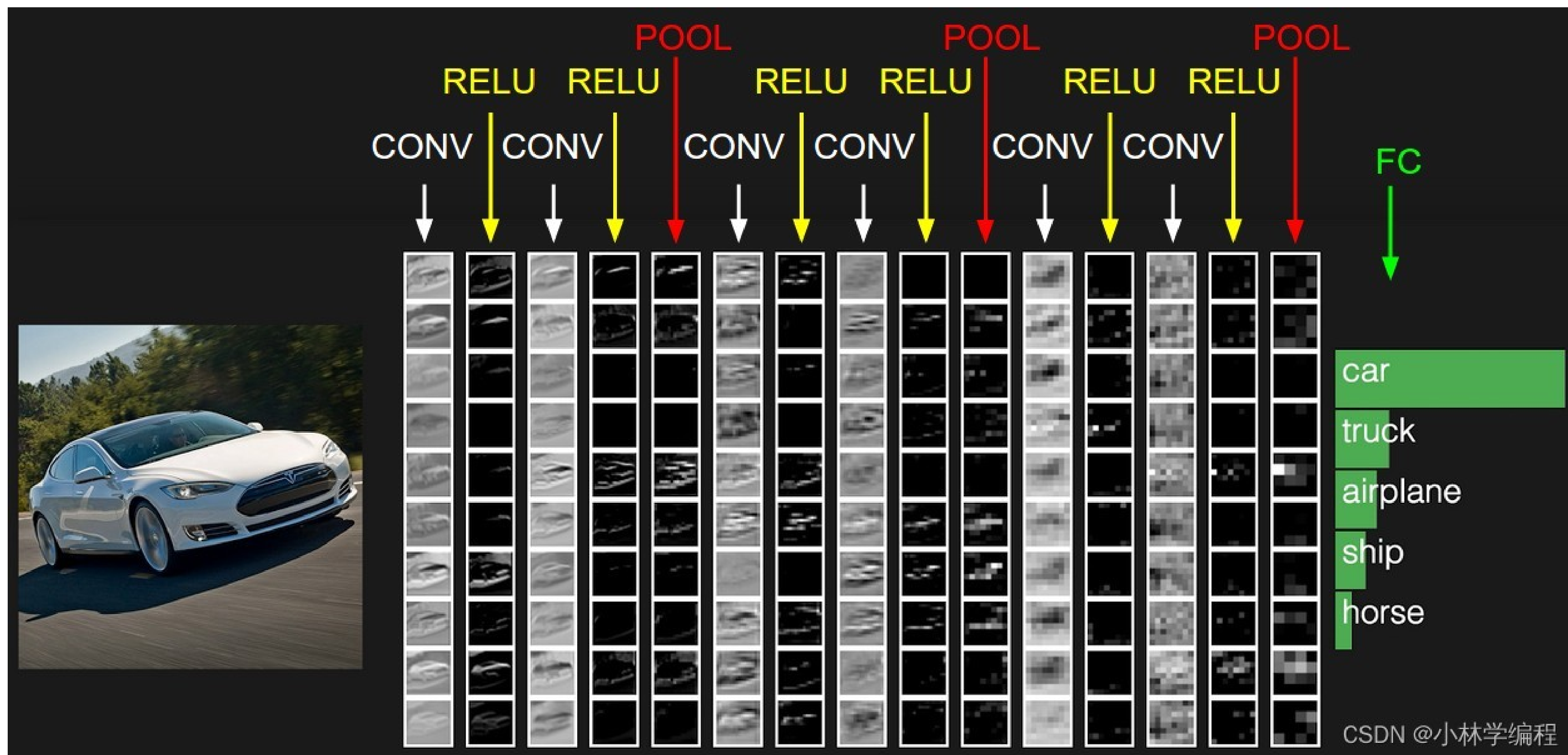


(d) 전체 연결층 (FC) :

전체 연결층은 마지막에 한 층 또는 여러 층의 전통적인 신경망(NN)층을 추가하는 것과 같다.

연결하기 전에 CNN의 3차원 매트릭스를 2차원으로 평평하게 해야 합니다.

예를 들어, 풀링 레이어에 크기가 $5 \times 5 \times 3$ 인 이미지가 입력된 경우, 풀링 연산을 통해 결과가 1×75 형태로 변환되어 풀링 레이어의 입력 조건을 충족하게 된다.



Wild Life Monitoring Environment And Data

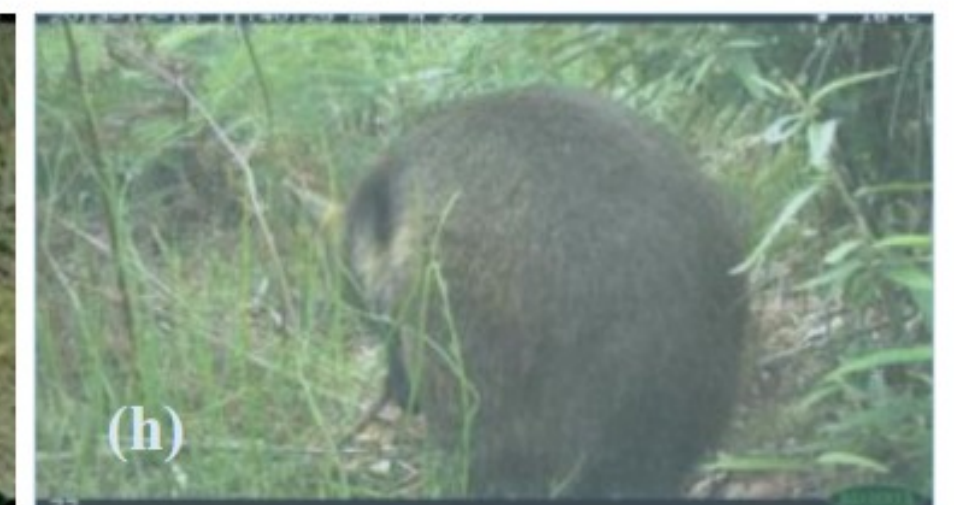
(a) 야생동물 모니터링 환경 특징:

서식지의 복잡성과 다양성

야생동물의 행위 복잡성

환경별 기후 변화와 지형 차이가 설비에 미치는 영향

(b) 카메라 트랩, 드론 등에 딥러닝 기술을 접목하여 효과:



Deep Learning In Wildlife Animal Detection And Recognition

Norouzzadeh 등은 심층 컨볼루션 신경망과 카메라 트랩을 사용하여 야생동물을 식별하고 계수했다.

그림에 포함된 종을 식별합니다

그림에 존재하는 종의 수를 결정합니다

이미지 속의 동물, 동물의 행동 특징을 묘사합니다.

모델의 정확도 향상

.

Evaluation And Discussion Of The Results

모형성능 비교

- (a)단일 단계 표적 탐지 모델(YOLO 및 SSD) :
속도는 빠르고 실시간 모니터링에 적합하지만, 복잡한 배경에서는 정확도가 낮다.
- (b)2단계 표적 탐지 모델(예: Faster R-CNN 및 Mask R-CNN):
검출 정확도는 높지만, 속도는 비교적 느리고 계산 요구가 상대적으로 높다.
- (c)딥러닝의 분류모델(예: ResNet, Inception):
대형 이미지 분류 작업에서 뛰어난 성능을 보여 동물 종을 정확하게 식별할 수 있음.

Conclusions

딥러닝 기술은 야생동물 탐지 및 식별에 매우 중요한 역할을 한다.

이를 통해 생태학 분야의 모니터링 효율이 크게 향상되었다.

딥러닝 모델을 활용하면 야생동물을 더욱 빠르고 정확하게 식별하고 감시할 수 있다.